

IBM SPSS Advanced Statistics 29

IBM

Poznámka

Než začnete používat tyto informace a produkt, který podporují, přečtěte si informace, které uvádí [“Poznámky” na stránce 133](#).

Informace o produktu

Toto vydání se vztahuje na verzi 29, vydání 0, úpravu 2 produktu IBM® SPSS Statistika a na všechna následná vydání a úpravy, dokud nebude v nových vydáních uvedeno jinak.

© **Copyright International Business Machines Corporation .**

Kapitola 1. Rozšířené statistiky.....	1
Úvod do pokročilé statistiky.....	1
Vícerozměrná analýza GLM.....	1
Vícevariantní model GLM.....	2
Vícerozměrné kontrasty GLM.....	4
GLM Multivariate profilové partie.....	4
Vícerozměrná posthoc porovnání GLM.....	5
GLM-Odhadované mezní střední hodnota.....	6
GLM-uložení.....	6
Volby vícevariantní správy GLM.....	7
Další funkce příkazu GLM.....	8
Opakovaná opatření GLM.....	9
Faktory definování opakovaných ukazatelů GLM.....	11
Model opakovaných ukazatelů GLM.....	11
GLM-opakovaná měření kontrastů.....	12
Grafy profilů opakovaných měření GLM.....	13
Porovnání následných opatření GLM-opakovaná.....	13
GLM-Odhadované mezní střední hodnota.....	14
Uložení opakovaných opatření GLM.....	15
Volby opakovaných ukazatelů GLM.....	16
Další funkce příkazu GLM.....	16
Analýza komponent odchylky.....	17
Model komponent odchylky.....	18
Volby komponent odchylky.....	19
Uložit komponenty odchylky do nového souboru.....	20
Další funkce příkazu VARCOMP.....	20
Lineární smíšené modely.....	20
Lineární smíšené modely: předměty a opakující se.....	21
Fixní efekty lineárních smíšených modelů.....	23
Náhodné efekty lineárních smíšených modelů.....	24
Odhad lineárních smíšených modelů.....	24
Statistika lineárních smíšených modelů.....	25
Lineární smíšené modely EM.....	26
Lineární smíšené modely Uložit.....	26
Lineární smíšené modely-Export.....	27
Další funkce příkazu MIXED.....	27
Zobecněné lineární modely.....	27
Zobecněná odezva lineárních modelů.....	30
Prediktory generalizovaných lineárních modelů.....	30
Zobecněný model lineárních modelů.....	31
Zobecněný odhad lineárních modelů.....	32
Statistika zobecněných lineárních modelů.....	33
Zobecněné lineární modely EM.....	34
Zobecněné uložení lineárních modelů.....	35
Export zobecněných lineárních modelů.....	36
Další funkce příkazu GENLIN.....	37
Zobecněné odhadovací rovnice.....	37
Typ modelu zobecněných odhadových rovnic.....	39
Odpověď na zobecněný odhad rovnic.....	40
Generalizované odhadování prediktorů rovnic.....	41
Model zobecněných odhadových rovnic.....	42

Zobecněný odhad rovnic.....	42
Statistika zobecněných odhadovaných rovnic.....	44
Zobecněné odhadování rovnic EM prostředky.....	45
Uložení generalizovaných odhadovaných rovnic.....	46
Export generalizovaných odhadových rovnic.....	46
Další funkce příkazu GENLIN.....	47
Zobecněné lineární smíšené modely.....	47
Získání zobecněného lineárního smíšeného modelu.....	49
Cíl	50
Pevné efekty	52
Náhodné efekty	53
Váha a odsazení	54
Obecné volby sestavení	54
Odhad	55
Odhadované prostředky	55
Uložit	56
Export	57
Zobrazení modelu	57
Výběr modelu Loglinear Analysis.....	60
Definice rozsahu logliniové analýzy.....	61
Model loglineární analýzy.....	61
Výběr modelu Loglinear Analysis Options.....	62
Další funkce příkazu HILOGLINEAR.....	62
Obecná logliniová analýza.....	62
Model obecné loglineární analýzy.....	63
Obecné volby logliniové analýzy.....	64
Uložení obecné logliniové analýzy.....	64
Dodatečné funkce příkazu GENLOG.....	65
Logit Loglineární analýza.....	65
Logit Loglinear Analysis Model (lineární analýza logit).....	66
Volby Logit Loglinear Analysis.....	67
Logit Loglinear Analysis Uložit.....	67
Dodatečné funkce příkazu GENLOG.....	68
Životní tabulky.....	68
Rozsah definice tabulek životnosti.....	69
Volby životních tabulek.....	69
Další funkce příkazu SURVIVAL.....	70
Parametrické časové modely urychleného selhání.....	70
Parametrické časové modely urychleného selhání: Kritéria.....	71
Parametrické akcelerované časové modely selhání: Model.....	71
Parametrické časové modely urychleného selhání: odhad.....	72
Parametrické časové modely urychleného selhání: Tisk.....	73
Parametrické časové modely akcelerovaného selhání: Předpovídat.....	74
Parametrické akcelerované časové modely selhání: zakreslení.....	74
Parametrické akcelerované časové modely selhání: Export.....	75
Survival AFT Definovat události pro proměnné stavu.....	75
Parametrické časové modely akcelerovaného selhání: Vybrat kategorii.....	75
Parametrické sdílené modely křehkosti.....	76
Parametrické sdílené modely křehkosti: kritéria.....	77
Parametrické sdílené modely křehkosti: Model.....	78
Parametrické sdílené modely křehkosti: odhad.....	79
Parametrické sdílené modely křehkosti: tisk.....	80
Parametrické sdílené modely Frailty: Predikce.....	80
Parametrické sdílené modely křehkosti: Plot.....	81
Parametrické sdílené modely křehkosti: Export.....	82
Parametrické sdílené modely křehkosti: definování událostí.....	82
Parametrické sdílené modely křehkosti-příklady.....	83
Parametrické sdílené modely křehkosti-případová studie pro rekurní data.....	85

Kaplan-Meier analýza přežití.....	87
Kaplan-Meier Definovat událost pro proměnnou stavu.....	88
Kaplan-Meier Porovnat úrovně faktoru.....	88
Kaplan-Meier Uložit nové proměnné.....	88
Možnosti společnosti Kaplan-Meier.....	88
Další funkce příkazu KM.....	89
Regresní analýza Cox.....	89
Cox regression-definice kategorických proměnných.....	90
Regresní grafy Cox.....	90
Cox regression Uložit nové proměnné.....	91
Volby regrese Cox.....	91
Událost definování regrese Cox pro proměnnou stavu.....	92
Další funkce příkazu COXREG.....	92
Výpočet časově závislých krytů.....	92
Výpočet časově závislého Covariate.....	93
Kategoriální schémata kódování proměnných.....	93
Odchylka.....	93
Jednoduché.....	93
Pomocník.....	94
Rozdíl.....	94
Polynomní.....	94
Opakováno.....	95
Speciální.....	95
Indikátor.....	96
Kovarianční struktury.....	96
Bayesovská statistika.....	100
Bayesian jeden vzorek odvozování: Normální.....	101
Bayesian jeden vzorek Inference: Binomické.....	105
Bayesian jeden vzorek Inference: Poisson.....	107
Bayesian související ukázkové odvozování: Normální.....	109
Bayesian Independent-Vzorové odvozování.....	110
Bayesian Inference o Pearsonově korelaci.....	114
Bayesian Inference o lineárních regresních modelech.....	116
Bayesovská jednosměrná ANOVA.....	121
Bayesovské loglineární modely.....	124
Bayesian jednosměrná opakovaná opatření AMNOVA modely.....	127
Regrese hřebenu jádra.....	129
parametry jádra.....	130
Regrese hřebenu jádra: Volby.....	131
Poznámky.....	133
Ochranné známky.....	134
Rejstřík.....	137

Kapitola 1. Rozšířené statistiky

Následující rozšířené statistické funkce jsou zahrnuty v části SPSS Statistika Standard Edition nebo volba Rozšířená statistika.

Úvod do pokročilé statistiky

Volba Rozšířená statistika obsahuje procedury, které nabízejí pokročilejší možnosti modelování, než jsou k dispozici prostřednictvím volby Statistická báze.

- GLM Multivariate rozšiřuje obecný lineární model poskytovaný GLM Univariate tak, aby umožňoval více závislých proměnných. Další rozšíření, GLM Repeated Measures, umožňuje opakované měření více závislých proměnných.
- Analýza komponent rozptylu je specifický nástroj pro rozkládání variability v závislé proměnné na pevné a náhodné komponenty.
- Lineární smíšené modely rozšiřují obecný lineární model tak, aby data mohla vykazovat korelované a nekonstantní variabilitu. Smíšený lineární model proto poskytuje flexibilitu modelování nejen prostředků dat, ale i rozptylů a kovariancí.
- Zobecněné lineární modely (GZLM) uvolňuje předpoklad normality pro chybový výraz a vyžaduje pouze to, aby závislá proměnná byla lineárně závislá na prediktorech prostřednictvím transformace nebo funkce odkazu. Zobecněný odhad rovnic (GEE) rozšiřuje GZLM o možnost opakovaného měření.
- Obecná logliniová analýza vám umožňuje přizpůsobit modely pro křížově klasifikovaná data počtu a Logliniová analýza výběru modelu vám může pomoci vybrat si mezi modely.
- Logit Loglinear Analysis vám umožňuje přizpůsobit logliniové modely pro analýzu vztahu mezi kategoriálním závislým a jedním nebo více kategoriálními prediktory.
- Analýza přežití je k dispozici prostřednictvím Life Tables pro zkoumání rozdělení proměnných času na události, případně podle úrovní proměnné faktoru; Kaplan-Meier pro analýzu přežití pro zkoumání rozdělení proměnných času na událost, případně podle úrovní proměnné faktoru nebo pro vytváření samostatných analýz podle úrovní proměnné stratifikace; a Cox Regression pro modelování času na určitou událost na základě hodnot daných kovariátů.

Vícerozměrná analýza GLM

Procedura GLM Multivariate poskytuje regresní analýzu a analýzu rozptylu pro více závislých proměnných pomocí jedné nebo více faktorovými proměnnými nebo kovariáty. Faktorové proměnné rozdělují populaci do skupin. Pomocí této obecné procedury lineárního modelu můžete testovat nulové hypotézy o účincích faktorových proměnných na prostředky různých seskupení společného rozdělení závislých proměnných. Můžete zkoumat interakce mezi faktory, stejně jako účinky jednotlivých faktorů. Kromě toho mohou být zahrnuty účinky kovariátů a kovariantní interakce s faktory. Pro regresní analýzu jsou nezávislé proměnné (prediktor) určeny jako kovariáty.

Lze testovat vyvážené i nevyvážené modely. Návrh je vyvážený, pokud každá buňka v modelu obsahuje stejný počet případů. V vícerozměrném modelu jsou součty čtverců v důsledku účinků v modelu a chybové součty čtverců v maticové formě spíše než skalární forma nalezená v jednorozměrné analýze. Tyto matice se nazývají SSCP (součty čtverců a křížových produktů) matice. Je-li zadána více než jedna závislá proměnná, je poskytnuta vícerozměrná analýza rozptylu pomocí Pillaiho stopy, Wilksovy lambdy, Hotellingova trasování a Royova největšího kořenového kritéria s přibližnou statistikou F , stejně jako jednorozměrná analýza rozptylu pro každou závislou proměnnou. Kromě testování hypotéz GLM Multivariate vytváří odhady parametrů.

Běžně používané *a priori* kontrasty jsou k dispozici pro provádění testování hypotéz. Poté, co celkový test F zobrazí významnost, můžete pomocí post hoc testů vyhodnotit rozdíly mezi specifickými prostředky. Odhadnuté mezní střední hodnoty poskytují odhady předpokládaných středních hodnot pro buňky

v modelu a profilové grafy (grafy interakce) těchto prostředků vám umožňují snadno vizualizovat některé vztahy. Post hoc vícenásobné porovnávací testy se provádějí pro každou závislou proměnnou zvlášť.

Rezidua, predikované hodnoty, Cook's distance a pákové hodnoty lze uložit jako nové proměnné ve vašem datovém souboru pro kontrolu předpokladů. K dispozici je také zbytková matice SSCP, což je čtvercová matice součtů čtverců a křížových produktů zbytků, zbytková kovarianční matice, která je zbytková matice SSCP dělená stupněm volnosti zbytkových chyb, a zbytková korelační matice, která je standardizovanou formou zbytkové kovarianční matice.

Váha WLS vám umožňuje určit proměnnou, která se používá k pozorování různých vah pro analýzu vážených nejmenších čtverců (WLS), případně kompenzovat různou přesnost měření.

Příklad. Výrobce plastů měří tři vlastnosti plastového filmu: odolnost proti roztržení, lesk a zákal. Zkouší se dvě rychlosti vytlačování a dvě různá množství aditiva a tři vlastnosti se měří podle každé kombinace rychlosti vytlačování a množství aditiva. Výrobce zjistí, že rychlost vytlačování a množství aditiva jednotlivě produkuje významné výsledky, ale že interakce těchto dvou faktorů není významná.

Metody. K vyhodnocení různých hypotéz lze použít součty čtverců typu I, typu II, typu III a typu IV. Typ III je výchozí.

Statistika. Post hoc testy rozsahu a vícenásobná srovnání: nejméně významný rozdíl, Bonferroni, Sidak, Scheffé, Ryan-Einot-Gabriel-Welsch více F , Ryan-Einot-Gabriel-Welsch více rozsahů, Student-Newman-Keuls, Tukey je upřímně významný rozdíl, Tukey je b , Duncan, Hochberg je GT_2 , Gabriel, Waller Duncan t test, Dunnett (jednostranný a oboustranný), Tamhane's T_2 , a Dunnettovy C . Deskriptivní statistika: pozorované střední hodnoty, směrodatné odchylky a počty všech závislých proměnných ve všech buňkách; Levenův test homogenity rozptylu; Box M test homogenity kovariančních matic závislých proměnných a Bartlettův test sféricity.

Záloty. Spread-versus-level, reziduum a profil (interakce).

Aspekty vícerozměrných dat GLM

Data. Závislé proměnné by měly být kvantitativní. Faktory jsou kategorické a mohou mít číselné hodnoty nebo řetězcové hodnoty. Kovariáty jsou kvantitativní proměnné, které souvisí se závislou proměnnou.

Předpoklady. Pro závislé proměnné jsou data náhodným vzorkem vektorů z vícerozměrné normální populace; v populaci jsou matice rozptylu a kovariance pro všechny buňky stejné. Analýza rozptylu je robustní vůči odchylce od normality, i když data by měla být symetrická. Chcete-li zkontrolovat hypotézy, můžete použít homogenitu testů variancí (včetně produktu Box M) a grafů typu spread-versus-level. Můžete také zkontrolovat zbytkové chyby a zbytkové pozemky.

Související procedury. Před provedením analýzy rozptylu prozkoumejte data pomocí procedury Prozkoumání. Pro jednu závislou proměnnou použijte GLM Univariate. Pokud jste při několika příležitostech měřili stejné závislé proměnné pro každý předmět, použijte měřítka GLM Repeated Measures.

Získání vícerozměrných tabulek GLM

1. Z nabídek vyberte:

Analyzovat > Obecný lineární model > Vícevariantní ...

2. Vyberte alespoň dvě závislé proměnné.

Volitelně můžete uvést pevné faktory, Covariate (s) a WLS Weight.

Vícevariantní model GLM

Určit model. Úplný faktoriální model obsahuje všechny hlavní efekty faktoru, všechny hlavní kovariantní účinky a všechny interakce faktoru podle faktoru. Neobsahuje kovariantní interakce. Volbu **Vlastní** vyberte, chcete-li určit pouze podmnožinu interakcí nebo interakce faktoru podle kovariátu. Musíte označit všechny podmínky, které mají být zahrnuty do modelu.

Faktory a kryty. Jsou uvedeny faktory a kovariáty.

Model. Model závisí na povaze vašich dat. Po výběru volby **Vlastní** můžete vybrat hlavní efekty a interakce, které vás zajímají v analýze.

Součet čtverců. Způsob výpočtu součtů čtverců. U vyvážených nebo nevyvážených modelů bez chybějících buněk se nejčastěji používá metoda součtu čtverců typu III.

Zahrnout zachycení v modelu. Zachycení je obvykle zahrnuto v modelu. Pokud můžete předpokládat, že data procházejí původem, můžete toto zachycení vyloučit.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Součet čtverců

Pro model si můžete vybrat typ součtů čtverců. Typ III je nejčastěji používaný a je výchozí.

Typ I. Tato metoda je také známá jako hierarchická dekompozice metody součtu čtverců. Každý termín je upraven pouze pro termín, který jej předchází v modelu. Typ I součty čtverců se běžně používají pro:

- Vyvážený model ANOVA, ve kterém jsou všechny hlavní účinky specifikovány před účinky interakce prvního řádu, jakékoli účinky interakce prvního řádu jsou specifikovány před účinky interakce druhého řádu atd.
- Polynomiální regresní model, v němž jsou všechny výrazy nižšího řádu specifikovány před výrazy vyššího řádu.
- Čistě vnořený model, ve kterém je první zadaný efekt vnořen do druhého zadaného efektu, druhý zadaný efekt je vnořen do třetího a tak dále. (Tuto formu vnoření lze zadat pouze pomocí syntaxe.)

Typ II. Tato metoda vypočítá součty čtverců efektu v modelu upravené pro všechny ostatní "vhodné" efekty. Vhodným účinkem je účinek, který odpovídá všem účinkům, které neobsahují zkoumaný účinek. Metoda součtu čtverců typu II se běžně používá pro:

- Vyvážený model ANOVA.
- Jakýkoli model, který má pouze hlavní faktory.
- Jakýkoli regresní model.
- Čistě vnořený návrh. (Tuto formu vnoření lze určit pomocí syntaxe.)

Typ III. Výchozí hodnota Tato metoda vypočítá součty čtverců efektu v designu jako součty čtverců, upravené pro všechny ostatní účinky, které neobsahují efekt, a ortogonální na jakékoliv účinky (pokud existují), které obsahují účinek. Typ III částky čtverců mají jednu hlavní výhodu v tom, že jsou invariantní s ohledem na buněčné frekvence tak dlouho, jak je obecná forma odhadu zůstává konstantní. Proto je tento typ součtů čtverců často považován za užitečný pro nevyvážený model bez chybějících buněk. V faktoriálním návrhu bez chybějících buněk je tato metoda ekvivalentní technice Yates 'weighted-of-means' (vážené čtverce středních hodnot). Metoda součtu čtverců typu III se běžně používá pro:

- Všechny modely uvedené v typech I a II.
- Jakýkoli vyvážený nebo nevyvážený model bez prázdných buněk.

Typ IV. Tato metoda je určena pro situaci, kdy chybí buňky. Pro jakýkoli efekt F v návrhu platí, že pokud F není obsažen v žádném jiném efektu, pak typ IV = Typ III = Typ II. Je-li hodnota F obsažena v jiných efektech, typ IV rozdělí rozdíly mezi parametry v F na všechny efekty vyšší úrovně spravedlivě. Metoda součtu čtverců typu IV se běžně používá pro:

- Všechny modely uvedené v typech I a II.
- Jakýkoli vyvážený model nebo nevyvážený model s prázdnými buňkami.

Vícerozměrné kontrasty GLM

Kontrasty se používají k testování, zda se úrovně účinku od sebe výrazně liší. Můžete určit kontrast pro každý faktor v modelu. Kontrasty představují lineární kombinace parametrů.

Testování hypotéz je založeno na nulové hypotéze $\mathbf{LBM} = \mathbf{0}$, kde \mathbf{L} je matice koeficientů kontrastu, \mathbf{M} je matice identity (která má dimenzi rovnající se počtu závislých proměnných) a \mathbf{B} je vektor parametru. Je-li zadán kontrast, vytvoří se matice \mathbf{L} tak, aby se sloupce odpovídající danému faktoru shodovaly s kontrastem. Zbývající sloupce jsou upraveny tak, aby bylo možné odhadnout matici \mathbf{L} .

Kromě jednorozměrného testu používajícího statistiku F a souběžné intervaly spolehlivosti typu Bonferroni založené na Studentově t distribuci pro rozdíly kontrastu mezi všemi závislými proměnnými jsou poskytnuty vícerozměrné testy využívající Pillaiho trasování, Wilksovu lambda, Hotellingovu stopu a Royova největší základní kritéria.

Dostupné kontrasty jsou odchylka, jednoduché, rozdíl, Helmert, opakované a polynom. Pro odchylky kontrastů a jednoduché kontrasty můžete zvolit, zda je referenční kategorie poslední nebo první kategorie.

Typy kontrastů

Odchylka. Porovná střední hodnotu každé úrovně (s výjimkou referenční kategorie) se střední hodnotou všech úrovní (grand mean). Úrovně faktoru mohou být v libovolném pořadí.

Jednoduché. Porovná střední hodnotu každé úrovně se střední hodnotou určené úrovně. Tento typ kontrastu je užitečný v případě, že existuje řídicí skupina. Jako odkaz můžete zvolit první nebo poslední kategorii.

Rozdíl. Porovná střední hodnotu každé úrovně (kromě první) se střední hodnotou předchozích úrovní. (Někdy nazývané reverzní Helmert kontrasty.)

Helmert. Porovná střední hodnotu každé úrovně faktoru (s výjimkou poslední úrovně) se střední hodnotou následných úrovní.

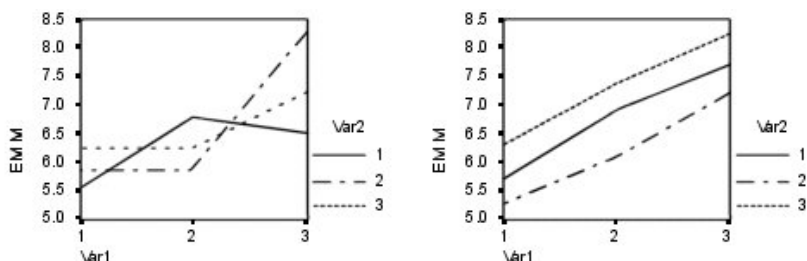
Opakovat. Porovná střední hodnotu každé úrovně (kromě poslední) se střední hodnotou následující úrovně.

Polynomiální. Porovnává lineární efekt, kvadratický efekt, kubický efekt a tak dále. První stupeň volnosti obsahuje lineární efekt ve všech kategoriích; druhý stupeň volnosti, kvadratický efekt; a tak dále. Tyto kontrasty se často používají k odhadu polynomiálních trendů.

GLM Multivariate profilové partie

Profilové grafy (interakční grafy) jsou užitečné pro porovnání mezních hodnot ve vašem modelu. Profilový graf je čárový graf, ve kterém každý bod označuje odhadovaný mezní průměr závislé proměnné (upravený pro všechny kovariáty) na jedné úrovni faktoru. Úrovně druhého faktoru lze použít k vytvoření oddělených čar. Každou úroveň ve třetím faktoru lze použít k vytvoření samostatného výkresu. Pro pozemky jsou k dispozici všechny faktory. Pro každou závislou proměnnou jsou vytvořeny profilové grafy.

Profilový graf s jedním faktorem ukazuje, zda se odhadované mezní střední hodnoty zvyšují nebo snižují napříč úrovněmi. Pro dva nebo více faktorů paralelní čáry označují, že neexistuje žádná interakce mezi faktory, což znamená, že můžete zkoumat úrovně pouze jednoho faktoru. Neparalelní čáry označují interakci.



Obrázek 1. Neparalelní vykreslení (vlevo) a paralelní vykreslení (vpravo)

Po zadání výkresu výběrem faktorů pro vodorovnou osu a volitelně faktorů pro samostatné čáry a samostatné grafy musí být graf přidán do seznamu Zákresy.

Vícerozměrná posthoc porovnání GLM

Post hoc vícenásobné porovnávací testy. Jakmile zjistíte, že existují rozdíly mezi prostředky, mohou post hoc testy rozsahu a párování více porovnání určit, které střední hodnoty se liší. Porovnání se provádí na neupravených hodnotách. Post hoc testy se provádějí pro každou závislou proměnnou zvlášť.

Bonferroniho a Tukeyho čestně významné rozdílové testy se běžně používají v několika srovnávacích testech. **Bonferroni test**, založený na Studentově statistice t , upravuje pozorovanou úroveň významnosti pro skutečnost, že bylo provedeno více porovnání. **Sidakův t test** také upravuje úroveň významnosti a poskytuje přísnější meze než Bonferroniho test. **Test upřímně významného rozdílu společnosti Tukey** používá statistiku rozsahu Studentized k provedení všech porovnání po dvojicích mezi skupinami a k nastavení míry chyb experimentálně na poměr chyb pro kolekci pro všechna porovnání po dvojicích. Při testování velkého počtu párů prostředků, Tukey je upřímně významný rozdíl test je silnější než Bonferroni test. Pro malý počet párů je Bonferroni silnější.

Hochbergův GT2 je podobný testu Tukeyho upřímně významného rozdílu, ale používá se Studentizovaný maximální modul. Obvykle je Tukeyho zkouška silnější. **Gabrielův porovnávací test po dvojicích** také používá Studentizovaný maximální modul a je obecně výkonnější než Hochbergův GT2, když jsou velikosti buněk nestejně. Gabrielův test se může stát liberálním, když se velikost buněk značně liší.

Dunnettův párování vícenásobný porovnávací t test porovnává sadu ošetření s jedinou střední hodnotou kontroly. Poslední kategorie je výchozí řídicí kategorie. Případně můžete zvolit první kategorii. Můžete si také vybrat oboustranný nebo jednostranný test. Chcete-li ověřit, že střední hodnota faktoru na jakékoli úrovni (s výjimkou kontrolní kategorie) se nerovná střední hodnotě kontrolní kategorie, použijte oboustrannou zkoušku. Chcete-li otestovat, zda je střední hodnota na libovolné úrovni faktoru menší než střední hodnota řídicí kategorie, vyberte volbu **< Řídicí prvek**. Podobně, chcete-li otestovat, zda je střední hodnota na libovolné úrovni faktoru větší než střední hodnota řídicí kategorie, vyberte volbu **> Řídicí prvek**.

Ryan, Einot, Gabriel a Welsch (R-E-G-W) vyvinuli dva testy s více stupňovitými rozsahy. Několik postupných kroků nejprve otestujte, zda jsou všechny prostředky stejné. Pokud všechny prostředky nejsou stejné, podskupiny prostředků jsou testovány na rovnost. **R-E-G-W F** je založeno na testu F a **R-E-G-W Q** je založeno na Studentizovaném rozsahu. Tyto testy jsou výkonnější než Duncanův test s vícenásobným rozsahem a Student-Newman-Keuls (což jsou také vícenásobné postupy krok dolů), ale nejsou doporučeny pro nerovné velikosti buněk.

Pokud jsou odchylky nestejně, použijte **Tamhanův T2** (konzervativní porovnávací test po dvojicích založený na t testu), **Dunnettův T3** (porovnávací test po dvojicích založený na Studentizovaném maximálním modulu), **Games-Howell porovnávací test po dvojicích** (někdy liberální) nebo **Dunnettův C** (porovnávací test po dvojicích založený na Studentizovaném rozsahu).

Duncan's multiple range test, Student-Newman-Keuls (**S-N-K**) a **Tukey's b** jsou testy rozsahu, které hodnotí střední hodnotu skupiny a vypočítávají hodnotu rozsahu. Tyto testy se nepoužívají tak často, jak se dříve diskutovalo.

Test Waller-Duncan t používá bayesovský přístup. Tato zkouška rozsahu používá harmonický průměr velikosti vzorku, pokud jsou velikosti vzorku nerovné.

Úroveň významnosti testu **Scheffé** je navržena tak, aby umožňovala testování všech možných lineárních kombinací středních hodnot skupiny, nikoli pouze porovnání po dvojicích, která jsou v této funkci k dispozici. Výsledkem je, že Scheffé test je často konzervativnější než jiné testy, což znamená, že pro významnost je nutný větší rozdíl mezi prostředky.

Nejméně významný rozdíl (**LSD**) porovnávacího testu po dvojicích je ekvivalentní několika individuálním testům t mezi všemi dvojicemi skupin. Nevýhodou tohoto testu je, že nebyl učiněn žádný pokus o úpravu pozorované úrovně významnosti pro vícenásobná porovnání.

Zobrazené testy. Porovnání po dvojicích jsou poskytována pro LSD, Sidak, Bonferroni, Games-Howell, Tamhane's T2 a T3, Dunnett's Ca Dunnett's T3. Homogenní dílčí sady pro testy rozsahu jsou poskytovány

pro S-N-K, Tukeyho b , Duncan, R-E-G-W F , R-E-G-W Qa Waller. Tukey je upřímně významný rozdíl test, Hochberg GT2, Gabriel je test, a Scheffé je test je jak více srovnávacích testů a testy rozsahu.

GLM-Odhadované mezní střední hodnota

Vyberte faktory a interakce, pro které chcete odhadnout mezní střední hodnotu populace v buňkách. Tyto prostředky jsou upraveny pro případné kovariáty.

Porovnat hlavní efekty

Poskytuje neopravené porovnávání po dvojicích mezi odhadovanými mezními hodnotami pro jakýkoli hlavní vliv v modelu, a to jak pro faktory mezi jednotlivými subjekty, tak pro faktory v rámci jednotlivých subjektů. Tato položka je k dispozici pouze v případě, že jsou v seznamu Zobrazit prostředky pro vybraný hlavní efekty.

Porovnat jednoduché hlavní efekty

Nastavení je povoleno vždy, když cílový seznam obsahuje jeden nebo více efektů produktu nebo interakce (například $A \times B$, $A \times B \times C$). Nastavení podporuje specifikaci porovnání mezi jednoduchými hlavními efekty, což jsou hlavní efekty vnořené do úrovní jiných faktorů.

Nastavení konfidenčního intervalu

Vyberte nejméně významný rozdíl (LSD), Bonferroni nebo Sidak nastavení intervalů spolehlivosti a významnosti. Tato položka je k dispozici pouze v případě, že je vybrána volba **Porovnat hlavní efekty** a/nebo **Porovnat jednoduché hlavní efekty**.

Určení odhadovaných mezních prostředků

1. Z nabídek vyberte jednu z procedur dostupných v části > **Analyzovat** > **Obecný lineární model**.
2. V hlavním dialogovém okně klepněte na volbu **EM Means**.

GLM-uložení

Hodnoty předpovězené modelem, reziduály a souvisejícími ukazateli můžete uložit jako nové proměnné v editoru dat. Mnoho z těchto proměnných lze použít pro zkoumání předpokladů o datech. Chcete-li uložit hodnoty pro použití v jiné relaci IBM SPSS Statistika, musíte uložit aktuální datový soubor.

Předpokládané hodnoty. Hodnoty, které model předpovídá pro každý případ.

- *Nestandardizovaný.* Hodnota, kterou model předpovídá pro závislou proměnnou.
- *Vážený.* Vážené nestandardizované předpovězené hodnoty. K dispozici pouze v případě, že byla dříve vybrána proměnná WLS.
- *Standardní chyba.* Odhad směrodatné odchylky průměrné hodnoty závislé proměnné pro případy, které mají stejné hodnoty nezávislých proměnných.

Diagnostika. Ukazatele pro identifikaci případů s neobvyklými kombinacemi hodnot pro nezávislé proměnné a případy, které mohou mít velký dopad na model.

- *Vzdálenost od kuchaře.* Měřítka toho, kolik by se změnilo rezidua všech případů, kdyby byl konkrétní případ vyloučen z výpočtu regresních koeficientů. Velký Cook's D označuje, že vyloučení případu z výpočtu regresní statistiky podstatně změní koeficienty.
- *Hodnoty pákového efektu.* Nesoustředěné hodnoty pákového efektu. Relativní vliv každého pozorování na vhodnost modelu.

Zbytky. Nestandardizovaný zbytek je skutečná hodnota závislé proměnné minus hodnota předpovězená modelem. K dispozici jsou také standardizované, studentizované a smazané rezidenční chyby. Pokud byla zvolena proměnná WLS, jsou k dispozici vážené nestandardizované zbytkové chyby.

- *Nestandardizovaný.* Rozdíl mezi pozorovanou hodnotou a hodnotou předpovězenou modelem.
- *Vážený.* Vážené nestandardizované zbytkové chyby. K dispozici pouze v případě, že byla dříve vybrána proměnná WLS.

- *Standardizované*. Zbytek vydělený odhadem směrodatné odchylky. Standardizované zbytkové chyby, které jsou také známé jako Pearsonovy rezidenční chyby, mají střední hodnotu 0 a směrodatnou odchylku 1.
- *Studovaný*. Zbytek děleno odhadem jeho směrodatné odchylky, která se liší případ od případu, v závislosti na vzdálenosti hodnot jednotlivých případů na nezávislých proměnných od prostředků nezávislých proměnných. Někdy označované jako vnitřně studentizované zbytkové chyby.
- *Odstraněná*. Zbytek pro případ, kdy je tento případ vyloučen z výpočtu regresních koeficientů. Jedná se o rozdíl mezi hodnotou závislé proměnné a upravenou předpovězenou hodnotou.

Statistika koeficientů. Zapiše matici proměnlivosti odhadů parametrů v modelu do nové datové sady v aktuální relaci nebo do externího datového souboru IBM SPSS Statistika . Také pro každou závislou proměnnou bude řádek odhadů parametrů, řádek standardních chyb odhadů parametrů, řádek hodnot významnosti pro statistiku t odpovídající odhadům parametrů a řádek zbytkových stupňů volnosti. Pro vícerozměrný model existují podobné řádky pro každou závislou proměnnou. Je-li vybrána statistika konzistentní s heteroskedasticitou (k dispozici pouze pro jednorozměrné modely), je matice rozptylu a kovariance vypočtena pomocí robustního odhadu, řádek standardních chyb zobrazuje robustní standardní chyby a hodnoty významnosti odrážejí robustní chyby. Tento soubor matice můžete použít v jiných procedurách, které čtou soubory matice.

Volby vícevariantní správy GLM

V tomto dialogovém okně jsou k dispozici volitelné statistiky. Statistiky se vypočítávají pomocí modelu s pevnými efekty.

Zobrazit. Vyberte volbu **Popisná statistika** , chcete-li vytvářet pozorované střední hodnoty, směrodatné odchylky a počty pro všechny závislé proměnné ve všech buňkách. **Odhady velikosti efektu** poskytují částečnou hodnotu eta kvadrát pro každý efekt a každý odhad parametru. Statistika eta-kvadrát popisuje podíl celkové variability, který lze přisoudit faktorů. Vyberte volbu **Sledovaná síla** , chcete-li získat sílu testu, když je nastavena alternativní hypotéza na základě pozorované hodnoty. Vyberte volbu **odhady parametrů** , chcete-li vytvořit odhady parametrů, standardní chyby, t testy, intervaly spolehlivosti a pozorovanou sílu pro každý test. Můžete zobrazit hypotézu a chybu **matice SSCP** a **Matici zbytkového SSCP** plus Bartlettův test sféricity matice zbytkového kovariance.

Testy homogenity vytvoří test Levene homogenity rozptylu pro každou závislou proměnnou napříč všemi kombinacemi úrovní faktorů mezi subjekty, pouze pro faktory mezi subjekty. Testy homogenity také zahrnují Box M test homogenity matic kovariance závislých proměnných napříč všemi kombinacemi úrovní faktorů mezi subjekty. Volby rozložení versus úroveň a zbytkové grafy jsou užitečné pro kontrolu předpokladů o datech. Tato položka je zakázána, pokud neexistují žádné faktory. Vyberte volbu **Zbytkové grafy** , chcete-li vytvořit pozorovaný-podle-předpovězených-standardizovaných zbytkových chyb pro každou závislou proměnnou. Tyto grafy jsou užitečné pro zkoumání předpokladu stejné odchylky. Vyberte volbu **Nedostatek vhodných testů** , chcete-li zkontrolovat, zda lze vztah mezi závislou proměnnou a nezávislými proměnnými odpovídajícím způsobem popsat modelem. **Obecné odhadnutelné funkce** vám umožňuje vytvářet vlastní hypotetické testy na základě obecných odhadovatelných funkcí. Řádky v jakékoliv matici koeficientu kontrastu jsou lineární kombinace obecně odhadnutelné funkce (funkcí).

Zobrazit

Deskriptivní statistika

Vytváří pozorované střední hodnoty, směrodatné odchylky a počty pro všechny závislé proměnné ve všech buňkách.

Odhady velikosti účinku

Poskytuje částečnou hodnotu eta-kvadrát pro každý efekt a každý odhad parametru. Statistika eta-kvadrát popisuje podíl celkové variability, který lze přisoudit faktorů.

Pozorovaný výkon

Získá sílu testu, když je alternativní hypotéza nastavena na základě pozorované hodnoty.

Odhady parametrů

Vytváří odhady parametrů, standardní chyby, t testy, intervaly spolehlivosti a pozorovanou sílu pro každý test.

Matice SSCP

Zobrazí hypotézu a matice SSCP chyb.

Zbytková matice SSCP

Zobrazuje hypotézu a chybovou zbytkovou matici SSCP.

Transformační matice

Zobrazuje Bartlettův test sféricity matice zbytkové kovariance.

Zkoušky homogenity

Vytvoří Levenův test homogenity rozptylu pro každou závislou proměnnou napříč všemi kombinacemi faktorů mezi subjekty, pouze pro faktory mezi subjekty. Testy homogenity také zahrnují Box *M* test homogenity matic kovariance závislých proměnných napříč všemi kombinacemi úrovní faktorů mezi subjekty.

Rozložení vs. úroňový graf

Užitečné pro kontrolu předpokladů o datech pro zkoumání předpokladu stejné odchylky. Tato položka je zakázána, pokud neexistují žádné faktory.

Zbývající pozemek

Vytvoří pro každou závislou proměnnou graf pozorovaných podle předpovědí podle standardizovaných zbytkových chyb. Pozemek je užitečný pro zkoumání předpokladu stejné odchylky.

Nedostatek fit

Zkontrolujte, zda může být vztah mezi závislou proměnnou a nezávislými proměnnými odpovídajícím způsobem popsán modelem.

Obecné odhadnutelné funkce

Umožňuje vytvářet vlastní hypotetické testy na základě obecných odhadovatelných funkcí. Řádky v jakékoliv matici koeficientu kontrastu jsou lineární kombinace obecně odhadnutelné funkce (funkcí).

Úroveň významnosti

Možná budete chtít upravit úroveň významnosti použitou v post-hoc testech a úroveň spolehlivosti použitou pro konstrukci intervalů spolehlivosti. Uvedená hodnota se také použije pro výpočet pozorovaného výkonu pro zkoušku. Zadáte-li úroveň významnosti, zobrazí se v dialogovém okně přidružená úroveň intervalů spolehlivosti.

Další funkce příkazu GLM

Tyto funkce se mohou vztahovat na analýzu jednorozměrných, vícerozměrných nebo opakovaných opatření. Jazyk syntaxe příkazu vám také umožňuje:

- Zadejte vnořené efekty v návrhu (pomocí podpříkazu DESIGN).
- Zadejte testy účinků oproti lineární kombinaci efektů nebo hodnoty (pomocí podpříkazu TEST).
- Zadejte více kontrastů (pomocí podpříkazu CONTRAST).
- Zahrnout chybějící hodnoty uživatele (pomocí podpříkazu MISSING).
- Zadejte kritéria EPS (pomocí podpříkazu CRITERIA).
- Vytvořte vlastní matici **L**, matici **M** nebo matici **K** (pomocí podpříkazů LMATRIX, MMATRIX nebo KMATRIX).
- Pro odchylky nebo jednoduché kontrasty zadejte přechodnou referenční kategorii (pomocí podpříkazu CONTRAST).
- Zadejte metriky pro polynomiální kontrasty (pomocí podpříkazu CONTRAST).
- Zadejte chybové termíny pro post hoc porovnání (pomocí podpříkazu POSTHOC).
- Vypočte odhadnuté mezní střední hodnotu pro jakoukoli interakci faktoru nebo faktoru mezi faktory v seznamu faktorů (pomocí podpříkazu EMMEANS).
- Zadejte názvy dočasných proměnných (pomocí podpříkazu SAVE).
- Vytvořte datový soubor matice korelace (pomocí podpříkazu OUTFILE).

- Vytvořte maticový datový soubor, který obsahuje statistiku z tabulky ANOVA mezi subjekty (pomocí podpříkazu `OUTFILE`).
- Uložte matici návrhu do nového datového souboru (pomocí podpříkazu `OUTFILE`).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Opakovaná opatření GLM

Postup opakovaných měření GLM poskytuje analýzu rozptylu, když je stejné měření provedeno několikrát na každém subjektu nebo případě. Pokud jsou specifikovány faktory mezi subjekty, rozdělí populaci do skupin. Pomocí této obecné lineární modelové procedury můžete testovat nulové hypotézy o účincích jak faktorů mezi subjekty, tak faktorů uvnitř subjektů. Můžete zkoumat interakce mezi faktory, stejně jako účinky jednotlivých faktorů. Kromě toho mohou být zahrnuty účinky konstantních kovariátů a kovarianční interakce s faktory mezi subjekty.

Při vícerozměrném návrhu opakovaných ukazatelů představují závislé proměnné měření více než jedné proměnné pro různé úrovně faktorů uvnitř subjektů. Například byste mohli měřit jak puls, tak dýchání ve třech různých časech na každé téma.

Procedura opakovaných měřítek GLM poskytuje pro data opakovaných měřítek jak jednorozměrné, tak vícerozměrné analýzy. Lze testovat vyvážené i nevyvážené modely. Návrh je vyvážený, pokud každá buňka v modelu obsahuje stejný počet případů. V vícerozměrném modelu jsou součty čtverců v důsledku účinků v modelu a chybové součty čtverců v maticové formě spíše než skalární forma nalezená v jednorozměrné analýze. Tyto matice se nazývají SSCP (součty čtverců a křížových produktů) matice. Kromě testovacích hypotéz GLM Repeated Measures vytváří odhady parametrů.

Běžně používané *a priori* kontrasty jsou k dispozici pro provádění testování hypotéz na faktory mezi subjekty. Poté, co celkový test *F* zobrazí významnost, můžete pomocí post hoc testů vyhodnotit rozdíly mezi specifickými prostředky. Odhadnuté mezní střední hodnoty poskytují odhady předpokládaných středních hodnot pro buňky v modelu a profilové grafy (grafy interakce) těchto prostředků vám umožňují snadno vizualizovat některé vztahy.

Rezidua, predikované hodnoty, Cook's distance a pákové hodnoty lze uložit jako nové proměnné ve vašem datovém souboru pro kontrolu předpokladů. K dispozici je také zbytková matice SSCP, což je čtvercová matice součtů čtverců a křížových produktů zbytků, zbytková kovarianční matice, která je zbytková matice SSCP dělená stupněm volnosti zbytkových chyb, a zbytková korelační matice, která je standardizovanou formou zbytkové kovarianční matice.

Váha WLS vám umožňuje určit proměnnou, která se používá k pozorování různých vah pro analýzu vážených nejmenších čtverců (WLS), případně kompenzovat různou přesnost měření.

Příklad. Dvanáct studentů je přiřazeno do skupiny s vysokou nebo nízkou úzkostí na základě jejich skóre na testu s hodnocením úzkosti. Hodnocení úzkosti se nazývá faktor mezi subjekty, protože rozděluje subjekty do skupin. Studenti jsou každý uveden čtyři zkoušky na učební úkol, a počet chyb pro každou zkoušku je zaznamenán. Chyby pro každou studii jsou zaznamenány v samostatných proměnných a faktor uvnitř subjektů (studie) je definován se čtyřmi úrovněmi pro čtyři studie. Bylo zjištěno, že účinek studie je významný, zatímco interakce studie podle úzkosti není významná.

Metody. K vyhodnocení různých hypotéz lze použít součty čtverců typu I, typu II, typu III a typu IV. Typ III je výchozí.

Statistika. Post hoc range tests and multiple comparisons (for between-subjects factors): least significant difference, Bonferroni, Sidak, Scheffé, Ryan-Einot-Gabriel-Welsch multiple *F*, Ryan-Einot-Gabriel-Welsch multiple range, Student-Newman-Keuls, Tukey's honestly significant difference, Tukey's *b*, Duncan, Hochberg's GT2, Gabriel, Waller Duncan *t* test, Dunnett (one-sided and two-sided), Tamhane's T2, Dunnett's T3, Games-Howell, and Dunnett's *C*. Deskriptivní statistika: pozorované střední hodnoty, směrodatné odchylky a počty všech závislých proměnných ve všech buňkách; Levenův test homogenity rozptylu; Box-Ma-Mauchlyův test sféricity.

Záloty. Spread-versus-level, reziduum a profil (interakce).

Aspekty dat opakovaných měřítek GLM

Data. Závislé proměnné by měly být kvantitativní. Faktory mezi subjekty rozdělují vzorek na diskrétní podskupiny, jako jsou muži a ženy. Tyto faktory jsou kategorické a mohou mít číselné hodnoty nebo řetězcové hodnoty. V dialogovém okně Opakované ukazatele definují faktory v rámci předmětů. Kovariáty jsou kvantitativní proměnné, které souvisí se závislou proměnnou. Pro analýzu opakovaných měření by tyto hodnoty měly zůstat konstantní na každé úrovni proměnné v rámci jednotlivých subjektů.

Datový soubor by měl obsahovat sadu proměnných pro každou skupinu měření na předmětech. Sada má jednu proměnnou pro každé opakování měření v rámci skupiny. Faktor uvnitř předmětu je definován pro skupinu s počtem úrovní rovnajícím se počtu opakování. Například měření hmotnosti může být provedeno v různých dnech. Pokud bylo měření stejné vlastnosti provedeno za pět dní, faktor uvnitř předmětu může být uveden jako *den* s pěti úrovněmi.

U více faktorů uvnitř subjektů se počet měření pro každý subjekt rovná součinu počtu úrovní každého faktoru. Například, pokud byla měření provedena ve třech různých časech každý den po dobu čtyř dnů, celkový počet měření je 12 pro každý subjekt. Faktory v rámci předmětů lze určit jako *den* (4) a *čas* (3).

Předpoklady. K analýze opakovaných ukazatelů lze přistupovat dvěma způsoby, jednorozměrně a vícerozměrně.

Jednorozměrný přístup (také známý jako dělicí graf nebo smíšený model) považuje závislé proměnné za reakce na úrovně faktorů uvnitř subjektů. Měření na subjektu by mělo být vzorkem z vícerozměrného normálního rozdělení a matice rozptylu a kovariance jsou stejné napříč buňkami vytvořenými účinky mezi jednotlivými subjekty. Na matici rozptylu-kovariance závislých proměnných jsou provedeny určité předpoklady. Platnost statistiky *F* použité v jednorozměrném přístupu lze zajistit, pokud je matice rozptylu a kovariance kruhová (Huynh a Mandeville, 1979).

K testování tohoto předpokladu lze použít Mauchlyho test sféricity, který provádí test sféricity na matici variance-kovariance na oronormalizovanou transformovanou závislou proměnnou. Mauchlyho test se automaticky zobrazí pro opakovanou analýzu ukazatelů. Pro malé velikosti vzorků není tento test příliš silný. U velkých velikostí vzorků může být zkouška významná, i když je dopad odchylky na výsledky malý. Pokud je význam zkoušky velký, lze předpokládat hypotézu sféricity. Je-li však význam malý a zdá se, že je porušen předpoklad sféricity, lze provést úpravu stupně volnosti čitatele a jmenovatele, aby se ověřila jednorozměrná statistika *F*. V proceduře GLM Repeated Measures jsou k dispozici tři odhady této úpravy, která se nazývá **epsilon**. Stupně volnosti čitatele i jmenovatele musí být vynásobeny epsilonem a významnost poměru *F* musí být vyhodnocena s novými stupni volnosti.

Vícerozměrný přístup považuje měření na subjektu za vzorek z vícerozměrného normálního rozdělení a matice rozptylu a kovariance jsou stejné napříč buňkami vytvořenými účinky mezi subjekty. Chcete-li otestovat, zda jsou matice rozptylu a kovariance v buňkách stejné, můžete použít test *M* produktu Box.

Související procedury. Před provedením analýzy rozptylu prozkoumejte data pomocí procedury Prozkoumání. Pokud u každého subjektu nejsou opakovaná měření, použijte GLM Univariate nebo GLM Multivariate. Pokud existují pouze dvě měření pro každý subjekt (například před testem a po testu) a neexistují žádné faktory mezi subjekty, můžete použít postup Paired-Samples T Test.

Získání opakovaných ukazatelů GLM

1. Z nabídek vyberte:

Analýza > Obecný lineární model > Opakované ukazatele ...

2. Zadejte název faktoru v rámci předmětu a jeho počet úrovní.

3. Klepněte na tlačítko **Přidat**.

4. Opakujte tyto kroky pro každý faktor uvnitř předmětu.

Chcete-li definovat faktory ukazatelů pro dvojnásob vícevariantní návrh opakovaných ukazatelů, postupujte takto:

5. Zadejte název ukazatele.

6. Klepněte na tlačítko **Přidat**.

Po definování všech vašich faktorů a měřitek:

7. Klepněte na tlačítko **Definovat**.

8. Vyberte závislou proměnnou, která odpovídá každé kombinaci faktorů uvnitř subjektů (a volitelně ukazatelů) v seznamu.

Chcete-li změnit umístění proměnných, použijte šipky nahoru a dolů.

Chcete-li provést změny v rámci faktorů předmětů, můžete znovu otevřít dialogové okno Definovat faktor (y) opakovaných ukazatelů, aniž byste zavřeli hlavní dialogové okno. Volitelně můžete určit faktor (y) mezi předměty a kovariáty.

Faktory definování opakovaných ukazatelů GLM

GLM-Opakované ukazatele analyzuje skupiny souvisejících závislých proměnných, které představují různá měření stejného atributu. Toto dialogové okno vám umožňuje definovat jeden nebo více faktorů v rámci předmětu pro použití v opakovaných ukazatelích GLM. Všimněte si, že pořadí, ve kterém určíte faktory v rámci předmětů, je důležité. Každý faktor představuje úroveň v rámci předchozího faktoru.

Chcete-li použít opakovaná měřítka, musíte správně nastavit svá data. V tomto dialogovém okně musíte definovat faktory v rámci předmětů. Všimněte si, že tyto faktory nejsou existujícími proměnnými ve vašich datech, ale spíše faktory, které zde definujete.

Příklad. Ve studii hubnutí předpokládejme, že hmotnost několika lidí se měří každý týden po dobu pěti týdnů. V datovém souboru je každá osoba subjektem nebo případem. Váhy týdnů se zaznamenávají do proměnných *weight1*, *weight2* atd. Pohlaví každé osoby je zaznamenáno v jiné proměnné. Váhy, měřené opakovaně pro každý subjekt, lze seskupit definováním faktoru uvnitř subjektu. Faktor může být nazýván *týden*, definovaný tak, aby měl pět úrovní. V hlavním dialogovém okně proměnné *weight1*, ..., *weight5* se používají k přiřazení pěti úrovní *týdne*. Proměnná v datovém souboru, která seskupuje muže a ženy (*pohlaví*), může být specifikována jako faktor mezi jednotlivými subjekty pro studium rozdílů mezi muži a ženami.

Ukazatele. Pokud byly subjekty testovány na více než jednom ukazateli v daném okamžiku, definujte opatření. Například puls a frekvence dýchání by mohly být měřeny na každém subjektu každý den po dobu jednoho týdne. Tyto ukazatele neexistují jako proměnné v datovém souboru, ale jsou zde definovány. Model s více než jedním měřítkem se někdy nazývá dvojnásob vícerozměrný model opakovaných ukazatelů.

Model opakovaných ukazatelů GLM

Určit model. Úplný faktoriální model obsahuje všechny hlavní efekty faktoru, všechny hlavní kovariantní účinky a všechny interakce faktoru podle faktoru. Neobsahuje kovariantní interakce. Volbu **Vlastní** vyberte, chcete-li určit pouze podmnožinu interakcí nebo interakce faktoru podle kovariátu. Musíte označit všechny podmínky, které mají být zahrnuty do modelu.

Mezi-předměty. Mezi subjekty jsou uvedeny faktory a kovariáty.

Model. Model závisí na povaze vašich dat. Po výběru volby **Vlastní** můžete vybrat efekty a interakce v rámci předmětu a efekty a interakce mezi subjekty, které jsou předmětem zájmu o vaši analýzu.

Součet čtverců. Způsob výpočtu součtů čtverců pro mezipředmětný model. U vyvážených nebo nevyvážených modelů mezi subjekty bez chybějících buněk se nejčastěji používá metoda součtu čtverců typu III.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Součet čtverců

Pro model si můžete vybrat typ součtů čtverců. Typ III je nejčastěji používaný a je výchozí.

Typ I. Tato metoda je také známá jako hierarchická dekompozice metody součtu čtverců. Každý termín je upraven pouze pro termín, který jej předchází v modelu. Typ I součty čtverců se běžně používají pro:

- Vyvážený model ANOVA, ve kterém jsou všechny hlavní účinky specifikovány před účinky interakce prvního řádu, jakékoli účinky interakce prvního řádu jsou specifikovány před účinky interakce druhého řádu atd.
- Polynomiální regresní model, v němž jsou všechny výrazy nižšího řádu specifikovány před výrazy vyššího řádu.
- Čistě vnořený model, ve kterém je první zadaný efekt vnořen do druhého zadaného efektu, druhý zadaný efekt je vnořen do třetího a tak dále. (Tuto formu vnoření lze zadat pouze pomocí syntaxe.)

Typ II. Tato metoda vypočítá součty čtverců efektu v modelu upravené pro všechny ostatní "vhodné" efekty. Vhodným účinkem je účinek, který odpovídá všem účinkům, které neobsahují zkoumaný účinek. Metoda součtu čtverců typu II se běžně používá pro:

- Vyvážený model ANOVA.
- Jakýkoli model, který má pouze hlavní faktory.
- Jakýkoli regresní model.
- Čistě vnořený návrh. (Tuto formu vnoření lze určit pomocí syntaxe.)

Typ III. Výchozí hodnota Tato metoda vypočítá součty čtverců efektu v designu jako součty čtverců, upravené pro všechny ostatní účinky, které neobsahují efekt, a ortogonální na jakékoliv účinky (pokud existují), které obsahují účinek. Typ III částky čtverců mají jednu hlavní výhodu v tom, že jsou invariantní s ohledem na buněčné frekvence tak dlouho, jak je obecná forma odhadu zůstává konstantní. Proto je tento typ součtů čtverců často považován za užitečný pro nevyvážený model bez chybějících buněk. V faktoriálním návrhu bez chybějících buněk je tato metoda ekvivalentní technice Yates 'weighted-of-means' (vážené čtverce středních hodnot). Metoda součtu čtverců typu III se běžně používá pro:

- Všechny modely uvedené v typech I a II.
- Jakýkoli vyvážený nebo nevyvážený model bez prázdných buněk.

Typ IV. Tato metoda je určena pro situaci, kdy chybí buňky. Pro jakýkoli efekt F v návrhu platí, že pokud F není obsažen v žádném jiném efektu, pak typ IV = Typ III = Typ II. Je-li hodnota F obsažena v jiných efektech, typ IV rozdělí rozdíly mezi parametry v F na všechny efekty vyšší úrovně spravedlivě. Metoda součtu čtverců typu IV se běžně používá pro:

- Všechny modely uvedené v typech I a II.
- Jakýkoli vyvážený model nebo nevyvážený model s prázdnými buňkami.

GLM-opakovaná měření kontrastů

Kontrasty se používají k testování rozdílů mezi úrovněmi faktoru mezi subjekty. Můžete určit kontrast pro každý faktor mezi subjekty v modelu. Kontrasty představují lineární kombinace parametrů.

Testování hypotéz je založeno na nulové hypotéze $\mathbf{L}\mathbf{B}\mathbf{M} = 0$, kde \mathbf{L} je matice koeficientů kontrastu, \mathbf{B} je vektor parametru a \mathbf{M} je průměrná matice, která odpovídá průměrné transformaci pro závislou proměnnou. Tuto transformační matici můžete zobrazit výběrem volby **Matice transformace** v dialogovém okně Volby opakovaných ukazatelů. Pokud například existují čtyři závislé proměnné, faktor v rámci předmětů se čtyřmi úrovněmi a polynomiální kontrasty (předvolba) se používají pro faktory v rámci předmětů, matice \mathbf{M} bude $(0.5 \ 0.5 \ 0.5 \ 0.5)'$. Je-li zadán kontrast, vytvoří se matice \mathbf{L} tak, aby se sloupce odpovídající faktoru mezi subjekty shodovaly s kontrastem. Zbývající sloupce jsou upraveny tak, aby bylo možné odhadnout matici \mathbf{L} .

Dostupné kontrasty jsou odchylka, jednoduché, rozdíl, Helmert, opakované a polynom. Pro odchylky kontrastů a jednoduché kontrasty můžete zvolit, zda je referenční kategorie poslední nebo první kategorie.

Pro faktory v rámci předmětů musí být vybrán jiný kontrast než **Žádný**.

Typy kontrastů

Odchylka. Porovná střední hodnotu každé úrovně (s výjimkou referenční kategorie) se střední hodnotou všech úrovní (grand mean). Úrovně faktoru mohou být v libovolném pořadí.

Jednoduché. Porovná střední hodnotu každé úrovně se střední hodnotou určené úrovně. Tento typ kontrastu je užitečný v případě, že existuje řídicí skupina. Jako odkaz můžete zvolit první nebo poslední kategorii.

Rozdíl. Porovná střední hodnotu každé úrovně (kromě první) se střední hodnotou předchozích úrovní. (Někdy nazývané reverzní Helmert kontrasty.)

Helmert. Porovná střední hodnotu každé úrovně faktoru (s výjimkou poslední úrovně) se střední hodnotou následných úrovní.

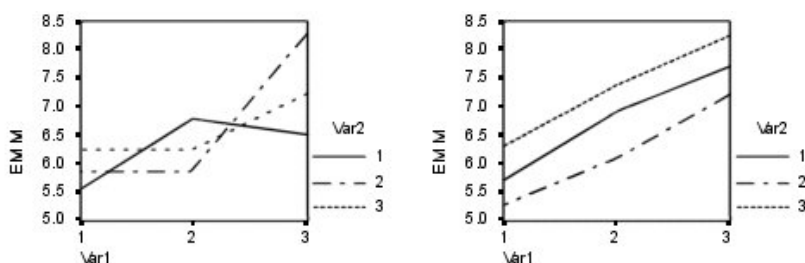
Opakovat. Porovná střední hodnotu každé úrovně (kromě poslední) se střední hodnotou následující úrovně.

Polynomiální. Porovnává lineární efekt, kvadratický efekt, kubický efekt a tak dále. První stupeň volnosti obsahuje lineární efekt ve všech kategoriích; druhý stupeň volnosti, kvadratický efekt; a tak dále. Tyto kontrasty se často používají k odhadu polynomiálních trendů.

Grafy profilů opakovaných měření GLM

Profilové grafy (interakční grafy) jsou užitečné pro porovnání mezních hodnot ve vašem modelu. Profilový graf je čárový graf, ve kterém každý bod označuje odhadovaný mezní průměr závislé proměnné (upravený pro všechny kovariáty) na jedné úrovni faktoru. Úrovně druhého faktoru lze použít k vytvoření oddělených čar. Každou úroveň ve třetím faktoru lze použít k vytvoření samostatného výkresu. Pro pozemky jsou k dispozici všechny faktory. Pro každou závislou proměnnou jsou vytvořeny profilové grafy. V profilových zákresech lze použít jak faktory mezi subjekty, tak i faktory uvnitř subjektů.

Profilový graf s jedním faktorem ukazuje, zda se odhadované mezní střední hodnoty zvyšují nebo snižují napříč úrovněmi. Pro dva nebo více faktorů paralelní čáry označují, že neexistuje žádná interakce mezi faktory, což znamená, že můžete zkoumat úrovně pouze jednoho faktoru. Neparalelní čáry označují interakci.



Obrázek 2. Neparalelní vykreslení (vlevo) a paralelní vykreslení (vpravo)

Po zadání výkresu výběrem faktorů pro vodorovnou osu a volitelně faktorů pro samostatné čáry a samostatné grafy musí být graf přidán do seznamu Zákresy.

Porovnání následných opatření GLM-opakovaná

Post hoc vícenásobné porovnávací testy. Jakmile zjistíte, že existují rozdíly mezi prostředky, mohou post hoc testy rozsahu a párování více porovnání určit, které střední hodnoty se liší. Porovnání se provádí na neupravených hodnotách. Tyto testy nejsou k dispozici, pokud nejsou žádné faktory mezi subjekty, a post hoc vícenásobné porovnávací testy jsou prováděny pro průměr napříč úrovněmi faktorů v rámci subjektů.

Bonferroniho a Tukeyho čestně významné rozdílové testy se běžně používají v několika srovnávacích testech. **Bonferroni test**, založený na Studentově statistice t , upravuje pozorovanou úroveň významnosti pro skutečnost, že bylo provedeno více porovnání. **Sidakův t test** také upravuje úroveň významnosti a poskytuje přísnější meze než Bonferroniho test. **Test upřímně významného rozdílu společnosti Tukey** používá statistiku rozsahu Studentized k provedení všech porovnání po dvojicích mezi skupinami

a k nastavení míry chyb experimentálně na poměr chyb pro kolekci pro všechna porovnávání po dvojicích. Při testování velkého počtu párů prostředků, Tukey je upřímně významný rozdíl test je silnější než Bonferroni test. Pro malý počet párů je Bonferroni silnější.

Hochbergův GT2 je podobný testu Tukeyho upřímně významného rozdílu, ale používá se Studentizovaný maximální modul. Obvykle je Tukeyho zkouška silnější. **Gabrielův porovnávací test po dvojicích** také používá Studentizovaný maximální modul a je obecně výkonnější než Hochbergův GT2, když jsou velikosti buněk nestejně. Gabrielův test se může stát liberálním, když se velikost buněk značně liší.

Dunnettův párování vícenásobný porovnávací t test porovnává sadu ošetření s jedinou střední hodnotou kontroly. Poslední kategorie je výchozí řídicí kategorie. Případně můžete zvolit první kategorii. Můžete si také vybrat oboustranný nebo jednostranný test. Chcete-li ověřit, že střední hodnota faktoru na jakékoli úrovni (s výjimkou kontrolní kategorie) se nerovná střední hodnotě kontrolní kategorie, použijte oboustrannou zkoušku. Chcete-li otestovat, zda je střední hodnota na libovolné úrovni faktoru menší než střední hodnota řídicí kategorie, vyberte volbu **< Řídicí prvek**. Podobně, chcete-li otestovat, zda je střední hodnota na libovolné úrovni faktoru větší než střední hodnota řídicí kategorie, vyberte volbu **> Řídicí prvek**.

Ryan, Einot, Gabriel a Welsch (R-E-G-W) vyvinuli dva testy s více stupňovitými rozsahy. Několik postupných kroků nejprve otestujte, zda jsou všechny prostředky stejné. Pokud všechny prostředky nejsou stejné, podskupiny prostředků jsou testovány na rovnost. **R-E-G-W F** je založeno na testu F a **R-E-G-W Q** je založeno na Studentizovaném rozsahu. Tyto testy jsou výkonnější než Duncanův test s vícenásobným rozsahem a Student-Newman-Keuls (což jsou také vícenásobné postupy krok dolů), ale nejsou doporučeny pro nerovné velikosti buněk.

Pokud jsou odchylky nestejně, použijte **Tamhanův T2** (konzervativní porovnávací test po dvojicích založený na t testu), **Dunnettův T3** (porovnávací test po dvojicích založený na Studentizovaném maximálním modulu), **Games-Howell porovnávací test po dvojicích** (někdy liberální) nebo **Dunnettův C** (porovnávací test po dvojicích založený na Studentizovaném rozsahu).

Duncan's multiple range test, Student-Newman-Keuls (**S-N-K**) a **Tukey's b** jsou testy rozsahu, které hodnotí střední hodnotu skupiny a vypočítávají hodnotu rozsahu. Tyto testy se nepoužívají tak často, jak se dříve diskutovalo.

Test Waller-Duncan t používá bayesovský přístup. Tato zkouška rozsahu používá harmonický průměr velikosti vzorku, pokud jsou velikosti vzorku nerovné.

Úroveň významnosti testu **Scheffé** je navržena tak, aby umožňovala testování všech možných lineárních kombinací středních hodnot skupiny, nikoli pouze porovnávání po dvojicích, která jsou v této funkci k dispozici. Výsledkem je, že Scheffé test je často konzervativnější než jiné testy, což znamená, že pro významnost je nutný větší rozdíl mezi prostředky.

Nejméně významný rozdíl (**LSD**) porovnávacího testu po dvojicích je ekvivalentní několika individuálním testům t mezi všemi dvojicemi skupin. Nevýhodou tohoto testu je, že nebyl učiněn žádný pokus o úpravu pozorované úrovně významnosti pro vícenásobná porovnání.

Zobrazené testy. Porovnávání po dvojicích jsou poskytována pro LSD, Sidak, Bonferroni, Games-Howell, Tamhane's T2 a T3, Dunnett's Ca Dunnett's T3. Homogenní dílčí sady pro testy rozsahu jsou poskytovány pro S-N-K, Tukeyho b , Duncan, R-E-G-W F , R-E-G-W Q a Waller. Tukey je upřímně významný rozdíl test, Hochberg GT2, Gabriel je test, a Scheffé je test je jak více srovnávacích testů a testy rozsahu.

GLM-Odhadované mezní střední hodnota

Vyberte faktory a interakce, pro které chcete odhadnout mezní střední hodnotu populace v buňkách. Tyto prostředky jsou upraveny pro případné kovariáty.

Porovnat hlavní efekty

Poskytuje neopravené porovnávání po dvojicích mezi odhadovanými mezními hodnotami pro jakýkoli hlavní vliv v modelu, a to jak pro faktory mezi jednotlivými subjekty, tak pro faktory v rámci jednotlivých subjektů. Tato položka je k dispozici pouze v případě, že jsou v seznamu Zobrazit prostředky pro vybraný hlavní efekty.

Porovnat jednoduché hlavní efekty

Nastavení je povoleno vždy, když cílový seznam obsahuje jeden nebo více efektů produktu nebo interakce (například $A*B$, $A*B*C$). Nastavení podporuje specifikaci porovnání mezi jednoduchými hlavními efekty, což jsou hlavní efekty vnořené do úrovní jiných faktorů.

Nastavení konfidenčního intervalu

Vyberte nejméně významný rozdíl (LSD), Bonferroni nebo Sidak nastavení intervalů spolehlivosti a významnosti. Tato položka je k dispozici pouze v případě, že je vybrána volba **Porovnat hlavní efekty** a/nebo **Porovnat jednoduché hlavní efekty**.

Určení odhadovaných mezních prostředků

1. Z nabídek vyberte jednu z procedur dostupných v části > **Analyzovat** > **Obecný lineární model**.
2. V hlavním dialogovém okně klepněte na volbu **EM Means**.

Uložení opakovaných opatření GLM

Hodnoty předpovězené modelem, reziduály a souvisejícími ukazateli můžete uložit jako nové proměnné v editoru dat. Mnoho z těchto proměnných lze použít pro zkoumání předpokladů o datech. Chcete-li uložit hodnoty pro použití v jiné relaci IBM SPSS Statistika, musíte uložit aktuální datový soubor.

Předpokládané hodnoty. Hodnoty, které model předpovídá pro každý případ.

- *Nestandardizovaný.* Hodnota, kterou model předpovídá pro závislou proměnnou.
- *Standardní chyba.* Odhad směrodatné odchylky průměrné hodnoty závislé proměnné pro případy, které mají stejné hodnoty nezávislých proměnných.

Diagnostika. Ukazatele pro identifikaci případů s neobvyklými kombinacemi hodnot pro nezávislé proměnné a případy, které mohou mít velký dopad na model. K dispozici jsou Cookovy vzdálenosti a nevystředěné pákové hodnoty.

- *Vzdálenost od kuchaře.* Měřítka toho, kolik by se změnilo rezidua všech případů, kdyby byl konkrétní případ vyloučen z výpočtu regresních koeficientů. Velký Cook's D označuje, že vyloučení případu z výpočtu regresní statistiky podstatně změní koeficienty.
- *Hodnoty pákového efektu.* Nesoustředěné hodnoty pákového efektu. Relativní vliv každého pozorování na vhodnost modelu.

Zbytky. Nestandardizovaný zbytek je skutečná hodnota závislé proměnné minus hodnota předpovězená modelem. K dispozici jsou také standardizované, studentizované a smazané rezidenční chyby.

- *Nestandardizovaný.* Rozdíl mezi pozorovanou hodnotou a hodnotou předpovězenou modelem.
- *Standardizované.* Zbytek vydělený odhadem směrodatné odchylky. Standardizované zbytkové chyby, které jsou také známé jako Pearsonovy rezidenční chyby, mají střední hodnotu 0 a směrodatnou odchylku 1.
- *Studovaný.* Zbytek děleno odhadem jeho směrodatné odchylky, která se liší případ od případu, v závislosti na vzdálenosti hodnot jednotlivých případů na nezávislých proměnných od prostředků nezávislých proměnných. Někdy označované jako vnitřně studentizované zbytkové chyby.
- *Odstraněná.* Zbytek pro případ, kdy je tento případ vyloučen z výpočtu regresních koeficientů. Jedná se o rozdíl mezi hodnotou závislé proměnné a upravenou předpovězenou hodnotou.

Statistika koeficientů. Uloží matici variance-kovariance odhadů parametrů do datové sady nebo datového souboru. Pro každou závislou proměnnou bude také uveden řádek odhadů parametrů, řádek hodnot významnosti pro statistiku t odpovídající odhadům parametrů a řádek se zbývajícími stupněmi volnosti. Pro vícerozměrný model existují podobné řádky pro každou závislou proměnnou. Tato data matice můžete použít v jiných procedurách, které čtou soubory matice. Datové sady jsou k dispozici pro následné použití ve stejné relaci, ale nejsou uloženy jako soubory, pokud nejsou výslovně uloženy před koncem relace. Názvy datových sad musí odpovídat pravidlům pojmenování proměnných.

Volby opakovaných ukazatelů GLM

V tomto dialogovém okně jsou k dispozici volitelné statistiky. Statistiky se vypočítávají pomocí modelu s pevnými efekty.

Zobrazit

Deskriptivní statistika

Vytváří pozorované střední hodnoty, směrodatné odchylky a počty pro všechny závislé proměnné ve všech buňkách.

Odhady velikosti účinku

Poskytuje částečnou hodnotu eta-kvadrát pro každý efekt a každý odhad parametru. Statistika eta-kvadrát popisuje podíl celkové variability, který lze přisoudit faktoru.

Pozorovaný výkon

Získá sílu testu, když je alternativní hypotéza nastavena na základě pozorované hodnoty.

Odhady parametrů

Vytváří odhady parametrů, standardní chyby, t testy, intervaly spolehlivosti a pozorovanou sílu pro každý test.

Matice SSCP

Zobrazení hypotézy a chybových SSCP matic.

Zbytková matice SSCP

Zobrazí reziduální matici SSCP.

Transformační matice

Zobrazuje Bartlettův test sféricity matice zbytkové kovariance.

Zkoušky homogenity

Vytvoří Levenův test homogenity rozptylu pro každou závislou proměnnou napříč všemi kombinacemi faktorů mezi subjekty, pouze pro faktory mezi subjekty. Testy homogenity také zahrnují Box M test homogenity matic kovariance závislých proměnných napříč všemi kombinacemi úrovní faktorů mezi subjekty.

Rozložení vs. úroňový graf

Užitečné pro kontrolu předpokladů o datech. Tato volba je zakázána, pokud neexistují žádné faktory.

Zbývající pozemek

Vytvoří pro každou závislou proměnnou graf pozorovaných podle předpovědí podle standardizovaných zbytkových chyb. Tyto grafy jsou užitečné pro zkoumání předpokladu stejné odchylky. Tato volba je zakázána, pokud neexistují žádné faktory.

Nedostatek fit

Zkontroluje, zda model může odpovídajícím způsobem popsat vztah mezi závislou proměnnou a nezávislými proměnnými.

Obecné odhadnutelné funkce

Umožňuje vytvářet vlastní hypotetické testy na základě obecných odhadovatelných funkcí. Řádky v jakékoliv matici koeficientu kontrastu jsou lineární kombinace obecně odhadnutelné funkce (funkcí).

Úroveň významnosti

Možná budete chtít upravit úroveň významnosti použitou v post-hoc testech a úroveň spolehlivosti použitou pro konstrukci intervalů spolehlivosti. Uvedená hodnota se také použije pro výpočet pozorovaného výkonu pro zkoušku. Zadáte-li úroveň významnosti, zobrazí se v dialogovém okně přidružená úroveň intervalů spolehlivosti.

Další funkce příkazu GLM

Tyto funkce se mohou vztahovat na analýzu jednorozměrných, vícerozměrných nebo opakovaných opatření. Jazyk syntaxe příkazu vám také umožňuje:

- Zadejte vnořené efekty v návrhu (pomocí podpříkazu DESIGN).

- Zadejte testy účinků oproti lineární kombinaci efektů nebo hodnoty (pomocí podpříkazu TEST).
- Zadejte více kontrastů (pomocí podpříkazu CONTRAST).
- Zahrnout chybějící hodnoty uživatele (pomocí podpříkazu MISSING).
- Zadejte kritéria EPS (pomocí podpříkazu CRITERIA).
- Vytvořte vlastní matici **L** , matici **M** nebo matici **K** (pomocí podpříkazů LMATRIX, MMATRIX a KMATRIX).
- Pro odchylky nebo jednoduché kontrasty zadejte přechodnou referenční kategorii (pomocí podpříkazu CONTRAST).
- Zadejte metriky pro polynomiální kontrasty (pomocí podpříkazu CONTRAST).
- Zadejte chybové termíny pro post hoc porovnání (pomocí podpříkazu POSTHOC).
- Vypočte odhadnuté mezní střední hodnotu pro jakoukoli interakci faktoru nebo faktoru mezi faktory v seznamu faktorů (pomocí podpříkazu EMMEANS).
- Zadejte názvy dočasných proměnných (pomocí podpříkazu SAVE).
- Vytvořte datový soubor matice korelace (pomocí podpříkazu OUTFILE).
- Vytvořte maticový datový soubor, který obsahuje statistiku z tabulky ANOVA mezi subjekty (pomocí podpříkazu OUTFILE).
- Uložte matici návrhu do nového datového souboru (pomocí podpříkazu OUTFILE).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference* .

Analýza komponent odchylky

Procedura Variance Components pro modely se smíšenými efekty odhaduje podíl každého náhodného efektu na rozptylu závislé proměnné. Tento postup je zvláště zajímavý pro analýzu smíšených modelů, jako je dělicí graf, jednorozměrná opakovaná opatření a náhodné blokové návrhy. Výpočtem komponent rozptylu můžete určit, kam zaměřit pozornost, aby se rozptyl snížil.

Pro odhad komponent rozptylu jsou k dispozici čtyři různé metody: minimální kvadratický nezaujatý odhad (MINQUE), analýza rozptylu (ANOVA), maximální pravděpodobnost (ML) a omezená maximální pravděpodobnost (REML). Pro různé metody jsou k dispozici různé specifikace.

Výchozí výstup pro všechny metody zahrnuje odhady komponent rozptylu. Je-li použita metoda ML nebo metoda REML, zobrazí se také asymptotická maticová tabulka kovariance. Další dostupný výstup zahrnuje tabulku ANOVA a očekávané střední čtverce pro metodu ANOVA a historii iterací pro metody ML a REML. Procedura Variance Components je plně kompatibilní s procedurou GLM Univariate.

Váha WLS vám umožňuje určit proměnnou používanou k pozorování různých vah pro váženou analýzu, případně kompenzovat rozdíly v přesnosti měření.

Příklad. V zemědělské škole se přírůstky hmotnosti prasat v šesti různých vrzích měří po jednom měsíci. Proměnná vrhu je náhodný faktor se šesti úrovněmi. (Šest zkoumaných vrhů je náhodný vzorek z velké populace prasečích vrhů.) Vyšetřovatel zjišťuje, že rozdíl v přírůstku hmotnosti je způsoben rozdílem ve vrzích mnohem více než rozdílem u prasat ve vrhu.

Aspekty dat komponent odchylky

Data. Závislá proměnná je kvantitativní. Faktory jsou kategorické. Mohou mít číselné hodnoty nebo řetězcové hodnoty až osm bajtů. Alespoň jeden z faktorů musí být náhodný. To znamená, že úroveň faktoru musí být náhodným vzorkem možných úrovní. Kovariáty jsou kvantitativní proměnné, které souvisí se závislou proměnnou.

Předpoklady. Všechny metody předpokládají, že parametry modelu náhodného efektu mají nulové střední hodnoty a konečné konstantní rozptyly a jsou vzájemně nekorelované. Parametry modelu z různých náhodných efektů jsou také nekorelované.

Zbývající termín má také nulový průměr a konečný konstantní rozptyl. Nekoreluje s parametry modelu jakéhokoli náhodného efektu. Zbytkové termíny z různých pozorování se považují za nekorelované.

Na základě těchto předpokladů jsou korelovány pozorování ze stejné úrovně náhodného faktoru. Tato skutečnost odlišuje rozptylový model od obecného lineárního modelu.

ANOVA a MINQUE nevyžadují předpoklady o normálnosti. Oba jsou robustní až mírné odchylky od předpokladu normality.

ML a REML vyžadují, aby byl parametr modelu a zbytkový termín normálně distribuován.

Související procedury. Před provedením analýzy komponent odchylky prozkoumejte data pomocí procedury Explore. Pro testování hypotéz použijte měřítka GLM Univariate, GLM Multivariate a GLM Repeated.

Získání tabulek komponent odchylky

1. Z nabídek vyberte:

Analyzovat > Obecný lineární model > Komponenty odchylky ...

2. Vyberte závislou proměnnou.

3. Vyberte proměnné pro fixní faktor (y), náhodný faktor (y) a Covariate (y), které odpovídají vašim datům. Chcete-li zadat proměnnou váhy, použijte volbu Váha WLS.

Model komponent odchylky

Určit model. Úplný faktoriální model obsahuje všechny hlavní efekty faktoru, všechny hlavní kovariantní účinky a všechny interakce faktoru podle faktoru. Neobsahuje kovariantní interakce. Volbu **Vlastní** vyberte, chcete-li určit pouze podmnožinu interakcí nebo interakce faktoru podle kovariátu. Musíte označit všechny podmínky, které mají být zahrnuty do modelu.

Faktory & kryty. Jsou uvedeny faktory a kovariáty.

Model. Model závisí na povaze vašich dat. Po výběru volby **Vlastní** můžete vybrat hlavní efekty a interakce, které vás zajímají v analýze. Model musí obsahovat náhodný faktor.

Pro vybrané faktory a kovariáty:

Interakce

Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné. Jedná se o výchozí nastavení.

Hlavní účinky

Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Všechny 2-way

Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3-way

Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny 4-way

Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Vše 5-ti cestné

Vytvoří všechny možné pěticestné interakce vybraných proměnných.

Zahrnout zachycení v modelu. Obvykle je zachycení zahrnuto v modelu. Pokud můžete předpokládat, že data procházejí původem, můžete toto zachycení vyloučit.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Volby komponent odchylky

- Můžete zvolit jednu ze čtyř metod pro odhad komponent rozptylu.
- **MINQUE** (minimální kvadratický nezaujatý odhad) vytváří odhady, které jsou invariantní s ohledem na pevné účinky. Pokud jsou data běžně distribuována a odhady jsou správné, tato metoda vytvoří nejmenší rozptyl mezi všemi nezaujatými odhady. Můžete zvolit metodu pro předchozí váhy s náhodným efektem.
- **ANOVA** (analýza rozptylu) vypočítává nezkrácené odhady pomocí součtů čtverců typu I nebo typu III pro každý efekt. Metoda ANOVA někdy vytváří negativní odhady rozptylu, které mohou indikovat nesprávný model, nevhodnou metodu odhadu nebo potřebu dalších údajů.
- **Maximální pravděpodobnost** (ML) vytváří odhady, které by byly nejvíce konzistentní s daty, která byla skutečně sledována, pomocí iterací. Tyto odhady mohou být zkrácené. Tato metoda je asymptoticky normální. ML a REML odhady jsou invariantní v překladu. Tato metoda nebere v úvahu stupně volnosti použité k odhadu pevných účinků.
- **Omezená maximální pravděpodobnost** (REML) snižuje odhady ANOVA pro mnoho (ne-li všechny) případů vyvážených dat. Vzhledem k tomu, že tato metoda je upravena pro fixní efekty, měla by mít menší standardní chyby než metoda ML. Tato metoda bere v úvahu stupně volnosti použité pro odhad pevných účinků.

náhodný efekt priors. Jednotný znamená, že všechny náhodné účinky a zbytkový termín mají stejný dopad na pozorování. Schéma **Nulová** je ekvivalentní předpokladu nulových odchylek náhodného efektu. K dispozici pouze pro metodu MINQUE.

Součet čtverců. Typ I součty čtverců se používají pro hierarchický model, který se často používá v literatuře komponenty rozptylu. Pokud zvolíte **Typ III**, výchozí nastavení v GLM, odhady rozptylu lze použít v GLM Univariate pro testování hypotéz s součty čtverců typu III. K dispozici pouze pro metodu ANOVA.

Kritéria. Můžete určit kritérium konvergence a maximální počet iterací. K dispozici pouze pro metody ML nebo REML.

Zobrazit. Pro metodu ANOVA můžete zvolit zobrazení součtů čtverců a očekávaných středních čtverců. Pokud jste vybrali volbu **Maximální pravděpodobnost** nebo **Omezená maximální pravděpodobnost**, můžete zobrazit historii iterací.

Součet čtverců (komponenty odchylky)

Pro model si můžete vybrat typ součtu čtverců. Typ III je nejčastěji používaný a je výchozí.

Typ I. Tato metoda je také známá jako hierarchická dekompozice metody součtu čtverců. Každý termín je upraven pouze pro termín, který jej předchází v modelu. Metoda součtu čtverců typu I se běžně používá pro:

- Vyvážený model ANOVA, ve kterém jsou všechny hlavní účinky specifikovány před účinky interakce prvního řádu, jakékoli účinky interakce prvního řádu jsou specifikovány před účinky interakce druhého řádu atd.
- Polynomiální regresní model, v němž jsou všechny výrazy nižšího řádu specifikovány před výrazy vyššího řádu.
- Čistě vnořený model, ve kterém je první zadaný efekt vnořen do druhého zadaného efektu, druhý zadaný efekt je vnořen do třetího a tak dále. (Tuto formu vnoření lze zadat pouze pomocí syntaxe.)

Typ III. Výchozí hodnota Tato metoda vypočítá součty čtverců efektu v designu jako součty čtverců upravené pro všechny ostatní účinky, které neobsahují, a ortogonální na jakékoliv účinky (pokud existují), které obsahují. Typ III částky čtverců mají jednu hlavní výhodu v tom, že jsou invariantní s ohledem na buněčné frekvence tak dlouho, jak je obecná forma odhadu zůstává konstantní. Proto je tento typ často považován za užitečný pro nevyvážený model bez chybějících buněk. V faktoriálním návrhu bez chybějících buněk je tato metoda ekvivalentní technice Yates 'weighted-of-means' (vážené čtverce středních hodnot). Metoda součtu čtverců typu III se běžně používá pro:

- Všechny modely uvedené v typu I.

- Všechny vyvážené nebo nevyvážené modely bez prázdných buněk.

Uložit komponenty odchylky do nového souboru

Některé výsledky této procedury můžete uložit do nového datového souboru IBM SPSS Statistika .

Odhady odchylky komponenty. Uloží odhady komponent rozptylu a popisků odhadu do datového souboru nebo datové sady. Ty lze použít pro výpočet více statistik nebo pro další analýzu v procedurách GLM. Můžete je například použít k výpočtu intervalů spolehlivosti nebo k testování hypotéz.

kovariance komponenty. Uloží matici variance-kovariance nebo korelační matici do datového souboru nebo datové sady. K dispozici pouze v případě, že byla zadána volba **Maximální pravděpodobnost** nebo **Omezená maximální pravděpodobnost** .

Cíl pro vytvořené hodnoty. Umožňuje vám uvést název datové sady nebo externí název souboru pro soubor obsahující odhady komponenty rozptylu a/nebo matici. Datové sady jsou k dispozici pro následné použití ve stejné relaci, ale nejsou uloženy jako soubory, pokud nejsou výslovně uloženy před koncem relace. Názvy datových sad musí odpovídat pravidlům pojmenování proměnných.

Pomocí příkazu MATRIX můžete extrahovat potřebná data z datového souboru a poté vypočítat intervaly spolehlivosti nebo provést testy.

Další funkce příkazu VARCOMP

Jazyk syntaxe příkazu vám také umožňuje:

- Zadejte vnořené efekty v návrhu (pomocí podpříkazu DESIGN).
- Zahrnout chybějící hodnoty uživatele (pomocí podpříkazu MISSING).
- Zadejte kritéria EPS (pomocí podpříkazu CRITERIA).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference* .

Lineární smíšené modely

Procedura lineárních smíšených modelů rozšiřuje obecný lineární model tak, aby data mohla vykazovat korelované a nekonstantní variabilitu. Smíšený lineární model proto poskytuje flexibilitu modelování nejen prostředků dat, ale i jejich rozptylů a kovariancí.

Procedura lineárních smíšených modelů je také flexibilní nástroj pro montáž jiných modelů, které mohou být formulovány jako smíšené lineární modely. Takové modely zahrnují víceúrovňové modely, hierarchické lineární modely a modely s náhodným koeficientem.

Příklad

Řetězec obchodu s potravinami se zajímá o účinky různých kupónů na výdaje zákazníků. Vezmeme-li náhodný vzorek svých stálých zákazníků, sledují výdaje každého zákazníka po dobu 10 týdnů. V každém týdnu, jiný kupón je poslán zákazníkům. Lineární smíšené modely se používají k odhadu účinku různých kupónů na výdaje a zároveň k přizpůsobení korelace v důsledku opakovaných pozorování na každé téma během 10 týdnů.

metody

Odhad maximální pravděpodobnosti (ML) a omezené maximální pravděpodobnosti (REML).

Statistika

Popisná statistika: velikost vzorku, střední hodnoty a směrodatné odchylky závislé proměnné a kovariátů pro každou jednotlivou kombinaci úrovní faktorů. Informace na úrovni faktorů: tříděné hodnoty úrovní jednotlivých faktorů a jejich frekvencí. Také odhady parametrů a intervaly spolehlivosti pro pevné efekty a Wald testy a intervaly spolehlivosti pro parametry kovariančních matic. Pro vyhodnocení různých hypotéz lze použít součty čtverců typu I a typu III. Typ III je výchozí.

Aspekty dat lineárních smíšených modelů

Data

Závislá proměnná by měla být kvantitativní. Faktory by měly být kategorické a mohou mít číselné hodnoty nebo řetězcové hodnoty. Proměnné kovariáty a hmotnostní proměnná by měly být kvantitativní. Předměty a opakované proměnné mohou být jakéhokoli typu.

Předpoklady

Předpokládá se, že závislá proměnná lineárně souvisí s fixními faktory, náhodnými faktory a kovariáty. Model fixních efektů modeluje střední hodnotu závislé proměnné. Náhodné efekty modelují kovarianční strukturu závislé proměnné. Vícenásobné náhodné efekty jsou považovány za nezávislé na sobě a pro každou z nich budou vypočteny samostatné kovarianční matice; avšak modelové termíny uvedené na stejném náhodném efektu mohou být korelovány. Opakovaná opatření modelují kovarianční strukturu zbytků. Předpokládá se také, že závislá proměnná pochází z normálního rozdělení.

Související postupy

Před spuštěním analýzy prozkoumejte data pomocí procedury Explore. Pokud nemáte podezření, že existuje korelovaná nebo nekonstantní variabilita, můžete použít proceduru GLM Univariate nebo GLM Repeated Measures. Alternativně můžete použít postup Analýza komponent rozptylu, pokud náhodné efekty mají strukturu kovariance komponent rozptylu a neexistují žádná opakovaná opatření.

Získání analýzy lineárních smíšených modelů

1. Z nabídek vyberte:

Analyzovat > Smíšené modely > Lineární ...

2. Volitelně vyberte jednu nebo více proměnných předmětu.
3. Volitelně vyberte jednu nebo více opakovaných proměnných. Jsou-li definovány opakované proměnné, vyberte z rozevíracího seznamu volbu **Typ opakovaného Covariance**.
4. Volitelně vyberte jednu nebo více proměnných ukazatele Kroneckerovo.
5. Volitelně vyberte zbytkovou kovarianční strukturu.
6. Klepněte na tlačítko **Pokračovat**.
7. Vyberte závislou proměnnou.
8. Vyberte alespoň jeden faktor nebo kovariát.
9. Klepněte na volbu **Opraveno** nebo **Náhodný** a zadejte alespoň model s pevnými nebo náhodnými efekty.

Volitelně vyberte proměnnou váhy.

Lineární smíšené modely: předměty a opakující se

Toto dialogové okno vám umožňuje vybrat proměnné, které definují předměty, opakovaná pozorování, kroneckerovo ukazatele a zvolit kovarianční strukturu pro rezidua.

Předměty

Předmět je pozorovací jednotka, kterou lze považovat za nezávislou na jiných předmětech. Například údaje o krevním tlaku od pacienta v lékařské studii lze považovat za nezávislé na hodnotách od jiných pacientů. Definování subjektů je důležité zejména v případě, že se provádí opakované měření na jeden subjekt a vy chcete modelovat korelaci mezi těmito pozorováními. Například můžete očekávat, že hodnoty krevního tlaku od jednoho pacienta během následných návštěv u lékaře jsou korelované.

Předměty lze také definovat pomocí kombinace více proměnných na úrovni faktoru. Můžete například zadat *Pohlaví* a *Věkovou kategorii* jako proměnné předmětu, abyste modelovali přesvědčení, že *muži starší 65 let* jsou si navzájem podobní, ale nezávislí na *mužích mladších 65 let* a *ženách*.

Všechny proměnné uvedené v seznamu **Subjekty** se používají k definování předmětů pro strukturu zbytkové kovariance. Můžete použít některé nebo všechny proměnné k definování předmětů pro strukturu kovariance náhodných efektů.

Opakováno

Proměnné uvedené v tomto seznamu se používají k identifikaci opakovaných pozorování. Například jedna proměnná *Týden* může identifikovat 10 týdnů pozorování v lékařské studii, nebo *Měsíc a Den* mohou být použity společně k identifikaci denních pozorování v průběhu roku.

Typ opakovaného Covariance

Tato volba určuje strukturu kovariance pro rezidua. Dostupné struktury jsou následující:

- Ante-závislost: První objednávka
- Účet pohledávek (1)
- Přímý produkt AR1 (UN_AR1)
- Nestrukturovaný přímý produkt (UN_UN)
- Přímá složená symetrie produktu (UN_CS)
- AR (1): Heterogenní
- ARMA (1, 1)
- Složená symetrie
- Složená symetrie: Korelační metrika
- Sloučenina symetrie: Heterogenní
- Diagonála
- Faktorová analýza: První objednávka
- Faktorová analýza: První objednávka, Heterogenní
- Huynh-Feldt
- Škálovaná identita
- Toeplitz
- Toeplitz: Heterogenní
- Nestrukturované
- Nestrukturovaná: Korelační metrika
- Prostorové: Napájení
- Prostorové: exponenciální
- Prostorové: Gaussovský
- Prostorové: Lineární
- Prostorový: Lineární-log
- Prostorové: Sférické

Kroneckerovo opatření

Vyberte proměnné, které určují strukturu předmětu pro měření kovariance Kronecker, a určete, jak jsou korelovány chyby měření. Pole je k dispozici pouze v případě, že je vybrán jeden z následujících **opakujících se typů Covariance** :

- Přímý produkt AR1 (UN_AR1)
- Nestrukturovaný přímý produkt (UN_UN)
- Přímá složená symetrie produktu (UN_CS)

Prostorová kovariance souřadnice

Proměnné v tomto seznamu určují souřadnice opakovaných pozorování, je-li pro opakovaný typ kovariance vybrán jeden z prostorových typů kovariance.

Další informace naleznete v tématu [“Kovarianční struktury” na stránce 96](#).

Fixní efekty lineárních smíšených modelů

Pevné efekty. Neexistuje žádný výchozí model, takže musíte explicitně určit pevné efekty. Případně můžete sestavit vnořené nebo nevnořené výrazy.

Zahrnout zachycení. Zachycení je obvykle zahrnuto v modelu. Pokud můžete předpokládat, že data procházejí výchozím bodem, můžete toto zachycení vyloučit.

Součet čtverců. Způsob výpočtu součtů čtverců. U modelů bez chybějících buněk se nejčastěji používá metoda Type III.

Sestavit nevnořené podmínky

Pro vybrané faktory a kovariáty:

Faktoriál. Vytvoří všechny možné interakce a hlavní efekty vybraných proměnných. Jedná se o výchozí nastavení.

Interakce. Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné.

Hlavní efekty. Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Všechny obousměrné. Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny trojcestné. Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny čtyřcestné. Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Všechny 5-cestné. Vytvoří všechny možné pěticestné interakce vybraných proměnných.

Sestavit vnořené výrazy

V této proceduře můžete sestavit vnořené výrazy pro svůj model. Vnořené výrazy jsou užitečné pro modelování účinku faktoru nebo kovariátu, jehož hodnoty neinteragují s úrovněmi jiného faktoru. Například řetězec obchodu s potravinami může sledovat výdaje svých zákazníků na několika místech obchodu. Protože každý zákazník navštívuje pouze jednu z těchto lokalit, efekt *Zákazník* může být **vnořen do** efektu *Umístění úložiště*.

Dále můžete do vnořeného výrazu zahrnout efekty interakce nebo přidat více úrovní vnoření.

Omezení. Vnořené výrazy mají následující omezení:

- Všechny faktory v rámci interakce musí být jedinečné. Je-li tedy faktor A , je zadání $A^* A$ neplatné.
- Všechny faktory ve vnořeném efektu musí být jedinečné. Je-li tedy faktor A , je zadání $A(A)$ neplatné.
- Žádný efekt nemůže být vnořen do kovariátu. Je-li tedy A faktorem a X je kovariantní, pak určení $A(X)$ je neplatné.

Součet čtverců

Pro model si můžete vybrat typ součtů čtverců. Typ III je nejčastěji používaný a je výchozí.

Typ I. Tato metoda je také známá jako hierarchická dekompozice metody součtu čtverců. Každý termín je upraven pouze pro termín, který jej předchází v modelu. Typ I součty čtverců se běžně používají pro:

- Vyvážený model ANOVA, ve kterém jsou všechny hlavní účinky specifikovány před účinky interakce prvního řádu, jakékoli účinky interakce prvního řádu jsou specifikovány před účinky interakce druhého řádu atd.
- Polynomiální regresní model, v němž jsou všechny výrazy nižšího řádu specifikovány před výrazy vyššího řádu.
- Čistě vnořený model, ve kterém je první zadaný efekt vnořen do druhého zadaného efektu, druhý zadaný efekt je vnořen do třetího a tak dále. (Tuto formu vnoření lze zadat pouze pomocí syntaxe.)

Typ III. Výchozí hodnota Tato metoda vypočítá součty čtverců efektu v designu jako součty čtverců upravené pro všechny ostatní účinky, které neobsahují, a ortogonální na jakékoliv účinky (pokud existují),

kteří obsahují. Typ III částky čtverců mají jednu hlavní výhodu v tom, že jsou invariantní s ohledem na buněčné frekvence tak dlouho, jak je obecná forma odhadu zůstává konstantní. Proto je tento typ součtů čtverců často považován za užitečný pro nevyvážený model bez chybějících buněk. V faktoriálním návrhu bez chybějících buněk je tato metoda ekvivalentní technice Yates 'weighted-of-means' (vážené čtverce středních hodnot). Metoda součtu čtverců typu III se běžně používá pro:

- Všechny modely uvedené v typu I.
- Všechny vyvážené nebo nevyvážené modely bez prázdných buněk.

Náhodné efekty lineárních smíšených modelů

Typ smlouvy. To vám umožňuje určit strukturu kovariance pro model náhodných efektů. Pro každý náhodný efekt se odhaduje samostatná kovarianční matice. Dostupné struktury jsou následující:

- Ante-závislost: První objednávka
- Účet pohledávek (1)
- AR (1): Heterogenní
- ARMA (1, 1)
- Složená symetrie
- Složená symetrie: Korelační metrika
- Sloučenina symetrie: Heterogenní
- Diagonála
- Faktorová analýza: První objednávka
- Faktorová analýza: První objednávka, Heterogenní
- Huynh-Feldt
- Škálovaná identita
- Toeplitz
- Toeplitz: Heterogenní
- Nestrukturované
- Nestrukturovaná: Korelační metrika
- Komponenty odchylky

Další informace naleznete v tématu [“Kovarianční struktury”](#) na stránce 96 .

Náhodné efekty. Neexistuje žádný výchozí model, takže musíte explicitně určit náhodné efekty. Případně můžete sestavit vnořené nebo nevnořené výrazy. Můžete se také rozhodnout zahrnout termín zachycení do modelu náhodných efektů.

Můžete zadat více modelů náhodných efektů. Po sestavení prvního modelu klepněte na tlačítko **Další** a sestavte další model. Klepnutím na tlačítko **Předchozí** se můžete posouvat zpět mezi existujícími modely. Předpokládá se, že každý model s náhodným efektem je nezávislý na každém jiném modelu s náhodným efektem; to znamená, že pro každý z nich se vypočítají samostatné kovarianční matice. Termíny uvedené ve stejném modelu náhodného efektu lze korelovat.

Seskupení předmětů. Uvedené proměnné jsou ty, které jste vybrali jako proměnné předmětu v dialogovém okně Vybrat předměty/Opakované proměnné. Vyberte některé nebo všechny z nich, abyste definovali předměty pro model náhodných efektů.

Zobrazit předpovědi parametrů pro tuto sadu náhodných efektů. Určuje, že se mají zobrazit odhady parametrů náhodných efektů.

Odhad lineárních smíšených modelů

Metoda

Vyberte maximální pravděpodobnost nebo omezený odhad maximální pravděpodobnosti.

Stupně volnosti

Poskytuje volby pro definování stupňů volnosti pro všechny testy.

Zbytková metoda

Zbytková metoda má pevné stupně volnosti pro všechny zkoušky. Je užitečné, pokud je velikost vzorku dostatečně velká, nebo pokud jsou data vyvážená, nebo pokud model používá jednodušší typ kovariance (například škálovaná identita nebo diagonální).

Satterthwaitová aproximace

Metoda Satterthwaite má v testech stupeň volnosti pole. Je užitečné, pokud je velikost vzorku malá, nebo pokud jsou data nevyvážená, nebo pokud model používá složitý typ kovariance (například nestrukturovaný).

Kenward-Rogerova aproximace

Kenward-Rogerova metoda nabízí přesnější odhad malých vzorků pro rozptyl-kovarianci fixních efektových parametrů a přibližný stupeň volnosti jmenovatele v t-testech a F-testech. Tato metoda zavádí faktor měřítka pro statistiku F a odhaduje ji a stupeň volnosti jmenovatele pomocí rozšíření Taylorovy řady pro odhadovanou náhodnou strukturu v datech.

Poznámka: Kenward-Roger metoda se používá v modelu založené kovariance (místo robustní kovariance). Je-li vybrána metoda Kenward-Roger i robustní kovariance, použije se metoda Kenward-Roger na kovarianci založenou na modelu a zobrazí se následující varování: "Vzhledem k tomu, že je vybrána metoda Kenward-Roger, je robustní kovarianční metoda změněna na kovarianční metodu založenou na modelu."

Iterace

K dispozici jsou následující volby:

Maximální počet iterací

Zadejte nezáporné celé číslo.

Maximální počet nášlapných půlení

Při každé iteraci se velikost kroku sníží o faktor 0.5, dokud se nezvýší pravděpodobnost protokolu nebo dokud se nedosáhne maximální poloviny kroku. Zadejte kladné celé číslo.

Tisknout historii iterací pro každý n krok (y)

Zobrazuje tabulku obsahující hodnotu funkce pravděpodobnosti protokolu a odhady parametrů v každé n iteraci začínající 0^{th} iterací (počáteční odhady). Pokud se rozhodnete vytisknout historii iterací, vždy se vytiskne poslední iterace bez ohledu na hodnotu n .

Konvergence pravděpodobnosti protokolu

Konvergence se předpokládá, pokud je absolutní změna nebo relativní změna ve funkci log-Pravděpodobnost menší než zadaná hodnota, která musí být nezáporná. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Konvergence parametrů

Konvergence se předpokládá, pokud je maximální absolutní změna nebo maximální relativní změna v odhadech parametrů menší než zadaná hodnota, která musí být nezáporná. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Hesovská konvergence

Pro specifikaci **Absolutní** se předpokládá konvergence, pokud je statistika založená na hesenské hodnotě menší než zadaná hodnota. Pro specifikaci **Relativní** se předpokládá konvergence, pokud je statistika menší než součin zadané hodnoty a absolutní hodnoty pravděpodobnosti protokolu. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Maximální počet kroků přidělení skóre

Požadavky na použití algoritmu přidělení skóre Fisher až do počtu iterací n . Zadejte nezáporné celé číslo.

Tolerance singularity

Tato hodnota se používá jako tolerance při kontrole singularity. Zadejte kladnou hodnotu.

Statistika lineárních smíšených modelů

Souhrnná statistika. Vytvoří tabulky pro:

- **Popisná statistika.** Zobrazí velikost vzorku, střední hodnoty a směrodatné odchylky závislé proměnné a kovariátů (jsou-li zadány). Tyto statistiky jsou zobrazeny pro každou jednotlivou kombinaci úrovní faktorů.
- **Souhrn zpracování případu.** Zobrazuje seřazené hodnoty faktorů, proměnné opakovaného měření, předměty opakovaného měření a předměty náhodných efektů a jejich frekvence.

Statistika modelu. Vytvoří tabulky pro:

- **odhady parametrů pro pevné efekty.** Zobrazí odhady parametrů s fixními efekty a jejich přibližné standardní chyby.
- **Testuje parametry kovariance.** Zobrazí asymptotické standardní chyby a Wald testy pro parametry kovariance.
- **Korelace odhadů parametrů.** Zobrazí asymptotickou korelační matici odhadů parametrů fixních efektů.
- **Vztahy odhadů parametrů.** Zobrazí asymptotickou matici kovariance odhadů parametrů s pevnými efekty.
- **Covariance náhodných efektů.** Zobrazí odhadovanou matici kovariance náhodných efektů. Tato volba je k dispozici pouze v případě, že je zadán alespoň jeden náhodný efekt. Je-li pro náhodný efekt zadána proměnná předmětu, zobrazí se společný blok.
- **Covariance zbytků.** Zobrazí odhadovanou matici zbytkové kovariance. Tato volba je k dispozici pouze v případě, že byla zadána opakovaná proměnná. Je-li zadána předmětná proměnná, zobrazí se společný blok.
- **Matice koeficientů kontrastu.** Tato volba zobrazuje odhadnutelné funkce použité pro testování pevných efektů a vlastních hypotéz.

Interval spolehlivosti. Tato hodnota se používá při každém vytváření konfidenčního intervalu. Zadejte hodnotu větší nebo rovnou 0 a menší než 100. Výchozí hodnota je 95.

Lineární smíšené modely EM

Odhadovaná mezní střední hodnota postavených modelů. Tato skupina umožňuje vyžádat si modelové odhadované mezní střední hodnoty závislé proměnné v buňkách a jejich standardní chyby pro určené faktory. Kromě toho můžete požadovat, aby byly porovnány úrovně faktorů hlavních účinků.

- **Faktor (y) a Interakce faktoru.** Tento seznam obsahuje faktory a interakce faktorů, které byly uvedeny v dialogovém okně Pevná, plus termín OVERALL . Modelové termíny sestavené z kovariátů jsou vyloučeny z tohoto seznamu.
- **Zobrazit prostředky pro.** Procedura vypočítá odhadované mezní prostředky pro faktory a interakce faktorů vybrané do tohoto seznamu. Je-li vybrána volba OVERALL , zobrazí se odhadované mezní hodnoty závislé proměnné, které se sbalí nad všemi faktory. Všimněte si, že všechny vybrané faktory nebo interakce faktorů zůstávají vybrány, pokud nebyla přidružená proměnná odebrána ze seznamu Faktory v hlavním dialogovém okně.
- **Porovnání hlavních efektů.** Tato volba vám umožňuje požadovat porovnávání po dvojicích úrovní vybraných hlavních efektů. Úprava konfidenčního intervalu vám umožňuje použít úpravu intervalů konfidance a hodnot významnosti pro zohlednění více porovnání. Dostupné metody jsou LSD (bez úpravy), Bonferroni a Sidak. Nakonec můžete pro každý faktor vybrat referenční kategorii, do které se budou porovnávat. Není-li vybrána žádná referenční kategorie, budou vytvořena všechna porovnávání po dvojicích. Volby pro referenční kategorii jsou první, poslední nebo vlastní (v takovém případě zadáte hodnotu referenční kategorie).

Lineární smíšené modely Uložit

Toto dialogové okno vám umožňuje uložit různé výsledky modelu do pracovního souboru.

Pevné předpovězené hodnoty. Uloží proměnné související s regresními prostředky bez efektů.

- **Předpokládané hodnoty.** Regrese znamená bez náhodných účinků.
- **Standardní chyby.** Standardní chyby odhadů.

- **Stupně volnosti.** Stupně volnosti spojené s odhady.

Předpokládané hodnoty & rezidua. Uloží proměnné související s nastavenou hodnotou modelu.

- **Předpokládané hodnoty.** Hodnota osazeného modelu.
- **Standardní chyby.** Standardní chyby odhadů.
- **Stupně volnosti.** Stupně volnosti spojené s odhady.
- **Zbytky.** Datová hodnota minus předpovězená hodnota.

Lineární smíšené modely-Export

Toto dialogové okno vám umožňuje exportovat obsah výstupních tabulek EBLUPS (Empirical Best Linear Unbiased Predictions) do datových sad nebo souborů .sav . Tlačítko Exportovat v dialogovém okně Lineární smíšené modely je povoleno, pokud je v dialogovém okně Náhodný uveden alespoň jeden náhodný efekt a je zaškrtnuto políčko "Zobrazit předpovědi parametrů pro tento blok".

Exportovat EBLUPS

Vyberte cíl: **Datová sada** nebo **Datový soubor**

Volitelně zadejte název.

Pokud bylo zadáno více náhodných efektů s vytvořenými EBLUPS, zrušte zaškrtnutí políčka, aby se každá tabulka výsledků zobrazila v samostatné datové sadě nebo souboru.

Další funkce příkazu MIXED

Jazyk syntaxe příkazu vám také umožňuje:

- Zadejte testy účinků oproti lineární kombinaci efektů nebo hodnoty (pomocí podpříkazu TEST).
- Zahrnout chybějící hodnoty uživatele (pomocí podpříkazu MISSING).
- Vypočte odhadované mezní hodnoty pro určené hodnoty kovariátů (pomocí klíčového slova WITH podpříkazu EMMEANS).
- Porovnejte jednoduché hlavní účinky interakcí (pomocí podpříkazu EMMEANS).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference* .

Zobecněné lineární modely

Zobecněný lineární model rozšiřuje obecný lineární model tak, aby závislá proměnná byla lineárně závislá na faktorech a kovariátech prostřednictvím specifikované funkce spoje. Kromě toho model umožňuje, aby závislá proměnná měla nenormální rozdělení. Zahrnuje široce používané statistické modely, jako je lineární regrese pro normálně distribuované odpovědi, logistické modely pro binární data, loglineární modely pro data počítání, komplementární log-log modely pro intervalově cenzurované údaje o přežití a mnoho dalších statistických modelů prostřednictvím jeho velmi obecné modelové formulace.

Příklady. Lodní společnost může použít zobecněné lineární modely, aby se vešly Poissonova regrese k počtu poškození pro několik typů lodí postavených v různých časových obdobích, a výsledný model může pomoci určit, které typy lodí jsou nejvíce náchylné k poškození.

Automobilová pojišťovna může použít zobecněné lineární modely, aby se vešly gama regrese k poškození nároků na auta, a výsledný model může pomoci určit faktory, které nejvíce přispívají k velikosti nároku.

Zdravotníci výzkumníci mohou použít generalizované lineární modely, aby se vešly do komplementární regrese log-log do intervalově cenzurovaných dat o přežití, aby předpovídali čas do recidivy pro zdravotní stav.

Aspekty zobecněných dat lineárních modelů

Data. Odpověď může být škálovatelnost, počet, binární nebo události ve zkouškách. Předpokládá se, že faktory jsou kategorické. Za měřítko se považují kovariáty, váha stupnice a posunutí.

Předpoklady. Případy se považují za nezávislé připomínky.

Získání generalizovaného lineárního modelu

Z nabídek vyberte:

Analyzovat > Zobecněné lineární modely > Zobecněné lineární modely ...

1. Uved'te funkci distribuce a propojení (podrobnosti o různých volbách viz níže).
2. Na kartě Odezva vyberte závislou proměnnou.
3. Na kartě Prediktory vyberte faktory a kovariáty, které se mají použít při předpovídání závislé proměnné.
4. Na kartě Model zadejte efekty modelu pomocí vybraných faktorů a kovariátů.

Karta Typ modelu vám umožňuje určit funkci distribuce a propojení pro váš model a poskytuje krátké řezy pro několik běžných modelů, které jsou kategorizovány podle typu odezvy.

Typy modelů

Měřítko odezvy. K dispozici jsou následující volby:

- **Lineární.** Uvádí Normální jako distribuci a Identitu jako funkci odkazu.
- **Gamma s odkazem na protokol.** Uvádí gama jako distribuci a protokol jako funkci odkazu.

Ordinální odpověď. K dispozici jsou následující volby:

- **ordinální logistika.** uvádí multinomiální (ordinální) jako distribuční a kumulativní logit jako funkci odkazu.
- **Ordinální zkouška.** uvádí multinomiální (ordinální) jako distribuci a kumulativní probit jako funkci odkazu.

Počty. K dispozici jsou následující volby:

- **Poisson loglinear.** Uvádí Poisson jako distribuci a Log jako funkci odkazu.
- **Negativní binomický s odkazem na protokol.** Uvádí záporný binomický (s hodnotou 1 pro pomocný parametr) jako distribuční a log jako funkci odkazu. Chcete-li, aby procedura odhadla hodnotu pomocného parametru, uveďte vlastní model s negativním binomickým rozdělením a vyberte volbu **Odhadnout hodnotu** ve skupině Parametr.

Binární odezva nebo data událostí/hodnocení. K dispozici jsou následující volby:

- **Binární logistika.** Uvádí Binomické jako distribuci a Logit jako funkci propojení.
- **Binární probit.** Uvádí Binomické jako distribuci a Probit jako funkci propojení.
- **Interval cenzurovaného přežití.** Uvádí Binomický jako distribuci a komplementární log-log jako funkci odkazu.

Směs. K dispozici jsou následující volby:

- **Tweedie s odkazem na protokol.** Určuje Tweedie jako distribuci a Log jako funkci odkazu.
- **Tweedie s odkazem na identitu.** Určuje Tweedie jako funkci distribuce a Identita jako funkci odkazu.

Vlastní. Určete vlastní kombinaci funkce distribuce a funkce propojení.

Rozdělení

Tato volba určuje rozdělení závislé proměnné. Schopnost určit nenormální rozdělení a neidentické spojovací funkce je základním zlepšením zobecněného lineárního modelu nad obecným lineárním modelem. Existuje mnoho možných kombinací funkcí distribuce a propojení a několik může být vhodné pro danou datovou sadu, takže vaše volba může být vedena a priori teoretickými úvahami nebo která kombinace se zdá, že nejlépe vyhovuje.

- **Binomické.** Toto rozdělení je vhodné pouze pro proměnné, které představují binární odezvu nebo počet událostí.
- **Gamma.** Toto rozdělení je vhodné pro proměnné s kladnými hodnotami měřítka, které jsou posunovány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.

- **Inverse Gaussian.** Toto rozdělení je vhodné pro proměnné s kladnými hodnotami měřítka, které jsou posouvány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.
- **Negativní binomické.** Toto rozdělení lze považovat za počet pokusů potřebných ke sledování úspěchů k a je vhodné pro proměnné s nezápornými celočíselnými hodnotami. Pokud je hodnota dat jiná než celé číslo, menší než 0 nebo chybí, pak se v analýze nepoužije odpovídající velikost písmen. Hodnota pomocného parametru záporného binomického rozdělení může být libovolné číslo větší nebo rovno 0; můžete jej nastavit na pevnou hodnotu nebo povolit, aby byl procedurou odhadnut. Je-li pomocný parametr nastaven na hodnotu 0, je použití tohoto rozdělení ekvivalentní použití Poissonova rozdělení.
- **Normální.** To je vhodné pro proměnné měřítka, jejichž hodnoty mají symetrickou, zvonovitou distribuci o centrální (střední) hodnotě. Závislá proměnná musí být číselná.
- **Poissonovo.** Toto rozdělení lze považovat za počet výskytů události, která je předmětem zájmu, v pevném časovém období a je vhodné pro proměnné s nezápornými celočíselnými hodnotami. Pokud je hodnota dat jiná než celé číslo, menší než 0 nebo chybí, pak se v analýze nepoužije odpovídající velikost písmen.
- **Tweedie.** Toto rozdělení je vhodné pro proměnné, které mohou být zastoupeny Poissonovými směsmi gama rozdělení; rozdělení je "smíšené" v tom smyslu, že kombinuje vlastnosti spojitých (bere nezáporné reálné hodnoty) a diskrétních rozdělení (kladná pravděpodobnostní hmotnost při jediné hodnotě, 0). Závislá proměnná musí být číselná, s datovými hodnotami většími nebo rovnými nule. Pokud je hodnota dat menší než nula nebo chybí, pak se v analýze nepoužije odpovídající případ. Pevná hodnota parametru distribuce Tweedie může být libovolné číslo větší než jedna a menší než dvě.
- **Multinomiální.** Toto rozdělení je vhodné pro proměnné, které představují ordinální odezvu. Závislá proměnná může být číselná nebo řetězcová a musí mít alespoň dvě odlišné platné datové hodnoty.

Funkce propojení

Funkce odkazu je transformací závislé proměnné, která umožňuje odhad modelu. K dispozici jsou následující funkce:

- **Identita.** $f(x) = x$. Závislá proměnná není transformována. Tento odkaz lze použít s jakoukoli distribucí.
- **Doplňkový protokol-log.** $f(x) = \log(-\log(1-x))$. To je vhodné pouze při binomické distribuci.
- **Kumulativní kavchit.** $f(x) = \tan(\pi(x - 0.5))$, použito pro kumulativní pravděpodobnost každé kategorie odpovědi. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní komplementární protokol-log.** $f(x) = \ln(-\ln(1-x))$, použitý na kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní protokol.** $f(x) = \ln(x/(1-x))$, použitý na kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní negativní protokol protokolu.** $f(x) = -\ln(-\ln(x))$, použito pro kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní zkouška.** $f(x) = \Phi^{-1}(x)$, použité na kumulativní pravděpodobnost každé kategorie odezvy, kde Φ^{-1} je inverzní standardní funkce normálního součtového rozdělení. To je vhodné pouze s multinomiální distribucí.
- **log.** $f(x) = \log(x)$. Tento odkaz lze použít s jakoukoli distribucí.
- **Doplňek protokolu.** $f(x) = \log(1-x)$. To je vhodné pouze při binomické distribuci.
- **Přihlásit se.** $f(x) = \log(x/(1-x))$. To je vhodné pouze při binomické distribuci.
- **Negativní binomický.** $f(x) = \log(x/(x+k^{-1}))$, kde k je pomocný parametr negativního binomického rozdělení. To je vhodné pouze s negativním binomickým rozdělením.
- **Protokol negativního protokolu.** $f(x) = -\log(-\log(x))$. To je vhodné pouze při binomické distribuci.
- **Kurzová síla.** $f(x) = [(x/(1-x))^{\alpha} - 1]/\alpha$, pokud $\alpha \neq 0$. $f(x) = \log(x)$, pokud $\alpha = 0$. α je požadované číslo specifikace a musí být reálné číslo. To je vhodné pouze při binomické distribuci.
- **Probit.** $f(x) = \Phi^{-1}(x)$, kde Φ^{-1} je inverzní standardní funkce normální kumulativní distribuce. To je vhodné pouze při binomické distribuci.

- **Power.** $f(x) = x^\alpha$, pokud $\alpha \neq 0$. $f(x) = \log(x)$, pokud $\alpha = 0$. α je požadované číslo specifikace a musí být reálné číslo. Tento odkaz lze použít s jakoukoli distribucí.

Zobecněná odezva lineárních modelů

V mnoha případech můžete jednoduše určit závislou proměnnou; avšak proměnné, které mají pouze dvě hodnoty a odezvy, které zaznamenávají události ve zkouškách, vyžadují zvýšenou pozornost.

- **Binární odpověď.** Když má závislá proměnná pouze dvě hodnoty, můžete uvést referenční kategorii pro odhad parametru. Binární proměnná odezva může být řetězcová nebo číselná.
- **Počet událostí, které se vyskytují v sadě pokusů.** Pokud je odpovědí řada událostí, které se vyskytují v sadě pokusů, závislá proměnná obsahuje počet událostí a můžete vybrat další proměnnou obsahující počet pokusů. Alternativně, pokud je počet studií stejný pro všechny subjekty, pak mohou být studie specifikovány pomocí pevné hodnoty. Počet pokusů by měl být větší nebo roven počtu událostí pro každý případ. Události by měly být nezáporná celá čísla a zkoušky by měly být kladná celá čísla.

Pro ordinální víceúčelové modely můžete určit pořadí kategorií odezvy: vzestupně, sestupně nebo data (pořadí dat znamená, že první hodnota zjištěná v datech definuje první kategorii, poslední zjištěná hodnota definuje poslední kategorii).

Váha měřítka. Parametr měřítka je odhadovaný parametr modelu související s odchylkou odezvy. Váhy stupnice jsou "známé" hodnoty, které se mohou lišit od pozorování k pozorování. Je-li zadána proměnná váhy stupnice, parametr stupnice, který se vztahuje k rozptylu odezvy, se vydělí pro každé pozorování. Případy s hodnotami váhy měřítka, které jsou menší nebo rovny 0 nebo chybí, nejsou v analýze použity.

Zobecněná referenční kategorie lineárních modelů

V případě binární odezvy můžete zvolit referenční kategorii pro závislou proměnnou. To může ovlivnit určitý výstup, jako např. odhady parametrů a uložené hodnoty, ale nemělo by měnit vhodnost modelu. Pokud například vaše binární odezva vezme hodnoty 0 a 1:

- Ve výchozím nastavení procedura vytvoří poslední kategorii (s nejvyšší hodnotou), nebo 1 kategorii odkazu. V této situaci odhadují pravděpodobnosti ušetřené modelem pravděpodobnost, že daný případ vezme hodnotu 0, a odhady parametrů by měly být interpretovány tak, že se vztahují k pravděpodobnosti kategorie 0.
- Zadáte-li jako referenční kategorii první kategorii (s nejnižší hodnotou) nebo hodnotu 0, budou pravděpodobnosti uložené v modelu odhadovat pravděpodobnost, že daný případ bude mít hodnotu 1.
- Pokud určíte vlastní kategorii a vaše proměnná má definované popisky, můžete nastavit referenční kategorii výběrem hodnoty ze seznamu. To může být výhodné, když uprostřed určení modelu si přesně nepamatujete, jak byla konkrétní proměnná kódována.

Prediktory generalizovaných lineárních modelů

Karta Prediktory vám umožňuje určit faktory a kovariáty použité k sestavení efektů modelu a určit volitelné posunutí.

Faktory. Faktory jsou kategorické prediktory; mohou být číselné nebo řetězcové.

Kryty. Kovariáty jsou prediktory měřítka; musí být číselné.

Poznámka: Je-li odezva binomická s binárním formátem, procedura vypočítá odchylku a statistiku dobrého jména chí-kvadrát podle dílčích populací, které jsou založeny na křížové klasifikaci pozorovaných hodnot vybraných faktorů a kovariátů. Měli byste zachovat stejnou sadu prediktorů v rámci více spuštění procedury, abyste zajistili konzistentní počet dílčích populací.

Posunutí. Termín offsetu je "strukturální" prediktor. Jeho koeficient není odhadnut modelem, ale předpokládá se, že má hodnotu 1; proto se hodnoty offsetu jednoduše přičítají k lineárnímu prediktoru cíle. To je zvláště užitečné v Poissonových regresních modelech, kde každý případ může mít různé úrovně expozice vůči události, která je předmětem zájmu.

Například při modelování nehodovosti pro jednotlivé řidiče existuje důležitý rozdíl mezi řidičem, který byl v jedné nehodě za tři roky zkušeností a řidičem, který byl v jedné nehodě za 25 let! Počet nehod může být modelován jako Poissonova nebo negativní binomická odpověď s odkazem na protokol, pokud je přirozený záznam zkušeností řidiče zahrnut jako offset.

Jiné kombinace rozdělení a typů vazeb by vyžadovaly jiné transformace proměnné offsetu.

Volby zobecněných lineárních modelů

Tyto volby se použijí na všechny faktory uvedené na kartě Prediktory.

Uživatel-Chybějící hodnoty. Faktory musí mít platné hodnoty pro případ, který má být zahrnut do analýzy. Tyto ovládací prvky vám umožňují rozhodnout, zda jsou chybějící hodnoty uživatele považovány za platné mezi proměnnými faktoru.

Pořadí kategorií. To je důležité pro určení poslední úrovně faktoru, která může být spojena s redundantním parametrem v algoritmu odhadu. Změna pořadí kategorií může změnit hodnoty efektů na úrovni faktoru, protože tyto odhady parametrů jsou vypočteny vzhledem k "poslední" úrovni. Faktory lze řadit ve vzestupném pořadí od nejnižší k nejvyšší hodnotě, v sestupném pořadí od nejvyšší k nejnižší hodnotě nebo v "datovém pořadí". To znamená, že první hodnota zjištěná v datech definuje první kategorii a poslední zjištěná jedinečná hodnota definuje poslední kategorii.

Zobecněný model lineárních modelů

Určit efekty modelu

Výchozí model je pouze zachycení, takže musíte explicitně určit jiné efekty modelu. Případně můžete sestavit vnořené nebo nevnořené výrazy.

Nevnořené výrazy

Pro vybrané faktory a kovariáty:

Hlavní účinky

Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Interakce

Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné.

Faktoriál

Vytvoří všechny možné interakce a hlavní efekty vybraných proměnných.

Všechny 2-way

Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3-way

Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny 4-way

Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Vše 5-ti cestné

Vytvoří všechny možné pěticečné interakce vybraných proměnných.

Vnořené výrazy

V této proceduře můžete sestavit vnořené výrazy pro svůj model. Vnořené výrazy jsou užitečné pro modelování účinku faktoru nebo kovariátu, jehož hodnoty neinteragují s úrovněmi jiného faktoru. Například řetězec obchodu s potravinami může sledovat výdajové návyky svých zákazníků na několika místech obchodu. Vzhledem k tomu, že každý zákazník navštěvuje pouze jednu z těchto lokalit, lze efekt *Zákazník říci*, že je **vnořený v** efektu *Umístění úložiště*.

Dále můžete zahrnout efekty interakce, například polynomiální výrazy zahrnující stejné kovariát, nebo přidat více úrovní vnoření do vnořeného výrazu.

Omezení: Vnořené výrazy mají následující omezení:

- Všechny faktory v rámci interakce musí být jedinečné. Je-li tedy faktor A , je zadání $A^* A$ neplatné.
- Všechny faktory ve vnořeném efektu musí být jedinečné. Je-li tedy faktor A , je zadání $A (A)$ neplatné.
- Žádný efekt nemůže být vnořen do kovariátu. Je-li tedy A faktorem a X je kovariantní, pak určení $A (X)$ je neplatné.

Zobecněný odhad lineárních modelů

Odhadování parametrů. Ovládací prvky v této skupině vám umožňují uvést metody odhadu a poskytnout počáteční hodnoty pro odhady parametrů.

- . Můžete vybrat metodu odhadu parametru. Vyberte si mezi Newton-Raphson, Fisher bodování, nebo hybridní metoda, ve které Fisher bodování iterace jsou prováděny před přepnutím na Newton-Raphson metody. Je-li dosaženo konvergence během fáze Fisher bodování hybridní metody před maximálním počtem Fisher iterací je dosaženo, algoritmus pokračuje Newton-Raphson metody.
- **Metoda parametru Měřítka.** Můžete vybrat metodu odhadu parametru měřítka. Maximální pravděpodobnost společně odhaduje parametr měřítka s efekty modelu; Všimněte si, že tato volba není platná, pokud má odpověď negativní binomické, Poissonovo, binomické nebo multinomiální rozdělení. Odchylky a Pearsonovy volby chý kvadrát odhadují parametr měřítka z hodnoty těchto statistik. Alternativně můžete zadat pevnou hodnotu pro parametr měřítka.
- **Počáteční hodnoty.** Procedura automaticky vypočítá počáteční hodnoty pro parametry. Případně můžete zadat počáteční hodnoty pro odhady parametrů.

Maticе Covariance. Odhad založený na modelu je negativní z generalizované inverzní funkce k hesenské matici. Robustní (také nazývaný Huber/White/sandwich) odhad je "korigovaný" model-based estimator, který poskytuje konzistentní odhad kovariance, i když specifikace rozptylu a spojovacích funkcí je nesprávné.

Iterace. K dispozici jsou následující volby:

- **Maximální počet iterací.** Maximální počet iterací, které bude algoritmus provádět. Zadejte nezáporné celé číslo.
- **Maximální počet kroků na polovinu.** Při každé iteraci se velikost kroku sníží o faktor 0.5, dokud se nezvýší pravděpodobnost protokolu nebo dokud se nedosáhne maximální poloviny kroku. Zadejte kladné celé číslo.
- **Zkontrolovat oddělení datových bodů.** Je-li vybrána tato volba, algoritmus provede testy, aby se ujistil, že odhady parametrů mají jedinečné hodnoty. K oddělení dochází, když procedura může vytvořit model, který správně klasifikuje každý případ. Tato volba je k dispozici pro multinomiální odpovědi a binomické odpovědi s binárním formátem.

Kritéria konvergence. K dispozici jsou následující volby

- **Konvergence parametrů.** Je-li vybráno, algoritmus se zastaví po iteraci, ve které je absolutní nebo relativní změna v odhadech parametrů menší než zadaná hodnota, která musí být kladná.
- **Konvergence pravděpodobnosti protokolu.** Je-li vybrána tato volba, algoritmus se zastaví po iteraci, ve které je absolutní nebo relativní změna funkce log-Pravděpodobnost menší než zadaná hodnota, která musí být kladná.
- **Hessovská konvergence.** Pro absolutní specifikaci se předpokládá konvergence, pokud je statistika založená na hesenské konvergenci menší než uvedená kladná hodnota. Pro relativní specifikaci se předpokládá konvergence, pokud je statistika menší než součin zadané kladné hodnoty a absolutní hodnoty logaritmické pravděpodobnosti.

Tolerance singularity. Singulární (nebo neinvertibilní) matice mají lineárně závislé sloupce, což může způsobit vážné problémy pro algoritmus odhadu. Dokonce i téměř singulární matice mohou vést ke špatným výsledkům, takže postup bude zacházet s maticí, jejíž determinant je menší než tolerance jako singulární. Zadejte kladnou hodnotu.

Zobecněné počáteční hodnoty lineárních modelů

Jsou-li zadány počáteční hodnoty, musí být zadány pro všechny parametry (včetně redundantních parametrů) v modelu. V datové sadě musí být pořadí proměnných zleva doprava: *RowType_*, *VarName_*, *P1*, *P2*, ..., kde *RowType_* a *VarName_* jsou řetězcové proměnné a *P1*, *P2* jsou číselné proměnné odpovídající seřazenému seznamu parametrů.

- Počáteční hodnoty jsou zadány v záznamu s hodnotou *EST* pro proměnnou *RowType_*; skutečné počáteční hodnoty jsou uvedeny v proměnných *P1*, *P2*, Procedura ignoruje všechny záznamy, pro které má *RowType_* jinou hodnotu než *EST*, a také všechny záznamy nad prvním výskytem *RowType_* rovnající se *EST*.
- Zachycení, pokud je zahrnuto v modelu, nebo parametry prahové hodnoty, pokud má odezva multinomiální distribuci, musí být první počáteční hodnoty vypsány.
- Parametr měřítka, a má-li odezva záporné binomické rozdělení, záporný binomický parametr, musí být poslední zadané počáteční hodnoty.
- Pokud je v platnosti rozdělený soubor, proměnné musí začínat proměnnou nebo proměnnými rozděleného souboru v pořadí určeném při vytváření rozděleného souboru, následované *RowType_*, *VarName_*, *P1*, *P2*, ... jak je uvedeno výše. K rozdělení musí dojít v zadané datové sadě ve stejném pořadí jako v původní datové sadě.

Poznámka: Názvy proměnných *P1*, *P2*, ... nejsou povinné; procedura přijme všechny platné názvy proměnných pro parametry, protože mapování proměnných na parametry je založeno na pozici proměnné, ne na názvu proměnné. Všechny proměnné za posledním parametrem jsou ignorovány.

Struktura souboru pro počáteční hodnoty je stejná jako ta, která se používá při exportu modelu jako data; proto můžete použít konečné hodnoty z jednoho spuštění procedury jako vstup v následném spuštění.

Statistika zobecněných lineárních modelů

Efekty modelu. K dispozici jsou následující volby:

- **Typ analýzy.** Zadejte typ analýzy, která se má vytvořit. Analýza typu I je obecně vhodná, pokud máte a priori důvody pro řazení prediktorů v modelu, zatímco typ III je obecněji použitelný. Statistika Wald nebo poměr pravděpodobnosti se vypočítá na základě výběru ve skupině Statistika chí-kvadrát.
- **Intervaly hodnověrnosti.** Zadejte úroveň spolehlivosti větší než 50 a menší než 100. Wald intervaly jsou založeny na předpokladu, že parametry mají asymptotické normální rozdělení; profil pravděpodobnosti intervaly jsou přesnější, ale může být výpočetně drahé. Úroveň tolerance pro intervaly pravděpodobnosti profilu je kritérium použité k zastavení iteračního algoritmu použitého k výpočtu intervalů.
- **Funkce Log-Pravděpodobnost.** Tato volba řídí formát zobrazení funkce log-Pravděpodobnost. Plná funkce zahrnuje další termín, který je konstantní s ohledem na odhady parametrů; nemá žádný vliv na odhad parametrů a je vynechán ze zobrazení v některých softwarových produktech.

Tisk. K dispozici je následující výstup:

- **Souhrn zpracování případů.** Zobrazuje počet a procento případů zahrnutých a vyloučených z analýzy a tabulky souhrnu korelovaných dat.
- **Popisná statistika.** Zobrazuje popisnou statistiku a souhrnné informace o závislé proměnné, kovariátech a faktorech.
- **Informace o modelu.** Zobrazuje název datové sady, závislou proměnnou nebo události a zkušební proměnné, proměnnou posunutí, proměnnou váhy, rozdělení pravděpodobnosti a funkci odkazu.
- **Statistické údaje o dobré kondici.** Zobrazuje odchylky a škálované odchylky, Pearson chí-kvadrát a škálovaný Pearson chí-kvadrát, log-pravděpodobnost, Akaike je informační kritérium (AIC), konečný vzorek korigovaný AIC (AICC), Bayesian informační kritérium (BIC), a konzistentní AIC (CAIC).
- **Souhrnná statistika modelu.** Zobrazuje testy fit modelu, včetně statistik poměru pravděpodobnosti pro test fit modelu omnibus a statistiky pro kontrasty typu I nebo III pro každý efekt.

- **odhady parametrů.** Zobrazuje odhady parametrů a odpovídající statistiku testů a intervaly spolehlivosti. Kromě prvotních odhadů parametrů můžete volitelně zobrazit i odhady exponenciováných parametrů.
- **Matice Covariance pro odhady parametrů.** Zobrazí matici kovariance odhadovaného parametru.
- **Matice korelace pro odhady parametrů.** Zobrazí odhadovanou matici korelace parametrů.
- **Matice koeficientu kontrastu (L).** Zobrazí koeficienty kontrastu pro výchozí efekty a pro odhadované mezní střední hodnoty, pokud jsou požadovány na kartě EM střední hodnoty.
- **Obecné odhadnutelné funkce.** Zobrazí matice pro generování matic kontrastního koeficientu (L).
- **Historie iterací.** Zobrazí historii iterací pro odhady parametrů a pravděpodobnost protokolu a vytiskne poslední vyhodnocení vektoru gradientu a hesenské matice. Tabulka historie iterací zobrazuje odhady parametrů pro každých n th iterací začínajících 0th iterací (počáteční odhady), kde n je hodnota intervalu tisku. Je-li požadována historie iterací, vždy se zobrazí poslední iterace bez ohledu na n .
- **Lagrangeova multiplikační zkouška parametru stupnice nebo negativního binomického pomocného parametru.** Zobrazí statistiku testu multiplikátoru Lagrange pro posouzení platnosti parametru měřítka, který je vypočítán pomocí odchylky nebo Pearson chi-square nebo nastaven na pevné číslo pro normální, gama, inverzní Gaussian a Tweedie distribuce. Pro negativní binomické rozdělení se testuje fixní pomocný parametr.

Zobecněné lineární modely EM

Tato karta vám umožňuje zobrazit odhadované mezní střední hodnoty pro úrovně faktorů a interakce faktorů. Můžete také požadovat, aby se zobrazila celková odhadovaná střední hodnota. Pro ordinální multinomiální modely nejsou k dispozici odhadované mezní střední hodnoty.

Faktory a interakce. Tento seznam obsahuje faktory uvedené na kartě Prediktory a interakce faktorů uvedené na kartě Model. Kovariáty jsou vyloučeny z tohoto seznamu. Výrazy lze vybrat přímo z tohoto seznamu nebo je lze sloučit do výrazu interakce pomocí tlačítka **Podle ***.

Zobrazit střední hodnotu pro. Pro vybrané faktory a interakce faktorů se vypočtou odhadované střední hodnoty. Kontrast určuje, jak jsou testy hypotéz nastaveny pro porovnání odhadovaných středních hodnot. Jednoduchý kontrast vyžaduje referenční kategorii nebo úroveň faktoru, se kterou jsou ostatní porovnávány.

- **Po dvojicích.** Porovnávání po dvojicích se vypočítávají pro kombinace všech úrovní uvedených nebo odvozených faktorů. Toto je jediný dostupný kontrast pro interakce faktorů.
- **Jednoduché.** Porovná střední hodnotu každé úrovně se střední hodnotou určené úrovně. Tento typ kontrastu je užitečný v případě, že existuje řídicí skupina.
- **Odchylka.** Každá úroveň faktoru je porovnána s velkou střední hodnotou. Odchylky kontrasty nejsou ortogonální.
- **Rozdíl.** Porovná střední hodnotu každé úrovně (kromě první) se střední hodnotou předchozích úrovní. Někdy se jim říká reverzní Helmert kontrasty.
- **Helmert.** Porovná střední hodnotu každé úrovně faktoru (kromě poslední) se střední hodnotou následných úrovní.
- **Opakované.** Porovná střední hodnotu každé úrovně (kromě poslední) se střední hodnotou následující úrovně.
- **Polynomiální.** Porovná lineární efekt, kvadratický efekt, kubický efekt a tak dále. První stupeň volnosti obsahuje lineární efekt ve všech kategoriích; druhý stupeň volnosti, kvadratický efekt; a tak dále. Tyto kontrasty se často používají k odhadu polynomiálních trendů.

Měřítka. Odhadnuté mezní střední hodnoty lze vypočítat pro odezvu na základě původního měřítka závislé proměnné nebo pro lineární prediktor na základě závislé proměnné transformované spojovou funkcí.

Úprava pro více porovnání. Při provádění testů hypotéz s více kontrasty lze celkovou úroveň významnosti upravit z úrovní významnosti pro zahrnuté kontrasty. Tato skupina vám umožňuje zvolit metodu úpravy.

- **Nejméně významný rozdíl.** Tato metoda nekontroluje celkovou pravděpodobnost odmítnutí hypotéz, že některé lineární kontrasty se liší od hodnot nulové hypotézy.

- *Bonferroni*. Tato metoda upraví pozorovanou úroveň významnosti pro skutečnost, že se testuje více kontrastů.
- *Sekvenční bonferroni*. Jedná se o postup postupného odmítání Bonferroni, který je mnohem méně konzervativní, pokud jde o odmítání jednotlivých hypotéz, ale zachovává stejnou celkovou úroveň významnosti.
- *Sidak*. Tato metoda poskytuje přísnější meze než Bonferroni přístup.
- *Sekvenční sidak*. Jedná se o postupně odstupňovaný odkladný postup Sidak, který je z hlediska odmítnutí jednotlivých hypotéz mnohem méně konzervativní, ale zachovává stejnou celkovou úroveň významnosti.

Zobecněné uložení lineárních modelů

Zaškrtnuté položky jsou uloženy se zadaným názvem; můžete se rozhodnout přepsat existující proměnné se stejným názvem jako nové proměnné nebo se vyhnout konfliktům názvů pomocí přípon příloh, aby se nové názvy proměnných proměnily v jedinečné.

Předpokládaná hodnota střední hodnoty odezvy

Uloží hodnoty předpovězené modelem pro každý případ v původní metrice odezvy. Pokud je distribuce odpovědi binomická a závislá proměnná je binární, procedura uloží předpovězené pravděpodobnosti. Je-li distribuce odezvy multinomiální, popisek položky se stane **Kumulativní predikovaná pravděpodobnost** procedura uloží kumulativní predikovanou pravděpodobnost pro každou kategorii odpovědi, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.

Dolní mez intervalu spolehlivosti pro střední hodnotu odpovědi

Uloží dolní mez intervalu spolehlivosti pro střední hodnotu odezvy. Je-li distribuce odezvy víceúčelová, popisek položky se změní na **Dolní mez intervalu spolehlivosti pro kumulativní předpovědanou pravděpodobnost** procedura uloží dolní mez pro každou kategorii odezvy, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.

Horní mez intervalu spolehlivosti pro střední hodnotu odpovědi

Uloží horní mez intervalu spolehlivosti pro střední hodnotu odezvy. Je-li distribuce odezvy multinomiální, stane se popisek položky **Horní mez intervalu spolehlivosti pro kumulativní předpovědanou pravděpodobnost** procedura uloží horní mez pro každou kategorii odpovědi, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.

Předpokládaná kategorie

U modelů s binomickým rozdělením a binárně závislou proměnnou nebo s multinomiálním rozdělením se pro každý případ uloží kategorie předpovězené odezvy. Tato volba není k dispozici pro jiné distribuce odpovědí.

Předpokládaná hodnota lineárního prediktoru

Uloží modelově predikované hodnoty pro každý případ v metrice lineárního prediktoru (transformovaná odezva přes zadanou funkci spoje). Je-li distribuce odezvy multinomiální, procedura uloží předpovězenou hodnotu pro každou kategorii odezvy, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.

Odhadovaná standardní chyba predikované hodnoty lineárního prediktoru

Když je distribuce odezvy multinomiální, procedura uloží odhadovanou standardní chybu pro každou kategorii odezvy, kromě poslední, až do počtu určených kategorií, které se mají uložit.

Následující položky nejsou k dispozici, pokud je distribuce odpovědi multinomiální.

Vzdálenost kuchaře

Měřitko toho, kolik by se změnilo rezidua všech případů, kdyby byl konkrétní případ vyloučen z výpočtu regresních koeficientů. Velký Cook's D označuje, že vyloučení případu z výpočtu regresní statistiky podstatně změní koeficienty.

Hodnota pákového efektu

Měří vliv bodu na vhodnost regrese. Střední páka se pohybuje od 0 (bez vlivu na přizpůsobení) do $(N-1)/N$.

Zbývající

Rozdíl mezi pozorovanou hodnotou a hodnotou předpovězenou modelem.

Pearsonův zbytek

Druhá odmocnina příspěvku případu ke statistice Pearsonova chý kvadrát se znaménkem nezpracovaného zbytku.

Standardizovaný Pearsonův zbytkový

Zbytková hodnota Pearsonovy hodnoty vynásobená druhou odmocninou inverzní hodnoty parametru scale a hodnota 1-leverage pro daný případ.

Zbytková odchylka

Druhá odmocnina příspěvku případu ke statistice Deviance se znaméním nezpracovaného zbytku.

Zbytková normalizovaná odchylka

Zbytková odchylka vynásobená druhou odmocninou inverzní funkce k součinu parametru stupnice a 1-leverage pro případ.

Zbytková pravděpodobnost

Druhá odmocnina váženého průměru (na základě pákového efektu případu) čtverců standardizovaných Pearsonových a standardizovaných zbytkových zbytků Deviance se znaméním nezpracovaných zbytků.

Export zobecněných lineárních modelů

Exportovat model jako data. Zapiše datovou sadu ve formátu IBM SPSS Statistika obsahující matici korelace parametru nebo kovariance s odhady parametrů, standardními chybami, hodnotami významnosti a stupni volnosti. Pořadí proměnných v maticovém souboru je následující.

- **Rozdělit proměnné.** Jsou-li použity, všechny proměnné definující rozdělení.
- **RowType_.** Přijímá hodnoty (a popisky hodnot) *COV* (kovariance), *CORR* (korelace), *EST* (odhady parametrů), *SE* (standardní chyby), *SIG* (úroveň významnosti) a *DF* (stupně volnosti návrhu vzorku). Pro každý parametr modelu existuje samostatný případ s řádkovým typem *COV* (nebo *CORR*) plus samostatný případ pro každý z ostatních řádkových typů.
- **Název_proměnné.** Vstupem jsou hodnoty *P1*, *P2*, ..., odpovídající seřazenému seznamu všech odhadovaných parametrů modelu (s výjimkou parametrů měřítka nebo negativních binomických parametrů) pro řádkové typy *COV* nebo *CORR*s popisky hodnot odpovídajícími řetězcům parametrů zobrazeným v tabulce odhadů parametrů. Buňky jsou prázdné pro ostatní typy řádků.
- **P1, P2, ...** Tyto proměnné odpovídají uspořádanému seznamu všech parametrů modelu (případně včetně popisků stupnice a záporných binomických parametrů) s popisky proměnných, které odpovídají řetězcům parametrů uvedeným v tabulce odhadů parametrů, a přijímají hodnoty podle řádkového typu.

Pro redundantní parametry jsou všechny kovariance nastaveny na nulu, korelace jsou nastaveny na chybějící hodnotu systému; všechny odhady parametrů jsou nastaveny na nulu; a všechny standardní chyby, úroveň významnosti a zbytkové stupně volnosti jsou nastaveny na chybějící hodnotu systému.

Pro parametr měřítka jsou kovariance, korelace, úroveň významnosti a stupně volnosti nastaveny na chybějící hodnotu systému. Je-li parametr scale odhadnut na základě maximální pravděpodobnosti, je uvedena standardní chyba; jinak je nastaven na chybějící hodnotu systému.

Pro záporný binomický parametr jsou kovariance, korelace, úroveň významnosti a stupně volnosti nastaveny na chybějící hodnotu systému. Je-li záporný binomický parametr odhadnut pomocí maximální pravděpodobnosti, je uvedena standardní chyba; jinak je nastaven na systémovou chybějící hodnotu.

Pokud existují rozdělení, pak seznam parametrů musí být nashromážděn přes všechna rozdělení. V daném rozdělení mohou být některé parametry irelevantní; to není stejné jako redundantní. V případě irelevantních parametrů jsou všechny kovariance nebo korelace, odhady parametrů, standardní chyby, úroveň významnosti a stupně volnosti nastaveny na systémovou chybějící hodnotu.

Tento maticový soubor můžete použít jako počáteční hodnoty pro další odhad modelu. Všimněte si, že tento soubor není okamžitě použitelný pro další analýzy v jiných procedurách, které čtou maticový soubor, pokud tyto procedury neakceptují všechny zde exportované typy řádků. I tehdy byste měli dbát na to, aby všechny parametry v tomto maticovém souboru měly stejný význam pro proceduru čtení souboru.

Exportovat model jako XML. Uloží odhady parametrů a matici kovariance parametrů, je-li vybrána, ve formátu XML (PMML). Tento soubor modelu můžete použít k použití informací o modelu na jiné datové soubory pro účely přidělení skóre.

Další funkce příkazu GENLIN

Jazyk syntaxe příkazu vám také umožňuje:

- Počáteční hodnoty pro odhady parametrů zadejte jako seznam čísel (pomocí podpříkazu CRITERIA).
- Při výpočtu odhadovaných mezních hodnot (pomocí podpříkazu EMMEANS) opravte kovariáty na jiné hodnoty, než jsou jejich střední hodnoty.
- Zadejte vlastní polynomiální kontrasty pro odhadované mezní střední hodnoty (pomocí podpříkazu EMMEANS).
- Určete podmnožinu faktorů, pro které jsou zobrazeny odhadované mezní střední hodnoty pro porovnání s použitím určeného typu kontrastu (pomocí klíčových slov TABLES a COMPARE podpříkazu EMMEANS).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Zobecněné odhadovací rovnice

Procedura zobecněných odhadovacích rovnic rozšiřuje zobecněný lineární model tak, aby umožňovala analýzu opakovaných měření nebo jiných korelovaných pozorování, například klastrovaných dat.

Příklad. Veřejní zdravotničtí úředníci mohou použít zobecněné odhadové rovnice, aby se přizpůsobili opakovaným opatřením logistické regrese ke studiu účinků znečištění ovzduší na děti.

Aspekty zobecněného odhadu dat rovnic

Data. Odpověď může být škálovatelnost, počet, binární nebo události ve zkouškách. Předpokládá se, že faktory jsou kategorické. Za měřítka se považují kovariáty, váha stupnice a posunutí. Proměnné použité k definování předmětů nebo opakovaných měření v rámci předmětu nelze použít k definování odezvy, ale mohou sloužit jiným rolím v modelu.

Předpoklady. Předpokládá se, že případy jsou závislé na subjektech a nezávislé na subjektech. Korelační matice, která představuje závislosti uvnitř předmětu, se odhaduje jako součást modelu.

Získání generalizovaných odhadovacích rovnic

Z nabídek vyberte:

Analýza > Zobecněné lineární modely > Zobecněné odhadování rovnic ...

1. Vyberte jednu nebo více proměnných předmětu (další volby viz níže).

Kombinace hodnot určených proměnných by měla jedinečně definovat **předměty** v rámci datové sady. Například jedna proměnná *ID pacienta* by měla být dostatečná pro definování subjektů v jedné nemocnici, ale kombinace *ID nemocnice* a *ID pacienta* může být nezbytná, pokud nejsou identifikační čísla pacientů v nemocnicích jedinečná. V nastavení opakovaných opatření se pro každý předmět zaznamenávají vícenásobná pozorování, takže každý subjekt může v datovém souboru zaujímat více případů.

2. Na kartě Typ modelu uveďte funkci distribuce a propojení.
3. Na kartě Odezva vyberte závislou proměnnou.
4. Na kartě Prediktory vyberte faktory a kovariáty, které se mají použít při předpovídání závislé proměnné.
5. Na kartě Model zadejte efekty modelu pomocí vybraných faktorů a kovariátů.

Volitelně můžete na kartě Opakovaný zadat:

Proměnné v rámci předmětu. Kombinace hodnot proměnných v rámci předmětu definuje pořadí měření v rámci subjektů, a proto kombinace proměnných v rámci předmětu a předmětu jednoznačně definuje každé měření. Například kombinace *Období*, *ID nemocnice* a *ID pacienta* definuje pro každý případ konkrétní návštěvu kanceláře pro konkrétního pacienta v rámci konkrétní nemocnice.

Pokud je datová sada již setříděna tak, aby se opakovaná měření každého subjektu vyskytovala v souvislých blocích případů a ve správném pořadí, není nezbytně nutné specifikovat proměnnou v rámci předmětů, a můžete zrušit výběr volby **Seřadit případy podle předmětu a proměnných v rámci předmětu** a uložit dobu zpracování potřebnou k provedení (dočasného) řazení. Obecně platí, že je dobré používat proměnné v rámci předmětu, aby se zajistilo správné řazení měření.

Proměnné předmětu a v rámci předmětu nelze použít k definování odezvy, ale mohou provádět další funkce v modelu. Například *ID nemocnice* lze použít jako faktor v modelu.

Maticové kovariance. Odhad založený na modelu je negativní z generalizované inverzní funkce k hesenské matici. Robustní odhad (nazývaný také Huber/White/sandwich estimator) je "opravený" model-based estimator, který poskytuje konzistentní odhad kovariance, i když je pracovní korelační matice nesprávně specifikována. Tato specifikace platí pro parametry v části lineárního modelu generalizovaných odhadovaných rovnic, zatímco specifikace na kartě Odhad platí pouze pro počáteční generalizovaný lineární model.

Pracovní korelační matice. Tato korelační matice představuje závislosti v rámci předmětu. Jeho velikost je určena počtem měření a tím i kombinací hodnot proměnných v rámci předmětu. Můžete uvést jednu z následujících struktur:

- **Nezávislé.** Opakovaná měření nejsou korelována.
- **AR (1).** Opakovaná měření mají autoregresivní vztah prvního řádu. Korelace mezi libovolnými dvěma prvky se rovná ρ pro sousední prvky, ρ^2 pro prvky oddělené třetinou atd. je omezeno tak, aby $-1 < \rho < 1$.
- **Vyměnitelné.** Tato struktura má homogenní korelace mezi prvky. To je také známé jako sloučenina symetrie struktury.
- **M-dependentní.** Po sobě jdoucí měření mají společný korelační koeficient, páry měření oddělené třetinou mají společný korelační koeficient, a tak dále, prostřednictvím dvojic měření oddělených $m-1$ dalšími měřeními. Například, pokud dáte studentům standardizované testy každý rok od 3rd do 7th třídy. Tato struktura předpokládá, že skóre 3rd a 4th, 4th a 5th, 5th a 6th a 6th a 7th budou mít stejnou korelaci; 3rd a 5th, 4th a 6th a 5th a 7th budou mít stejnou korelaci; 3rd a 6th a 4th a 7th budou mít stejnou korelaci. Měření s oddělováním větším než m se považují za nekorelována. Při výběru této struktury zadejte hodnotu m menší než pořadí pracovní korelační matice.
- **Nestrukturovaný.** Jedná se o zcela obecnou korelační matici.

Při výchozím nastavení procedura upraví odhady korelace podle počtu neredundantních parametrů. Odebrání této úpravy může být žádoucí, pokud chcete, aby odhady byly invariantní na změny replikace na úrovni předmětu v datech.

- **Maximální počet iterací.** Maximální počet iterací, které provede algoritmus generalizovaných odhadovaných rovnic. Zadejte nezáporné celé číslo. Tato specifikace platí pro parametry v části lineárního modelu generalizovaných odhadovaných rovnic, zatímco specifikace na kartě Odhad platí pouze pro počáteční generalizovaný lineární model.
- **Aktualizovat matici.** Prvky v pracovní korelační matici jsou odhadnuty na základě odhadů parametrů, které jsou aktualizovány v každé iteraci algoritmu. Není-li pracovní korelační matice vůbec aktualizována, použije se počáteční pracovní korelační matice během celého procesu odhadu. Pokud je matice aktualizována, můžete určit interval iterace, ve kterém se mají aktualizovat pracovní prvky matice korelace. Uvedení hodnoty větší než 1 může zkrátit dobu zpracování.

Kritéria konvergence. Tyto specifikace platí pro parametry v části lineárního modelu generalizovaných odhadovaných rovnic, zatímco specifikace na kartě Odhad platí pouze pro počáteční generalizovaný lineární model.

- **Konvergence parametrů.** Je-li vybráno, algoritmus se zastaví po iteraci, ve které je absolutní nebo relativní změna v odhadech parametrů menší než zadaná hodnota, která musí být kladná.
- **Hesovská konvergence.** Konvergence se předpokládá, pokud je statistika založená na hesenské hodnotě menší než určená hodnota, která musí být kladná.

Typ modelu zobecněných odhadových rovnic

Karta Typ modelu vám umožňuje určit funkci distribuce a propojení pro váš model a poskytuje zástupce pro několik běžných modelů, které jsou kategorizovány podle typu odezvy.

Typy modelů

Měřítko odezvy. K dispozici jsou následující volby:

- **Lineární.** Uvádí Normální jako distribuci a Identitu jako funkci odkazu.
- **Gamma s odkazem na protokol.** Uvádí gama jako distribuci a protokol jako funkci odkazu.

Ordinální odpověď. K dispozici jsou následující volby:

- **ordinální logistika.** uvádí multinomiální (ordinální) jako distribuční a kumulativní logit jako funkci odkazu.
- **Ordinální zkouška.** uvádí multinomiální (ordinální) jako distribuci a kumulativní probit jako funkci odkazu.

Počty. K dispozici jsou následující volby:

- **Poisson loglinear.** Uvádí Poisson jako distribuci a Log jako funkci odkazu.
- **Negativní binomický s odkazem na protokol.** Uvádí záporný binomický (s hodnotou 1 pro pomocný parametr) jako distribuční a log jako funkci odkazu. Chcete-li, aby procedura odhadla hodnotu pomocného parametru, uveďte vlastní model s negativním binomickým rozdělením a vyberte volbu **Odhadnout hodnotu** ve skupině Parametr.

Binární odezva nebo data událostí/hodnocení. K dispozici jsou následující volby:

- **Binární logistika.** Uvádí Binomické jako distribuci a Logit jako funkci propojení.
- **Binární probit.** Uvádí Binomické jako distribuci a Probit jako funkci propojení.
- **Interval cenzurovaného přežití.** Uvádí Binomický jako distribuci a komplementární log-log jako funkci odkazu.

Směs. K dispozici jsou následující volby:

- **Tweedie s odkazem na protokol.** Určuje Tweedie jako distribuci a Log jako funkci odkazu.
- **Tweedie s odkazem na identitu.** Určuje Tweedie jako funkci distribuce a Identita jako funkci odkazu.

Vlastní. Určete vlastní kombinaci funkce distribuce a funkce propojení.

Rozdělení

Tato volba určuje rozdělení závislé proměnné. Schopnost určit nenormální rozdělení a neidentické spojovací funkce je základním zlepšením zobecněného lineárního modelu nad obecným lineárním modelem. Existuje mnoho možných kombinací funkcí distribuce a propojení a několik může být vhodné pro danou datovou sadu, takže vaše volba může být vedena a priori teoretickými úvahami nebo která kombinace se zdá, že nejlépe vyhovuje.

- **Binomické.** Toto rozdělení je vhodné pouze pro proměnné, které představují binární odezvu nebo počet událostí.
- **Gamma.** Toto rozdělení je vhodné pro proměnné s kladnými hodnotami měřítka, které jsou posunovány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.
- **Inverse Gaussian.** Toto rozdělení je vhodné pro proměnné s kladnými hodnotami měřítka, které jsou posunovány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.
- **Negativní binomické.** Toto rozdělení lze považovat za počet pokusů potřebných ke sledování úspěchů k a je vhodné pro proměnné s nezápornými celočíselnými hodnotami. Pokud je hodnota dat jiná než celé číslo, menší než 0 nebo chybí, pak se v analýze nepoužije odpovídající velikost písmen. Hodnota pomocného parametru záporného binomického rozdělení může být libovolné číslo větší nebo rovno

0; můžete jej nastavit na pevnou hodnotu nebo povolit, aby byl procedurou odhadnut. Je-li pomocný parametr nastaven na hodnotu 0, je použití tohoto rozdělení ekvivalentní použití Poissonova rozdělení.

- **Normální.** To je vhodné pro proměnné měřítka, jejichž hodnoty mají symetrickou, zvonovitou distribuci o centrální (střední) hodnotě. Závislá proměnná musí být číselná.
- **Poissonovo.** Toto rozdělení lze považovat za počet výskytů události, která je předmětem zájmu, v pevném časovém období a je vhodné pro proměnné s nezápornými celočíselnými hodnotami. Pokud je hodnota dat jiná než celé číslo, menší než 0 nebo chybí, pak se v analýze nepoužije odpovídající velikost písmen.
- **Tweedie.** Toto rozdělení je vhodné pro proměnné, které mohou být zastoupeny Poissonovými směsmi gama rozdělení; rozdělení je "smíšené" v tom smyslu, že kombinuje vlastnosti spojitých (bere nezáporné reálné hodnoty) a diskrétních rozdělení (kladná pravděpodobnostní hmotnost při jediné hodnotě, 0). Závislá proměnná musí být číselná, s datovými hodnotami většími nebo rovnými nule. Pokud je hodnota dat menší než nula nebo chybí, pak se v analýze nepoužije odpovídající případ. Pevná hodnota parametru distribuce Tweedie může být libovolné číslo větší než jedna a menší než dvě.
- **Multinomiální.** Toto rozdělení je vhodné pro proměnné, které představují ordinální odezvu. Závislá proměnná může být číselná nebo řetězcová a musí mít alespoň dvě odlišné platné datové hodnoty.

Funkce propojení

Funkce odkazu je transformací závislé proměnné, která umožňuje odhad modelu. K dispozici jsou následující funkce:

- **Identita.** $f(x) = x$. Závislá proměnná není transformována. Tento odkaz lze použít s jakoukoli distribucí.
- **Doplňkový protokol-log.** $f(x) = \log(-\log(1-x))$. To je vhodné pouze při binomické distribuci.
- **Kumulativní kavchit.** $f(x) = \tan(\pi(x - 0.5))$, použito pro kumulativní pravděpodobnost každé kategorie odpovědi. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní komplementární protokol-log.** $f(x) = \ln(-\ln(1-x))$, použitý na kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní protokol.** $f(x) = \ln(x/(1-x))$, použitý na kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní negativní protokol protokolu.** $f(x) = -\ln(-\ln(x))$, použito pro kumulativní pravděpodobnost každé kategorie odezvy. To je vhodné pouze s multinomiální distribucí.
- **Kumulativní zkouška.** $f(x) = \Phi^{-1}(x)$, použité na kumulativní pravděpodobnost každé kategorie odezvy, kde Φ^{-1} je inverzní standardní funkce normálního součtového rozdělení. To je vhodné pouze s multinomiální distribucí.
- **log.** $f(x) = \log(x)$. Tento odkaz lze použít s jakoukoli distribucí.
- **Doplňek protokolu.** $f(x) = \log(1-x)$. To je vhodné pouze při binomické distribuci.
- **Přihlásit se.** $f(x) = \log(x/(1-x))$. To je vhodné pouze při binomické distribuci.
- **Negativní binomický.** $f(x) = \log(x/(x+k^{-1}))$, kde k je pomocný parametr negativního binomického rozdělení. To je vhodné pouze s negativním binomickým rozdělením.
- **Protokol negativního protokolu.** $f(x) = -\log(-\log(x))$. To je vhodné pouze při binomické distribuci.
- **Kurzová síla.** $f(x) = [(x/(1-x))^{\alpha} - 1]/\alpha$, pokud $\alpha \neq 0$. $f(x) = \log(x)$, pokud $\alpha = 0$. α je požadované číslo specifikace a musí být reálné číslo. To je vhodné pouze při binomické distribuci.
- **Probit.** $f(x) = \Phi^{-1}(x)$, kde Φ^{-1} je inverzní standardní funkce normální kumulativní distribuce. To je vhodné pouze při binomické distribuci.
- **Power.** $f(x) = x^{\alpha}$, pokud $\alpha \neq 0$. $f(x) = \log(x)$, pokud $\alpha = 0$. α je požadované číslo specifikace a musí být reálné číslo. Tento odkaz lze použít s jakoukoli distribucí.

Odpověď na zobecněný odhad rovnic

V mnoha případech můžete jednoduše určit závislou proměnnou; avšak proměnné, které mají pouze dvě hodnoty a odezvy, které zaznamenávají události ve zkouškách, vyžadují zvýšenou pozornost.

- **Binární odpověď.** Když má závislá proměnná pouze dvě hodnoty, můžete uvést referenční kategorii pro odhad parametru. Binární proměnná odezvy může být řetězcová nebo číselná.
- **Počet událostí, které se vyskytují v sadě pokusů.** Pokud je odpovědí řada událostí, které se vyskytují v sadě pokusů, závislá proměnná obsahuje počet událostí a můžete vybrat další proměnnou obsahující počet pokusů. Alternativně, pokud je počet studií stejný pro všechny subjekty, pak mohou být studie specifikovány pomocí pevné hodnoty. Počet pokusů by měl být větší nebo roven počtu událostí pro každý případ. Události by měly být nezáporná celá čísla a zkoušky by měly být kladná celá čísla.

Pro ordinální víceúčelové modely můžete určit pořadí kategorií odezvy: vzestupně, sestupně nebo data (pořadí dat znamená, že první hodnota zjištěná v datech definuje první kategorii, poslední zjištěná hodnota definuje poslední kategorii).

Váha měřítka. Parametr měřítka je odhadovaný parametr modelu související s odchylkou odezvy. Váhy stupnice jsou "známé" hodnoty, které se mohou lišit od pozorování k pozorování. Je-li zadána proměnná váhy stupnice, parametr stupnice, který se vztahuje k rozptylu odezvy, se vydělí pro každé pozorování. Případy s hodnotami váhy měřítka, které jsou menší nebo rovny 0 nebo chybí, nejsou v analýze použity.

Referenční kategorie zobecněných odhadových rovnic

V případě binární odezvy můžete zvolit referenční kategorii pro závislou proměnnou. To může ovlivnit určitý výstup, jako např. odhady parametrů a uložené hodnoty, ale nemělo by měnit vhodnost modelu. Pokud například vaše binární odezva vezme hodnoty 0 a 1:

- Ve výchozím nastavení procedura vytvoří poslední kategorii (s nejvyšší hodnotou), nebo 1 kategorii odkazu. V této situaci odhadují pravděpodobnosti ušetřené modelem pravděpodobnost, že daný případ vezme hodnotu 0, a odhady parametrů by měly být interpretovány tak, že se vztahují k pravděpodobnosti kategorie 0.
- Zadáte-li jako referenční kategorii první kategorii (s nejnižší hodnotou) nebo hodnotu 0, budou pravděpodobnosti uložené v modelu odhadovat pravděpodobnost, že daný případ bude mít hodnotu 1.
- Pokud určíte vlastní kategorii a vaše proměnná má definované popisky, můžete nastavit referenční kategorii výběrem hodnoty ze seznamu. To může být výhodné, když uprostřed určení modelu si přesně nepamatujete, jak byla konkrétní proměnná kódována.

Generalizované odhadování prediktorů rovnic

Karta Prediktory vám umožňuje určit faktory a kovariáty použité k sestavení efektů modelu a určit volitelné posunutí.

Faktory. Faktory jsou kategoričké prediktory; mohou být číselné nebo řetězcové.

Kryty. Kovariáty jsou prediktory měřítka; musí být číselné.

Poznámka: Je-li odezva binomická s binárním formátem, procedura vypočítá odchylku a statistiku dobrého jména chí-kvadrát podle dílčích populací, které jsou založeny na křížové klasifikaci pozorovaných hodnot vybraných faktorů a kovariátů. Měli byste zachovat stejnou sadu prediktorů v rámci více spuštění procedury, abyste zajistili konzistentní počet dílčích populací.

Posunutí. Termín offsetu je "strukturální" prediktor. Jeho koeficient není odhadnut modelem, ale předpokládá se, že má hodnotu 1; proto se hodnoty offsetu jednoduše přičítají k lineárnímu prediktoru cíle. To je zvláště užitečné v Poissonových regresních modelech, kde každý případ může mít různé úrovně expozice vůči události, která je předmětem zájmu.

Například při modelování nehodovosti pro jednotlivé řidiče existuje důležitý rozdíl mezi řidičem, který byl v jedné nehodě za tři roky zkušeností a řidičem, který byl v jedné nehodě za 25 let! Počet nehod může být modelován jako Poissonova nebo negativní binomická odpověď s odkazem na protokol, pokud je přirozený záznam zkušeností řidiče zahrnut jako offset.

Jiné kombinace rozdělení a typů vazeb by vyžadovaly jiné transformace proměnné offsetu.

Volby zobecněných odhadovaných rovnic

Tyto volby se použijí na všechny faktory uvedené na kartě Prediktory.

Uživatel-Chybějící hodnoty. Faktory musí mít platné hodnoty pro případ, který má být zahrnut do analýzy. Tyto ovládací prvky vám umožňují rozhodnout, zda jsou chybějící hodnoty uživatele považovány za platné mezi proměnnými faktoru.

Pořadí kategorií. To je důležité pro určení poslední úrovně faktoru, která může být spojena s redundantním parametrem v algoritmu odhadu. Změna pořadí kategorií může změnit hodnoty efektů na úrovni faktoru, protože tyto odhady parametrů jsou vypočteny vzhledem k "poslední" úrovni. Faktory lze řadit ve vzestupném pořadí od nejnižší k nejvyšší hodnotě, v sestupném pořadí od nejvyšší k nejnižší hodnotě nebo v "datovém pořadí". To znamená, že první hodnota zjištěná v datech definuje první kategorii a poslední zjištěná jedinečná hodnota definuje poslední kategorii.

Model zobecněných odhadových rovnic

Určete efekty modelu. Výchozí model je pouze zachycení, takže musíte explicitně určit jiné efekty modelu. Případně můžete sestavit vnořené nebo nevnořené výrazy.

Nevnořené podmínky

Pro vybrané faktory a kovariáty:

Hlavní efekty. Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Interakce. Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné.

Faktoriál. Vytvoří všechny možné interakce a hlavní efekty vybraných proměnných.

Všechny obousměrné. Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3 cesty. Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Celé 4-cestné. Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Všechny 5-cestné. Vytvoří všechny možné pěticečné interakce vybraných proměnných.

Vnořené výrazy

V této proceduře můžete sestavit vnořené výrazy pro svůj model. Vnořené výrazy jsou užitečné pro modelování účinku faktoru nebo kovariátu, jehož hodnoty neinteragují s úrovněmi jiného faktoru. Například řetězec obchodu s potravinami může sledovat výdajové návyky svých zákazníků na několika místech obchodu. Vzhledem k tomu, že každý zákazník navštěvuje pouze jednu z těchto lokalit, lze efekt *Zákazník* říci, že je **vnořený v** efektu *Umístění úložiště*.

Dále můžete do vnořeného výrazu zahrnout efekty interakce nebo přidat více úrovní vnoření.

Omezení. Vnořené výrazy mají následující omezení:

- Všechny faktory v rámci interakce musí být jedinečné. Je-li tedy faktor A , je zadání $A * A$ neplatné.
- Všechny faktory ve vnořeném efektu musí být jedinečné. Je-li tedy faktor A , je zadání $A (A)$ neplatné.
- Žádný efekt nemůže být vnořen do kovariátu. Je-li tedy A faktorem a X je kovariantní, pak určení $A (X)$ je neplatné.

Zachycení. Zachycení je obvykle zahrnuto v modelu. Pokud můžete předpokládat, že data procházejí výchozím bodem, můžete toto zachycení vyloučit.

Modely s multinomálním pořadovým rozdělením nemají jediný termín zachycení; místo toho existují parametry prahové hodnoty, které definují přechodové body mezi sousedními kategoriemi. Prahové hodnoty jsou vždy zahrnuty do modelu.

Zobecněný odhad rovnic

Odhadování parametrů. Ovládací prvky v této skupině vám umožňují uvést metody odhadu a poskytnout počáteční hodnoty pro odhady parametrů.

- Můžete si vybrat metodu odhadu parametrů; vyberte mezi Newton-Raphson, Fisher skóre, nebo hybridní metodu, ve které Fisher bodování iterace jsou prováděny před přepnutím na Newton-Raphson metody. Je-li dosaženo konvergence během fáze Fisher bodování hybridní metody před maximálním počtem Fisher iterací je dosaženo, algoritmus pokračuje Newton-Raphson metody.
- **Metoda parametru měřítka.** Můžete vybrat metodu odhadu parametru měřítka.

Maximální pravděpodobnost společně odhaduje parametr stupnice s efekty modelu; Všimněte si, že tato možnost není platná, pokud má odpověď negativní binomické, Poissonovo nebo binomické rozdělení. Vzhledem k tomu, že pojem pravděpodobnosti nevstupuje do zobecněných odhadových rovnic, tato specifikace se vztahuje pouze na počáteční zobecněný lineární model; tento odhad parametrů stupnice je pak předán zobecněným odhadujícím rovnicím, které aktualizují parametr stupnice Pearsonovým chí-kvadratem děleným jeho stupněm volnosti.

Možnosti deviance a Pearson chí-kvadrát odhadují parametr stupnice z hodnoty těchto statistik v počátečním zobecněném lineárním modelu; tento odhad parametru stupnice je pak předán zobecněným odhadovaným rovnicím, které jej považují za fixní.

Případně zadejte pevnou hodnotu pro parametr měřítka. Bude se s ním zacházet jako s fixním odhadem počátečního zobecněného lineárního modelu a zobecněných odhadových rovnic.

- **Počáteční hodnoty.** Procedura automaticky vypočítá počáteční hodnoty pro parametry. Případně můžete zadat počáteční hodnoty pro odhady parametrů.

Iterace a kritéria konvergence uvedená na této kartě jsou použitelná pouze pro počáteční generalizovaný lineární model. Kritéria odhadu použítá při přizpůsobení generalizovaných rovnic odhadu viz karta Opakovaný .

Iterace. K dispozici jsou následující volby:

- **Maximální počet iterací.** Maximální počet iterací, které bude algoritmus provádět. Zadejte nezáporné celé číslo.
- **Maximální počet kroků na polovinu.** Při každé iteraci se velikost kroku sníží o faktor 0.5 , dokud se nezvýší pravděpodobnost protokolu nebo dokud se nedosáhne maximální poloviny kroku. Zadejte kladné celé číslo.
- **Zkontrolovat oddělení datových bodů.** Je-li vybrána tato volba, algoritmus provede testy, aby se ujistil, že odhady parametrů mají jedinečné hodnoty. K oddělení dochází, když procedura může vytvořit model, který správně klasifikuje každý případ. Tato volba je k dispozici pro multinomiální odpovědi a binomické odpovědi s binárním formátem.

Kritéria konvergence. K dispozici jsou následující volby

- **Konvergence parametrů.** Je-li vybráno, algoritmus se zastaví po iteraci, ve které je absolutní nebo relativní změna v odhadech parametrů menší než zadaná hodnota, která musí být kladná.
- **Konvergence pravděpodobnosti protokolu.** Je-li vybrána tato volba, algoritmus se zastaví po iteraci, ve které je absolutní nebo relativní změna funkce log-Pravděpodobnost menší než zadaná hodnota, která musí být kladná.
- **Hesovská konvergence.** Pro absolutní specifikaci se předpokládá konvergence, pokud je statistika založená na hesenské konvergenci menší než uvedená kladná hodnota. Pro relativní specifikaci se předpokládá konvergence, pokud je statistika menší než součin zadané kladné hodnoty a absolutní hodnoty logaritmicke pravděpodobnosti.

Tolerance singularity. Singulární (nebo neinvertibilní) matice mají lineárně závislé sloupce, což může způsobit vážné problémy pro algoritmus odhadu. Dokonce i téměř singulární matice mohou vést ke špatným výsledkům, takže postup bude zacházet s maticí, jejíž determinant je menší než tolerance jako singulární. Zadejte kladnou hodnotu.

Zobecněné odhadování počátečních hodnot rovnic

Procedura odhaduje počáteční zobecněný lineární model a odhady z tohoto modelu se používají jako počáteční hodnoty pro odhady parametrů v části lineárního modelu zobecněných odhadovaných rovnic. Počáteční hodnoty nejsou potřebné pro pracovní korelační matici, protože maticové prvky jsou založeny

na odhadech parametrů. Počáteční hodnoty uvedené v tomto dialogovém okně se použijí jako počáteční bod pro počáteční zobecněný lineární model, nikoli zobecněný odhad rovnic, pokud není maximální počet iterací na kartě Odhad nastaven na hodnotu 0.

Jsou-li zadány počáteční hodnoty, musí být zadány pro všechny parametry (včetně redundantních parametrů) v modelu. V datové sadě musí být pořadí proměnných zleva doprava: *RowType_*, *VarName_*, *P1*, *P2*, ..., kde *RowType_* a *VarName_* jsou řetězcové proměnné a *P1*, *P2* jsou číselné proměnné odpovídající seřazenému seznamu parametrů.

- Počáteční hodnoty jsou zadány v záznamu s hodnotou *EST* pro proměnnou *RowType_*; skutečné počáteční hodnoty jsou uvedeny v proměnných *P1*, *P2*, Procedura ignoruje všechny záznamy, pro které má *RowType_* jinou hodnotu než *EST*, a také všechny záznamy nad prvním výskytem *RowType_* rovnající se *EST*.
- Zachycení, pokud je zahrnuto v modelu, nebo parametry prahové hodnoty, pokud má odezva multinomiální distribuci, musí být první počáteční hodnoty vypsány.
- Parametr měřítka, a má-li odezva záporné binomické rozdělení, záporný binomický parametr, musí být poslední zadané počáteční hodnoty.
- Pokud je v platnosti rozdělený soubor, proměnné musí začínat proměnnou nebo proměnnými rozděleného souboru v pořadí určeném při vytváření rozděleného souboru, následované *RowType_*, *VarName_*, *P1*, *P2*, ... jak je uvedeno výše. K rozdělení musí dojít v zadané datové sadě ve stejném pořadí jako v původní datové sadě.

Poznámka: Názvy proměnných *P1*, *P2*, ... nejsou povinné; procedura přijme všechny platné názvy proměnných pro parametry, protože mapování proměnných na parametry je založeno na pozici proměnné, ne na názvu proměnné. Všechny proměnné za posledním parametrem jsou ignorovány.

Struktura souboru pro počáteční hodnoty je stejná jako ta, která se používá při exportu modelu jako data; proto můžete použít konečné hodnoty z jednoho spuštění procedury jako vstup v následném spuštění.

Statistika zobecněných odhadovaných rovnic

Efekty modelu. K dispozici jsou následující volby:

- **Typ analýzy.** Zadejte typ analýzy, která se má vytvořit pro efekty testovacího modelu. Analýza typu I je obecně vhodná, pokud máte a priori důvody pro řazení prediktorů v modelu, zatímco typ III je obecněji použitelný. Wald nebo generalizované statistiky skóre se vypočítávají na základě výběru ve skupině Statistiky chí-kvadrát.
- **Intervaly hodnověrnosti.** Zadejte úroveň spolehlivosti větší než 50 a menší než 100. Wald intervaly jsou vždy vyrobeny bez ohledu na typ chí-kvadrát statistiky vybraných, a jsou založeny na předpokladu, že parametry mají asymptotické normální rozdělení.
- **Protokolovat funkci kvazi-pravděpodobnosti.** Tato volba řídí formát zobrazení funkce kvazi-pravděpodobnosti protokolu. Plná funkce zahrnuje další termín, který je konstantní s ohledem na odhady parametrů; nemá žádný vliv na odhad parametrů a je vynechán ze zobrazení v některých softwarových produktech.

Tisk. K dispozici je následující výstup.

- **Souhrn zpracování případů.** Zobrazuje počet a procento případů zahrnutých a vyloučených z analýzy a tabulky souhrnu korelovaných dat.
- **Popisná statistika.** Zobrazuje popisnou statistiku a souhrnné informace o závislé proměnné, kovariátech a faktorech.
- **Informace o modelu.** Zobrazuje název datové sady, závislou proměnnou nebo události a zkušební proměnné, proměnnou posunutí, proměnnou váhy, rozdělení pravděpodobnosti a funkci odkazu.
- **Statistické údaje o dobré kondici.** Zobrazí dvě rozšíření kritéria Akaike's Information Criterion pro výběr modelu: Kvazi-pravděpodobnost podle kritéria modelu nezávislosti (QIC) pro výběr nejlepší korelační struktury a další ukazatel QIC pro výběr nejlepší podmnožiny prediktorů.
- **Souhrnná statistika modelu.** Zobrazuje testy fit modelu, včetně statistik poměru pravděpodobnosti pro test fit modelu omnibus a statistiky pro kontrasty typu I nebo III pro každý efekt.

- **odhady parametrů.** Zobrazuje odhady parametrů a odpovídající statistiku testů a intervaly spolehlivosti. Kromě prvotních odhadů parametrů můžete volitelně zobrazit i odhady exponenciováných parametrů.
- **Matice Covariance pro odhady parametrů.** Zobrazí matici kovariance odhadovaného parametru.
- **Matice korelace pro odhady parametrů.** Zobrazí odhadovanou matici korelace parametrů.
- **Matice koeficientu kontrastu (L).** Zobrazí koeficienty kontrastu pro výchozí efekty a pro odhadované mezní střední hodnoty, pokud jsou požadovány na kartě EM střední hodnoty.
- **Obecné odhadnutelné funkce.** Zobrazí matice pro generování matic kontrastního koeficientu (L).
- **Historie iterací.** Zobrazí historii iterací pro odhady parametrů a pravděpodobnost protokolu a vytiskne poslední vyhodnocení vektoru gradientu a hesenské matice. Tabulka historie iterací zobrazuje odhady parametrů pro každých n^{th} iterací začínajících 0th iterací (počáteční odhady), kde n je hodnota intervalu tisku. Je-li požadována historie iterací, vždy se zobrazí poslední iterace bez ohledu na n .
- **Pracovní korelační matice.** Zobrazí hodnoty matice představující závislosti v rámci předmětu. Jeho struktura závisí na specifikacích na kartě Opakovaný.

Zobecněné odhadování rovnic EM prostředky

Tato karta vám umožňuje zobrazit odhadované mezní střední hodnoty pro úroveň faktorů a interakce faktorů. Můžete také požadovat, aby se zobrazila celková odhadovaná střední hodnota. Pro ordinální multinomiální modely nejsou k dispozici odhadované mezní střední hodnoty.

Faktory a interakce. Tento seznam obsahuje faktory uvedené na kartě Prediktory a interakce faktorů uvedené na kartě Model. Kovariáty jsou vyloučeny z tohoto seznamu. Výrazy lze vybrat přímo z tohoto seznamu nebo je lze sloučit do výrazu interakce pomocí tlačítka **Podle ***.

Zobrazit střední hodnotu pro. Pro vybrané faktory a interakce faktorů se vypočtou odhadované střední hodnoty. Kontrast určuje, jak jsou testy hypotéz nastaveny pro porovnání odhadovaných středních hodnot. Jednoduchý kontrast vyžaduje referenční kategorii nebo úroveň faktoru, se kterou jsou ostatní porovnávány.

- **Po dvojicích.** Porovnávání po dvojicích se vypočítávají pro kombinace všech úrovní uvedených nebo odvozených faktorů. Toto je jediný dostupný kontrast pro interakce faktorů.
- **Jednoduché.** Porovná střední hodnotu každé úrovně se střední hodnotou určené úrovně. Tento typ kontrastu je užitečný v případě, že existuje řídicí skupina.
- **Odchylka.** Každá úroveň faktoru je porovnána s velkou střední hodnotou. Odchylky kontrasty nejsou ortogonální.
- **Rozdíl.** Porovná střední hodnotu každé úrovně (kromě první) se střední hodnotou předchozích úrovní. Někdy se jim říká reverzní Helmert kontrasty.
- **Helmert.** Porovná střední hodnotu každé úrovně faktoru (kromě poslední) se střední hodnotou následných úrovní.
- **Opakované.** Porovná střední hodnotu každé úrovně (kromě poslední) se střední hodnotou následující úrovně.
- **Polynomiální.** Porovná lineární efekt, kvadratický efekt, kubický efekt a tak dále. První stupeň volnosti obsahuje lineární efekt ve všech kategoriích; druhý stupeň volnosti, kvadratický efekt; a tak dále. Tyto kontrasty se často používají k odhadu polynomiálních trendů.

Měřitko. Odhadnuté mezní střední hodnoty lze vypočítat pro odezvu na základě původního měřítka závislé proměnné nebo pro lineární prediktor na základě závislé proměnné transformované spojovou funkcí.

Úprava pro více porovnání. Při provádění testů hypotéz s více kontrasty lze celkovou úroveň významnosti upravit z úrovní významnosti pro zahrnuté kontrasty. Tato skupina vám umožňuje zvolit metodu úpravy.

- **Nejméně významný rozdíl.** Tato metoda nekontroluje celkovou pravděpodobnost odmítnutí hypotéz, že některé lineární kontrasty se liší od hodnot nulové hypotézy.
- **Bonferroni.** Tato metoda upraví pozorovanou úroveň významnosti pro skutečnost, že se testuje více kontrastů.

- *Sekvenční bonferroni*. Jedná se o postup postupného odmítání Bonferroni, který je mnohem méně konzervativní, pokud jde o odmítání jednotlivých hypotéz, ale zachovává stejnou celkovou úroveň významnosti.
- *Sidak*. Tato metoda poskytuje přísnější meze než Bonferroni přístup.
- *Sekvenční sidak*. Jedná se o postupně odstupňovaný odkladný postup Sidak, který je z hlediska odmítnutí jednotlivých hypotéz mnohem méně konzervativní, ale zachovává stejnou celkovou úroveň významnosti.

Uložení generalizovaných odhadovaných rovnic

Zaškrtnuté položky jsou uloženy se zadaným názvem; můžete se rozhodnout přepsat existující proměnné se stejným názvem jako nové proměnné nebo se vyhnout konfliktům názvů pomocí přípon příloh, aby se nové názvy proměnných proměnily v jedinečné.

- **Předpokládaná hodnota střední hodnoty odezvy**. Uloží hodnoty předpovězené modelem pro každý případ v původní matici odezvy. Pokud je distribuce odpovědi binomická a závislá proměnná je binární, procedura uloží předpovězené pravděpodobnosti. Je-li distribuce odezvy multinomiální, popisek položky se stane **Kumulativní predikovaná pravděpodobnost** procedura uloží kumulativní predikovanou pravděpodobnost pro každou kategorii odpovědi, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.
- **Dolní mez intervalu spolehlivosti pro střední hodnotu odezvy**. Uloží dolní mez intervalu spolehlivosti pro střední hodnotu odezvy. Je-li distribuce odezvy víceúčelová, popisek položky se změní na **Dolní mez intervalu spolehlivosti pro kumulativní předpovědanou pravděpodobnost** procedura uloží dolní mez pro každou kategorii odezvy, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.
- **Horní mez intervalu spolehlivosti pro střední hodnotu odezvy**. Uloží horní mez intervalu spolehlivosti pro střední hodnotu odezvy. Je-li distribuce odezvy multinomiální, stane se popisek položky **Horní mez intervalu spolehlivosti pro kumulativní předpovědanou pravděpodobnost** procedura uloží horní mez pro každou kategorii odpovědi, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.
- **Předpovězená kategorie**. U modelů s binomickým rozdělením a binárně závislou proměnnou nebo s multinomiálním rozdělením se pro každý případ uloží kategorie předpovězené odezvy. Tato volba není k dispozici pro jiné distribuce odpovědi.
- **Předpovědaná hodnota lineárního prediktoru**. Uloží modelově predikované hodnoty pro každý případ v matici lineárního prediktoru (transformovaná odezva přes zadanou funkci spoje). Je-li distribuce odezvy multinomiální, procedura uloží předpovězenou hodnotu pro každou kategorii odezvy, s výjimkou poslední, až do počtu určených kategorií, které se mají uložit.
- **Odhadovaná standardní chyba předpovězené hodnoty lineárního prediktoru**. Když je distribuce odezvy multinomiální, procedura uloží odhadovanou standardní chybu pro každou kategorii odezvy, kromě poslední, až do počtu určených kategorií, které se mají uložit.

Následující položky nejsou k dispozici, pokud je distribuce odpovědi multinomiální.

- *Nezpracovaná zbytková*. Rozdíl mezi pozorovanou hodnotou a hodnotou předpovězenou modelem.
- **Pearsonův zbytek**. Druhá odmocnina příspěvku případu ke statistice Pearsonova chí kvadrát se znaménkem nezpracovaného zbytku.

Export generalizovaných odhadových rovnic

Exportovat model jako data. Zapiše datovou sadu ve formátu IBM SPSS Statistika obsahující matici korelace parametru nebo kovariance s odhady parametrů, standardními chybami, hodnotami významnosti a stupni volnosti. Pořadí proměnných v maticovém souboru je následující.

- **Rozdělit proměnné**. Jsou-li použity, všechny proměnné definující rozdělení.
- **RowType_**. Přijímá hodnoty (a popisky hodnot) *COV* (kovariance), *CORR* (korelace), *EST* (odhady parametrů), *SE* (standardní chyby), *SIG* (úroveň významnosti) a *DF* (stupně volnosti návrhu vzorku). Pro každý parametr modelu existuje samostatný případ s řádkovým typem *COV* (nebo *CORR*) plus samostatný případ pro každý z ostatních řádkových typů.

- **Název_proměnné.** Vstupem jsou hodnoty $P1, P2, \dots$, odpovídající seřazenému seznamu všech odhadovaných parametrů modelu (s výjimkou parametrů měřítka nebo negativních binomických parametrů) pro řádkové typy *COV* nebo *CORR*s popisky hodnot odpovídajícími řetězcům parametrů zobrazeným v tabulce odhadů parametrů. Buňky jsou prázdné pro ostatní typy řádků.
- **P1, P2, ...** Tyto proměnné odpovídají uspořádanému seznamu všech parametrů modelu (případně včetně popisků stupnice a záporných binomických parametrů) s popisky proměnných, které odpovídají řetězcům parametrů uvedeným v tabulce odhadů parametrů, a přijímají hodnoty podle řádkového typu.

Pro redundantní parametry jsou všechny kovariance nastaveny na nulu, korelace jsou nastaveny na chybějící hodnotu systému; všechny odhady parametrů jsou nastaveny na nulu; a všechny standardní chyby, úroveň významnosti a zbytkové stupně volnosti jsou nastaveny na chybějící hodnotu systému.

Pro parametr měřítka jsou kovariance, korelace, úroveň významnosti a stupně volnosti nastaveny na chybějící hodnotu systému. Je-li parametr scale odhadnut na základě maximální pravděpodobnosti, je uvedena standardní chyba; jinak je nastaven na chybějící hodnotu systému.

Pro záporný binomický parametr jsou kovariance, korelace, úroveň významnosti a stupně volnosti nastaveny na chybějící hodnotu systému. Je-li záporný binomický parametr odhadnut pomocí maximální pravděpodobnosti, je uvedena standardní chyba; jinak je nastaven na systémovou chybějící hodnotu.

Pokud existují rozdělení, pak seznam parametrů musí být nashromážděn přes všechna rozdělení. V daném rozdělení mohou být některé parametry irelevantní; to není stejné jako redundantní. V případě irelevantních parametrů jsou všechny kovariance nebo korelace, odhady parametrů, standardní chyby, úroveň významnosti a stupně volnosti nastaveny na systémovou chybějící hodnotu.

Tento maticový soubor můžete použít jako počáteční hodnoty pro další odhad modelu. Všimněte si, že tento soubor není okamžitě použitelný pro další analýzy v jiných procedurách, které čtou maticový soubor, pokud tyto procedury neakceptují všechny zde exportované typy řádků. I tehdy byste měli dbát na to, aby všechny parametry v tomto maticovém souboru měly stejný význam pro proceduru čtení souboru.

Exportovat model jako XML. Uloží odhady parametrů a matici kovariance parametrů, je-li vybrána, ve formátu XML (PMML). Tento soubor modelu můžete použít k použití informací o modelu na jiné datové soubory pro účely přidělení skóre.

Další funkce příkazu GENLIN

Jazyk syntaxe příkazu vám také umožňuje:

- Počáteční hodnoty pro odhady parametrů zadejte jako seznam čísel (pomocí podpříkazu *CRITERIA*).
- Určete pevnou pracovní korelační matici (pomocí podpříkazu *REPEATED*).
- Při výpočtu odhadovaných mezních hodnot (pomocí podpříkazu *EMMEANS*) opravte kovariáty na jiné hodnoty, než jsou jejich střední hodnoty.
- Zadejte vlastní polynomiální kontrasty pro odhadované mezní střední hodnoty (pomocí podpříkazu *EMMEANS*).
- Určete podmnožinu faktorů, pro které jsou zobrazeny odhadované mezní střední hodnoty pro porovnání s použitím určeného typu kontrastu (pomocí klíčových slov *TABLES* a *COMPARE* podpříkazu *EMMEANS*).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Zobecněné lineární smíšené modely

Zobecněné lineární smíšené modely rozšiřují lineární model tak, aby:

- Cíl je lineárně spojen s faktory a kovariáty prostřednictvím specifikované funkce odkazu.
- Cíl může mít jinou než normální distribuci.
- Pozorování lze korelovat.

Zobecněné lineární smíšené modely pokrývají širokou škálu modelů, od jednoduché lineární regrese až po složité víceúrovňové modely pro nenormální podélná data.

Příklady

Okresní školní rada může použít zobecněný lineární smíšený model k určení, zda experimentální výuková metoda je účinná při zlepšování matematických skóre. Studenti ze stejné třídy by měli být korelováni, protože jsou vyučováni stejným učitelem, a třídy ve stejné škole mohou být také korelované, takže můžeme zahrnout náhodné efekty na úrovni školy a třídy, aby se zohlednily různé zdroje variability.

Lékařští výzkumníci mohou použít generalizovaný lineární smíšený model k určení, zda nový antikonvulzivní lék může snížit míru epileptických záchvatů pacienta. Opakovaná měření od stejného pacienta jsou obvykle pozitivně korelována, takže by měl být vhodný smíšený model s některými náhodnými účinky. Cílové pole, počet záchvatů, má kladné celé hodnoty, takže může být vhodný generalizovaný lineární smíšený model s Poissonovou distribucí a logem.

Manažeři u poskytovatele kabelové televize, telefonu a internetových služeb mohou používat zobecněný lineární smíšený model, aby se dozvěděli více o potenciálních zákaznících. Vzhledem k tomu, že možné odpovědi mají nominální úroveň měření, analytik společnosti používá generalizovaný logit smíšený model s náhodným zachycením korelace mezi odpověďmi na otázky týkající se využití služeb napříč typy služeb (tv, telefon, internet) v rámci odpovědí respondenta daného průzkumu.

Karta Datová struktura vám umožňuje uvést strukturální vztahy mezi záznamy v datové sadě, když jsou pozorování korelována. Pokud záznamy v datové sadě představují nezávislá pozorování, nemusíte na této kartě nic uvádět.

Volby efektů

Předměty

Kombinace hodnot určených kategorických polí by měla jedinečně definovat předměty v rámci datové sady. Například jedno pole *ID pacienta* by mělo být dostatečné pro definování subjektů v jedné nemocnici, ale kombinace *ID nemocnice* a *ID pacienta* může být nezbytná, pokud identifikační čísla pacientů nejsou v nemocnicích jedinečná. Při nastavení opakovaných opatření se pro každý předmět zaznamenávají vícenásobná pozorování, takže každý subjekt může v datové sadě obsadit více záznamů.

Předmět je pozorovací jednotka, kterou lze považovat za nezávislou na jiných předmětech. Například údaje o krevním tlaku od pacienta v lékařské studii lze považovat za nezávislé na hodnotách od jiných pacientů. Definování subjektů je důležité zejména v případě, že se provádí opakované měření na jeden subjekt a vy chcete modelovat korelaci mezi těmito pozorováními. Například můžete očekávat, že hodnoty krevního tlaku od jednoho pacienta během následných návštěv u lékaře jsou korelované.

Všechna pole zadaná jako **Subjekty** v dialogovém okně Proměnné se používají k definování předmětů pro strukturu zbytkových kovariancí a poskytují seznam možných polí pro definování předmětů pro strukturu kovariance náhodných efektů v Bloku náhodných efektů.

Opakovaná měřítka

Zde uvedená pole se používají k identifikaci opakovaných pozorování. Například jedna proměnná *Týden* může identifikovat 10 týdnů pozorování v lékařské studii, nebo *Měsíc* a *Den* mohou být použity společně k identifikaci denních pozorování v průběhu roku.

Volby kovariance

Definovat kovarianční skupiny podle

Zde uvedená kategorická pole definují nezávislé sady parametrů kovariance opakovaných efektů; jedna pro každou kategorii definovanou křížovou klasifikací polí seskupení. Všechny subjekty mají stejný typ kovariance; předměty v rámci stejného kovariance seskupení budou mít stejné hodnoty pro parametry.

Typ opakované kovariance

Tato volba určuje strukturu kovariance pro rezidua. Na základě vybraného **typu opakovaného kovariance** jsou k dispozici různé volby kovariance. Dostupné struktury jsou:

- Autoregrese prvního pořadí (AR1)

- Přímý produkt AR1 (UN_AR1)
- Nestrukturovaný přímý produkt (UN_UN)
- Přímá složená symetrie produktu (UN_CS)
- Heterogenní sloučenina symetrie (CSH)
- Heterogenní autoregrese (ARH1)
- Autoregresivní klouzavý průměr (1, 1) (ARMA11)
- Složená symetrie
- Diagonála
- Škálovaná identita
- Toeplitz
- Nestrukturované
- Komponenty rozptylu
- Prostorové: Napájení
- Prostorové: exponenciální
- Prostorové: Gaussovský
- Prostorové: Lineární
- Prostorový: Lineární-log
- Prostorové: Sférické

Kroneckerovo opatření

Vyberte proměnné, které určují strukturu předmětu pro měření kovariance Kronecker, a určete, jak jsou korelovány chyby měření. Pole je k dispozici pouze v případě, že je vybrán jeden z následujících **opakujících se typů Covariance** :

- Přímý produkt AR1 (UN_AR1)
- Nestrukturovaný přímý produkt (UN_UN)
- Přímá složená symetrie produktu (UN_CS)

Prostorové kovarianční souřadnice

Proměnné v tomto seznamu určují souřadnice opakovaných pozorování, je-li pro opakovaný typ kovariance vybrán jeden z prostorových typů kovariance.

Další informace naleznete v tématu [“Kovarianční struktury”](#) na stránce 96 .

Ukazatele pseudoR²

Pseudo-R² ukazatele a korelační koeficient uvnitř třídy jsou zahrnuty do výstupu GLMM (je-li to vhodné). Ukazatele pseudoR² jsou zcela založeny na konečných odhadech a jsou vytvořeny po dokončení odhadu. Koeficient určení R² je běžně vykazovaný statistický údaj, protože představuje podíl rozptylu vysvětlený lineárním modelem. Intratřídní korelační koeficient (ICC) je související statistika, která kvantifikuje podíl rozptylu vysvětlený seskupujícím (náhodným) faktorem ve víceúrovňových/hierarchických datech.

Získání zobecněného lineárního smíšeného modelu

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Z nabídek vyberte:

Analyzovat > Smíšené modely > Zobecněné lineární ...

1. Definujte strukturu předmětu datové sady na kartě **Datová struktura** . Karta poskytuje volby pro určení efektů modelu i opakovaného typu kovariance.
2. Na kartě **Pole a efekty** musí existovat jediný cíl, který může mít jakoukoli úroveň měření, nebo specifikace události/zkoušek, v takovém případě musí být specifikace událostí a zkoušek souvislé.

Volitelně můžete určit její distribuční a spojovací funkci, pevné efekty a libovolné bloky náhodných efektů, offsety nebo váhy analýzy.

3. Klepnutím na volbu **Volby sestavení** určete volitelné nastavení sestavení.
4. Klepnutím na volbu **Volby modelu** uložíte skóre do aktivní datové sady a exportujete model do externího souboru.
5. Klepnutím na tlačítko **Spustit** spustíte proceduru a vytvoříte objekty modelu.

Cíl

Tato nastavení definují cíl, jeho distribuci a jeho vztah k prediktorům prostřednictvím funkce odkazu.

Cíl. Cíl je povinný. Může mít libovolnou úroveň měření a úroveň měření cíle omezuje, které distribuce a funkce propojení jsou vhodné.

- **Použít počet pokusů jako jmenovatel.** Když je cílovou odpovědí řada událostí, které se vyskytují v sadě pokusů, cílové pole obsahuje počet událostí a můžete vybrat další pole obsahující počet pokusů. Například při testování nového pesticidu můžete vystavit vzorky mravenců různým koncentracím pesticidů a poté zaznamenat počet usmrcených mravenců a počet mravenců v každém vzorku. V tomto případě by pole zaznamenávající počet usmrcených mravenců mělo být specifikováno jako cílové pole (události) a pole zaznamenávající počet mravenců v každém vzorku by mělo být specifikováno jako pole pokusů. Je-li počet mravenců pro každý vzorek stejný, lze počet pokusů určit pomocí pevné hodnoty.

Počet pokusů by měl být větší nebo roven počtu událostí pro každý záznam. Události by měly být nezáporná celá čísla a zkoušky by měly být kladná celá čísla.

- **Upravit referenční kategorii.** Pro kategoričtý cíl můžete zvolit referenční kategorii. To může ovlivnit určitý výstup, například odhady parametrů, ale neměl by měnit vhodnost modelu. Pokud například váš cíl přebírá hodnoty 0, 1 a 2, procedura standardně vytvoří poslední kategorii (s nejvyšší hodnotou), nebo 2 kategorii odkazu. V této situaci by měly být odhady parametrů interpretovány jako související s pravděpodobností kategorie 0 nebo 1 *relativní* vzhledem k pravděpodobnosti kategorie 2. Pokud určíte vlastní kategorii a váš cíl má definované popisky, můžete nastavit referenční kategorii výběrem hodnoty ze seznamu. To může být výhodné, když uprostřed zadávání modelu si nepamätujete přesně, jak bylo určité pole kódováno.

Cílová distribuce a relace (odkaz) s lineárním modelem. Vzhledem k hodnotám prediktorů model očekává, že rozdělení hodnot cíle bude odpovídat určenému tvaru a že cílové hodnoty budou lineárně souviset s prediktory prostřednictvím určené funkce odkazu. K dispozici jsou krátké řezy pro několik běžných modelů, nebo zvolte nastavení **Vlastní**, pokud existuje konkrétní kombinace funkcí distribuce a propojení, kterou chcete přizpůsobit a která není uvedena v seznamu krátkých položek.

- **Lineární model.** Uvádí normální distribuci s vazbou identity, která je užitečná, když lze cíl předpovědět pomocí lineární regrese nebo modelu ANOVA.
- **Gamma regrese.** Určuje distribuci gama s odkazem protokolu, který by měl být použit, když cíl obsahuje všechny kladné hodnoty a je posouván směrem k větším hodnotám.
- **Loglinear.** Určuje Poissonovu distribuci s odkazem na protokol, který by měl být použit, když cíl představuje počet výskytů v pevném časovém období.
- **Negativní binomická regrese.** Určuje negativní binomické rozdělení s odkazem na protokol, které by se mělo použít, když cíl a jmenovatel představují počet pokusů požadovaných pro sledování úspěchů k .
- **Multinomiální logistická regrese.** Uvádí multinomiální distribuci, která by se měla použít, když je cílem odezva s více kategoriemi. Používá buď kumulativní logit link (ordinal výsledky) nebo generalized logit link (multi-category nominální odezvy).
- **Binární logistická regrese.** Určuje binomickou distribuci s odkazem logit, který by měl být použit v případě, že cílem je binární odezva předpovězená modelem logistické regrese.
- **Binární probit.** Určuje binomické rozdělení s odkazem probit, který by měl být použit v případě, že cílem je binární odezva se základní normální distribucí.
- **Interval cenzurovaného přežití.** Určuje binomickou distribuci s doplňkovým odkazem log-log, který je užitečný v analýze přežití, když některá pozorování nemají žádnou událost ukončení.

Rozdělení

Tato volba určuje distribuci cíle. Schopnost specifikovat nenormální rozdělení a funkci neidentických vazeb je základním zlepšením zobecněného lineárního smíšeného modelu oproti lineárnímu smíšenému modelu. Existuje mnoho možných kombinací funkcí distribuce a propojení a několik může být vhodné pro danou datovou sadu, takže vaše volba může být vedena a priori teoretickými úvahami nebo která kombinace se zdá, že nejlépe vyhovuje.

- **Binomické.** Tato distribuce je vhodná pouze pro cíl, který představuje binární odezvu nebo počet událostí.
- **Gamma.** Toto rozdělení je vhodné pro cíl s kladnými hodnotami měřítka, které jsou posunovány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.
- **Inverse Gaussian.** Toto rozdělení je vhodné pro cíl s kladnými hodnotami měřítka, které jsou posunovány směrem k větším kladným hodnotám. Pokud je hodnota dat menší nebo rovna 0 nebo chybí, pak se v analýze nepoužije odpovídající případ.
- **Multinomiální.** Tato distribuce je vhodná pro cíl, který představuje odezvu s více kategoriemi. Forma modelu bude záviset na úrovni měření cíle.

Nominální cíl bude mít za následek nominální multinomiální model, ve kterém se pro každou kategorii cíle odhaduje samostatná sada parametrů modelu (s výjimkou referenční kategorie). Odhady parametrů pro daný prediktor zobrazují vztah mezi tímto prediktorem a pravděpodobností každé kategorie cíle vzhledem k referenční kategorii.

Pořadový cíl bude mít za následek ordinální víceúčelový model, ve kterém je tradiční výraz zachycení nahrazen sadou **prahových hodnot** parametrů, které se vztahují k kumulativní pravděpodobnosti cílových kategorií.

- **Negativní binomické.** Negativní binomická regrese používá negativní binomické rozdělení s odkazem protokolu, který by měl být použit, když cíl představuje počet výskytů s vysokou rozptylem.
- **Normální.** To je vhodné pro spojitý cíl, jehož hodnoty mají symetrické, zvonovité rozdělení kolem centrální (střední) hodnoty.
- **Poissonovo.** Toto rozdělení lze považovat za počet výskytů události, která je předmětem zájmu, v pevném časovém období a je vhodné pro proměnné s nezápornými celočíselnými hodnotami. Pokud je hodnota dat jiná než celé číslo, menší než 0 nebo chybí, pak se v analýze nepoužije odpovídající velikost písmen.

Funkce propojení

Funkce odkazu je transformace cíle, která umožňuje odhad modelu. K dispozici jsou následující funkce:

- **Identita.** $f(x) = x$. Cíl není transformován. Tento odkaz lze použít s jakoukoli distribucí, s výjimkou multinomiální.
- **Doplňkový protokol-log.** $f(x) = \log(-\log(1-x))$. To je vhodné pouze s binomickým nebo multinomiálním rozdělením.
- **Cauchit.** $f(x) = \tan(\pi(x - 0.5))$. To je vhodné pouze s binomickým nebo multinomiálním rozdělením.
- **log.** $f(x) = \log(x)$. Tento odkaz lze použít s jakoukoli distribucí, s výjimkou multinomiální.
- **Doplňek protokolu.** $f(x) = \log(1-x)$. To je vhodné pouze při binomické distribuci.
- **Přihlásit se.** $f(x) = \log(x/(1-x))$. To je vhodné pouze s binomickým nebo multinomiálním rozdělením.
- **Protokol negativního protokolu.** $f(x) = -\log(-\log(x))$. To je vhodné pouze s binomickým nebo multinomiálním rozdělením.
- **Probit.** $f(x) = \Phi^{-1}(x)$, kde Φ^{-1} je inverzní standardní funkce normální kumulativní distribuce. To je vhodné pouze s binomickým nebo multinomiálním rozdělením.
- **Power.** $f(x) = x^\alpha$, pokud $\alpha \neq 0$. $f(x) = \log(x)$, pokud $\alpha=0$. α je požadované číslo specifikace a musí být reálné číslo. Tento odkaz lze použít s jakoukoli distribucí, s výjimkou multinomiální.

Pevné efekty





Faktory fixních efektů jsou obecně považovány za pole, jejichž hodnoty zájmu jsou všechny zastoupeny v datové sadě, a lze je použít pro hodnocení. Standardně se pole s předdefinovanou vstupní rolí, která nejsou určena jinde v dialogovém okně, zadávají do části modelu s pevnými efekty. Kategorická (nominální a ordinální) pole se používají jako faktory v modelu a spojitá pole se používají jako kovariáty.

Zadejte efekty do modelu výběrem jednoho nebo více polí ve zdrojovém seznamu a přetažením do seznamu efektů. Typ vytvořeného efektu závisí na tom, který aktivní bod chcete zrušit.

- **Hlavní.** Vypuštěná pole se zobrazují jako samostatné hlavní efekty v dolní části seznamu efektů.
- **2-cestný.** Všechny možné dvojice zahozených polí se zobrazí jako dvoucestné interakce v dolní části seznamu efektů.
- **3-cestný.** Všechny možné trojice zahozených polí se zobrazují jako třícestné interakce v dolní části seznamu efektů.
- *****. Kombinace všech zrušených polí se zobrazí jako jedna interakce v dolní části seznamu efektů.

Tlačítka napravo od tvůrce efektů vám umožňují provádět různé akce.

Tabulka 1. Popisy tlačítek tvůrce efektů

Ikona	Popis
	Odstraňte termíny z modelu pevných efektů výběrem podmínek, které chcete odstranit, a klepnutím na tlačítko Odstranit.
 	Změňte pořadí termínů v rámci modelu pevných efektů výběrem termínů, které chcete změnit, a klepnutím na šipku nahoru nebo dolů.
	Vnořené výrazy přidejte do modelu pomocí dialogového okna <u>“Přidat vlastní výraz”</u> na stránce 52 klepnutím na tlačítko Přidat vlastní výraz.

Zahrnout zachycení. Zachycení je obvykle zahrnuto v modelu. Pokud můžete předpokládat, že data procházejí výchozím bodem, můžete toto zachycení vyloučit.

Přidat vlastní výraz

V této proceduře můžete sestavit vnořené výrazy pro svůj model. Vnořené výrazy jsou užitečné pro modelování účinku faktoru nebo kovariátu, jehož hodnoty neinteragují s úrovněmi jiného faktoru. Například řetězec obchodu s potravinami může sledovat výdajové návyky svých zákazníků na několika místech obchodu. Vzhledem k tomu, že každý zákazník navštívuje pouze jednu z těchto lokalit, lze efekt *Zákazník* říci, že je **vnořený v** efektu *Umístění úložiště*.

Dále můžete zahrnout efekty interakce, například polynomiální výrazy zahrnující stejné kovariát, nebo přidat více úrovní vnoření do vnořeného výrazu.

Omezení. Vnořené výrazy mají následující omezení:

- Všechny faktory v rámci interakce musí být jedinečné. Je-li tedy faktor A , je zadání $A^* A$ neplatné.
- Všechny faktory ve vnořeném efektu musí být jedinečné. Je-li tedy faktor A , je zadání $A(A)$ neplatné.
- Žádný efekt nemůže být vnořen do kovariátu. Je-li tedy A faktorem a X je kovariantní, pak určení $A(X)$ je neplatné.

Vytvoření vnořeného výrazu

1. Vyberte faktor nebo kovariát, který je vnořen do jiného faktoru, a poté klepněte na tlačítko se šipkou.
2. Klepněte na volbu **(v rámci)**.
3. Vyberte faktor, v němž je vnořen předchozí faktor nebo kovariát, a poté klepněte na tlačítko se šipkou.

4. Klepněte na volbu **Přidat výraz**.

Volitelně můžete do vnořeného výrazu zahrnout efekty interakce nebo přidat více úrovní vnoření.

Náhodné efekty

Náhodné faktory efektů jsou pole, jejichž hodnoty v datovém souboru lze považovat za náhodný vzorek z větší populace hodnot. Jsou užitečné pro vysvětlení nadměrné variability v cíli. Pokud jste na kartě Datová struktura vybrali více než jeden předmět, bude standardně vytvořen blok Náhodný efekt pro každý předmět mimo nejvnitřnější předmět. Pokud jste například jako předměty na kartě Struktura dat vybrali volbu Škola, Třída a Student, automaticky se vytvoří následující náhodné bloky efektů:

- Náhodný efekt 1: předmět je škola (bez efektů, pouze zachycení)
- Náhodný efekt 2: předmět je školní * třída (bez efektů, pouze zachycení)

S bloky náhodných efektů můžete pracovat následujícími způsoby:

1. Chcete-li přidat nový blok, klepněte na volbu **Přidat blok ...** Tím se otevře dialogové okno "Blok náhodných efektů" na stránce 53 .
2. Chcete-li upravit existující blok, vyberte blok, který chcete upravit, a klepněte na volbu **Upravit blok ...** Tím se otevře dialogové okno "Blok náhodných efektů" na stránce 53 .
3. Chcete-li odstranit jeden nebo více bloků, vyberte bloky, které chcete odstranit, a klepněte na tlačítko Odstranit.

Blok náhodných efektů

Zadejte efekty do modelu výběrem jednoho nebo více polí ve zdrojovém seznamu a jejich přidáním do seznamu **Tvůrce efektů** .

Typ vytvořeného efektu závisí na tom, který **Typ** vyberete. Kategorická (nominální a ordinální) pole se používají jako faktory v modelu a spojitá pole se používají jako kovariáty.

Interakce

Kombinace všech polí se zobrazí jako jedna interakce v dolní části seznamu efektů.

Hlavní účinky

Vypuštěná pole se zobrazují jako samostatné hlavní efekty v dolní části seznamu efektů.

Všechny 2-way

Všechny možné dvojice zahozených polí se zobrazí jako dvoucestné interakce v dolní části seznamu efektů.

Všechny 3-way

Všechny možné trojice zahozených polí se zobrazují jako třícestné interakce v dolní části seznamu efektů.

Všechny 4-way

Všechny možné trojice zahozených polí se zobrazují jako 4cestné interakce v dolní části seznamu efektů.

Vše 5-ti cestné

Všechny možné trojice zahozených polí se zobrazují jako 5cestné interakce v dolní části seznamu efektů.

Zahrnout zachycení

Zachycení není standardně zahrnuto v modelu náhodných efektů. Pokud můžete předpokládat, že data procházejí výchozím bodem, můžete toto zachycení vyloučit.

Zobrazit předpovědi parametrů pro tento blok

Určuje, že se mají zobrazit odhady parametrů náhodných efektů.

Kombinace předmětů

To vám umožňuje určit předměty s náhodným efektem z přednastavených kombinací předmětů v dialogovém okně Proměnné. Pokud jsou například *Škola*, *Třída* a *Student* definovány jako předměty

v dialogovém okně Proměnné a v tomto pořadí bude mít rozevírací seznam Kombinace předmětu jako volby **Žádná, Škola, Škola * Třída, Škola * Třída * * Student**.

Typ kovariance náhodného efektu

Tato volba určuje strukturu kovariance pro rezidua. Dostupné struktury jsou:

- Autoregrese prvního pořadí (AR1)
- Autoregresivní klouzavý průměr (1, 1) (ARMA11)
- Složená symetrie
- Diagonála
- Škálovaná identita
- Toeplitz
- Nestrukturované
- Komponenty rozptylu

Definovat kovarianční skupiny podle

Zde uvedená kategorická pole definují nezávislé sady parametrů kovariance náhodných efektů; jedna pro každou kategorii definovanou křížovou klasifikací polí seskupení. Pro každý náhodný blok efektů lze zadat jinou sadu seskupovacích polí. Všechny subjekty mají stejný typ kovariance; předměty v rámci stejného kovariance seskupení budou mít stejné hodnoty pro parametry.

Váha a odsazení

Váha analýzy. Parametr měřítka je odhadovaný parametr modelu související s odchylkou odezvy. Váhy analýzy jsou "známé" hodnoty, které se mohou lišit od pozorování k pozorování. Je-li uvedeno pole váhy analýzy, parametr měřítka, který souvisí s odchylkou odezvy, se vydělí hodnotami váhy analýzy pro každé pozorování. Záznamy s hodnotami váhy analýzy, které jsou menší nebo rovny 0 nebo chybí, se v analýze nepoužívají.

Posunutí. Termín offsetu je "strukturální" prediktor. Jeho koeficient není odhadnut modelem, ale předpokládá se, že má hodnotu 1; proto se hodnoty offsetu jednoduše přičítají k lineárnímu prediktoru cíle. To je zvláště užitečné v Poissonových regresních modelech, kde každý případ může mít různé úrovně expozice vůči události, která je předmětem zájmu.

Například při modelování nehodovosti pro jednotlivé řidiče existuje důležitý rozdíl mezi řidičem, který byl v jedné nehodě za tři roky zkušeností a řidičem, který byl v jedné nehodě za 25 let! Počet nehod může být modelován jako Poissonova nebo negativní binomická odpověď s odkazem na protokol, pokud je přirozený záznam zkušeností řidiče zahrnut jako offset.

Jiné kombinace rozdělení a typů vazeb by vyžadovaly jiné transformace proměnné offsetu.

Obecné volby sestavení

Tyto výběry určují některá pokročilejší kritéria použitá k sestavení modelu.

Pořadí řazení

Tyto ovládací prvky určují pořadí kategorií pro cíl a faktory (kategorické vstupy) pro účely určení kategorie "poslední". Nastavení pořadí řazení cíle je ignorováno, pokud cíl není kategorický nebo pokud je v nastavení "[Cíl](#)" na stránce 50 určena vlastní referenční kategorie.

Pravidla zastavení

Můžete určit maximální počet iterací, které bude algoritmus provádět. Algoritmus používá dvojitý iterační proces, který se skládá z vnitřní smyčky a vnější smyčky. Hodnota, která je určena pro maximální počet iterací, platí pro oba cykly. Zadejte nezáporné celé číslo. Výchozí hodnota je 100.

Nastavení následného odhadu

Tato nastavení určují, jak jsou některé výstupy modelu vypočítány pro zobrazení.

Úroveň důvěry (%)

Jedná se o úroveň spolehlivosti používanou pro výpočet odhadů intervalů koeficientů modelu. Zadejte hodnotu větší než 0 a menší než 100. Výchozí hodnota je 95.

Stupně volnosti

Určuje způsob výpočtu stupňů volnosti pro testy významnosti. Volbu **Zbytková metoda** vyberte, pokud je velikost vzorku dostatečně velká nebo pokud jsou data vyvážená, nebo pokud model používá jednodušší typ kovariance (například škálovaná identita nebo diagonální). Toto je výchozí nastavení. Pokud je velikost vzorku malá nebo jsou data nevyvážená, nebo pokud model používá složitý typ kovariance (například nestrukturovaný), vyberte volbu **Stěrná aproximace**. Vyberte volbu **Kenward-Roger aproximace**, pokud je velikost vzorku malá a máte model REML (Restricted Maximum Pravděpodobnost).

Zkoušky pevných účinků a koeficientů

Jedná se o metodu výpočtu matice odhadu kovariance parametru. Vyberte robustní odhad, pokud máte obavy, že jsou porušeny předpoklady modelu.

Odhad

Algoritmus sestavení modelu používá dvojnásob iterativní proces, který se skládá z vnitřní smyčky a vnější smyčky. Následující nastavení platí pro vnitřní smyčku.

Konvergence parametrů.

Konvergence se předpokládá, pokud je maximální absolutní změna nebo maximální relativní změna v odhadech parametrů menší než zadaná hodnota, která musí být nezáporná. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Konvergence pravděpodobnosti protokolu.

Konvergence se předpokládá, pokud je absolutní změna nebo relativní změna ve funkci log-Pravděpodobnost menší než zadaná hodnota, která musí být nezáporná. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Hessovská konvergence.

Pro specifikaci **Absolutní** se předpokládá konvergence, pokud je statistika založená na hesenské hodnotě menší než zadaná hodnota. Pro specifikaci **Relativní** se předpokládá konvergence, pokud je statistika menší než součin zadané hodnoty a absolutní hodnoty pravděpodobnosti protokolu. Kritérium se nepoužije, pokud je zadaná hodnota rovna 0.

Maximální Fisher bodování kroky.

Zadejte nezáporné celé číslo. Hodnota 0 určuje metodu Newton-Raphson. Hodnoty větší než 0 určují použití algoritmu přidělení skóre Fisher až do čísla iterace n , kde n je zadané celé číslo a poté Newton-Raphson.

Singularitní tolerance.

Tato hodnota se používá jako tolerance při kontrole singularity. Zadejte kladnou hodnotu.

Poznámka: Standardně se používá Konvergence parametrů, kde je zaškrtnuta maximální změna **Absolutní** s tolerancí $1E-6$. Toto nastavení může vést k výsledkům, které se liší od výsledků získaných ve verzích před verzí 22. Chcete-li reprodukovat výsledky z verzí pre-22, použijte volbu **Relativní** pro kritérium konvergence parametrů a ponechte výchozí hodnotu tolerance $1E-6$.

Odhadované prostředky

Tato karta vám umožňuje zobrazit odhadované mezní střední hodnoty pro úrovně faktorů a interakce faktorů. Pro multinomiální modely nejsou k dispozici odhadované mezní střední hodnoty.

Podmínky. Zde jsou uvedeny modelové termíny v části Pevné efekty, které se zcela skládají z kategoriálních polí. Zaškrtněte každý termín, pro který chcete, aby model vytvořil odhadované mezní prostředky.

- **Typ kontrastu.** Určuje typ kontrastu, který má být použit pro úrovně pole kontrastu. Je-li vybrána volba **Není**, nejsou vytvářeny žádné kontrasty. **Porovnávání po dvojicích** vytváří porovnávání po dvojicích pro všechny kombinace úrovní uvedených faktorů. Toto je jediný dostupný kontrast pro interakce faktorů. **Odchylka** porovnává každou úroveň faktoru s velkou střední hodnotou. **Jednoduché** kontrasty porovnávají každou úroveň faktoru, s výjimkou poslední, s poslední úrovní. "Poslední" úroveň je určena pořadím řazení pro faktory určené ve volbách sestavení. Všimněte si, že všechny tyto typy kontrastu nejsou ortogonální.

- **Kontrastní pole.** Tato volba určuje faktor, jehož úrovně jsou porovnávány s použitím vybraného typu kontrastu. Je-li jako typ kontrastu vybrána volba **Není**, nelze vybrat žádné pole kontrastu.

Spojité pole. Uvedená souvislá pole jsou extrahována z termínů v části Pevné efekty, které používají souvislá pole. Při výpočtu odhadovaných mezních hodnot jsou kovariáty fixovány na zadané hodnoty. Vyberte střední hodnotu nebo zadejte vlastní hodnotu.

Zobrazit odhadované střední hodnoty z hlediska. Tato volba určuje, zda se mají vypočítat odhadované mezní střední hodnoty na základě původního měřítka cíle nebo na základě transformace funkce odkazu.

Původní měřítko cíle vypočítá odhadované mezní střední hodnoty pro cíl. Všimněte si, že když je cíl specifikován pomocí volby příhody/pokusy, dává to odhadované mezní střední hodnoty pro poměr příhody/pokusy spíše než pro počet událostí. **Transformace spojové funkce** vypočítá odhadované mezní střední hodnotu pro lineární prediktor.

Upravit pro použití více porovnání. Při provádění testů hypotéz s více kontrasty lze celkovou úroveň významnosti upravit z úrovně významnosti pro zahrnuté kontrasty. To vám umožní zvolit způsob nastavení.

- **Nejméně významný rozdíl.** Tato metoda nekontroluje celkovou pravděpodobnost odmítnutí hypotéz, že některé lineární kontrasty se liší od hodnot nulové hypotézy.
- *Sekvenční bonferroni.* Jedná se o postup postupného odmítání Bonferroni, který je mnohem méně konzervativní, pokud jde o odmítání jednotlivých hypotéz, ale zachovává stejnou celkovou úroveň významnosti.
- *Sekvenční sidak.* Jedná se o postupně odstupňovaný odkladný postup Sidak, který je z hlediska odmítnutí jednotlivých hypotéz mnohem méně konzervativní, ale zachovává stejnou celkovou úroveň významnosti.

Nejméně signifikantní rozdíl metoda je méně konzervativní než sekvenční Sidak metoda, která je zase méně konzervativní než sekvenční Bonferroni; to znamená, že nejméně významný rozdíl bude odmítnout alespoň tolik jednotlivých hypotéz jako sekvenční Sidak, který zase odmítne alespoň tolik jednotlivých hypotéz jako sekvenční Bonferroni.

Uložit

Zaškrtnuté položky jsou uloženy s uvedeným názvem; konflikty s existujícími názvy polí nejsou povoleny.

Předpokládané hodnoty- Uloží předpovězenou hodnotu cíle. Výchozí název pole je *PredictedValue*.

Předpokládaná pravděpodobnost pro kategorické cíle- Pokud je cíl kategorický, toto klíčové slovo uloží předpokládané pravděpodobnosti prvních n kategorií až do hodnoty určené jako **Maximální počet kategorií k uložení**. Vypočtené hodnoty jsou kumulativní pravděpodobnosti pro ordinální cíle. Výchozí kořenový název je *PredictedProbability*. Chcete-li uložit predikovanou pravděpodobnost predikované kategorie, uložte hodnověrnost (viz níže).

Intervaly spolehlivosti- Uloží horní a dolní meze konfidenčního intervalu pro predikovanou hodnotu nebo predikovanou pravděpodobnost. Pro všechny distribuce kromě multinomiálních vytvoří dvě proměnné a výchozí kořenový název je *CI*, s příponami *_Lower* a *_Upper*.

Pro multinomiální rozdělení a nominální cíl se vytvoří jedno pole pro každou závislou kategorii proměnných. Tím se uloží dolní a horní hranice predikované pravděpodobnosti pro první n kategorie až do hodnoty uvedené jako **Maximální počet kategorií k uložení**. Výchozí kořenový název je *CI* a výchozí názvy polí jsou *CI_Lower_1*, *CI_Upper_1*, *CI_Lower_2*, *CI_Upper_2* atd. odpovídající pořadí cílových kategorií.

Pro multinomiální distribuci a ordinální cíl je vytvořeno jedno pole pro každou kategorii závislých proměnných kromě poslední (další informace viz téma "Obecné volby sestavení" na stránce 54.). Tím se uloží dolní a horní hranice kumulativní predikované pravděpodobnosti pro první n kategorie, až do posledního, ale ne včetně, a až do hodnoty uvedené jako **Maximální počet kategorií k uložení**. Výchozí kořenový název je *CI* a výchozí názvy polí jsou *CI_Lower_1*, *CI_Upper_1*, *CI_Lower_2*, *CI_Upper_2* atd. odpovídající pořadí cílových kategorií.

Pearsonovy zbytkové chyby- Uloží Pearsonův zbytek pro každý záznam, který lze použít v poodhadové diagnostice vhodné pro model. Výchozí název pole je *PearsonResidual*.

Důvěrné- Ukládá důvěryhodnost předpovězené hodnoty pro kategorický cíl. Vypočtená hodnověrnost může být založena na pravděpodobnosti predikované hodnoty (nejvyšší predikované pravděpodobnosti) nebo na rozdílu mezi nejvyšší predikovanou pravděpodobností a druhou nejvyšší predikovanou pravděpodobností. Výchozí název pole je *Důvěra*.

Export

Exportovat model. Tím se model zapíše do externího souboru *.zip*. Tento soubor modelu můžete použít k použití informací o modelu na jiné datové soubory pro účely přidělení skóre. Další informace naleznete v tématu [Průvodce skóre](#). Zadejte jedinečný platný název souboru. Pokud specifikace souboru odkazuje na existující soubor, bude soubor přepsán.

Exportovat výstup

Exportovat EBLUPS. Tato sekce je povolena, pokud je vytvořen alespoň jeden blok náhodných efektů prostřednictvím dialogového okna položky Náhodné efekty na kartě Pole & Efekty a je vybráno zaškrtačkové políčko pro zobrazení předpovědi parametrů pro tento blok. Pokud bylo zadáno více takových bloků, můžete zvolit, zda mají být výsledky uloženy do samostatných datových sad nebo souborů, nebo zda mají být sloučeny do jedné datové sady nebo souboru.

Zobrazení modelu

Procedura vytvoří objekt modelu v prohlížeči. Aktivací (poklepaním) tohoto objektu získáte interaktivní pohled na model.

Standardně se zobrazí pohled Souhrn modelu. Chcete-li zobrazit jiný pohled modelu, vyberte jej z miniatur pohledu.

Jako alternativu k objektu Model můžete generovat kontingenční tabulky a grafy výběrem volby **Otočit tabulky a grafy** ve skupině Zobrazení výstupu na kartě Výstup dialogového okna Volby (Upravit > Volby). Následující témata popisují objekt Model.

Souhrn modelu

Tento pohled je snímek, stručný souhrn modelu a jeho přizpůsobení.

tabulka: Tabulka identifikuje cíl, rozdělení pravděpodobnosti a funkci odkazu uvedenou v poli Nastavení cíle. Pokud je cíl definován událostmi a zkouškami, buňka je rozdělena tak, aby zobrazovala pole událostí a pole pokusů nebo pevný počet pokusů. Kromě toho se zobrazí kritérium pro konečné informace Akaike (AICC) a Bayesovské informační kritérium (BIC).

- *Opraveno akaike.* Opatření pro výběr a porovnání smíšených modelů na základě pravděpodobnosti protokolu -2 (s omezením). Menší hodnoty označují lepší modely. AICC "opravuje" AIC pro malé velikosti vzorků. Jak se velikost vzorku zvětšuje, konverguje AICC k AIC.
- *Bayesian.* Ukazatel pro výběr a porovnání modelů na základě pravděpodobnosti protokolu -2. Menší hodnoty označují lepší modely. BIC také "penalizuje" overparameterizované modely (například komplexní modely s velkým počtem vstupů), ale striktněji než AIC.

Graf. Pokud je cíl kategorický, graf zobrazí přesnost konečného modelu, což je procentní část správných klasifikací.

Datová struktura

Tento pohled poskytuje souhrn datové struktury, kterou jste zadali, a pomáhá vám zkontrolovat, zda byly správně zadány předměty a opakované ukazatele. Pozorované informace pro první předmět se zobrazí pro každé pole předmětu a pole opakovaných ukazatelů a cíl. Dále se zobrazí počet úrovní pro každé pole předmětu a pole opakovaných ukazatelů.

Předpovězeno pozorovaným

U spojených cílů, včetně cílů určených jako události/pokusy, se zobrazí spojitý bodový graf předpovězených hodnot na svislé ose podle pozorovaných hodnot na vodorovné ose. V ideálním případě by měly body ležet na přímce o 45 stupních; tento pohled vám může říci, zda jsou některé záznamy modelem obzvláště špatně předpovězeny.

Klasifikace

V případě kategorických cílů se zobrazí křížová klasifikace pozorovaných a předpokládaných hodnot v teplotní mapě plus celkové procento správnosti.

Styly tabulek. Existuje několik různých stylů zobrazení, které jsou přístupné z rozevíracího seznamu **Styl**.

- **Procent řádků.** Zobrazí procentní části řádků (počet buněk vyjádřený jako procentní část součtu řádků) v buňkách. Jedná se o výchozí nastavení.
- **Počty buněk.** Tato volba zobrazí počet buněk v buňkách. Stínování pro mapu využití je stále založeno na procentech řádků.
- **Mapa využití.** Tato volba nezobrazuje v buňkách žádné hodnoty, pouze stínování.
- **Komprimováno.** Tato volba nezobrazuje žádné záhlaví řádků či sloupců ani hodnoty v buňkách. To může být užitečné, když cíl má mnoho kategorií.

Chybí. Pokud některé záznamy mají chybějící hodnoty v cíli, zobrazí se v řádku (**Chybějící**) pod všemi platnými řádky. Záznamy s chybějícími hodnotami nepřispívají k celkové procentní správnosti.

Více cílů. Pokud existuje více kategorických cílů, pak se každý cíl zobrazí v samostatné tabulce a existuje rozevírací seznam **Cíl**, který řídí, který cíl se má zobrazit.

Velké tabulky. Pokud má zobrazený cíl více než 100 kategorií, nezobrazí se žádná tabulka.

Pevné efekty

Tento pohled zobrazuje velikost každého pevného efektu v modelu.

Styly. Existují různé styly zobrazení, které jsou přístupné z rozevíracího seznamu **Styl**.

- **Diagram.** Toto je graf, ve kterém jsou efekty seřazeny shora dolů v pořadí, ve kterém byly zadány v nastavení Pevné efekty. Spojovací čáry v diagramu jsou váženy na základě významnosti efektu, s větší šířkou čáry odpovídající významnějším efektům (menší p -values). Jedná se o výchozí nastavení.
- **tabulka:** Toto je tabulka ANOVA pro celkový model a jednotlivé modelové efekty. Jednotlivé efekty jsou seřazeny shora dolů v pořadí, v jakém byly uvedeny v nastavení Pevné efekty.

Význam. Existuje posuvný ovladač významnosti, který řídí, které efekty se zobrazí v pohledu. Efekty s hodnotami významnosti většími než hodnota posuvného ovladače jsou skryté. To nemění model, ale jednoduše vám umožní zaměřit se na nejdůležitější efekty. Standardně je hodnota 1.00, takže se na základě významnosti nefiltrují žádné efekty.

Pevné koeficienty

Tento pohled zobrazuje hodnotu každého pevného koeficientu v modelu. Všimněte si, že faktory (kategoriální prediktory) jsou v rámci modelu kódovány podle indikátorů, takže **efekty** obsahující faktory budou mít obecně více přidružených **koeficientů**; jeden pro každou kategorii kromě kategorie odpovídající redundantnímu koeficientu.

Styly. Existují různé styly zobrazení, které jsou přístupné z rozevíracího seznamu **Styl**.

- **Diagram.** Toto je graf, který nejprve zobrazí zachycení a poté seřadí efekty shora dolů v pořadí, ve kterém byly zadány v nastavení Pevné efekty. V rámci efektů obsahujících faktory jsou koeficienty řazeny vzestupně podle datových hodnot. Spojovací čáry v diagramu jsou barevné a vážené na základě významnosti koeficientu, s větší šířkou čáry odpovídající významnějším koeficientům (menší p -values). Jedná se o výchozí styl.

- **tabulka:** Zobrazuje hodnoty, testy významnosti a intervaly spolehlivosti pro jednotlivé koeficienty modelu. Po zachycení jsou efekty seřazeny shora dolů v pořadí, v jakém byly uvedeny v nastavení Pevné efekty. V rámci efektů obsahujících faktory jsou koeficienty řazeny vzestupně podle datových hodnot.

Multinomiální. Pokud je v platnosti multinomiální distribuce, pak rozevírací seznam Multinomial řídí, která cílová kategorie se má zobrazit. Pořadí řazení hodnot v seznamu je určeno specifikací v nastavení Volby sestavení.

Exponenciální. Zobrazí odhady exponenciálních koeficientů a intervaly spolehlivosti pro určité typy modelů, včetně binární logistické regrese (binomická distribuce a logitová vazba), Nominální logistické regrese (multinomiální distribuce a logitová vazba), Negativní binomické regrese (negativní binomická distribuce a logitová vazba) a Log-lineární model (Poissonova distribuce a logitová vazba).

Význam. Existuje posuvný ovladač významnosti, který řídí, které koeficienty jsou zobrazeny v pohledu. Koeficienty s hodnotami významnosti většími než hodnota posuvného ovladače jsou skryty. To nemění model, ale jednoduše umožňuje zaměřit se na nejdůležitější koeficienty. Standardně je hodnota 1.00, takže na základě významnosti nejsou filtrovány žádné koeficienty.

Náhodný efekt Covariance

Tento pohled zobrazuje matici kovariance náhodných efektů (**G**).

Styly. Existují různé styly zobrazení, které jsou přístupné z rozevíracího seznamu **Styl**.

- **Hodnoty covariance.** Toto je teplotní mapa matice kovariance, ve které jsou efekty seřazeny shora dolů v pořadí, ve kterém byly uvedeny v nastavení Pevné efekty. Barvy v kogramu odpovídají hodnotám buněk, jak je uvedeno v klíči. Jedná se o výchozí nastavení.
- **Corrgram.** Toto je teplotní mapa matice kovariance.
- **Komprimováno.** Toto je teplotní mapa matice kovariance bez záhlaví řádků a sloupců.

Bloky. Pokud existuje více náhodných bloků efektů, pak je k dispozici rozevírací seznam bloků pro výběr bloku, který se má zobrazit.

Skupiny. Pokud má blok náhodných efektů specifikaci skupiny, pak existuje rozevírací seznam skupin pro výběr úrovně skupiny, která se má zobrazit.

Multinomiální. Pokud je v platnosti multinomiální distribuce, pak rozevírací seznam Multinomial řídí, která cílová kategorie se má zobrazit. Pořadí řazení hodnot v seznamu je určeno specifikací v nastavení Volby sestavení.

Parametry kovariance

Tento pohled zobrazuje odhady parametrů kovariance a související statistiky pro reziduální a náhodné efekty. Jedná se o pokročilé, ale zásadní výsledky, které poskytují informace o tom, zda je kovarianční struktura vhodná.

Souhrnná tabulka. Toto je rychlý odkaz pro počet parametrů ve zbytkových maticích (**R**) a náhodných maticích kovariance (**G**), ohodnocení důležitosti (počet sloupců) v maticích návrhu s pevným efektem (**X**) a náhodným efektem (**Z**) a počet subjektů definovaných poli předmětu, která definují datovou strukturu.

Tabulka parametrů Covariance. Pro vybraný efekt se pro každý parametr kovariance zobrazí odhad, standardní chyba a interval spolehlivosti. Počet zobrazených parametrů závisí na kovarianční struktuře efektu a u náhodných bloků efektů na počtu efektů v bloku. Pokud vidíte, že off-diagonální parametry nejsou významné, můžete být schopni použít jednodušší kovariance strukturu.

Efekty. Pokud existují náhodné bloky efektů, pak je rozevírací seznam efektů pro výběr zbytkového nebo náhodného bloku efektů, který se má zobrazit. Zbytkový efekt je vždy k dispozici.

Skupiny. Pokud má reziduální nebo náhodný blok efekt specifikaci skupiny, pak je zde rozevírací seznam skupin pro výběr úrovně skupiny, která se má zobrazit.

Multinomiální. Pokud je v platnosti multinomiální distribuce, pak rozevírací seznam Multinomial řídí, která cílová kategorie se má zobrazit. Pořadí řazení hodnot v seznamu je určeno specifikací v nastavení Volby sestavení.

Odhadované prostředky: Významné efekty

Jedná se o grafy zobrazené pro 10 "nejvýznamnějších" fixních all-factor efektů, počínaje trojcestnými interakcemi, pak obousměrnými interakcemi a nakonec hlavními efekty. Graf zobrazuje modelově odhadovanou hodnotu cíle na svislé ose pro každou hodnotu hlavního efektu (nebo prvního efektu uvedeného v interakci) na vodorovné ose; pro každou hodnotu druhého efektu uvedeného v interakci se vytvoří samostatná čára; pro každou hodnotu třetího efektu uvedeného v trojcestné interakci se vytvoří samostatný graf; všechny ostatní prediktory jsou udržovány konstantní. Poskytuje užitečnou vizualizaci účinků koeficientů jednotlivých prediktorů na cíl. Všimněte si, že pokud nejsou žádné prediktory významné, nevytvoří se žádné odhadované prostředky.

Důvěryhodnost. Tato volba zobrazí horní a dolní limity spolehlivosti pro mezní střední hodnoty s použitím úrovně spolehlivosti určené jako součást voleb sestavení.

Odhadované prostředky: Vlastní efekty

Jedná se o tabulky a grafy pro uživatelem požadované pevné efekty všech faktorů.

Styly. Existují různé styly zobrazení, které jsou přístupné z rozevřacího seznamu **Styl**.

- **Diagram.** Tento styl zobrazuje spojnicový graf odhadované hodnoty modelu cíle na svislé ose pro každou hodnotu hlavního efektu (nebo prvního uvedeného efektu v interakci) na vodorovné ose; pro každou hodnotu druhého uvedeného efektu v interakci se vytvoří samostatná čára; pro každou hodnotu třetího uvedeného efektu v trojcestné interakci se vytvoří samostatný graf; všechny ostatní prediktory jsou udržovány konstantní.

Pokud byly požadovány kontrasty, zobrazí se další graf pro porovnání úrovní kontrastního pole; pro interakce se zobrazí graf pro každou kombinaci úrovní jiných efektů než kontrastní pole. V případě kontrastů **po dvojicích** se jedná o graf vzdálené sítě, tj. grafickou reprezentaci tabulky porovnání, ve které vzdálenosti mezi uzly v síti odpovídají rozdílům mezi vzorky. Žluté čáry odpovídají statisticky významným rozdílům; černé čáry odpovídají nevýznamným rozdílům. Ponecháte-li ukazatel myši nad linkou v síti, zobrazí se popis s upravenou významností rozdílu mezi uzly spojenými linkou.

V případě **odchylek** kontrastů se zobrazí pruhový graf s odhadovanou hodnotou modelu cíle na svislé ose a hodnotami pole kontrastu na vodorovné ose; v případě interakcí se zobrazí graf pro každou kombinaci úrovní jiných efektů než pole kontrastu. Pruhy ukazují rozdíl mezi každou úrovní pole kontrastu a celkovou střední hodnotou, která je znázorněna černou vodorovnou čarou.

Pro **jednoduché** kontrasty se zobrazí pruhový graf s odhadovanou hodnotou modelu cíle na svislé ose a hodnotami pole kontrastu na vodorovné ose; pro interakce se zobrazí graf pro každou kombinaci úrovní jiných efektů než pole kontrastu. Pruhy ukazují rozdíl mezi každou úrovní pole kontrastu (kromě poslední) a poslední úrovní, která je znázorněna černou vodorovnou čarou.

- **tabulka:** Tento styl zobrazuje tabulku odhadované hodnoty modelu cíle, jeho standardní chybu a interval spolehlivosti pro každou kombinaci úrovní polí v efektu; všechny ostatní prediktory jsou udržovány konstantní.

Pokud byly požadovány kontrasty, zobrazí se další tabulka s odhadem, standardní chybou, testem významnosti a intervalem spolehlivosti pro každý kontrast; pro interakce existuje samostatná sada řádků pro každou kombinaci úrovní jiných efektů než pole kontrastu. Kromě toho se zobrazí tabulka s celkovými výsledky testu; pro interakce existuje samostatný celkový test pro každou kombinaci úrovní účinků jiných než kontrastní pole.

Důvěryhodnost. Tato volba přepíná zobrazení horních a dolních mezí spolehlivosti pro mezní střední hodnoty s použitím úrovně spolehlivosti určené jako součást voleb sestavení.

Rozvržení. Tato volba přepíná rozvržení diagramu s dvojitými kontrasty. Rozvržení kruhu méně odhaluje kontrasty než rozvržení sítě, ale vyhýbá se překrývajícím se čarům.

Výběr modelu Loglinear Analysis

Procedura Logineární analýza výběru modelu analyzuje vícecestné křížové tabulky (kontingenční tabulky). Přizpůsobuje hierarchické loglineární modely vícerozměrným křížovým slovům pomocí iterativního

algoritmu proporcionálního přizpůsobení. Tento postup vám pomůže zjistit, které kategorické proměnné jsou přidruženy. Pro sestavení modelů jsou k dispozici metody vynuceného vstupu a zpětné eliminace. U nasycených modelů můžete požadovat odhady parametrů a testy částečného přidružení. Nasycený model přidá hodnotu 0.5 ke všem buňkám.

Příklad. Ve studii uživatelských preferencí pro jeden ze dvou pracích prostředků vědci spočítali lidi v každé skupině, kombinovali různé kategorie měkkosti vody (měkké, střední nebo tvrdé), předchozí použití jedné značky a teplotu praní (studené nebo teplé). Zjistili, jak teplota souvisí s měkkostí vody a také s preferencemi značky.

Statistika. Frekvence, rezidua, odhady parametrů, standardní chyby, intervaly spolehlivosti a testy částečného přidružení. Pro vlastní modely, grafy zbytků a normální pravděpodobnostní grafy.

Aspekty dat loglineární analýzy výběru modelu

Data. Faktorové proměnné jsou kategorické. Všechny proměnné, které mají být analyzovány, musí být číselné. Kategoriální řetězcové proměnné lze před spuštěním analýzy výběru modelu překódovat na číselné proměnné.

Vyhnete se zadávání mnoha proměnných s mnoha úrovněmi. Tyto specifikace mohou vést k situaci, kdy mnoho buněk má malý počet pozorování a hodnoty chí-kvadrát nemusí být užitečné.

Související procedury. Procedura výběru modelu může pomoci identifikovat termíny potřebné v modelu. Pak můžete pokračovat v vyhodnocování modelu pomocí General Loglinear Analysis nebo Logit Loglinear Analysis. K překódování řetězcových proměnných můžete použít Autorecode. Pokud má číselná proměnná prázdné kategorie, použijte Recode k vytvoření následných celočíselných hodnot.

Získání lineární analýzy výběru modelu

Z nabídek vyberte:

Analyzovat > Loglinear > Výběr modelu ...

1. Vyberte dva nebo více číselných kategorických faktorů.
2. Vyberte jednu nebo více proměnných faktorů v seznamu Faktor (y) a klepněte na volbu **Definovat rozsah**.
3. Definujte rozsah hodnot pro každou proměnnou faktorů.
4. Vyberte volbu ve skupině Sestavení modelu.

Volitelně můžete vybrat proměnnou váhy buňky pro zadání strukturálních nul.

Definice rozsahu logliniové analýzy

Musíte označit rozsah kategorií pro každou proměnnou faktorů. Hodnoty pro Minimum a Maximum odpovídají nejnižším a nejvyšším kategoriím faktorové proměnné. Obě hodnoty musí být celá čísla a minimální hodnota musí být menší než maximální hodnota. Případy s hodnotami mimo rozsah jsou vyloučeny. Zadáte-li například minimální hodnotu 1 a maximální hodnotu 3, budou použity pouze hodnoty 1, 2 a 3. Opakujte tento proces pro každou proměnnou faktorů.

Model loglineární analýzy

Určit model. Nasycený model obsahuje všechny hlavní účinky faktorů a všechny interakce mezi faktory. Chcete-li určit třídu generování pro nenasycený model, vyberte volbu **Vlastní**.

Generování tříd. Třída generování je seznam termínů s nejvyšším pořadím, ve kterých se objevují faktory. Hierarchický model obsahuje výrazy, které definují třídu generování a všechny příbuzné s nižším pořadím. Předpokládejme, že v seznamu Faktory vyberete proměnné A, B a C a poté v rozevřacím seznamu Podmínky sestavení vyberete volbu **Interakce**. Výsledný model bude obsahovat zadanou trojcestnou interakci $A^* B^* C$, obousměrné interakce $A^* B$, $A^* C$ a $B^* C$ a hlavní efekty pro A, B a C. Neuvádějte příbuzné nižší pořadí v generující třídě.

Pro vybrané faktory:

Interakce

Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné. Jedná se o výchozí nastavení.

Hlavní účinky

Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Všechny 2-way

Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3-way

Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny 4-way

Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Vše 5-ti cestné

Vytvoří všechny možné pěticečné interakce vybraných proměnných.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Výběr modelu Loglinear Analysis Options

Zobrazit. Můžete zvolit **Frekvence**, **Zbytky**, nebo obojí. V nasyceném modelu jsou pozorované a očekávané frekvence stejné a rezidua jsou rovna 0.

Vykreslit. Pro vlastní modely můžete zvolit jeden nebo oba typy grafů, **Zbytky** a **Normální pravděpodobnost**. Ty pomohou určit, jak dobře model vyhovuje datům.

Zobrazit pro sytost modelu. Pro nasycený model můžete zvolit **odhady parametrů**. Odhady parametrů mohou pomoci určit, které termíny lze z modelu zrušit. K dispozici je také tabulka přidružení, která vypisuje testy částečného přidružení. Tato volba je výpočetně nákladná pro tabulky s mnoha faktory.

Kritéria modelu. K získání odhadů parametrů se používá iterativní algoritmus proporcionálního přizpůsobení. Můžete přepsat jedno nebo více kritérií odhadu zadáním **Maximální počet iterací**, **Konvergenční** nebo **Rozdílová data** (hodnota přidaná ke všem frekvencím buněk pro nasycené modely).

Další funkce příkazu HILOGLINEAR

Jazyk syntaxe příkazu vám také umožňuje:

- Určete váhy buněk ve tvaru matice (pomocí podpříkazu **CWEIGHT**).
- Generujte analýzy několika modelů pomocí jediného příkazu (pomocí podpříkazu **DESIGN**).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Obecná loglineární analýza

Procedura obecné loglineární analýzy analyzuje četnost pozorování, která spadají do každé kategorie křížové klasifikace v křížové nebo pohotovostní tabulce. Každá křížová klasifikace v tabulce představuje buňku a každá kategorická proměnná se nazývá faktor. Závislá proměnná je počet případů (frekvence) v buňce křížové a vysvětlující proměnné jsou faktory a kovariáty. Tento postup odhaduje parametry maximální pravděpodobnosti hierarchických a nehierarchických loglineárních modelů pomocí Newton-Raphsonovy metody. Lze analyzovat buď Poissonovo, nebo multinomální rozdělení.

Chcete-li definovat buňky tabulky, můžete vybrat až 10 faktorů. Proměnná struktury buňky vám umožňují definovat strukturální nuly pro neúplné tabulky, zahrnout do modelu výraz offsetu, přizpůsobit model log-rate nebo implementovat metodu úpravy okrajových tabulek. Kontrastní proměnné umožňují výpočet generalizovaných poměrů log-odds (GLOR).

Automaticky se zobrazí informace o modelu a statistické údaje o dobré kondici. Můžete také zobrazit různé statistiky a grafy nebo uložit rezidua a předpokládané hodnoty v aktivní datové sadě.

Příklad. Údaje ze zprávy o automobilových nehodách na Floridě se používají k určení vztahu mezi nošením bezpečnostního pásu a tím, zda bylo zranění smrtelné nebo nesmrtelné. Poměr kurzů naznačuje významné důkazy o vztahu.

Statistika. Pozorované a očekávané frekvence; surové, upravené a odchylovající se rezidua; návrhové matice; odhady parametrů; poměr kurzů; poměr log-odds; GLOR; Waldova statistika; a intervaly spolehlivosti. Grafy: upravené zbytky, odchylky zbytků a normální pravděpodobnost.

Aspekty dat obecné loglineární analýzy

Data. Faktory jsou kategorické a buněčné kovariáty jsou spojité. Je-li kovariát v modelu, použije se na tuto buňku střední kovariantní hodnota pro případy v buňce. Kontrastní proměnné jsou spojité. Používají se k výpočtu generalizovaných poměrů log-odds. Hodnoty kontrastní proměnné jsou koeficienty pro lineární kombinaci protokolů očekávaných počtů buněk.

Proměnná struktury buňky přiřazuje váhy. Pokud jsou například některé buňky strukturální nuly, má proměnná struktury buňky hodnotu 0 nebo 1. K vážení agregovaných dat nepoužívejte proměnnou struktury buňky. Místo toho vyberte volbu **Případy váhy** z nabídky Data.

Předpoklady. V General Loglinear Analysis jsou k dispozici dvě distribuce: Poisson a multinomial.

Podle Poissonova předpokladu rozdělení:

- Celková velikost vzorku není před studií stanovena nebo analýza není podmíněna celkovou velikostí vzorku.
- Událost pozorování v buňce je statisticky nezávislá na počtu buněk jiných buněk.

Za předpokladu multinomiální distribuce:

- Celková velikost vzorku je pevná nebo je analýza podmíněna celkovou velikostí vzorku.
- Počty buněk nejsou statisticky nezávislé.

Související procedury. Pomocí procedury kontingenčních tabulek prozkoumejte křížové tabulky. Proceduru Logit Loglinear použijte, když je přirozené považovat jednu nebo více kategorických proměnných za proměnné odezvy a ostatní za vysvětlující proměnné.

Získání obecné loglineární analýzy

1. Z nabídek vyberte:

Analyzovat > Loglinear > Obecné ...

2. V dialogovém okně Obecná loglineární analýza vyberte až 10 proměnných faktorů.

Volitelně můžete:

- Vyberte buňky kovariáty.
- Chcete-li definovat strukturální nuly nebo zahrnout výraz offsetu, vyberte proměnnou struktury buňky.
- Vyberte kontrastní proměnnou.

Model obecné loglineární analýzy

Určit model. Nasycený model obsahuje všechny hlavní efekty a interakce zahrnující faktorové proměnné. Neobsahuje kovariantní výrazy. Volbu **Vlastní** vyberte, chcete-li určit pouze podmnožinu interakcí nebo interakce faktoru podle kovariátu.

Faktory & kryty. Jsou uvedeny faktory a kovariáty.

Podmínky v modelu. Model závisí na povaze vašich dat. Po výběru volby **Vlastní** můžete vybrat hlavní efekty a interakce, které vás zajímají v analýze. Musíte označit všechny podmínky, které mají být zahrnuty do modelu.

Pro vybrané faktory a kovariáty:

Interakce

Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné. Jedná se o výchozí nastavení.

Hlavní účinky

Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Všechny 2-way

Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3-way

Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny 4-way

Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Vše 5-ti cestné

Vytvoří všechny možné pěticečné interakce vybraných proměnných.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Obecné volby logliniové analýzy

Procedura obecné logliniové analýzy zobrazuje informace o modelu a statistiku dobré pověsti. Kromě toho si můžete vybrat jednu nebo více z následujících možností:

Zobrazit. Pro zobrazení je k dispozici několik statistik -- pozorované a očekávané frekvence buněk; nezpracované, upravené a odchylovající se rezidua; návrhové matice modelu; a odhady parametrů pro model.

Vykreslit. Grafy, které jsou k dispozici pouze pro vlastní modely, zahrnují dvě matice bodového grafu (upravené zbytkové chyby nebo odchylky zbytků vůči pozorovanému a očekávanému počtu buněk). Můžete také zobrazit normální pravděpodobnost a detrendované normální grafy upravených zbytkových chyb nebo odchylek zbytkových chyb.

Interval spolehlivosti. Interval spolehlivosti pro odhady parametrů lze upravit.

Kritéria. Metoda Newton-Raphson se používá k získání odhadů parametrů maximální pravděpodobnosti. Můžete zadat nové hodnoty pro maximální počet iterací, kritérium konvergence a rozdílová data (konstanta přidaná ke všem buňkám pro počáteční aproximace). Delta zůstává v buňkách pro nasycené modely.

Uložení obecné logliniové analýzy

Vyberte hodnoty, které chcete uložit jako nové proměnné v aktivní datové sadě. Přípona *n* v přírůstcích názvů nových proměnných vytvoří jedinečný název pro každou uloženou proměnnou.

Uložené hodnoty odkazují na agregovaná data (buňky v kontingenční tabulce), i když jsou data zaznamenána v jednotlivých pozorováních v editoru dat. Pokud uložíte rezidua nebo předpovězené hodnoty pro neagregovaná data, uložená hodnota pro buňku v kontingenční tabulce se zadá v editoru

dat pro každý případ v dané buňce. Chcete-li porozumět uloženým hodnotám, měli byste agregovat data, abyste získali počty buněk.

Lze uložit čtyři typy zbytkových chyb: surové, standardizované, upravené a odchylky. Predikované hodnoty lze také uložit.

- *Zbytky*. Také nazývaný jednoduchý nebo nezpracovaný reziduum, je to rozdíl mezi pozorovaným počtem buněk a jeho očekávaným počtem.
- *Standardizované zbytkové chyby*. Zbytek vydělený odhadem standardní chyby. Standardizované rezidenční chyby jsou také známé jako Pearsonovy rezidenční chyby.
- *Upravené zbytkové chyby*. Standardizovaný zbytek vydělený odhadovanou standardní chybou. Vzhledem k tomu, že upravené rezidua jsou asymptoticky standardní normální, když je vybraný model správný, upřednostňují se před standardizovanými reziduály pro kontrolu normality.
- *Zbytková odchylka*. Podepsaná druhá odmocnina individuálního příspěvku ke statistice pravděpodobného poměru chí-kvadrát (G kvadrát), kde znaménko je znaménkem zbytkového (pozorovaný počet minus očekávaný počet). Odchylující se rezidua mají asymptotické standardní normální rozdělení.

Dodatečné funkce příkazu GENLOG

Jazyk syntaxe příkazu vám také umožňuje:

- Vypočtete lineární kombinace pozorovaných frekvencí buněk a očekávaných frekvencí buněk a tiskových zbytků, standardizovaných zbytků a upravených zbytků této kombinace (pomocí podpříkazu GERESID).
- Změňte výchozí prahovou hodnotu pro kontrolu redundance (pomocí podpříkazu CRITERIA).
- Zobrazte standardizované zbytkové chyby (pomocí podpříkazu PRINT).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Logit Loglineární analýza

Procedura Logit Loglinear Analysis analyzuje vztah mezi závislými (nebo odezвовými) proměnnými a nezávislými (nebo vysvětlujícími) proměnnými. Závislé proměnné jsou vždy kategorické, zatímco nezávislé proměnné mohou být kategorické (faktory). Jiné nezávislé proměnné (buněčné kovariáty) mohou být spojité, ale nejsou aplikovány případ od případu. Vážený kovariantní průměr pro buňku se použije na tuto buňku. Logaritmus pravděpodobnosti závislých proměnných je vyjádřen jako lineární kombinace parametrů. Automaticky se předpokládá multinomiální distribuce; tyto modely se někdy nazývají multinomiální logitové modely. Tento postup odhaduje parametry logit loglinear modelů pomocí Newton-Raphson algoritmu.

Můžete si vybrat z 1 až 10 závislých a faktorovaných proměnných. Proměnná struktury buňky vám umožňuje definovat strukturální nuly pro neúplné tabulky, zahrnout do modelu výraz offsetu, přizpůsobit model log-rate nebo implementovat metodu úpravy okrajových tabulek. Kontrastní proměnné umožňují výpočet generalizovaných poměrů log-odds (GLOR). Hodnoty kontrastní proměnné jsou koeficienty pro lineární kombinaci protokolů očekávaných počtů buněk.

Automaticky se zobrazí informace o modelu a statistické údaje o dobré kondici. Můžete také zobrazit různé statistiky a grafy nebo uložit rezidua a předpokládané hodnoty v aktivní datové sadě.

Příklad. Studie na Floridě zahrnovala 219 aligátorů. Jak se liší typ potravin aligátorů podle jejich velikosti a čtyř jezer, ve kterých žijí? Studie zjistila, že pravděpodobnost, že menší aligátor upřednostňuje plazy před rybami, je 0.70 krát nižší než u větších aligátorů; také pravděpodobnost výběru především plazů místo ryb byla nejvyšší v jezeře 3.

Statistika. Pozorované a očekávané frekvence; nezpracované, upravené a odchylující se rezidua; návrhové matice; odhady parametrů; generalizovaný poměr log-kurz; Waldova statistika; a intervaly spolehlivosti. Grafy: upravené zbytkové chyby, odchylky zbytkové hodnoty a normální pravděpodobnostní grafy.

Aspekty dat Logit Loglinear Analysis

Data. Závislé proměnné jsou kategorické. Faktory jsou kategorické. Buněčné kovariáty mohou být spojité, ale když je kovariát v modelu, střední kovariantní hodnota pro případy v buňce se aplikuje na tuto buňku. Kontrastní proměnné jsou spojité. Používají se k výpočtu generalizovaných poměrů log-odds (GLOR). Hodnoty kontrastní proměnné jsou koeficienty pro lineární kombinaci protokolů očekávaných počtů buněk.

Proměnná struktury buňky přiřazuje váhy. Pokud jsou například některé buňky strukturální nuly, má proměnná struktury buňky hodnotu 0 nebo 1. K vážení agregovaných dat nepoužívejte proměnnou struktury buňky. Místo toho použijte volbu Případy váhy v nabídce Data.

Předpoklady. Předpokládá se, že počty v rámci každé kombinace kategorií vysvětlujících proměnných mají multinomiální rozdělení. Za předpokladu multinomiální distribuce:

- Celková velikost vzorku je pevná nebo je analýza podmíněna celkovou velikostí vzorku.
- Počty buněk nejsou statisticky nezávislé.

Související procedury. K zobrazení kontingenčních tabulek použijte proceduru kontingenčních tabulek. Proceduru obecné logineární analýzy použijte, když chcete analyzovat vztah mezi pozorovaným počtem a sadou vysvětlujících proměnných.

Získání Logit Loglinear Analysis

1. Z nabídek vyberte:

Analyzovat > Loglinear > Logit ...

2. V dialogovém okně Logit Loglinear Analysis vyberte jednu nebo více závislých proměnných.

3. Vyberte jednu nebo více proměnných faktorů.

Celkový počet závislých a faktorovacích proměnných musí být menší nebo roven 10.

Volitelně můžete:

- Vyberte buňky kovariáty.
- Chcete-li definovat strukturální nuly nebo zahrnout výraz offsetu, vyberte proměnnou struktury buňky.
- Vyberte jednu nebo více kontrastní proměnných.

Logit Loglinear Analysis Model (lineární analýza logit)

Určit model. Nasycený model obsahuje všechny hlavní efekty a interakce zahrnující faktorové proměnné. Neobsahuje kovariantní výrazy. Volbu **Vlastní** vyberte, chcete-li určit pouze podmnožinu interakcí nebo interakce faktoru podle kovariátu.

Faktory & kryty. Jsou uvedeny faktory a kovariáty.

Podmínky v modelu. Model závisí na povaze vašich dat. Po výběru volby **Vlastní** můžete vybrat hlavní efekty a interakce, které vás zajímají v analýze. Musíte označit všechny podmínky, které mají být zahrnuty do modelu.

Pro vybrané faktory a kovariáty:

Interakce

Vytvoří termín interakce na nejvyšší úrovni pro všechny vybrané proměnné. Jedná se o výchozí nastavení.

Hlavní účinky

Vytvoří výraz hlavních efektů pro každou vybranou proměnnou.

Všechny 2-way

Vytvoří všechny možné obousměrné interakce vybraných proměnných.

Všechny 3-way

Vytvoří všechny možné trojcestné interakce vybraných proměnných.

Všechny 4-way

Vytvoří všechny možné čtyřcestné interakce vybraných proměnných.

Vše 5-ti cestné

Vytvoří všechny možné pěticeinterakce vybraných proměnných.

Termíny se přidávají do návrhu tak, že se vezmou všechny možné kombinace závislých výrazů a každá kombinace se shoduje s každým výrazem v seznamu modelů. Je-li vybrána volba **Zahrnout konstantu pro závislé**, je do seznamu modelů přidán také výraz jednotky (1).

Předpokládejme například, že proměnné $D1$ a $D2$ jsou závislé proměnné. Seznam závislých výrazů je vytvořen procedurou Logit Loglinear Analysis ($D1$, $D2$, $D1*D2$). Pokud podmínky v seznamu modelů obsahují $M1$ a $M2$ a je zahrnuta konstanta, seznam modelů obsahuje 1, $M1$ a $M2$. Výsledný návrh zahrnuje kombinace jednotlivých termínů modelu s každým závislým termínem:

$D1$, $D2$, $D1*D2$

$M1*D1$, $M1*D2$, $M1*D1*D2$

$M2*D1$, $M2*D2$, $M2*D1*D2$

Zahrnout konstantu pro závislé. Obsahuje konstantu pro závislou proměnnou ve vlastním modelu.

Podmínky sestavení a vlastní podmínky

Podmínky sestavení

Tuto volbu použijte, chcete-li zahrnout nevnořené výrazy určitého typu (například hlavní efekty) pro všechny kombinace vybrané sady faktorů a kovariátů.

Sestavit vlastní podmínky

Tuto volbu použijte, chcete-li zahrnout vnořené výrazy nebo chcete-li explicitně sestavit libovolnou proměnnou výrazu podle proměnné. Sestavení vnořeného výrazu zahrnuje následující kroky:

Volby Logit Loglinear Analysis

Procedura Logit Loglinear Analysis zobrazuje informace o modelu a statistiku dobré pověsti. Kromě toho můžete zvolit jednu nebo více z následujících voleb:

Zobrazit. Pro zobrazení je k dispozici několik statistik: pozorované a očekávané frekvence buněk; nezpracované, upravené a odchylující se rezidua; návrhové matice modelu a odhady parametrů pro model.

Vykreslit. Grafy dostupné pro vlastní modely zahrnují dvě matice bodového grafu (upravené zbytkové chyby nebo zbytkové odchylky vůči pozorovaným a očekávaným počtům buněk). Můžete také zobrazit normální pravděpodobnost a detrendované normální grafy upravených zbytkových chyb nebo odchylek zbytkových chyb.

Interval spolehlivosti. Interval spolehlivosti pro odhady parametrů lze upravit.

Kritéria. Metoda Newton-Raphson se používá k získání odhadů parametrů maximální pravděpodobnosti. Můžete zadat nové hodnoty pro maximální počet iterací, kritérium konvergence a rozdílová data (konstanta přidaná ke všem buňkám pro počáteční aproximace). Delta zůstává v buňkách pro nasycené modely.

Logit Loglinear Analysis Uložit

Vyberte hodnoty, které chcete uložit jako nové proměnné v aktivní datové sadě. Přípona n v přírůstcích názvů nových proměnných vytvoří jedinečný název pro každou uloženou proměnnou.

Uložené hodnoty odkazují na agregovaná data (na buňky v kontingenční tabulce), i když jsou data zaznamenána v jednotlivých pozorováních v editoru dat. Pokud uložíte rezidua nebo předpovězené hodnoty pro neagregovaná data, uložená hodnota pro buňku v kontingenční tabulce se zadá v editoru dat pro každý případ v dané buňce. Chcete-li porozumět uloženým hodnotám, měli byste agregovat data, abyste získali počty buněk.

Lze uložit čtyři typy zbytkových chyb: surové, standardizované, upravené a odchylky. Predikované hodnoty lze také uložit.

- *Zbytky*. Také nazývaný jednoduchý nebo nezpracovaný reziduum, je to rozdíl mezi pozorovaným počtem buněk a jeho očekávaným počtem.
- *Standardizované zbytkové chyby*. Zbytek vydělený odhadem standardní chyby. Standardizované rezidenční chyby jsou také známé jako Pearsonovy rezidenční chyby.
- *Upravené zbytkové chyby*. Standardizovaný zbytek vydělený odhadovanou standardní chybou. Vzhledem k tomu, že upravené rezidua jsou asymptoticky standardní normální, když je vybraný model správný, upřednostňují se před standardizovanými reziduály pro kontrolu normality.
- *Zbytková odchylka*. Podepsaná druhá odmocnina individuálního příspěvku ke statistice pravděpodobného poměru chí-kvadrát (G kvadrát), kde znaménko je znaménkem zbytkového (pozorovaný počet minus očekávaný počet). Odchylující se rezidua mají asymptotické standardní normální rozdělení.

Dodatečné funkce příkazu GENLOG

Jazyk syntaxe příkazu vám také umožňuje:

- Vypočítejte lineární kombinace pozorovaných frekvencí buněk a očekávaných frekvencí buněk a vytiskněte rezidua, standardizované rezidua a upravené rezidua této kombinace (pomocí podpříkazu GERESID).
- Změňte výchozí prahovou hodnotu pro kontrolu redundance (pomocí podpříkazu CRITERIA).
- Zobrazte standardizované zbytkové chyby (pomocí podpříkazu PRINT).

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Životní tabulky

Existuje mnoho situací, ve kterých byste chtěli prozkoumat rozdělení časů mezi dvě události, jako je délka zaměstnání (doba mezi najatím a odchodem ze společnosti). Nicméně, tento druh dat obvykle zahrnuje některé případy, pro které druhá událost není zaznamenána (například, lidé stále pracují pro společnost na konci studie). To se může stát z několika důvodů: v některých případech se událost prostě nevyskytuje před koncem studie; v jiných případech ztrácíme přehled o svém stavu někdy před koncem studie; ještě jiné případy nemusí být schopny pokračovat z důvodů, které nesouvisejí se studií (například zaměstnanec onemocní a vezme si volno). Tyto případy jsou souhrnně označovány jako **cenzenované případy** a činí tento druh studie nevhodný pro tradiční techniky, jako jsou testy t nebo lineární regrese.

Statistická technika užitečná pro tento typ dat se nazývá následná **tabulka životnosti**. Základní myšlenkou životní tabulky je dále rozdělit období pozorování do menších časových intervalů. Pro každý interval jsou všichni lidé, kteří byli sledováni alespoň tak dlouho, použiti k výpočtu pravděpodobnosti události terminálu, která se v tomto intervalu vyskytne. Pravděpodobnosti odhadnuté z jednotlivých intervalů se pak použijí k odhadu celkové pravděpodobnosti výskytu události v různých časových bodech.

Příklad. Je nová léčba nikotinovou náplastí lepší než tradiční léčba náplastí, která pomáhá lidem přestat kouřit? Mohli byste provést studii s použitím dvou skupin kuřáků, z nichž jedna dostávala tradiční terapii a druhá experimentální terapii. Vytvoření životních tabulek z dat by vám umožnilo porovnat celkovou míru abstinence mezi oběma skupinami, abyste zjistili, zda je experimentální léčba lepší než tradiční léčba. Můžete také vykreslit funkce přežití nebo nebezpečí a vizuálně je porovnat, abyste získali podrobnější informace.

Statistika. Počet zadávaných, počet odcházejících, počet vystavených riziku, počet koncových událostí, ukončený podíl, přežívající podíl, kumulativní podíl přežívajících (a standardní chyba), hustota pravděpodobnosti (a standardní chyba) a míra rizika (a standardní chyba) pro každý časový interval pro každou skupinu; střední doba přežívání pro každou skupinu; a Wilcoxonův (Gehanův) test pro porovnání rozdělení přežití mezi skupinami. Grafy: funkční grafy pro přežití, log přežití, hustota, míra rizika a jedno mínus přežití.

Aspekty dat tabulek životnosti

Data. Vaše časová proměnná by měla být kvantitativní. Vaše stavová proměnná by měla být dichotomní nebo kategorická, kódovaná jako celá čísla, s událostmi kódovanými jako jedna hodnota nebo rozsah po sobě jdoucích hodnot. Proměnné faktoru by měly být kategorické, kódované jako celá čísla.

Předpoklady. Pravděpodobnost události, která je předmětem zájmu, by měla záviset pouze na čase po počáteční události-předpokládá se, že jsou stabilní s ohledem na absolutní čas. To znamená, že případy, které vstupují do studie v různých časech (například pacienti, kteří začínají léčbu v různých časech), by se měli chovat podobně. Neměly by také existovat žádné systematické rozdíly mezi cenzurovanými a necenzurovanými případy. Pokud například mnoho cenzurovaných případů jsou pacienti s vážnějšími podmínkami, vaše výsledky mohou být zkreslené.

Související procedury. Procedura Life Tables používá k tomuto druhu analýzy pojistně-matematický přístup (obecně známý jako analýza přežití). Postup Kaplan-Meier Survival Analysis používá mírně odlišnou metodu výpočtu životních tabulek, která nespočívá na rozdělení sledovaného období do menších časových intervalů. Tato metoda se doporučuje, pokud máte malý počet pozorování, tak, aby tam byl jen malý počet pozorování v každém časovém intervalu přežití. Pokud máte proměnné, u kterých máte podezření, že souvisejí s dobou přežití nebo s proměnnými, které chcete kontrolovat (kovariáty), použijte proceduru Cox Regression. Pokud mohou mít vaše kovariáty různé hodnoty v různých časových bodech pro stejný případ, použijte Cox Regression with Time-Dependent Covariates.

Vytváření životních tabulek

1. Z nabídek vyberte:

Analýza > Přežití > Tabulky života ...

2. Vyberte jednu *číselnou* proměnnou přežití.

3. Uveďte časové intervaly, které se mají zkontrolovat.

4. Vyberte proměnnou stavu, abyste definovali případy, pro které došlo k události terminálu.

5. Klepněte na volbu **Definovat událost**, chcete-li uvést hodnotu proměnné stavu, která označuje, že došlo k události.

Volitelně můžete vybrat proměnnou faktoru prvního pořadí. Pojistně matematické tabulky pro proměnnou přežití jsou generovány pro každou kategorii faktorové proměnné.

Můžete také vybrat proměnnou *podle faktoru* druhého pořadí. Pojistně matematické tabulky pro proměnnou přežití jsou generovány pro každou kombinaci faktorových proměnných prvního a druhého řádu.

Rozsah definice tabulek životnosti

Případy s hodnotami pro proměnnou faktoru v zadaném rozsahu budou zahrnuty do analýzy a samostatné tabulky (a grafy, pokud jsou požadovány) budou generovány pro každou jedinečnou hodnotu v rozsahu.

Volby životních tabulek

Můžete řídit různé aspekty analýzy životních tabulek.

Tabulky životnosti. Chcete-li potlačit zobrazení tabulek životnosti ve výstupu, zrušte výběr volby **Tabulky životnosti**.

Vykreslit. Umožňuje vám požádat o grafy funkcí přežití. Pokud jste definovali faktorovou proměnnou (proměnné), jsou grafy generovány pro každou podskupinu definovanou (i) faktorovou proměnnou (y). Dostupné pozemky jsou přežití, log přežití, nebezpečí, hustota, a jeden mínus přežití.

- *Přežití.* Zobrazuje funkci kumulativního přežití na lineárním měřítku.
- *Přežití protokolu.* Zobrazuje funkci kumulativního přežití na logaritmické stupnici.
- *Nebezpečí.* Zobrazuje funkci kumulativního rizika na lineárním měřítku.
- *Hustota.* Zobrazí funkci hustoty.
- *Jeden mínus přežití.* Plots one-minus funkce přežití na lineární stupnici.

Porovnat úroveň prvního faktoru. Máte-li řídicí proměnnou prvního řádu, můžete vybrat jednu z alternativ v této skupině k provedení testu Wilcoxon (Gehan), který porovnává přežití podskupin. Testy se provádějí s faktorem prvního řádu. Pokud jste definovali faktor druhého řádu, provedou se testy pro každou úroveň proměnné druhého řádu.

Další funkce příkazu SURVIVAL

Jazyk syntaxe příkazu vám také umožňuje:

- Zadejte více než jednu závislou proměnnou.
- Zadejte nerovnoměrně rozložené intervaly.
- Zadejte více než jednu proměnnou stavu.
- Uvedte porovnání, která nezahrnují všechny faktory a všechny řídicí proměnné.
- Vypočítejte přibližná, spíše než přesná porovnání.

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Parametrické časové modely urychleného selhání

Parametrická analýza modelu AFT (Parametric Accelerated Failure Time) vyvolává parametrický postup modelů přežití s neopakovanými daty doby života. Parametrické modely přežití předpokládají, že doba přežití se řídí známou distribucí, a tato analýza se hodí ke zrychleným časovým modelům selhání s jejich modelovými účinky proporcionálním s ohledem na dobu přežití.

Získání parametrické analýzy časových modelů urychleného selhání

1. Z nabídek vyberte:

Analýza > Přežití > Modely parametrické akcelerované doby selhání (AFT)

2. Vyberte zdrojovou proměnnou.

Čas

Přežití

Jednotlivá číselná proměnná označující dobu trvání přežití.

Začátek/konec

Číselné proměnné označující **Čas zahájení** a **Čas ukončení**.

Stav

Jeden volitelný řetězec nebo číselná proměnná, která určuje jedno z následujících nastavení stavu:

Selhání/událost

Mapuje záznam na kategorii selhání/události. Výchozí hodnota pro proměnnou stavu řetězce je F.

Správné hodnocení

Mapuje záznam na pravou kategorii cenzury. Výchozí hodnota pro proměnnou stavu řetězce je R.

Levé hodnocení

Mapuje záznam na kategorii cenzury vlevo. Výchozí hodnota pro proměnnou stavu řetězce je L.

Zajištění intervalu

Mapuje záznam na kategorii cenzury intervalu. Pouze pro **začátek/konec**. Výchozí hodnota pro proměnnou stavu řetězce je I.

Zacházení s nemapovanými hodnotami

Řídí, na kterou kategorii se mají mapovat nemapované záznamy. Chcete-li odstranit záznamy, jejichž mapování se nezdařilo, vyberte volbu **Vyloučit je z analýzy**.

Pro **Přežití** je výchozí stav pro všechny případy **Selhání/Událost**. Pro **Začátek/konec** je výchozí stav **Interval Censoring**. Klepněte na tlačítko **Definovat událost**, chcete-li definovat událost pro proměnnou stavu.

Kovariát (y)

Jedna nebo více volitelných číselných proměnných, které mají být považovány za kovariáty. Všimněte si, že proměnnou nelze zadat současně pomocí **Covariate (s)** a **Fixed Factor (s)**.

Pevný (é) faktor (y)

Jedna nebo více volitelných proměnných, které se mají považovat za faktory. Proměnnou nelze zadat současně pomocí **fixních faktorů** a **Covariate (s)**.

Oříznutí vlevo

Jediná volitelná číselná proměnná pro levé oříznutí pouze pro **Přežití**.

Parametrické časové modely urychleného selhání: Kritéria

Podmínky

Volitelný panel pro určení obecných kritérií.

Interval spolehlivosti

Volitelná procentní část pro určení úrovně intervalů spolehlivosti regresních parametrů. Musí to být jedna číselná hodnota mezi 0 a 100. Výchozí hodnota je 95.

Chybějící hodnoty

Volba pro řízení způsobu zacházení s chybějícími hodnotami uživatele:

Vyloučit chybějící i chybějící hodnoty uživatele i systému

Zachází s chybějícími hodnotami uživatele jako s platnými hodnotami. Jedná se o výchozí nastavení.

S chybějícími hodnotami uživatele se zachází jako s platnými

Ignoruje určení chybějících hodnot uživatele a považuje je za platné hodnoty.

Zacházení se stavem

Pouze pro **začátek/konec**. Volba pro řízení toho, jak zacházet se záznamy s chybnými poli stavu:

Vyřadit konfliktní záznam

Zruší konfliktní záznamy. Toto je výchozí nastavení.

Získat informace o čase podle stavu

Získá informace o čase podle stavu.

Odvodit stav podle informací o čase

Změní stav podle informací o čase.

Parametrické akcelerované časové modely selhání: Model

Model

Volitelný panel pro určení voleb a nastavení modelu.

Rozdělení doby přežití

Možnost určit rozdělení doby přežití.

WEIBULL

Určuje Weibullovo rozdělení. Toto je výchozí nastavení.

Exponenciální

Určuje exponenciální rozdělení.

Protokol-normální

Určuje normální distribuci protokolu.

Protokol-Logistika

Určuje logistickou distribuci protokolu.

Nastavení kovariace

Zadejte proměnné kovariate.

Nastavení faktoru

Zadejte proměnné faktoru.

Počáteční hodnota zachycení

Volba pro zadání počáteční hodnoty termínu zachycení. Je-li uveden, musí to být jedna číselná hodnota a nesmí být 0.

Počáteční hodnota parametru měřítka

Volba pro řízení nastavení parametru měřítka.

Standardní chyba odpovídající regrese OLS

Jako počáteční hodnotu použije standardní chybu odpovídající běžné regrese nejmenších čtverců.

Obrátí standardní chybu odpovídající regrese OLS.

Používá reciproční standardní chybu.

Hodnota zadaná uživatelem

Je-li zadána jedna číselná hodnota, použije se jako počáteční hodnota. Je-li uveden, musí být větší než 0.

Parametrické časové modely urychleného selhání: odhad

Odhad

Volitelný panel, který určuje nastavení pro řízení odhadu časových modelů zrychleného selhání a volitelného procesu výběru funkcí.

Metoda střídavého směru nebo multiplikátory (ADMM)

Rychle

Používá metodu rychlého střídavého směru multiplikátorů (ADMM). Jedná se o výchozí nastavení.

Tradiční

Použije tradiční algoritmus ADMM.

Použít regularizaci L-1

Provádí proces pro řízení výběru funkcí. Pole **Parametr trestu** uvádí parametr pokuty, který řídí proces regularizace. Musí být jedna hodnota větší než 0. Výchozí nastavení je 0.001.

Kritéria konvergence modelu

Konvergence parametrů

Určuje kritéria konvergence pro parametr. Musí se jednat o jedinou číselnou hodnotu patřící k [0, 1). Výchozí nastavení je 0.000001. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelné pole **Hodnota** určuje klíčové slovo.

Konvergence objektivních funkcí

Určuje kritéria konvergence pro funkci cíle. Musí se jednat o jedinou číselnou hodnotu patřící k [0, 1). Výchozí nastavení je 0, které neuplatňuje kritéria konvergence. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelné pole **Hodnota** určuje klíčové slovo.

Hessovská konvergence

Určuje kritéria konvergence pro hessovskou matici. Musí se jednat o jedinou číselnou hodnotu patřící k [0, 1). Výchozí nastavení je 0, které neuplatňuje kritéria konvergence. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelné pole **Hodnota** určuje klíčové slovo.

Kritéria zbytkové konvergence

Volba pro řízení procesu optimalizace.

Primární i duální rezidua

Uplatňuje se jak základní, tak i dvojí zbytkové konvergenční kritérium. Toto je výchozí nastavení.

Pouze primární rezidua

Použije se primární zbytkové konvergenční kritérium.

Pouze duální reziduální

Použije dvojí kritérium zbytkové konvergence.

Metoda

Nepovinný parametr pro určení metody odhadu.

Automaticky

Automaticky vybere metodu na základě ukázkové datové sady. Jedná se o výchozí nastavení. Pole **Prahová hodnota počtu prediktorů** určuje prahovou hodnotu počtu prediktorů a musí být celé číslo větší než 1. Výchozí hodnota je 1000.

Newton-Raphsonová

Používá Newton-Raphsonovu metodu.

L-BFGS

Použije algoritmus BFGS s omezenou pamětí. Pole **Aktualizace** uvádí počet minulých aktualizací udržovaných algoritmem BFGS s omezenou pamětí a musí být jediné celé číslo větší nebo rovno 1. Výchozí hodnota je 5.

Iterace**Maximální počet iterací**

Určuje maximální počet iterací. Musí se jednat o jedno celé číslo patřící do [1, 100]. Výchozí nastavení je 20.

Maximální půlení schodů

Určuje maximální počet kroků na polovinu. Musí to být jedno celé číslo patřící do [1, 20]. Výchozí nastavení je 5.

Maximální počet hledání řádků

Určuje maximální počet hledání řádků. Musí se jednat o jedno celé číslo patřící do [1, 100]. Výchozí nastavení je 20.

Absolutní konvergence pro iterační proces

Určuje absolutní konvergenci pro proces vnější iterace. Musí se jednat o jedinou číselnou hodnotu, která patří (0, 1). Výchozí nastavení je 0.0001.

Relativní konvergence pro proces iterace

Určuje relativní konvergenci pro proces vnější iterace. Musí se jednat o jedinou číselnou hodnotu, která patří (0, 1). Výchozí nastavení je 0.01.

Parametrické časové modely urychleného selhání: Tisk

Tisk

Volitelný panel pro řízení výstupů tabulky.

Podrobnosti kódování faktoru

Je-li vybráno, zobrazí a vytiskne podrobnosti kódování faktorů. Proces je ignorován, pokud neexistují žádné platné faktory.

Počáteční hodnoty přiřazené parametrům regrese

Je-li vybráno, zobrazí počáteční hodnoty použité v procesu odhadu.

Historie iterací modelu

Je-li vybráno, zobrazí se historie iterací analýzy přežití. Do pole **Počet kroků** zadejte počet kroků mezi 1 a 99999999. Výchozím nastavením je hodnota 1.

Výsledky výběru obsahující

Řídí zobrazení podrobností výběru funkce.

Vybrané i nevybrané proměnné

Zobrazit v tabulce vybrané i nevybrané proměnné.

Pouze vybrané proměnné

Zobrazit pouze vybrané proměnné.

Pouze nevybrané proměnné

Zobrazit pouze nevybrané proměnné. Pole **Maximální počet proměnných k zobrazení** určuje maximální počet proměnných vytištěných v tabulce. Výchozí nastavení je 30.

Parametrické časové modely akcelerovaného selhání: Předpovídat

Předvídat

Volitelný panel pro skóre a uložení předpovězených statistik do aktivní datové sady.

Časové hodnoty pro skórování

Časové hodnoty definované závislými proměnnými

Skóre **Predikce** na základě časové proměnné určené pro parametrický model přežití.

Pravidelné intervaly

Skóre **Predictions** na základě budoucích časových hodnot. Pole **Časový interval** uvádí časový interval a musí být jedna číselná hodnota větší než 0. Pole **Počet časových období** uvádí počet časových období a musí se jednat o jedno číselné celé číslo mezi 2 a 100.

Doba trvání

Skóre **Předpovědi** na základě doby trvání definuje budoucí časové hodnoty. Musí se jednat o jedinou číselnou proměnnou.

Předpovědi

Přežití

Skóre a uloží predikovanou statistiku přežití do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je PredSurvival.

Riziko

Skóre a uloží předpokládaná rizika do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je PredHazard.

Kumulativní riziko

Skóre a uloží předpokládaná kumulativní rizika do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je PredCumHazard.

Podmíněné přežití

Skóre a uloží predikovanou statistiku podmíněného přežití do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je PredConditionalSurvival. Není-li zadán parametr PASTTIME, bude proces ignorován. Je požadována hodnota **Doba po minulém přežití** a uvádí hodnoty uplynulého času pro přidělení skóre. Musí se jednat o jedinou číselnou proměnnou.

Parametrické akcelerované časové modely selhání: zakreslení

Graf

Cox-Snell zbytkový pozemek

Vyberte volbu **Zobrazit zákres**, chcete-li vytvořit zbytkový zákres Cox-Snell. Do pole **Počet bodů oříznutí** zadejte číslo od 1 do 10000. Výchozí nastavení je 100.

Funkční partie

Volba pro řízení vykreslení funkcí.

Typ

Přežití

Vytvoří děj pro funkce přežití.

Riziko

Vytvoří zákres pro rizikové funkce.

Hustota

Vytvoří graf pro funkce hustoty.

Počet bodů k zobrazení

Určuje počet funkčních bodů mezi 1 a 200. Výchozí nastavení je 100.

Kovariantní hodnoty pro zápočet

Volitelné určení uživatelem zadaných hodnot a jejich přiřazení k prediktorům. Standardně se určené grafy vytvoří na **Střední** každého kovariátu v platnosti a frekvenci kategorií každého faktoru v platnosti. Je-li uvedeno, určené grafy budou vytvořeny na základě nastavení vzoru. V přítomnosti libovolných duplicitních proměnných bude rozpoznána ta, která byla uvedena jako první, a zbytek bude ignorován. Platná proměnná musí být obsažena v efektu modelu. Pro kovariát musí být uživatelem zadaná hodnota číselná. Vynechání platné proměnné označuje, že frekvence kategorie a **Střední** by se standardně použily pro faktor a kovariát. Pokud je proměnné přiřazena neplatná hodnota, požadovaný vzor nebude zakreslen.

Hodnoty faktoru pro Plot

Volitelné určení uživatelem zadaných hodnot a jejich přiřazení k prediktorům. V přítomnosti libovolných duplicitních proměnných bude rozpoznána ta, která byla uvedena jako první, a zbytek bude ignorován. Platná proměnná musí být obsažena v efektu modelu. Vynechání účinné proměnné znamená, že pro daný faktor a pro kovariát by byla použita frekvence kategorie a střední hodnota. Pokud je proměnné přiřazena neplatná hodnota, požadovaný vzor nebude zakreslen.

Oddělené řádky pro

Volba pro určení kategorické proměnné, podle které budou vykresleny čárový grafy.

Maximální počet čar v grafu

Určuje maximální počet čar v grafu, je-li zadána volba **Samostatné čáry pro** . Výchozí nastavení je 10.

Parametrické akcelerované časové modely selhání: Export

Exportovat

Vyberte volbu **Exportovat informace o modelu do souboru XML** , chcete-li zapsat informace o modelu a parametru do souboru PMML pro přidělení skóre. Musíte uvést adresář a název souboru PMML, který se má uložit.

Survival AFT Definovat události pro proměnné stavu

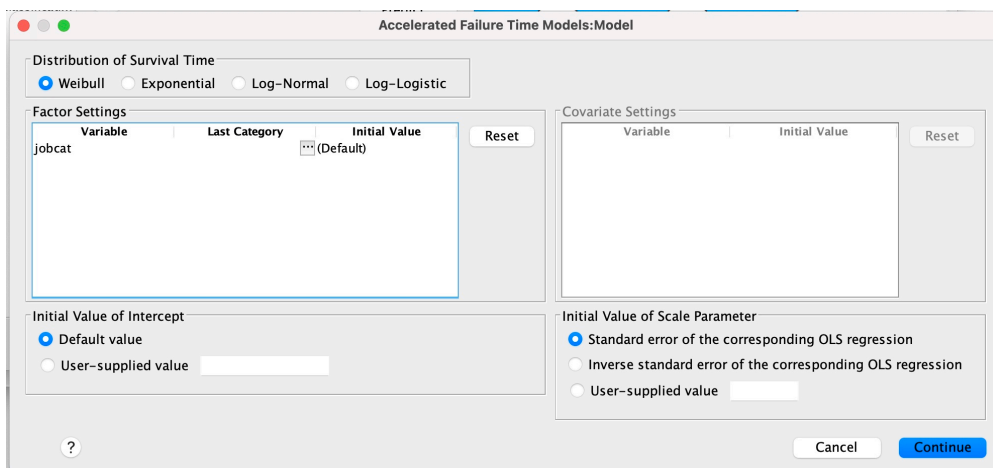
Výskyty vybrané hodnoty nebo hodnot pro proměnnou stavu označují, že pro tyto případy došlo k události terminálu. Všechny ostatní případy jsou považovány za cenzurované. Zadejte buď jednotlivou hodnotu, nebo rozsah hodnot, které identifikují událost, která je předmětem zájmu.

Parametrické časové modely akcelerovaného selhání: Vybrat kategorii

Nastavení Vybrat kategorii poskytuje volby pro výběr hodnoty, která označuje kategorii, která se má modelovat jako základna pro porovnání.

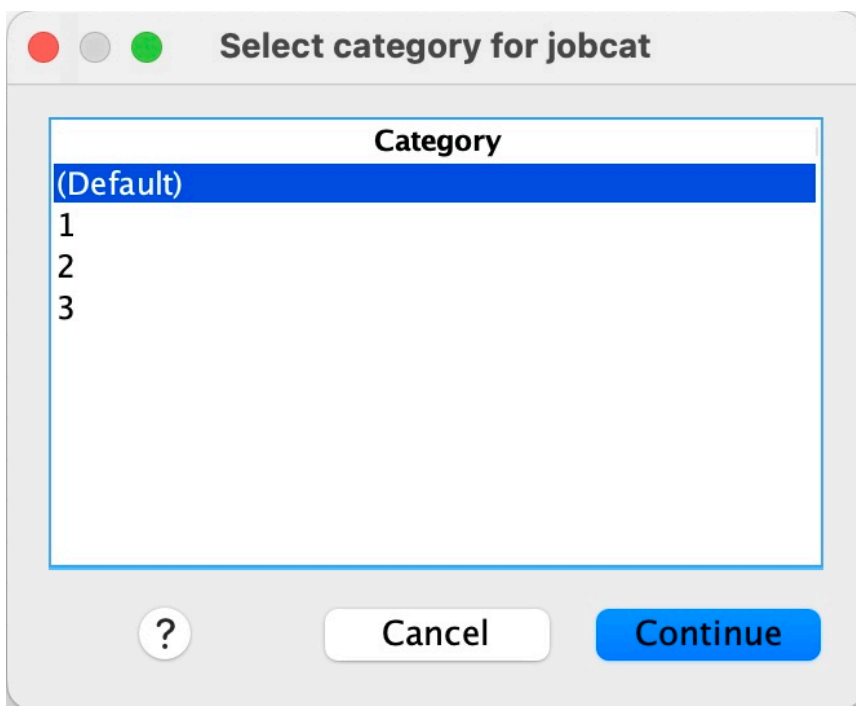
Výběr kategorie

Kliknutím na 'Poslední kategorii' otevřete dialogové okno 'Vybrat kategorii'.



Obrázek 3. Modely zrychleného života-dialogové okno-Kategorie

Chcete-li označit kategorii jako úroveň baseline, vyberte hodnotu z dialogového okna 'Vybrat kategorii'.



Obrázek 4. Dialogové okno Zrychlené modely života-Vybrat kategorii

Klepněte na tlačítko Pokračovat.

Parametrické sdílené modely křehkosti

Parametric Shared Frailty Models Survival analýza spouští parametrický postup pro modely přežití s opakovaným vstupem dat o době života. Parametrické modely přežití předpokládají, že doba přežití následuje po známé distribuci, a tato analýza zahrnuje křehkou dobu do parametrického modelu přežití. Je považována za náhodnou složku, která zohledňuje nepozorovaný účinek v důsledku individuálních nebo skupinových rozdílů.

Získání parametrické analýzy sdílených modelů křehkosti

1. Z nabídky vyberte:

Analýza > Přežití > parametrické sdílené modely křehkosti

2. Vyberte zdrojovou proměnnou.

Čas

Přežití

Doba přežití je reprezentována jednou proměnnou, která označuje čas ukončení. Čas zahájení bude nastaven na hodnotu 0.

Začátek/konec

Číselné proměnné, které označují **Čas zahájení** a **Čas ukončení**.

Předmět

Nezbytné pro spuštění procedury. Určuje jednu proměnnou pro ID subjektu.

Interval

Uvádí jednotlivou a číselnou proměnnou pro číslo intervalu, které se používá k identifikaci různých opakujících se záznamů, které sdílejí stejné ID předmětu.

Stav

Jeden volitelný řetězec nebo číselná proměnná, která určuje jedno z následujících nastavení stavu:

Selhání/událost

Mapuje záznam na kategorii selhání/události. Výchozí hodnota pro proměnnou stavu řetězce je F.

Správné hodnocení

Mapuje záznam na pravou kategorii cenzury. Výchozí hodnota pro proměnnou stavu řetězce je R.

Zacházení s nemapovanými hodnotami

Řídí, na kterou kategorii se mají mapovat nemapované záznamy. Chcete-li odstranit záznamy, jejichž mapování se nezdařilo, vyberte volbu **Vyloučit je z analýzy**.

Klepněte na tlačítko **Definovat událost**, chcete-li definovat událost pro proměnnou stavu.

Kovariát (y)

Jedna nebo více volitelných číselných proměnných, které mají být považovány za kovariáty. Všimněte si, že proměnnou nelze zadat současně pomocí **Covariate (s)** a **Fixed Factor (s)**.

Pevný (é) faktor (y)

Jedna nebo více volitelných proměnných, které se mají považovat za faktory. Proměnnou nelze zadat současně pomocí **fixních faktorů** a **Covariate (s)**.

Parametrické sdílené modely křehkosti: kritéria

Podmínky

Volitelný panel pro určení obecných kritérií.

Interval spolehlivosti

Volitelná procentní část pro určení úrovně intervalů spolehlivosti regresních parametrů. Musí to být jedna číselná hodnota mezi 0 a 100. Výchozí hodnota je 95.

úroveň významnosti

Volba pro určení úrovně významnosti testu poměru pravděpodobnosti pro komponentu křehkosti. Musí to být jedna číselná hodnota mezi 0 a 1. Výchozí nastavení je 0.05.

Chybějící hodnoty

Volba pro řízení způsobu zacházení s chybějícími hodnotami uživatele:

Vyloučit chybějící i chybějící hodnoty uživatele i systému

Zachází s chybějícími hodnotami uživatele jako s platnými hodnotami. Jedná se o výchozí nastavení.

S chybějícími hodnotami uživatele se zachází jako s platnými

Ignoruje určení chybějících hodnot uživatele a považuje je za platné hodnoty.

Intervalová úprava

Volba pro řízení způsobu práce se záznamy, jejichž interval je v konfliktu s časem zahájení a ukončení. Projeví se, pokud existují dvě časové proměnné s proměnnou Interval určenou v hlavním dialogu.

Vyřadit konfliktní záznamy

Vyřadí celé sériové záznamy předmětu, pokud je hodnota intervalu v konfliktu s časem zahájení a ukončení. Toto je výchozí nastavení.

Zjistit hodnoty intervalu na základě času zahájení a času ukončení

Zjišťuje hodnotu intervalu od času zahájení a času ukončení.

Parametrické sdílené modely křehkosti: Model

Model

Volitelný panel pro určení voleb a nastavení modelu.

Rozdělení doby přežití

Možnost určit rozdělení doby přežití.

WEIBULL

Určuje Weibullovo rozdělení. Toto je výchozí nastavení.

Exponenciální

Určuje exponenciální distribuci.

Protokol-normální

Určuje normální distribuci protokolu.

Protokol-Logistika

Určuje logistickou distribuci protokolu.

Nastavení kovariace

Zadejte proměnné kovariate.

Nastavení faktoru

Zadejte proměnné faktoru.

Počáteční hodnota zachycení

Volba pro zadání počáteční hodnoty termínu zachycení. Je-li uveden, musí to být jedna číselná hodnota a nesmí být 0.

Počáteční hodnota parametru měřítka

Volba pro řízení nastavení parametru měřítka.

Standardní chyba odpovídající regrese OLS

Jako počáteční hodnotu použije standardní chybu odpovídající běžné regrese nejmenších čtverců.

Obrátí standardní chybu odpovídající regrese OLS.

Používá reciproční standardní chybu.

Hodnota zadaná uživatelem

Je-li zadána jedna číselná hodnota, použije se jako počáteční hodnota. Je-li uveden, musí být větší než 0.

Křehčí komponenta

Nepovinný parametr pro uvedení **Distribuce** komponenty křehkosti.

Gama

Určuje distribuci gama. Toto je výchozí nastavení.

Inverzní-gaussština

Určuje inverzní Gaussovo rozdělení.

Počáteční hodnota rozptylu

Určuje počáteční hodnotu rozptylu komponenty křehkosti. Musí to být jedna číselná hodnota větší než 0. Výchozí hodnota je 1.0 pro distribuci gama a 0.1 pro inverzní Gaussovo rozdělení.

Parametrické sdílené modely křehkosti: odhad

Odhad

Volitelný panel, který určuje nastavení pro řízení odhadu sdílených modelů křehkosti a volitelného procesu výběru funkcí.

Metoda střídavého směru nebo multiplikátory (ADMM)

Rychle

Používá metodu rychlého střídavého směru multiplikátorů (ADMM). Jedná se o výchozí nastavení.

Tradiční

Použije tradiční algoritmus ADMM.

Použití regularizaci L-1

Provádí proces pro řízení výběru funkcí. Pole **Parametr trestu** uvádí parametr pokuty, který řídí proces regularizace. Musí být jedna hodnota větší než 0. Výchozí nastavení je 0.001.

Kritéria konvergence modelu

Konvergence parametrů

Určuje kritéria konvergence pro parametr. Musí se jednat o jedinou číselnou hodnotu patřící k $[0, 1)$. Výchozí nastavení je 0.000001. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelná **Hodnota** určuje číselnou prahovou hodnotu pro typ konvergence.

Konvergence objektivních funkcí

Určuje kritéria konvergence pro funkci cíle. Musí to být jedna číselná hodnota, která patří do $[0, 1)$. Výchozí nastavení je 0, které neuplatňuje kritéria konvergence. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelná **Hodnota** určuje číselnou prahovou hodnotu pro typ konvergence.

Hessovská konvergence

Určuje kritéria konvergence pro hessovskou matici. Musí to být jedna číselná hodnota, která patří do $[0, 1)$. Výchozí nastavení je 0, které neuplatňuje kritéria konvergence. Pro **Typ** můžete vybrat buď volbu **ABSOLUTE**, chcete-li použít absolutní konvergenci na vnitřní optimalizaci, nebo volbu **RELATIVE**, chcete-li použít relativní konvergenci na vnitřní optimalizaci. Volitelná **Hodnota** určuje číselnou prahovou hodnotu pro typ konvergence.

Kritéria zbytkové konvergence

Volba pro řízení procesu optimalizace.

Primární i duální rezidua

Uplatňuje se jak základní, tak i dvojí zbytkové konvergenční kritérium. Toto nastavení je standardně.

Pouze primární rezidua

Použije se primární zbytkové konvergenční kritérium.

Pouze duální reziduální

Použije dvojí kritérium zbytkové konvergence.

Metoda

Nepovinný parametr pro určení metody odhadu.

Automaticky

Automaticky vybere metodu na základě ukázkové datové sady. Tato metoda je standardně vybrána. Pole **Prahová hodnota počtu prediktorů** určuje prahovou hodnotu počtu prediktorů a musí být celé číslo větší než 1. Výchozí hodnota je 1000.

Newton-Raphsonová

Používá Newton-Raphsonovu metodu.

L-BFGS

Použije algoritmus BFGS s omezenou pamětí. Pole **Aktualizovat** uvádí počet předchozích aktualizací, které jsou udržovány algoritmem BFGS s omezenou pamětí a musí být jediné celé číslo větší nebo rovno 1. Výchozí hodnota je 5.

Iterace

Maximální počet iterací

Určuje maximální počet iterací. Musí to být jediné celé číslo, které patří do [1, 300]. Výchozí nastavení je 20.

Maximální půlení schodů

Určuje maximální počet kroků na polovinu. Musí to být jediné celé číslo, které patří do [1, 200]. Výchozí nastavení je 5.

Maximální počet hledání řádků

Určuje maximální počet hledání řádků. Musí to být jediné celé číslo, které patří do [1, 300]. Výchozí nastavení je 20.

Absolutní konvergence pro iterační proces

Určuje absolutní konvergenci pro proces vnější iterace. Musí to být jedna číselná hodnota, která patří do (0, 1). Výchozí nastavení je 0.0001.

Relativní konvergence pro proces iterace

Určuje relativní konvergenci pro proces vnější iterace. Musí to být jedna číselná hodnota, která patří do (0, 1). Výchozí nastavení je 0.01.

Parametrické sdílené modely křehkosti: tisk

Tisk

Volitelný panel, který řídí výstupy tabulky.

Podrobnosti kódování faktorů

Je-li vybráno, zobrazí a vytiskne podrobnosti kódování faktorů. Proces je ignorován, pokud neexistují žádné platné faktory.

Počáteční hodnoty přiřazené k parametrům regrese

Je-li vybráno, zobrazí počáteční hodnoty, které se použijí v procesu odhadu.

Historie iterací modelu

Je-li vybráno, zobrazí se historie iterací analýzy přežití. Do pole **Počet kroků** zadejte počet kroků mezi 1 a 99999999. Výchozím nastavením je hodnota 1.

Parametrické sdílené modely Frailty: Predikce

Předvídat

Volitelný panel pro skóre a uložení předpovězených statistik do aktivní datové sady.

Časové hodnoty pro skórování

Časové hodnoty definované závislými proměnnými

Skóre **Predikce** na základě časové proměnné určené pro parametrický model přežití.

Pravidelné intervaly

Skóre **Predictions** na základě budoucích časových hodnot. Pole **Časový interval** uvádí časový interval a musí být jedna číselná hodnota větší než 0. Pole **Počet časových období** uvádí počet časových období a musí se jednat o jedno číselné celé číslo mezi 2 a 100.

Doba trvání

Skóre **Předpovědi** na základě doby trvání definuje budoucí časové hodnoty. Musí se jednat o jedinou číselnou proměnnou.

Předpovědi

Přežití

Skóre a uloží predikovanou statistiku přežití do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je `PredSurvival`.

Riziko

Skóre a uloží předpokládaná rizika do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je `PredHazard`.

Kumulativní riziko

Skóre a uloží předpokládaná kumulativní rizika do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je `PredCumHazard`.

Podmíněné přežití

Skóre a uloží predikovanou statistiku podmíněného přežití do aktivní datové sady. Výchozí název vlastní proměnné (nebo kořenový název) je `PredConditionalSurvival`. Není-li zadán parametr `PASTTIME`, je proces ignorován. Je požadována hodnota **Doba po minulém přežití** a uvádí hodnoty uplynulého času pro přidělení skóre. Musí se jednat o jedinou číselnou proměnnou.

Bezpodmínečné přežití

Provede skóre a uloží predikovanou statistiku bezpodmínečného přežití do aktivní datové sady. Klíčové slovo je standardně potlačeno. Je-li uveden, může být následován volitelným uživatelem zadaným názvem proměnné (nebo kořenovým názvem) uvedeným v závorkách. Výchozí název je `PredUnCondSurvival`.

Bezpodmínečné riziko

Skóre a uložení předpovězených nepodmíněných statistik rizika do aktivní datové sady. Klíčové slovo je standardně potlačeno. Je-li uveden, může být následován volitelným uživatelem zadaným názvem proměnné (nebo kořenovým názvem) uvedeným v závorkách. Výchozí název je `PredUncondHazard`.

Bezpodmínečné nebezpečí cum

Skóre a uloží předpokládanou nepodmíněnou statistiku kumulativního rizika do aktivní datové sady. Klíčové slovo je standardně potlačeno. Je-li uveden, může být následován volitelným názvem uživatelem zadané proměnné (nebo kořenovým názvem, který je uveden v závorkách. Výchozí název je `PredUncondCumHazard`.

Parametrické sdílené modely křehkosti: Plot

Graf

Funkční partie

Volba pro řízení vykreslení funkcí.

Typ

Přežití

Vytvoří děj pro funkce bezpodmínečného přežití.

Riziko

Vytvoří zákresu pro nepodmíněné funkce rizika.

Hustota

Vytvoří graf pro funkce hustoty.

Počet bodů k zobrazení

Určuje počet funkčních bodů mezi 1 a 200. Výchozí nastavení je 100.

Kovariantní hodnoty pro zápočet

Volba pro určení uživatelem zadaných hodnot a jejich přiřazení k prediktorům. Ve výchozím nastavení budou určené grafy vytvořeny na střední hodnotě každého kovariátu v platnosti. Je-li uvedeno, určené grafy budou vytvořeny na základě nastavení vzoru. V přítomnosti libovolných duplicitních proměnných bude rozpoznána ta, která byla uvedena jako první, a zbytek bude ignorován. Platná proměnná musí být obsažena v efektu modelu. Pro kovariát musí být uživatelem

zadaná hodnota číselná. Vynechání proměnné, která je v platnosti, označuje, že by se pro kovariát standardně použila střední hodnota. Pokud je proměnné přiřazena neplatná hodnota, požadovaný vzor nebude zakreslen.

Hodnoty faktoru pro Plot

Volba pro určení uživatelem zadaných hodnot a jejich přiřazení k prediktorům. Standardně se určené grafy vytvoří na frekvenci kategorií každého činného faktoru. Je-li uvedeno, určené grafy budou vytvořeny na základě nastavení vzoru. V přítomnosti libovolných duplicitních proměnných bude rozpoznána ta, která byla uvedena jako první, a zbytek bude ignorován. Platná proměnná musí být obsažena v efektu modelu. Vynechání proměnné v platnosti znamená, že frekvence kategorie bude použita jako výchozí pro daný faktor. Pokud je proměnné přiřazena neplatná hodnota, požadovaný vzor nebude zakreslen.

Oddělené řádky pro

Volba pro určení kategorické proměnné, podle které budou vykresleny čárový grafy.

Maximální počet čar v grafu

Určuje maximální počet čar v grafu, je-li zadána volba **Samostatné čáry pro** . Výchozí nastavení je 10.

Parametrické sdílené modely křehkosti: Export

Exportovat

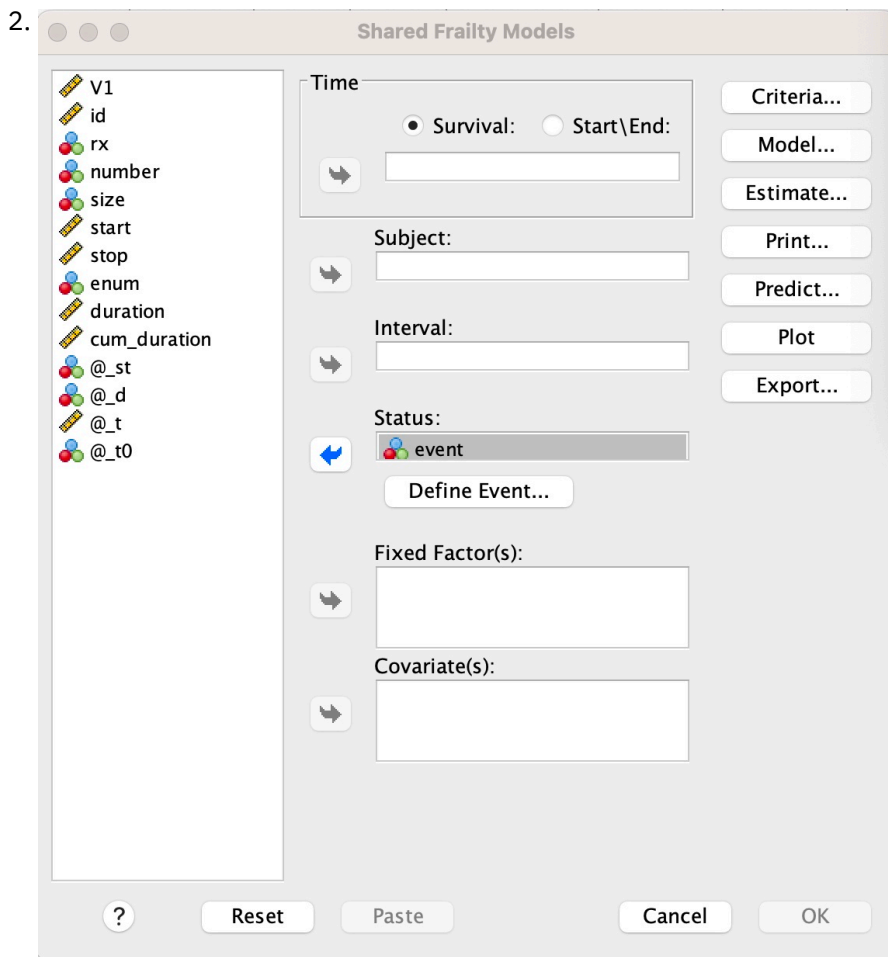
Vyberte volbu **Exportovat informace o modelu do souboru XML** , chcete-li zapsat informace o modelu a parametru do souboru PMML pro přidělení skóre. Musíte uvést adresář a název souboru PMML, který se má uložit.

Parametrické sdílené modely křehkosti: definování událostí

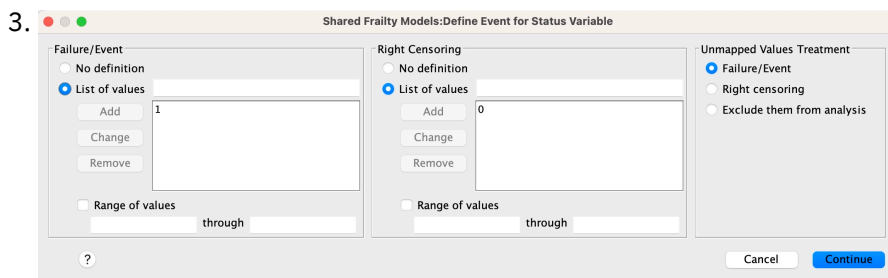
Volba pro definování stavu. Pokud je proměnná stavu vynechána, stane se selhání nebo událost výchozím stavem pro všechny případy.

1. Z nabídky vybrat,

Analýza > Přežití > parametrické sdílené modely křehkosti ...



Obrázek 5. Sdílené modely křehkosti-dialogové okno-Stav



Obrázek 6. Sdílené modely křehkosti-dialogové okno-Stav-Definovat událost

Parametrické sdílené modely křehkosti-příklady

Příklad 1

PŘEŽÍVAJÍCÍ RECURRENT y S x1 BY x2

/MODEL SUBJECT = ID FRAILTY=GAMMA DISTRIBUTION=WEIBULL.

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2.

Doba přežití je reprezentována jedinou proměnnou y.

Předměty jsou identifikovány pomocí ID proměnné.

Předpokládá se, že doba přežití bude následovat po Weibulovi distribuci.

Předpokládá se, že rozptyl křehkosti bude následovat po rozdělení gama.

Všechny platné záznamy jsou použity v analýze přežití.

Příklad 2

PŘEŽÍVAJÍCÍ RECURRENT y S x1 BY x2

```
/MODEL SUBJECT = id FRAILITY=INV_GAUSSIAN DISTRIBUTION=LOG_NORMAL INTERVAL=z.
```

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2.

Doba přežití je reprezentována dvěma proměnnými y1 a y2 , které označují počáteční a koncový čas.

Předměty jsou identifikovány pomocí ID proměnné.

Předpokládá se, že doba přežití bude následovat po log-normální distribuci.

Předpokládá se, že rozptyl křehkosti je následován inverzní Gaussovou distribucí.

Časové intervaly jsou definovány proměnnou z. Pro každý předmět procedura použije pouze nekonfliktní záznamy a vyloučí z analýzy všechny záznamy po prvním stavu poruchy.

Příklad 3

ZBÝVAJÍCÍ RECURRENT y1 y2 S x1 BY x2(1)

```
/MODEL SUBJECT = ID FRAILITY=INV_GAUSSIAN DISTRIBUTION=LOG_LOGISTIC
```

```
/STATUS VARIABLE=událost FAILURE=1 RIGHT=0.
```

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2. Přežití

čas představuje dvě proměnné y1 a y2 označující čas zahájení a čas ukončení. Pro faktor x2 je kategorie "1" označena jako úroveň baseline, která má být modelována.

Předměty jsou identifikovány pomocí ID proměnné.

Předpokládá se, že doba přežití bude následovat logistické rozdělení.

Předpokládá se, že rozptyl křehkosti je následován inverzní Gaussovou distribucí.

Událost proměnné je určena tak, aby definovala stav s hodnotou 1 a 0 označující selhání a právo na cenzuru.

Příklad 4

PŘEŽÍVAJÍCÍ RECURRENT y S x1 BY x2

```
/MODEL SUBJECT = ID
```

```
/STATUS VARIABLE=událost FAILURE=1 RIGHT=0
```

```
/PŘEDPOVÍDAT UNCONDSURVIVAL UNCONDHAZARD UNCONDHAZARD UNCONDCUMHAZARD
```

```
/FUNCTIONPLOT SURVIVAL HAZARD DENSITY PLOTBY (x2).
```

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2. Doba přežití je reprezentována jedinou proměnnou y.

Předměty jsou identifikovány pomocí ID proměnné.

Nepodmíněné nebo populační přežití, riziko a kumulativní riziko jsou hodnoceny a uloženy do aktivní datové sady.

Nepodmíněné nebo populační křivky přežití a nebezpečné křivky jsou vyneseny odděleně podle kategorií v poli x2.

Příklad 5

PŘEŽÍVAJÍCÍ RECURRENT y S x1 BY x2

/MODEL SUBJECT = ID FRAILITY=GAMMA DISTRIBUTION=WEIBULL

/STATUS VARIABLE=událost FAILURE=1 RIGHT=0

/ESTIMATION HCONVERGE=1e-12(RELATIVE) PCONVERGE=0 FCONVERGE=0 SELECTFEATURES=TRUE PENALTY=0.01.

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2. Doba přežití je reprezentována jedinou proměnnou y.

Předměty jsou identifikovány pomocí ID proměnné.

Předpokládá se, že doba přežití bude následovat po Weibullovi distribuci.

Předpokládá se, že rozptyl křehkosti bude následovat po rozdělení gama.

Kritéria konvergence jsou založena na hesenské matici. Používá 1e-12 jako relativní konvergenci.

Model obsahuje sankční termín pro řízení procesu regularizace. Parametr penalty je nastaven na hodnotu 0.01.

Příklad 6

PŘEŽÍVAJÍCÍ RECURRENT y S x1 BY x2

/MODEL SUBJECT = ID

/STATUS VARIABLE=infikovat FAILURE=1 RIGHT=0

/ESTIMATION MAXLINESEARCH=100 MAXITER=50 MAXSTEPHALVING=20.

Parametrický model přežití sdílené křehkosti je vybaven y na kovariátu x1 a faktoru x2. Doba přežití je reprezentována jedinou proměnnou y.

Předměty jsou identifikovány pomocí ID proměnné.

Procedura uvádí maximální počet hledání řádku na 100, maximální počet iterací na 50 a maximální počet kroků na polovinu na 20.

Parametrické sdílené modely křehkosti-případová studie pro rekurní data

Parametrické sdílené modely křehkosti-případová studie pro rekurní data

Název příkladu použití-vedlejší efekt léčby.

Herci-vyšetřovatel veřejného zdraví a praktický lékař.

Předběžné podmínky-Vyčištěná datová sada, která je k dispozici na základě doby přežití, stavu vedlejších účinků a prediktorů, které mají být upraveny.

Popis-Patrick, vyšetřovatel veřejného zdraví, vyšetřuje vzorek dat, který zahrnuje 20 účastníků. Tito účastníci jsou přijati do studie o mírném vedlejších účinků, který je potenciálně způsoben novou léčbou. Návrhář léčby tvrdí, že by nebyly žádné rozdíly mezi muži a ženami, pokud jde o vedlejší účinek. Patrick by rád zhodnotil takovou hypotézu. Proměnné, které jsou zahrnuty do ukázky dat, jsou uvedeny v následujícím seznamu:

- patID: Číslo ID pro identifikaci jedinečného účastníka.

- endTime: Doba přežití (ve dnech) vedlejších účinků po léčbě, která se měří od začátku léčby k hlášeným nežádoucím účinkům nebo cenzurování během 60 dnů.

- sideEffect: Stav vedlejších účinků, stav = 0, pokud je cenzurován, a stav = 1, pokud je hlášen mírný oboustranný efekt.

-věk: věk účastníka ve výzkumném období.

-samice: samice = 0, pokud samec a samice = 1, pokud samice.

Může se použít více procedur, což vede k vícenásobným záznamům opakovaných časů, které jsou měřeny pro určitého účastníka. Počáteční čas je vždy 0 pro každý záznam, který je vynechán ve vzorku dat. Patrick se zajímá o vizualizaci funkcí přežití a nebezpečí pro srovnání mezi mužem a ženou tím, že kontroluje jejich věk a křehkost. Je si vědom toho, že ty procedury, které jsou podávány stejnému účastníkovi, jsou více korelované. Předpokládá-li se, že doba přežití následuje po Weibullovi rozdělení, rozhodne se Patrick sestavit parametrický model přežití sdílené křehkosti v SPSS Statistics, který bude zohledňovat závislost na léčbě pro stejného účastníka.

Syntaxe-

```
DATA LIST FREE
/patID(F5.0) endTime(F5.0) sideEffect(F2.0) age(F5.2) female(F2.0) .
BEGIN DATA .
1 45 0 38.00 0
2 26 1 20.00 1
3 58 0 53.00 0
4 31 1 37.00 1
4 24 0 37.00 1
4 50 0 37.00 1
5 20 1 51.00 0
5 38 1 51.00 0
6 30 0 35.00 1
7 22 1 58.00 1
8 53 1 29.00 1
8 49 1 29.00 1
9 25 0 45.00 0
9 25 0 45.00 0
10 27 0 33.00 1
11 34 1 21.00 1
11 40 0 21.00 1
11 49 0 21.00 1
12 42 1 26.00 0
13 25 0 40.00 0
14 21 1 52.00 0
14 32 1 52.00 0
15 56 0 28.00 1
15 34 0 28.00 1
16 30 0 41.00 0
16 29 0 41.00 0
17 25 1 27.00 0
18 26 1 54.00 1
18 36 1 54.00 1
19 27 0 39.00 0
20 58 1 22.00 1
20 54 0 22.00 1
20 43 1 22.00 1
END DATA.
SURVREG RECURRENT endTime WITH age BY female
/MODEL SUBJECT=patID FRAILTY=GAMMA DISTRIBUTION=WEIBULL
/ESTIMATION HCONVERGE=1e-12 PCONVERGE=0 FCONVERGE=0
/STATUS VARIABLE=sideEffect FAILURE=1 RIGHT=0
/FUNCTIONPLOT SURVIVAL HAZARD PLOTBY(female) .
```

Synopse:

Syntaxe určená Patrickem označuje endTime jako jednotlivou závislou časovou proměnnou. Procedura automaticky předpokládá, že počáteční čas je 0 pro každý záznam. Proměnné věk a žena jsou modelovány jako kovariát a faktor, příslušně. Předpokládá se, že doba přežití recidivy bude následovat po Weibullově distribuci. Předpokládá se, že nepozorovaný termín křehkosti sleduje distribuci gama a jeho složka rozptylu je modelována. Pokud jde o výstupy, tabulka Souhrn modelu poskytuje informace o proceduře a modelu. Tabulka Souhrn zpracování případů poskytuje souhrnný seznam stavu selhání/cenzury a také těch případů, které jsou vyloučeny z analýzy.

V Patrickově vzorku dat jsou všechny záznamy platné a zahrnuty do analýzy. Porovnáním pravděpodobnosti protokolu s pravděpodobností odpovídajícího modelu bez komponenty frailty se modelu shared-frailty nepodaří dosáhnout významné úrovně (p -value = 0.168). Patrick se zajímá, zda je nutné zahrnout do modelu termín sdíleného křehkosti. Odhadovaný faktor zrychlení mužského účastníka je 1.017, který se získá výpočtem exponentu odhadovaného regresního koeficientu 0.017 [žena = 0.0]. Jeho přidružený 95% interval spolehlivosti je (.688, 1.504). Tyto výsledky naznačují, že mužský jedinec má téměř stejný akcelerační faktor jako ženský jedinec se stejným věkem a křehkou. Na úrovni

populace Patrick vykresluje bezpodmínečné křivky přežití a nebezpečí odděleně pro muže a ženy, kteří jsou hodnoceni ve věku vzorku (37.45 let).

Patrick potvrzuje, že u každé pevné hodnoty doby přežití se předpokládá, že samec a samice mají v průměru stejnou pravděpodobnost přežití. Je zajímavé, že i přes neimodální tvar, který je uveden v bezpodmínečném grafu nebezpečí, Patrick zjistí, že během 60 dnů se riziko populace skutečně zvyšuje. Toto chování by mohlo znamenat existenci slabého efektu. Chcete-li dále zkoumat vedlejší účinek způsobený léčbou, Patrick může pokračovat v modelu bez křehkosti složky a porovnávat chování mužů a žen. Kromě toho může zvážit sledování s účastníky po dobu delší než 60 dnů, aby shromáždil více údajů.

Kaplan-Meier analýza přežití

Existuje mnoho situací, ve kterých byste chtěli prozkoumat rozdělení časů mezi dvě události, jako je délka zaměstnání (doba mezi najatím a odchodem ze společnosti). Tento druh dat však obvykle zahrnuje některé cenzurované případy. Cenzurované případy jsou případy, pro které není druhá událost zaznamenána (například lidé stále pracují pro společnost na konci studie). Postup Kaplan-Meier je metoda odhadu časových modelů událostí v přítomnosti cenzurovaných případů. Model Kaplan-Meier je založen na odhadu podmíněných pravděpodobností v každém časovém bodě, kdy dojde k události, a na stanovení produktového limitu těchto pravděpodobností pro odhad míry přežití v každém časovém bodě.

Příklad. Má nová léčba AIDS nějaký terapeutický přínos při prodlužování života? Dalo by se provést studii s využitím dvou skupin pacientů s AIDS, jeden dostává tradiční léčbu a druhý dostávají experimentální léčbu. Vytvoření Kaplan-Meierova modelu z dat by vám umožnilo porovnat celkovou míru přežití mezi oběma skupinami, abyste zjistili, zda je experimentální léčba lepší než tradiční léčba. Můžete také vykreslit funkce přežití nebo nebezpečí a vizuálně je porovnat, abyste získali podrobnější informace.

Statistika. Tabulka přežití, včetně času, stavu, kumulativního přežití a standardní chyby, kumulativních příhod a zbývajícího počtu; a střední a střední doba přežití se standardní chybou a 95% intervalem spolehlivosti. Pozemky: přežití, hazard, log přežití, a jeden mínus přežití.

Pokyny pro Kaplan-Meier Data

Data. Časová proměnná by měla být spojitá, stavová proměnná může být kategorická nebo spojitá a faktorové a strátové proměnné by měly být kategorické.

Předpoklady. Pravděpodobnost události, která je předmětem zájmu, by měla záviset pouze na čase po počáteční události-předpokládá se, že jsou stabilní s ohledem na absolutní čas. To znamená, že případy, které vstupují do studie v různých časech (například pacienti, kteří začínají léčbu v různých časech), by se měli chovat podobně. Neměly by také existovat žádné systematické rozdíly mezi cenzurovanými a necenzurovanými případy. Pokud například mnoho cenzurovaných případů jsou pacienti s vážnějšími podmínkami, vaše výsledky mohou být zkreslené.

Související procedury. Kaplan-Meierova procedura používá metodu výpočtu životních tabulek, která odhaduje funkci přežití nebo nebezpečí v době každé události. Postup Life Tables používá pojistně-matematický přístup k analýze přežití, který se spoléhá na rozdělení období pozorování do menších časových intervalů a může být užitečný pro práci s velkými vzorky. Pokud máte proměnné, u kterých máte podezření, že souvisejí s dobou přežití nebo s proměnnými, které chcete kontrolovat (kovariáty), použijte proceduru Cox Regression. Pokud mohou mít vaše kovariáty různé hodnoty v různých časových bodech pro stejný případ, použijte Cox Regression with Time-Dependent Covariates.

Získání Kaplan-Meierovy analýzy přežití

1. Z nabídek vyberte:

Analýza > Přežití > Kaplan-Meier ...

2. Vyberte proměnnou času.

3. Vyberte proměnnou stavu, abyste identifikovali případy, pro které došlo k události terminálu. Tato proměnná může být číselná nebo *krátký řetězec*. Poté klepněte na volbu **Definovat událost**.

Volitelně můžete vybrat faktorovou proměnnou, abyste prozkoumali rozdíly ve skupinách. Můžete také vybrat proměnnou vrstev, která vytvoří oddělené analýzy pro každou úroveň (vrstvu) proměnné.

Kaplan-Meier Definovat událost pro proměnnou stavu

Zadejte hodnotu nebo hodnoty označující, že došlo k události terminálu. Můžete zadat jednu hodnotu, rozsah hodnot nebo seznam hodnot. Volba Rozsah hodnot je k dispozici pouze v případě, že je proměnná stavu číselná.

Kaplan-Meier Porovnat úrovně faktorů

Můžete požádat o statistiku pro testování rovnosti rozdělení přežití pro různé úrovně faktorů. Dostupné statistiky jsou log rank, Breslow a Tarone-Ware. Vyberte jednu z alternativ, abyste určili porovnání, která se mají provést: sdružená nad vrstvami, pro každou vrstvu, po dvojicích nad vrstvami nebo po dvojicích pro každou vrstvu.

Lineární trend pro úrovně faktorů. Umožňuje testovat lineární trend napříč úrovněmi faktorů. Tato volba je k dispozici pouze pro celkové (spíše než párování) porovnání úrovní faktorů.

- *Ohodnocení důležitosti protokolu.* Test pro porovnání rovnosti rozdělení přežití. Všechny časové body jsou v tomto testu rovnoměrně váženy.
- *Breslow.* Test pro porovnání rovnosti rozdělení přežití. Časové body jsou váženy podle počtu rizikových případů v každém časovém bodě.
- *Tarone-Ware.* Test pro porovnání rovnosti rozdělení přežití. Časové body jsou váženy druhou odmocninou počtu rizikových případů v každém časovém bodě.
- *Sdružené přes vrstvy.* Porovná všechny úrovně faktorů v jednom testu, aby se testovala rovnost křivek přežití.
- *Párování nad vrstvami.* Porovná každou odlišnou dvojici úrovní faktorů. Testy trendu po dvojicích nejsou k dispozici.
- *Pro každou vrstvu.* Provede samostatný test rovnosti všech úrovní faktorů pro každou vrstvu. Pokud nemáte proměnnou stratifikace, testy se neprovedou.
- *Po dvojicích pro každou vrstvu.* Porovná každou odlišnou dvojici úrovní faktorů pro každou vrstvu. Testy trendu po dvojicích nejsou k dispozici. Pokud nemáte proměnnou stratifikace, testy se neprovedou.

Kaplan-Meier Uložit nové proměnné

Informace z tabulky Kaplan-Meier můžete uložit jako nové proměnné, které pak mohou být použity v následných analýzách k testování hypotéz nebo ke kontrole předpokladů. Můžete uložit přežití, standardní chybu přežití, riziko a kumulativní události jako nové proměnné.

- *Přežití.* Kumulativní odhad pravděpodobnosti přežití. Výchozí název proměnné je předpona sur _ s připojeným pořadovým číslem. Pokud již například sur_1 existuje, Kaplan-Meier přiřadí proměnné název sur_2.
- *Standardní chyba přežití.* Standardní chyba souhrnného odhadu přežití. Výchozí název proměnné je předpona se _ s připojeným pořadovým číslem. Pokud například se_1 již existuje, Kaplan-Meier přiřadí proměnné název se_2.
- *Nebezpečí.* Kumulativní odhad funkce rizika. Výchozí název proměnné je předpona haz _ s připojeným pořadovým číslem. Pokud například haz_1 již existuje, Kaplan-Meier přiřadí název proměnné haz_2.
- *Kumulativní události.* Kumulativní frekvence událostí, kdy jsou případy seřazeny podle doby přežití a stavových kódů. Výchozí název proměnné je předpona cum _ s připojeným pořadovým číslem. Pokud například cum_1 již existuje, Kaplan-Meier přiřadí název proměnné cum_2.

Možnosti společnosti Kaplan-Meier

Z Kaplan-Meierovy analýzy si můžete vyžádat různé typy výstupů.

Statistika. Můžete vybrat statistiky zobrazené pro výpočet funkcí přežití, včetně tabulky (tabulek) přežití, střední a střední doby přežití a kvartilů. Pokud jste zahrnuli faktorové proměnné, vygenerují se pro každou skupinu samostatné statistiky.

Záloty. Grafy vám umožňují zkoumat funkce přežití, one-minus-survival, hazard a log-survival vizuálně. Pokud jste zahrnuli faktorové proměnné, funkce se zakreslí pro každou skupinu.

- *Přežití.* Zobrazuje funkci kumulativního přežití na lineárním měřítku.
- *Jeden minus přežití.* Plots one-minus funkce přežití na lineární stupnici.
- *Nebezpečí.* Zobrazuje funkci kumulativního rizika na lineárním měřítku.
- *Přežití protokolu.* Zobrazuje funkci kumulativního přežití na logaritmické stupnici.

Další funkce příkazu KM

Jazyk syntaxe příkazu vám také umožňuje:

- Získejte frekvenční tabulky, které považují případy ztracené při následném zpracování za oddělenou kategorii od cenzurovaných případů.
- Zadejte nerovné mezery pro test lineárního trendu.
- Získejte percentily jiné než kvartil pro proměnnou doby přežití.

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference*.

Regresní analýza Cox

Cox Regression sestavuje prediktivní model pro data času do události. Model vytváří funkci přežití, která předpovídá pravděpodobnost, že k události, která je předmětem zájmu, došlo v daném čase t pro dané hodnoty proměnných prediktoru. Tvar funkce přežití a regresní koeficienty pro prediktory jsou odhadnuty z pozorovaných subjektů; model pak lze aplikovat na nové případy, které mají měření pro proměnné prediktoru. Všimněte si, že informace z cenzurovaných předmětů, tj. těch, které nezaznamenávají zajímavou událost v době pozorování, přispívají k odhadu modelu.

Příklad. Mají muži a ženy různá rizika vzniku rakoviny plic na základě kouření cigaret? Vytvořením modelu Cox Regression, s užíváním cigaret (cigarety uzené za den) a pohlaví zadané jako kovariáty, můžete testovat hypotézy týkající se účinků pohlaví a užívání cigaret na čas-to-počátek pro rakovinu plic.

Statistika. Pro každý model: $-2LL$, statistika poměru pravděpodobnosti a celkový chí-kvadrát. Pro proměnné v modelu: odhady parametrů, standardní chyby a Wald statistiky. Pro proměnné, které nejsou v modelu: statistika skóre a zbytkový chí-kvadrát.

Aspekty regresních dat Cox

Data. Vaše časová proměnná by měla být kvantitativní, ale vaše stavová proměnná může být kategorická nebo souvislá. Nezávislé proměnné (kovariáty) mohou být spojité nebo kategorické; pokud jsou kategorické, měly by být fiktivní-nebo indikátorem kódované (v proceduře je možnost automaticky překódovat kategorické proměnné). Strata proměnné by měly být kategorické, kódované jako celá čísla nebo krátké řetězce.

Předpoklady. Pozorování by měla být nezávislá a poměr nebezpečí by měl být v průběhu času konstantní; to znamená, že přiměřenost nebezpečí v jednotlivých případech by se neměla v průběhu času měnit. Druhý předpoklad je znám jako **předpoklad proporcionálních rizik**.

Související procedury. Pokud předpoklad proporcionálních rizik není v platnosti (viz výše), budete možná muset použít postup Cox s časově závislým krytím. Nemáte-li žádné kovariáty nebo máte-li pouze jeden kategorický kovariát, můžete použít postup Life Tables nebo Kaplan-Meier k prozkoumání funkcí přežití nebo nebezpečí pro váš vzorek (vzorky). Pokud ve vzorku nemáte žádná cenzurovaná data (to znamená, že v každém případě došlo k události terminálu), můžete použít proceduru lineární regrese k modelování vztahu mezi prediktory a časem k události.

Získání analýzy regrese Cox

1. Z nabídek vyberte:

Analýza > Přežití > Cox Regrese ...

2. Vyberte proměnnou času. Případy, jejichž časové hodnoty jsou záporné, se neanalyzují.

3. Vyberte proměnnou stavu a poté klepněte na volbu **Definovat událost**.
 4. Vyberte jeden nebo více kovariátů. Chcete-li zahrnout podmínky interakce, vyberte všechny proměnné zahrnuté v interakci a poté klepněte na volbu **> a*b>**.
- Volitelně můžete vypočítat samostatné modely pro různé skupiny definováním proměnné vrstvy.

Cox regression-definice kategorických proměnných

Můžete určit podrobnosti o tom, jak bude procedura Cox Regression zpracovávat kategorické proměnné.

Kryty. Vypíše všechny kovariáty uvedené v hlavním dialogovém okně, buď samy, nebo jako součást interakce, v libovolné vrstvě. Pokud jsou některé z nich řetězcové proměnné nebo jsou kategorické, můžete je použít pouze jako kategorické kovariáty.

Kategorické kovariáty. Vypíše proměnné identifikované jako kategorické. Každá proměnná obsahuje v závorkách notaci označující kontrastní kódování, které se má použít. Řetězcové proměnné (označené symbolem < za jejich názvy) jsou již uvedeny v seznamu Kategorické Covariates. Vyberte libovolné další kategorické kovariáty ze seznamu Covariates a přesuňte je do seznamu Categorical Covariates.

Změnit kontrast. Umožňuje změnit metodu kontrastu. Dostupné kontrastní metody jsou:

- **Indikátor.** Kontrasty naznačují přítomnost nebo nepřítomnost členství v kategoriích. Referenční kategorie je v kontrastní matici znázorněna jako řádek s nulami.
- **Jednoduché.** Každá kategorie proměnné prediktoru s výjimkou referenční kategorie je porovnána s referenční kategorií.
- **Rozdíl.** Každá kategorie proměnné prediktoru s výjimkou první kategorie je porovnána s průměrným efektem předchozích kategorií. Také známý jako reverzní Helmert kontrasty.
- **Helmert.** Každá kategorie proměnné prediktoru s výjimkou poslední kategorie je porovnána s průměrným efektem následujících kategorií.
- **Opakovat.** Každá kategorie proměnné prediktoru s výjimkou první kategorie je porovnána s kategorií, která ji předchází.
- **Polynomiální.** Ortogonální polynomiální kontrasty. Předpokládá se, že kategorie jsou rovnoměrně rozmístěny. Polynomiální kontrasty jsou k dispozici pouze pro číselné proměnné.
- **Odchylka.** Každá kategorie proměnné prediktoru s výjimkou referenční kategorie je porovnána s celkovým efektem.

Vyberete-li volbu **Odchylka**, **Jednoduchý** nebo **Indikátor**, vyberte jako referenční kategorii buď **První**, nebo **Poslední**. Všimněte si, že metoda se ve skutečnosti nezmění, dokud neklepnete na tlačítko **Změnit**.

Řetězcové kovariáty musí být kategorické kovariáty. Chcete-li odebrat řetězcovou proměnnou ze seznamu Kategorické Covariates, musíte odebrat všechny výrazy obsahující tuto proměnnou ze seznamu Covariates v hlavním dialogovém okně.

Regresní grafy Cox

Grafy vám mohou pomoci vyhodnotit odhadovaný model a interpretovat výsledky. Můžete vykreslit funkce přežití, nebezpečí, log-minus-log a jeden-minus-přežití.

Typ výkresu

Přežití

Zobrazuje funkci kumulativního přežití na lineárním měřítku.

Riziko

Zobrazuje funkci kumulativního rizika na lineárním měřítku.

Protokol minus protokol

Na odhad se použije kumulativní odhad přežití po transformaci $\ln(-\ln)$.

Jedno minus přežití

Plots one-minus funkce přežití na lineární stupnici.

Kovariantní hodnoty vykreslovány v

Můžete zakreslit samostatný řádek pro každou hodnotu kategoriálního kovariátu přesunutím tohoto kovariátu do textového pole **Oddělit řádky pro**. Tato volba je k dispozici pouze pro kategoriální kovariáty, které jsou označeny pomocí **(Cat)** za jejich názvy v seznamu **Covariate Values Plotted at**.

Změnit hodnotu

Vzhledem k tomu, že tyto funkce závisí na hodnotách kovariátů, musíte použít konstantní hodnoty pro kovariáty, abyste vykresli funkce v závislosti na čase. Výchozí nastavení je použít střední hodnotu každého kovariátu jako konstantní hodnotu, ale můžete zadat vlastní hodnoty pro vykreslení pomocí řídicí skupiny **Změnit hodnotu**.

Cox regression Uložit nové proměnné

Můžete uložit různé výsledky analýzy jako nové proměnné. Tyto proměnné pak mohou být použity v následných analýzách k testování hypotéz nebo ke kontrole předpokladů.

Uložit proměnné modelu

Umožňuje vám uložit funkci přežití a její standardní chybu, odhady log-minus-log, funkci nebezpečí, dílčí rezidua, DfBeta(y) pro regresi a lineární prediktor $X^* \text{Beta}$ jako nové proměnné.

Funkce přežití

Hodnota kumulativní funkce přežití pro daný čas. Rovná se pravděpodobnosti přežití tohoto časového období.

Standardní chyba funkce přežití

Standardní chyba odhadu kumulativního přežití.

Log minus funkce přežití logu

Kumulativní odhad přežití po transformaci $\ln(-\ln)$ se použije na odhad.

Funkce nebezpečí

Uloží kumulativní odhad funkce rizika (nazývaný také Cox-Snell reziduum).

Dílčí rezidua

Můžete zakreslit dílčí rezidua proti době přežití, abyste otestovali předpoklad proporcionálních rizik. Pro každý kovariát v konečném modelu se uloží jedna proměnná. Peritální rezidua jsou k dispozici pouze pro modely obsahující alespoň jeden kovariát.

DfBeta(s)

Odhadovaná změna koeficientu při odebrání případu. Pro každý kovariát v konečném modelu se uloží jedna proměnná. DfBetas jsou k dispozici pouze pro modely obsahující alespoň jeden kovariát.

$X^* \text{Beta}$

skóre lineárního prediktoru. Součet součinu kovariátních hodnot se středem a jejich odpovídajících odhadů parametrů pro každý případ.

Poznámka: Pokud spouštíte Cox s kovariátem závislým na čase, uloží se pouze DfBeta.

Exportovat informace o modelu do souboru XML

Odhady parametrů jsou exportovány do určeného souboru ve formátu XML. Tento soubor modelu můžete použít k použití informací o modelu na jiné datové soubory pro účely přidělení skóre.

Volby regrese Cox

Můžete řídit různé aspekty analýzy a výstupu.

Statistika modelu. Můžete získat statistiku pro parametry modelu, včetně intervalů spolehlivosti pro výraz (B) a korelace odhadů. Tyto statistiky můžete požadovat buď v každém kroku, nebo pouze v posledním kroku.

Pravděpodobnost pro Stepwise. Pokud jste vybrali postupnou metodu, můžete určit pravděpodobnost pro položku nebo odebrání z modelu. Proměnná se zadává, pokud je úroveň významnosti jejího parametru F -to-enter menší než hodnota položky, a proměnná se odebere, pokud je úroveň významnosti větší než hodnota odebrání. Hodnota položky musí být menší než hodnota odebrání.

Maximální počet iterací. Umožňuje vám určit maximální počet iterací pro model, který řídí, jak dlouho bude procedura hledat řešení.

Zobrazit funkci úrovně baseline. Umožňuje vám zobrazit základní funkci rizika a kumulativní přežití ve střední hodnotě kovariátů. Tato obrazovka není k dispozici, pokud jste zadali časové kovariáty.

Událost definování regrese Cox pro proměnnou stavu

Zadejte hodnotu nebo hodnoty označující, že došlo k události terminálu. Můžete zadat jednu hodnotu, rozsah hodnot nebo seznam hodnot. Volba Rozsah hodnot je k dispozici pouze v případě, že je proměnná stavu číselná.

Další funkce příkazu COXREG

Jazyk syntaxe příkazu vám také umožňuje:

- Získejte frekvenční tabulky, které považují případy ztracené při následném zpracování za oddělenou kategorii od cenzurovaných případů.
- Vyberte jinou referenční kategorii, než první nebo poslední, pro metody odchylky, jednoduché metody a metody kontrastu indikátorů.
- Pro metodu polynomiálního kontrastu zadejte nestejně mezery mezi kategoriemi.
- Zadejte další kritéria iterace.
- Řízení léčby chybějících hodnot.
- Zadejte názvy pro uložené proměnné.
- Zapsat výstup do externího datového souboru IBM SPSS Statistika .
- Během zpracování zadržte data pro každou skupinu rozdělených souborů v externím souboru pracovního prostředí. To může pomoci šetřit paměťové prostředky při spouštění analýz s velkými datovými sadami. Není k dispozici u kovariátů závislých na čase.

Úplné informace o syntaxi naleznete v příručce *Command Syntax Reference* .

Výpočet časově závislých krytů

V určitých situacích byste chtěli vypočítat model Cox Regression, ale předpoklad proporcionálních rizik se nedrží. To znamená, že poměry rizika se mění v průběhu času; hodnoty jednoho (nebo více) vašich kovariátů se liší v různých časových bodech. V takových případech musíte použít rozšířený model Cox Regression, ve kterém můžete zadat **časové kovariáty**.

Chcete-li takový model analyzovat, musíte nejprve definovat kovariát závislý na čase. Pro usnadnění je k dispozici 'systémová proměnná', která představuje čas. Tato proměnná se nazývá $T_.$ Tuto proměnnou můžete použít k definování kovariátů závislých na čase dvěma obecnými způsoby:

- Chcete-li testovat předpoklad proporcionálních rizik s ohledem na konkrétní kovariát nebo odhad rozšířeného Coxova regresního modelu, který umožňuje neproporcionální rizika, musíte definovat kovariát závislý na čase jako funkci časové proměnné $T_.$ a příslušného kovariátu. Běžným příkladem byl jednoduchý součin časové proměnné a kovariátu, ale složitější funkce mohou být specifikovány také. Testování významnosti koeficientu kovariátu závislého na čase ukazuje, zda je předpoklad přiměřeného rizika přiměřený.
- Některé proměnné mohou mít různé hodnoty v různých časových obdobích, ale nesouvisí systematicky s časem. V takových případech musíte definovat **segmentovaný kovariát závislý na čase**, který lze provést pomocí **logických výrazů**. Logické výrazy používají hodnotu 1, pokud je hodnota true, a 0, pokud je hodnota false. Pomocí řady logických výrazů můžete vytvořit kovariát závislý na čase ze sady měření. Máte-li například krevní tlak, který se měří jednou týdně po dobu čtyř týdnů vaší studie (identifikované jako $BP1$ až $BP4$), můžete definovat kovariát závislý na čase jako $(T_ < 1) * BP1 + (T_ >= 1 \& T_ < 2) * BP2 + (T_ >= 2 \& T_ < 3) * BP3 + (T_ >= 3 \& T_ < 4) * BP4$. Všimněte si, že přesně jeden z termínů v závorkách se rovná 1 pro každý konkrétní případ a zbytek se bude rovnat 0. Chcete-li to shrnout, tato

funkce znamená, že pokud je čas kratší než jeden týden, použijte *BP1*; pokud je delší než jeden týden, ale kratší než dva týdny, použijte *BP2atd*.

V dialogovém okně Compute Time-Dependent Covariate můžete použít ovládací prvky sestavení funkce k sestavení výrazu pro kovariát závislý na čase, nebo jej můžete zadat přímo do textové oblasti Název. Všimněte si, že řetězcové konstanty musí být uzavřeny v uvozovkách nebo apostrofech a numerické konstanty musí být zadány v americkém formátu s tečkou jako desetinným oddělovačem. Všechny výsledné kovariáty závislé na čase musí být zahrnuty jako kovariáty ve vašem modelu Cox Regression.

Výpočet časově závislého Covariate

1. Z nabídky vyberte

Analýza > Přežití > Cox w/Time-Dep Cov ...

2. Zadejte výraz pro kovariát závislý na čase.

3. Klepněte na volbu **Model** a pokračujte v Cox Regression.

Poznámka-Ujistěte se, že jste zahrnuli nové proměnné, které jsou přidány jako kovariáty ve vašem modelu Cox Regression.

Kategoriální schémata kódování proměnných

V mnoha procedurách můžete požádat o automatické nahrazení kategoričké nezávislé proměnné sadou kontrastních proměnných, které pak budou zadány nebo odebrány z rovnice jako blok. Můžete určit, jak má být sada kontrastních proměnných kódována, obvykle v podpříkazu CONTRAST . Tato příloha vysvětluje a ilustruje, jak různé typy kontrastu požadované na CONTRAST skutečně fungují.

Odchylka

Odchylka od velké střední hodnoty. V maticových pojmech mají tyto kontrasty formu:

```
mean ( 1/k 1/k ... 1/k 1/k)
dř(1) ( 1-1/k -1/k ... -1/k -1/k)
dř(2) ( -1/k 1-1/k ... -1/k -1/k)
.
.
.
dř(k-1) ( -1/k -1/k ... 1-1/k -1/k)
```

kde k je počet kategorií pro nezávislou proměnnou a poslední kategorie je standardně vynechána. Například odchylka kontrastuje pro nezávislou proměnnou se třemi kategoriemi:

```
( 1/3 1/3 1/3)
( 2/3 -1/3 -1/3)
(-1/3 2/3 -1/3)
```

Chcete-li vynechat kategorii jinou než poslední, uveďte číslo vynechané kategorie v závorkách za klíčovým slovem DEVIATION . Například následující podpříkaz získá odchylky pro první a třetí kategorii a vynechá druhou kategorii:

```
/CONTRAST (FACTOR)=DEVIATION(2)
```

Předpokládejme, že *factor* má tři kategorie. Výsledná kontrastní matice bude

```
( 1/3 1/3 1/3)
( 2/3 -1/3 -1/3)
(-1/3 -1/3 2/3)
```

Jednoduché

Jednoduché kontrasty. Porovná každou úroveň faktoru s poslední. Obecná maticová forma je

```
mean (1/k 1/k ... 1/k 1/k)
dř(1) ( 1 0 ... 0 -1)
dř(2) ( 0 1 ... 0 -1)
.
.
.
```

```
df(k-1) ( 0 0 ... 1 -1)
```

kde k je počet kategorií pro nezávislou proměnnou. Například jednoduché kontrasty pro nezávislou proměnnou se čtyřmi kategoriemi jsou následující:

```
(1/4 1/4 1/4 1/4)
( 1 0 0 -1)
( 0 1 0 -1)
( 0 0 1 -1)
```

Chcete-li použít jinou kategorii místo poslední jako referenční kategorii, uveďte v závorkách za klíčovým slovem SIMPLE pořadové číslo referenční kategorie, což není nutně hodnota přidružená k této kategorii. Například následující podpříkaz CONTRAST získá matici kontrastu, která vynechá druhou kategorii:

```
/CONTRAST(FACTOR) = SIMPLE(2)
```

Předpokládejme, že *faktor* má čtyři kategorie. Výsledná kontrastní matice bude

```
(1/4 1/4 1/4 1/4)
( 1 -1 0 0)
( 0 -1 1 0)
( 0 -1 0 1)
```

Pomocník

Helmert kontrastuje. Porovnává kategorie nezávislé proměnné se střední hodnotou následujících kategorií. Obecná maticová forma je

```
mean (1/k 1/k ... 1/k 1/k 1/k)
df(1) ( 1 -1/(k-1) ... -1/(k-1) -1/(k-1) -1/(k-1))
df(2) ( 0 1 ... -1/(k-2) -1/(k-2) -1/(k-2))
.
.
df(k-2) ( 0 0 ... 1 -1/2 -1/2)
df(k-1) ( 0 0 ... 0 1 -1)
```

kde k je počet kategorií nezávislé proměnné. Například nezávislá proměnná se čtyřmi kategoriemi má kontrastní matici Helmert v následujícím tvaru:

```
(1/4 1/4 1/4 1/4)
( 1 -1/3 -1/3 -1/3)
( 0 1 -1/2 -1/2)
( 0 0 1 -1)
```

Rozdíl

Rozdíl nebo reverzní Helmert kontrastuje. Porovnává kategorie nezávislé proměnné se střední hodnotou předchozích kategorií proměnné. Obecná maticová forma je

```
mean ( 1/k 1/k 1/k ... 1/k)
df(1) ( -1 1 0 ... 0)
df(2) ( -1/2 -1/2 1 ... 0)
.
.
df(k-1) (-1/(k-1) -1/(k-1) -1/(k-1) ... 1)
```

kde k je počet kategorií pro nezávislou proměnnou. Rozdíl mezi nezávislými proměnnými se čtyřmi kategoriemi je například následující:

```
( 1/4 1/4 1/4 1/4)
( -1 1 0 0)
(-1/2 -1/2 1 0)
(-1/3 -1/3 -1/3 1)
```

Polynomní

Ortogonalní polynomiální kontrasty. První stupeň volnosti obsahuje lineární účinek ve všech kategoriích; druhý stupeň volnosti, kvadratický účinek; třetí stupeň volnosti, krychlový; a tak dále, pro účinky vyššího řádu.

Můžete určit vzdálenost mezi úrovněmi léčby měřenou danou kategoričkou proměnnou. Rovné mezery, což je výchozí hodnota, pokud vynesete metriku, lze zadat jako následná celá čísla od 1 do k , kde k je počet kategorií. Pokud má proměnná *drug* tři kategorie, podpříkaz

```
/CONTRAST (DRUG)=POLYNOMIAL
```

je stejná jako

```
/CONTRAST (DRUG)=POLYNOMIAL (1, 2, 3)
```

Stejné mezery však nejsou vždy nutné. Předpokládejme například, že *droga* představuje různé dávky léku poskytnuté třem skupinám. Je-li dávka podávaná druhé skupině dvakrát vyšší než dávka podávaná první skupině a dávka podávaná třetí skupině je třikrát vyšší než dávka podávaná první skupině, jsou kategorie léčby rovnoměrně rozloženy a vhodná metrika pro tuto situaci se skládá z po sobě jdoucích celých čísel:

```
/CONTRAST (DRUG)=POLYNOMIAL (1, 2, 3)
```

Pokud je však dávka podávaná druhé skupině čtyřikrát vyšší než dávka podávaná první skupině a dávka podávaná třetí skupině je sedmkrát vyšší než dávka podávaná první skupině,

```
/CONTRAST (DRUG)=POLYNOMIAL (1, 4, 7)
```

V obou případech je výsledkem specifikace kontrastu, že první stupeň volnosti pro *lék* obsahuje lineární účinek úrovní dávkování a druhý stupeň volnosti obsahuje kvadratický účinek.

Polynomiální kontrasty jsou užitečné zejména při testech trendů a při zkoumání povahy povrchů odezvy. Můžete také použít polynomiální kontrasty k provedení nelineárního prokládání křivek, jako je například curvilineární regrese.

Opakováno

Porovná sousední úrovně nezávislé proměnné. Obecná maticová forma je

```
mean (1/k 1/k 1/k ... 1/k 1/k)
df(1) ( 1 -1 0 ... 0 0)
df(2) ( 0 1 -1 ... 0 0)
.
.
df(k-1) ( 0 0 0 ... 1 -1)
```

kde k je počet kategorií pro nezávislou proměnnou. Například opakované kontrasty pro nezávislou proměnnou se čtyřmi kategoriemi jsou následující:

```
(1/4 1/4 1/4 1/4)
( 1 -1 0 0)
( 0 1 -1 0)
( 0 0 1 -1)
```

Tyto kontrasty jsou užitečné v analýze profilu a všude tam, kde jsou potřeba rozdílové skóre.

Speciální

Uživatelsky definovaný kontrast. Umožňuje zadávání speciálních kontrastů v podobě čtvercových matic s tolika řádky a sloupci, jako jsou kategorie dané nezávislé proměnné. Pro MANOVA a LOGLINEAR je první zadaný řádek vždy střední hodnotou nebo konstantou a představuje sadu vah, která označuje, jak se mají ostatní nezávislé proměnné, pokud existují, průměrovat nad danou proměnnou. Obecně je tento kontrast vektorem těch.

Zbývající řádky matice obsahují speciální kontrasty, které označují porovnání mezi kategoriemi proměnné. Obvykle jsou nejužitečnější ortogonální kontrasty. Ortogonální kontrasty jsou statisticky nezávislé a nejsou redundantní. Kontrasty jsou ortogonální, pokud:

- Pro každý řádek, kontrastní koeficienty součet na 0.
- Součty odpovídajících koeficientů pro všechny dvojice disjunktní řádky také sečtou na 0.

Předpokládejme například, že léčba má čtyři úrovně a že chcete vzájemně porovnat různé úrovně léčby. Vhodným speciálním kontrastem je

```
(1 1 1 1) weights for mean calculation
(3 -1 -1 -1) compare 1st with 2nd through 4th
(0 2 -1 -1) compare 2nd with 3rd and 4th
(0 0 1 -1) compare 3rd with 4th
```

kteřé určité pomocí následujícího podpříkazu CONTRAST pro MANOVA, LOGISTIC REGRESSIONa COXREG:

```
/CONTRAST(TREATMNT)=SPECIAL( 1 1 1 1
                             3 -1 -1 -1
                             0 2 -1 -1
                             0 0 1 -1 )
```

Pro systém LOGLINEARje třeba zadat:

```
/CONTRAST(TREATMNT)=BASIS SPECIAL( 1 1 1 1
                                    3 -1 -1 -1
                                    0 2 -1 -1
                                    0 0 1 -1 )
```

Každý řádek kromě znamená řádek součty na 0. Produkty z každé dvojice disjunktní řádky součet na 0, jakož:

```
Rows 2 and 3: (3)(0) + (-1)(2) + (-1)(-1) + (-1)(-1) = 0
Rows 2 and 4: (3)(0) + (-1)(0) + (-1)(1) + (-1)(-1) = 0
Rows 3 and 4: (0)(0) + (2)(0) + (-1)(1) + (-1)(-1) = 0
```

Zvláštní kontrasty nemusí být ortogonální. Nesmí se však jednat o lineární vzájemné kombinace. Pokud ano, procedura ohlásí lineární závislost a ukončí zpracování. Helmert, rozdíl a polynomiální kontrasty jsou všechny ortogonální kontrasty.

Indikátor

kódování proměnné indikátoru. Také známý jako fiktivní kódování, toto není k dispozici v LOGLINEAR nebo MANOVA. Počet nových kódovaných proměnných je $k-1$. Případy v referenční kategorii jsou kódovány 0 pro všechny proměnné $k-1$. Příklad v kategorii i^{th} je kódován 0 pro všechny proměnné indikátoru kromě i^{th} , který je kódován 1.

Kovarianční struktury

Tento oddíl poskytuje další informace o kovarianních strukturách.

Ante-Dependence: První objednávka. Tato kovarianční struktura má heterogenní rozptyly a heterogenní korelace mezi sousedními prvky. Korelace mezi dvěma nesousedními prvky je součin korelací mezi prvky, které leží mezi prvky zájmu.

$(\sigma_1^2$	$\sigma_2\sigma_1P_1$	$\sigma_3\sigma_1P_1P_2$	$\sigma_4\sigma_1P_1P_2P_3$
$(\sigma_2\sigma_1P_1$	σ_2^2	$\sigma_3\sigma_2P_2$	$\sigma_4\sigma_2P_2P_3$
$(\sigma_3\sigma_1P_1P_2$	$\sigma_3\sigma_2P_2$	σ_3^2	$\sigma_4\sigma_3P_3$
$(\sigma_4\sigma_1P_1P_2P_3$	$\sigma_4\sigma_2P_2P_3$	$\sigma_4\sigma_3P_3$	σ_4^2

AR (1). Jedná se o autoregresní strukturu prvního řádu s homogenními odchylkami. Korelace mezi libovolnými dvěma prvky se rovná rho pro sousední prvky, rho² pro prvky oddělené třetinou atd. je omezeno tak, aby $-1 < \rho < 1$.

$(\sigma^2$	σ^2P	σ^2P^2	σ^2P^3
$(\sigma^2P$	σ^2	σ^2P	σ^2P^2
$(\sigma^2P^2$	σ^2P	σ^2	σ^2P
$(\sigma^2P^3$	σ^2P^2	σ^2P	σ^2

AR (1): Heterogenní. Jedná se o autoregresní strukturu prvního řádu s heterogenními odchylkami. Korelace mezi libovolnými dvěma prvky se rovná r pro sousední prvky, r^2 pro dva prvky oddělené třetinou atd. je nucen ležet mezi -1 a 1.

$(\sigma_1^2$	$\sigma_2\sigma_1$	$\sigma_3\sigma_1P^2$	$\sigma_4\sigma_1P^3$)
$(\sigma_2\sigma_1P$	σ_2^2	$\sigma_3\sigma_2P$	$\sigma_4\sigma_2P^2$)
$(\sigma_3\sigma_1P^2$	$\sigma_3\sigma_2P$	σ_3^2	$\sigma_4\sigma_3P$)
$(\sigma_4\sigma_1P^3$	$\sigma_4\sigma_2P^2$	$\sigma_4\sigma_3P$	σ_4^2)

ARMA (1, 1). Jedná se o autoregresivní strukturu klouzavého průměru prvního řádu. Má homogenní odchylky. Korelace mezi dvěma prvky se rovná ϕ pro sousední prvky, ϕ^2 pro prvky oddělené třetinou atd. a jsou parametry autoregrese a klouzavého průměru, v tomto pořadí, a jejich hodnoty jsou omezeny tak, aby ležely mezi -1 a 1, včetně.

$(\sigma^2$	$\sigma^2\phi P$	$\sigma^2\phi P^2$	$\sigma^2\phi P^3$)
$(\sigma^2\phi P$	σ^2	$\sigma^2\phi P$	$\sigma^2\phi P^2$)
$(\sigma^2\phi P^2$	$\sigma^2\phi P$	σ^2	$\sigma^2\phi P$)
$(\sigma^2\phi P^3$	$\sigma^2\phi P^2$	$\sigma^2\phi P$	σ^2)

Složená symetrie. Tato struktura má konstantní rozptyl a stálou kovarianci.

$(\sigma^2 + \sigma_1^2$	σ_1	σ_1	σ_1)
$(\sigma_1$	$\sigma^2 + \sigma_1^2$	σ_1	σ_1)
$(\sigma_1$	σ_1	$\sigma^2 + \sigma_1^2$	σ_1)
$(\sigma_1$	σ_1	σ_1	$\sigma^2 + \sigma_1^2$)

Složená symetrie: Korelační metrika. Tato kovarianční struktura má homogenní rozptyly a homogenní korelace mezi prvky.

$(\sigma^2$	σ^2P	σ^2P	σ^2P)
$(\sigma^2P$	σ^2	σ^2P	σ^2P)
$(\sigma^2P$	σ^2P	σ^2	σ^2P)
$(\sigma^2P$	σ^2P	σ^2P	σ^2)

Složená symetrie: heterogenní. Tato kovarianční struktura má heterogenní rozptyly a konstantní korelaci mezi prvky.

$(\sigma_1^2$	$\sigma_2\sigma_1$	$\sigma_3\sigma_1P$	$\sigma_4\sigma_1P$)
$(\sigma_2\sigma_1P$	σ_2^2	$\sigma_3\sigma_2P$	$\sigma_4\sigma_2P$)
$(\sigma_3\sigma_1P$	$\sigma_3\sigma_2P$	σ_3^2	$\sigma_4\sigma_3P$)
$(\sigma_4\sigma_1P$	$\sigma_4\sigma_2P$	$\sigma_4\sigma_3P$	σ_4^2)

Diagonální. Tato kovarianční struktura má heterogenní rozptyly a nulovou korelaci mezi prvky.

$(\sigma_1^2$	0	0	0)
(0	σ_2^2	0	0)
(0	0	σ_3^2	0)
(0	0	0	σ_4^2)

Přímý produkt AR1 (UN_AR1). Určuje produkt Kroneckerovo jedné nestrukturované matice a druhou matici kovariance autoregrese prvního řádu. První nestrukturovaná matice modeluje vícerozměrné pozorování a druhá struktura autoregresní kovariance prvního řádu modeluje datovou kovarianci v čase nebo jiném faktoru.

Přímý produkt nestrukturovaný (UN_UN). Určuje produkt Kroneckerovo dvou nestrukturovaných matic, přičemž první z nich modeluje vícerozměrné pozorování a druhá modeluje kovarianci dat v čase nebo jiném faktoru.

Složená symetrie přímého produktu (UN_CS). Určuje Kroneckerovo součin jedné nestrukturované matice a druhé matice kovariance sloučenina symetrie s konstantním rozptylem a kovariancí. První nestrukturovaná matice modeluje vícerozměrné pozorování a druhá sloučenina symetrie kovariance modeluje datovou kovarianci v čase nebo jiném faktoru.

Faktorová analýza: První objednávka. Tato kovarianční struktura má heterogenní rozptyly, které se skládají z výrazu, který je heterogenní napříč prvky, a z výrazu, který je homogenní napříč prvky. Kovariance mezi dvěma prvky je druhou odmocninou produktu jejich heterogenních rozptylu.

$$\begin{array}{cccc}
 (\lambda_1^2 + d & \lambda_2\lambda_1 & \lambda_3\lambda_1 & \lambda_4\lambda_1) \\
 (\lambda_2\lambda_1 & \lambda_2^2 + d & \lambda_3\lambda_2 & \lambda_4\lambda_2) \\
 (\lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2 + d & \lambda_4\lambda_3) \\
 (\lambda_4\lambda_1 & \lambda_4\lambda_2 & \lambda_4\lambda_3 & \lambda_4^2 + d)
 \end{array}$$

Faktorová analýza: První pořadí, heterogenní. Tato kovarianční struktura má heterogenní rozptyly, které se skládají ze dvou pojmů, které jsou heterogenní napříč prvky. Kovariance mezi dvěma prvky je druhou odmocninou produktu prvního z jejich heterogenních rozptylu.

$$\begin{array}{cccc}
 (\lambda_1^2 + d_1 & \lambda_2\lambda_1 & \lambda_3\lambda_1 & \lambda_4\lambda_1) \\
 (\lambda_2\lambda_1 & \lambda_2^2 + d_2 & \lambda_3\lambda_2 & \lambda_4\lambda_2) \\
 (\lambda_3\lambda_1 & \lambda_3\lambda_2 & \lambda_3^2 + d_3 & \lambda_4\lambda_3) \\
 (\lambda_4\lambda_1 & \lambda_4\lambda_2 & \lambda_4\lambda_3 & \lambda_4^2 + d_4)
 \end{array}$$

Huynh-Feldt. Jedná se o "kruhovou" matici, ve které se kovariance mezi dvěma prvky rovná průměru jejich rozptylu minus konstanta. Ani rozptyly, ani kovariance nejsou konstantní.

$$\begin{array}{cccc}
 (\sigma_1^2 & [\sigma_1^2 + \sigma_2^2]/2-\lambda & [\sigma_1^2 + \sigma_3^2]/2-\lambda & [\sigma_1^2 + \sigma_4^2]/2-\lambda) \\
 ([\sigma_1^2 + \sigma_2^2]/2-\lambda & \sigma_2^2 & [\sigma_2^2 + \sigma_3^2]/2-\lambda & [\sigma_2^2 + \sigma_4^2]/2-\lambda) \\
 ([\sigma_1^2 + \sigma_3^2]/2-\lambda & [\sigma_2^2 + \sigma_3^2]/2-\lambda & \sigma_3^2 & [\sigma_3^2 + \sigma_4^2]/2-\lambda) \\
 ([\sigma_1^2 + \sigma_4^2]/2-\lambda & [\sigma_2^2 + \sigma_4^2]/2-\lambda & [\sigma_3^2 + \sigma_4^2]/2-\lambda & \sigma_4^2)
 \end{array}$$

Zmenšená identita. Tato struktura má konstantní rozptyl. Předpokládá se, že neexistuje žádná korelace mezi žádnými prvky.

$$\begin{array}{cccc}
 (\sigma^2 & 0 & 0 & 0) \\
 (0 & \sigma^2 & 0 & 0) \\
 (0 & 0 & \sigma^2 & 0) \\
 (0 & 0 & 0 & \sigma^2)
 \end{array}$$

Prostorový: Napájení. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. d_{ij} je odhadovaná euklidovská vzdálenost mezi měřením i^{th} a j^{th} .

$$\begin{array}{cccc}
 (\sigma^2 & \sigma^2 p^{d_{12}} & \sigma^2 p^{d_{13}} & \sigma^2 p^{d_{14}}) \\
 (\sigma^2 p^{d_{12}} & \sigma^2 & \sigma^2 p^{d_{23}} & \sigma^2 p^{d_{24}})
 \end{array}$$

$(\sigma^2 P^{d_{13}}$	$\sigma^2 P^{d_{23}}$	σ^2	$\sigma^2 P^{d_{34}}$
$(\sigma^2 P^{d_{14}}$	$\sigma^2 P^{d_{24}}$	$\sigma^2 P^{d_{34}}$	σ^2

Prostorová: exponenciální. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. d_{ij} je odhadovaná euklidovská vzdálenost mezi měřením i^{th} a j^{th} .

$(\sigma^2$	σ^2 výraz $\{-d_{12}/\theta\}$	σ^2 výraz $\{-d_{13}/\theta\}$	σ^2 výraz $\{-d_{14}/\theta\}$
$(\sigma^2$ výraz $\{-d_{12}/\theta\}$	σ^2	σ^2 výraz $\{-d_{23}/\theta\}$	σ^2 výraz $\{-d_{24}/\theta\}$
$(\sigma^2$ výraz $\{-d_{13}/\theta\}$	σ^2 výraz $\{-d_{23}/\theta\}$	σ^2	σ^2 výraz $\{-d_{34}/\theta\}$
$(\sigma^2$ výraz $\{-d_{14}/\theta\}$	σ^2 výraz $\{-d_{24}/\theta\}$	σ^2 výraz $\{-d_{34}/\theta\}$	σ^2

Prostorový: gaussovský. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. d_{ij} je odhadovaná euklidovská vzdálenost mezi měřením i^{th} a j^{th} .

$(\sigma^2$	σ^2 výraz $\{-d_{12}/P^2\}$	σ^2 výraz $\{-d_{13}/P^2\}$	σ^2 výraz $\{-d_{14}/P^2\}$
$(\sigma^2$ výraz $\{-d_{12}/P^2\}$	σ^2	σ^2 výraz $\{-d_{23}/P^2\}$	σ^2 výraz $\{-d_{24}/P^2\}$
$(\sigma^2$ výraz $\{-d_{13}/P^2\}$	σ^2 výraz $\{-d_{23}/P^2\}$	σ^2	σ^2 výraz $\{-d_{34}/P^2\}$
$(\sigma^2$ výraz $\{-d_{14}/P^2\}$	σ^2 výraz $\{-d_{24}/P^2\}$	σ^2 výraz $\{-d_{34}/P^2\}$	σ^2

Prostorová: Lineární. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. d_{ij} je odhadovaná Euclidean vzdálenost mezi i^{th} a j^{th} měření a 1_{ij} je indikátorová funkce, která je 1, pokud $pd_{ij} \leq 0$ a 0 jinak.

$(\sigma^2$	$\sigma^2(1-pd_{12}) 1_{12}$	$\sigma^2(1-pd_{13}) 1_{13}$	$\sigma^2(1-pd_{14}) 1_{14}$
$(\sigma^2(1-pd_{12}) 1_{12}$	σ^2	$\sigma^2(1-pd_{23}) 1_{23}$	$\sigma^2(1-pd_{24}) 1_{24}$
$(\sigma^2(1-pd_{13}) 1_{13}$	$\sigma^2(1-pd_{23}) 1_{23}$	σ^2	$\sigma^2(1-pd_{34}) 1_{34}$
$(\sigma^2(1-pd_{14}) 1_{14}$	$\sigma^2(1-pd_{24}) 1_{24}$	$\sigma^2(1-pd_{34}) 1_{34}$	σ^2

Prostorový: Lineární-protokol. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. d_{ij} je odhadovaná euklidovská vzdálenost mezi i^{th} a j^{th} měření a 1_{ij} je indikátorová funkce, která je 1, pokud je hodnota $p \cdot \log(d_{ij}) \leq 0$ a 0 jinak.

$(\sigma^2$	$\sigma^2(1-P \text{ protokol } (d_{12})) 1_{12}$	$\sigma^2(1-P \text{ protokol } (d_{13})) 1_{13}$	$\sigma^2(1-P \text{ protokol } (d_{14})) 1_{14}$
$(\sigma^2(1-P \text{ protokol } (d_{12})) 1_{12}$	σ^2	$\sigma^2(1-P \text{ protokol } (d_{23})) 1_{23}$	$\sigma^2(1-P \text{ protokol } (d_{24})) 1_{24}$
$(\sigma^2(1-P \text{ protokol } (d_{13})) 1_{13}$	$\sigma^2(1-P \text{ protokol } (d_{23})) 1_{23}$	σ^2	$\sigma^2(1-P \log(d_{34})) 1_{34}$
$(\sigma^2(1-P \text{ protokol } (d_{14})) 1_{14}$	$\sigma^2(1-P \text{ protokol } (d_{24})) 1_{24}$	$\sigma^2(1-P \text{ protokol } (d_{34})) 1_{34}$	σ^2

Prostorová: sférická. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. $r_{ij} = d_{ij}/p$, kde d_{ij} je odhadovaná euklidovská vzdálenost mezi i^{th} a j^{th} měření. 1_{ij} je indikátorová funkce, která je 1, pokud $d_{ij} \leq P$ a 0 jinak.

$(\sigma^2$	$\sigma^2(1-3/2r_{12} + 1/2r_{12}^3) 1_{12}$	$\sigma^2(1-3/2r_{13} + 1/2r_{13}^3) 1_{13}$	$\sigma^2(1-3/2r_{14} + 1/2r_{14}^3) 1_{14}$
$(\sigma^2(1-3/2r_{12} + 1/2r_{12}^3) 1_{12}$	σ^2	$\sigma^2(1-3/2r_{23} + 1/2r_{23}^3) 1_{23}$	$\sigma^2(1-3/2r_{24} + 1/2r_{24}^3) 1_{24}$

$$\begin{pmatrix}
 \sigma^2(1-3/2r_{13} + 1/2r_{13}^3) & \sigma^2(1-3/2r_{23} + 1/2r_{23}^3) & 1_{23} & \sigma^2 & \sigma^2(1-3/2r_{34} + 1/2r_{34}^3) \\
 1_{13} & & & & 1_{34} \\
 \sigma^2(1-3/2r_{14} + 1/2r_{14}^3) & \sigma^2(1-3/2r_{24} + 1/2r_{24}^3) & 1_{24} & \sigma^2(1-3/2r_{34} + 1/2r_{34}^3) & 1_{34} \\
 1_{14} & & & & \sigma^2
 \end{pmatrix}$$

Toeplitz. Tato kovarianční struktura má homogenní rozptyly a heterogenní korelace mezi prvky. Korelace mezi sousedními prvky je homogenní napříč dvojicemi sousedních prvků. Korelace mezi prvky oddělenými třetinou je opět homogenní, a tak dále.

$$\begin{pmatrix}
 \sigma^2 & \sigma^2 P_1 & \sigma^2 P_2 & \sigma^2 P_3 \\
 \sigma^2 P_1 & \sigma^2 & \sigma^2 P_1 & \sigma^2 P_2 \\
 \sigma^2 P_2 & \sigma^2 P_1 & \sigma^2 & \sigma^2 P_1 \\
 \sigma^2 P_3 & \sigma^2 P_2 & \sigma^2 P_1 & \sigma^2
 \end{pmatrix}$$

Toeplitz: Heterogenní. Tato kovarianční struktura má heterogenní rozptyly a heterogenní korelace mezi prvky. Korelace mezi sousedními prvky je homogenní napříč dvojicemi sousedních prvků. Korelace mezi prvky oddělenými třetinou je opět homogenní, a tak dále.

$$\begin{pmatrix}
 \sigma_1^2 & \sigma_2 \sigma_1 P_1 & \sigma_3 \sigma_1 P_2 & \sigma_4 \sigma_1 P_3 \\
 \sigma_2 \sigma_1 P_1 & \sigma_2^2 & \sigma_3 \sigma_2 P_1 & \sigma_4 \sigma_2 P_2 \\
 \sigma_3 \sigma_1 P_2 & \sigma_3 \sigma_2 P_1 & \sigma_3^2 & \sigma_4 \sigma_3 P_1 \\
 \sigma_4 \sigma_1 P_3 & \sigma_4 \sigma_2 P_2 & \sigma_4 \sigma_3 P_1 & \sigma_4^2
 \end{pmatrix}$$

Nestrukturovaný. Jedná se o zcela obecnou kovarianční matici.

$$\begin{pmatrix}
 \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\
 \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\
 \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\
 \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2
 \end{pmatrix}$$

Nestrukturovaná: Korelační metrika. Tato kovarianční struktura má heterogenní rozptyly a heterogenní korelace.

$$\begin{pmatrix}
 \sigma_1^2 & \sigma_2 \sigma_1 P_{21} & \sigma_3 \sigma_1 P_{31} & \sigma_4 \sigma_1 P_{41} \\
 \sigma_2 \sigma_1 P_{21} & \sigma_2^2 & \sigma_3 \sigma_2 P_{32} & \sigma_4 \sigma_2 P_{42} \\
 \sigma_3 \sigma_1 P_{31} & \sigma_3 \sigma_2 P_{32} & \sigma_3^2 & \sigma_4 \sigma_3 P_{43} \\
 \sigma_4 \sigma_1 P_{41} & \sigma_4 \sigma_2 P_{42} & \sigma_4 \sigma_3 P_{43} & \sigma_4^2
 \end{pmatrix}$$

Komponenty odchylky. Tato struktura přiřazuje ke každému ze zadaných náhodných efektů škálovanou strukturu identity (ID).

Bayesovská statistika

Produkt IBM SPSS Statistika poskytuje podporu pro následující bayesovské statistiky.

Jeden vzorek a Paired vzorek T-testy

Bayesian One Sample Inference procedura poskytuje možnosti pro výrobu Bayesian inference na jeden vzorek a dva vzorky spárované t-test charakterizací posteriorních distribucí. Když máte normální data, můžete použít normální před získáním normální zadní.

Binomické poroční testy

Bayesian One Sample Inference: Binomial procedura poskytuje volby pro provedení Bayesian one-sample inference na binomické distribuci. Parametr zájmu je π , což označuje pravděpodobnost úspěchu v pevně stanoveném počtu pokusů, které mohou vést buď k úspěchu, nebo k selhání. Všimněte si, že každá studie je nezávislá na sobě, a pravděpodobnost π zůstává stejná v každé studii. Binomická náhodná proměnná může být považována za součet fixního počtu nezávislých Bernoulliho pokusů.

Poissonova distribuční analýza

Bayesian One Sample Inference: Poisson procedura poskytuje možnosti pro provedení Bayesian one-sample inference na Poissonovu distribuci. Poissonova distribuce, užitečný model pro vzácné události, předpokládá, že v malých časových intervalech je pravděpodobnost výskytu události úměrná délce čekací doby. Konjugát před v rámci rodiny distribuce Gamma se používá při kreslení Bayesovské statistické inference na Poissonovo rozdělení.

Související ukázky

Bayesovský návrh odvozování vzorků je velmi podobný Bayesovskému odvozování jednoho vzorku, pokud jde o manipulaci s párovanými vzorky. Názvy proměnných můžete zadat ve dvojicích a spustit bayesovskou analýzu rozdílu středních hodnot.

Nezávislé vzorky T-testy

Bayesovská procedura odvozování nezávislých vzorků poskytuje volby pro použití skupinové proměnné k definování dvou nesouvisejících skupin a k odvozování Bayesovského odvozování rozdílu dvou středních hodnot skupiny. Můžete odhadnout Bayesovy faktory pomocí různých přístupů a také charakterizovat požadované zadní rozdělení buď za předpokladu, že rozptyly jsou známé, nebo neznámé.

Korelace po dvojicích (Pearson)

Bayesův závěr o Pearsonově korelačním koeficientu měří lineární vztah mezi dvěma proměnnými stupnice společně po bivariátním normálním rozdělení. Konvenční statistické odvozování o korelačním koeficientu bylo široce diskutováno a jeho praxe je již dlouho nabízena v produktu IBM SPSS Statistics. Návrh Bayesovského odvozování o Pearsonově korelačním koeficientu vám umožňuje čerpat Bayesovské odvozování odhadem Bayesových faktorů a charakterizací posteriorních distribucí.

Lineární regrese

Bayesovská inference o lineární regrese je statistická metoda, která se široce používá v kvantitativním modelování. Lineární regrese je základní a standardní přístup, ve kterém výzkumníci používají hodnoty několika proměnných k vysvětlení nebo predikci hodnot výsledků stupnice. Bayesovská jednorozměrná lineární regrese je přístup k lineární regrese, kde se statistická analýza provádí v kontextu Bayesovské inference.

Jednosměrný test ANOVA

Bayesovská jednosměrná procedura ANOVA vytváří jednosměrnou analýzu rozptylu pro kvantitativní závislou proměnnou pomocí jediné (nezávislé) proměnné. Analýza rozptylu se používá k testování hypotézy, že několik střední hodnoty jsou stejné. SPSS Statistics podporuje Bayesovy faktory, konjugované hodnoty a neinformativní hodnoty.

Log-Modely lineární regrese

Návrh pro testování nezávislosti dvou faktorů vyžaduje dvě kategorické proměnné pro konstrukci pohotovostní tabulky a provádí Bayesovské odvozování na přidružení řádku-sloupce. Můžete odhadnout Bayesovy faktory předpokládáním různých modelů a charakterizovat požadovanou zadní distribuci simulací současného věrohodného intervalu pro podmínky interakce.

Jednosměrná opakovaná opatření AMNOVA

Bayesovská jednosměrná opakovaná opatření Postup ANOVA měří jeden faktor ze stejného subjektu v každém jednotlivém časovém bodu nebo podmínce a umožňuje, aby subjekty byly překročeny v rámci úrovně. Předpokládá se, že každý subjekt má jediné pozorování pro každý časový bod nebo stav (jako takový není interakce mezi předmětem a léčbou zohledněna).

Bayesian jeden vzorek odvozování: Normální

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Bayesian One Sample Inference: Normal postup poskytuje možnosti pro výrobu Bayesian inference na jeden-vzorek a dva-vzorek spárovaný t-test charakterizovat posterior distribucí. Když máte normální data, můžete použít normální před získáním normální zadní.

1. Z nabídek vyberte:

Analyzovat > Bayesovská statistika > Jedna ukázka normální

2. V seznamu **Dostupné proměnné** vyberte příslušné **Proměnné testu** . Musí být vybrána alespoň jedna proměnná.

Poznámka: Seznam dostupných proměnných poskytuje všechny proměnné kromě proměnných typu Date a String.

3. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 2. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

1
2

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků a Odhadnout faktor bayes** .

4. Vyberte a/nebo zadejte příslušná nastavení **Odchylka dat a Hypotéza** . Tabulka odráží proměnné, které jsou aktuálně v seznamu **Proměnné testu** . Při přidávání nebo odebrání proměnných ze seznamu **Proměnné testu** tabulka automaticky přidává nebo odebírá stejné proměnné ze svých sloupců proměnných.

- Je-li v seznamu **Proměnné testu** uvedena jedna nebo více proměnných, jsou povoleny sloupce **Proměnná známá a Hodnota odchylky** .

Známa odchylka

Tuto volbu vyberte pro každou proměnnou, je-li odchylka známa.

¹ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

² Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Hodnota odchytky

Volitelný parametr, který určuje hodnotu rozptylu, je-li známa, pro pozorovaná data.

- Když se v seznamu **Proměnné testu** nachází jedna nebo více proměnných a není vybrána volba **Distribuce posterioru znaků**, jsou povoleny sloupce **Hodnota testu Null** a **Hodnota g**.

Hodnota testu s hodnotou null

Povinný parametr, který uvádí hodnotu null v odhadu Bayesova faktoru. Je povolena pouze jedna hodnota a výchozí hodnota je 0.

g Hodnota

Určuje hodnotu, která má být definována $ps^2 = g\sigma_x^2$ v odhadu Bayesova faktoru. Je-li zadána volba **Hodnota odchytky**, je výchozí hodnota **g hodnota 1**. Není-li zadána hodnota **Odchytky**, můžete zadat pevnou hodnotu g nebo hodnotu vynechat a integrovat ji.

5. Volitelně můžete klepnout na volbu **Kritéria** a určit nastavení "Bayesian jeden vzorový odvozování: Kritéria" na stránce 103 (procento důvěryhodného intervalu, volby chybějících hodnot a nastavení numerické metody), nebo klepnout na volbu **Priory** a zadat nastavení "Bayesian jeden vzorek Inference: Normální Priory" na stránce 104 (typ položek, jako jsou parametry odvozování, daná odchytky střední hodnoty nebo přesnost).

Bayesian jeden vzorový odvozování: Kritéria

Pro Bayesian One-Sample Inference můžete určit následující kritéria analýzy:

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Chybějící hodnoty

Uveďte metodu, ve které se mají řídit chybějící hodnoty.

Vyloučit případy po dvojicích

Toto je výchozí nastavení a vylučuje záznamy s chybějícími hodnotami na základě analýzy. Záznamy, které obsahují chybějící hodnoty, pro pole, které se používá pro specifický test, jsou z testu vynechány.

Vyloučit případy podle seznamu

Toto nastavení vyloučí záznamy, které obsahují chybějící hodnoty v seznamu. Záznamy, které obsahují chybějící hodnoty pro jakékoli pole, které je pojmenováno v libovolném podpříkazu, jsou vyloučeny ze všech analýz.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody**.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Adaptivní Gauss-Lobatto kvadratura

Toto je výchozí nastavení a nazývá se Adaptive Gauss-Lobatto Quadrature.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001. Tato volba je k dispozici pouze v případě, že je vybráno nastavení **Adaptivní Gauss-Lobatto Quadrature**.

Maximální počet iterací

Zadejte maximální počet iterací metody Adaptive Gauss-Lobatto Quadrature. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 2000. Tato volba je k dispozici pouze v případě, že je vybráno nastavení **Adaptivní Gauss-Lobatto Quadrature**.

Přibližování Monte Carlo

Tato volba volá přibližný přístup Monte Carlo.

Nastavit vlastní počáteční hodnotu

Je-li tato volba vybrána, můžete do pole **Seed** zadat vlastní hodnotu náhodného řetězce.

Počáteční položka

Zadejte náhodnou sadu náhodných hodnot pro metodu přibližování Monte Carlo. Hodnotou musí být kladné celé číslo. Standardně je přiřazena náhodná počáteční hodnota.

Počet ukázek Monte Carlo

Uvedte počet bodů, které jsou vzorkovány pro aproximaci Monte Carlo. Hodnotou musí být kladné celé číslo. Výchozí hodnota je 1000000. Volba je k dispozici pouze v případě, že je vybráno nastavení **Sbližování Monte Carlo**.

Bayesian jeden vzorek Inference: Normální Priory

Pro Bayesian One-Sample Inference můžete určit následující předchozí distribuční kritéria:

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahrnut, jak data se zvyšuje. Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Předchozí odchylka/přesnost

Poskytuje volby pro definování rozptylu a hodnot přesnosti.

Odchylka

Tuto volbu vyberte, chcete-li pro parametr rozptylu určit předchozí rozdělení. Je-li vybrána tato volba, seznam **Předchozí distribuce** poskytuje následující volby:

Poznámka: Je-li pro některé proměnné již zadána odchylka dat, následující nastavení se pro tyto proměnné ignorují.

- **Diffuse** -výchozí nastavení. Uvádí difúzní před.
- **Inverzní chí-kvadrát** -Určuje rozdělení a parametry pro inverzní-volbu²(v_0, σ^2_0), kde $v_0 > 0$ je stupeň volnosti a $\sigma^2_0 > 0$ je parametr měřítka.
- **Inverzní gama** -Určuje rozdělení a parametry pro inverzní gama (α_0, β_0), kde $\alpha_0 > 0$ je parametr tvaru a $\beta_0 > 0$ je parametr měřítka.
- **Jeffreys S2** -Uvádí neinformativní předchozí hodnotu $1/\sigma^2_0$.
- **Jeffreys S4** -Určuje neinformativní předchozí hodnotu umístění $1/\sigma^4_0$.

Přesnost

Tuto volbu vyberte, chcete-li pro parametr přesnosti určit předchozí distribuci. Je-li vybrána tato volba, seznam **Předchozí distribuce** poskytuje následující volby:

- **Gamma** -Určuje rozdělení a parametry pro Gamma (α_0, β_0), kde $\alpha_0 > 0$ je parametr tvaru a $\beta_0 > 0$ je parametr měřítka.
- **Čchí-kvadrát** -Určuje rozdělení a parametry pro volbu²(v_0), kde $v_0 > 0$ je stupeň volnosti.

Parametr tvaru

Zadejte parametr tvaru a_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0.

Parametr měřítka

Zadejte parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0. Čím větší je parametr měřítka, tím více je rozložení rozloženo.

Před na střední danou odchylku/přesnost

Uvedte předchozí distribuci pro parametr mean, který je podmíněný rozptyl nebo parametrem precision.

Běžný

Určuje rozdělení a parametry pro hodnotu Normální ($\mu_0, K^{-1}_0\sigma^2_0$) na rozptyl nebo Normální ($\mu_0, K_0/\sigma^2_0$) na přesnost, kde μ_0 Umístění ($-\infty, \infty$) a $\sigma^2 > 0$.

Parametr umístění

Zadejte číselnou hodnotu, která uvádí parametr umístění pro distribuci.

Parametr měřítka

Zadejte parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0.

Kappa a další

Zadejte hodnotu K_0 v Normální ($\mu_0, K^{-1}_0\sigma^2_0$) nebo Normální ($\mu_0, K_0/\sigma^2_0$). Musíte zadat jednu hodnotu, která je větší než 0 (1 je výchozí hodnota).

Difuzní

Výchozí nastavení, které určuje difúzní předchozí hodnotu 2.

Bayesian jeden vzorek Inference: Binomické

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Bayesian One Sample Inference: Binomial procedura poskytuje volby pro provedení Bayesian one-sample inference na binomické distribuci. Parametr zájmu je π , což označuje pravděpodobnost úspěchu v pevně stanoveném počtu pokusů, které mohou vést buď k úspěchu, nebo k selhání. Všimněte si, že každá studie je nezávislá na sobě, a pravděpodobnost π zůstává stejná v každé studii. Binomická náhodná proměnná může být považována za součet fixního počtu nezávislých Bernoulliho pokusů.

Ačkoli to není nutné, při odhadu binomického parametru se obvykle volí předchozí z rodiny beta distribuce. Beta rodina je konjugát pro binomické rodiny, a jako takový vede k zadní distribuci s uzavřenou formou stále v beta distribuční rodině.

1. Z nabídek vyberte:

Analyzovat > Bayesovská statistika > Jedna ukázka binomické

2. V seznamu **Dostupné proměnné** vyberte příslušné **Proměnné testu**. Musí být vybrána alespoň jedna proměnná.

Poznámka: Seznam dostupných proměnných poskytuje všechny proměnné kromě proměnných typu Date a String.

3. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

3

4

- **Použít obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

4. Vyberte a/nebo zadejte odpovídající nastavení **Kategorie úspěchu a Hypotéza**. Tabulka odráží proměnné, které jsou aktuálně v seznamu **Proměnné testu**. Při přidávání nebo odebrání proměnných z **Proměnné testu** tabulka automaticky přidává nebo odebrá stejné proměnné ze svých sloupců dvojic proměnných.

- Je-li jako **Bayesovská analýza** vybrána volba **Znaková distribuce posterioru**, je sloupec **Kategorie úspěchu** povolen.
- Je-li jako **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo volba **Použít obě metody**, budou povoleny všechny upravitelné sloupce.

Hodnota Null bodu

Povolí a zakáže volbu **Podíl hodnoty Null**. Je-li nastavení povoleno, jsou zakázány volby **Předchozí tvar s hodnotou Null** i **Předchozí měřítko s hodnotou Null**.

Předchozí tvar s hodnotou null

Určuje parametr tvaru a_0 pod nulovou hypotézou binomického odvozování.

Předchozí měřítko s hodnotou null

Určuje parametr měřítka b_0 pod nulovou hypotézou binomického odvozování.

Nulový podíl

Uvádí parametr tvaru a_0 a parametr měřítka b_0 pod nulovou hypotézou pro konjugaci před distribucí (pro přizpůsobení beta a Haldaneových předností). Platný rozsah jsou číselné hodnoty mezi 0 a 1.

Alternativní předchozí tvar

Povinný parametr pro určení a_0 pod alternativní hypotézou binomického odvozování, pokud má být odhadnut Bayesovský faktor.

Alternativní předchozí měřítko

Povinný parametr pro určení b_0 pod alternativní hypotézou Binomického odvozování, pokud má být odhadnut Bayesovský faktor.

Kategorie úspěchu

Poskytuje volby pro definování konjugace předchozích distribucí. Poskytnuté volby určují, jak je definován úspěch pro číselné a řetězcové proměnné, když jsou datové hodnoty testovány na testovací hodnotě.

Poslední kategorie

Výchozí nastavení, které provede binomický test s použitím poslední číselné hodnoty nalezené v kategorii po jejím seřazení ve vzestupném pořadí.

První kategorie

Provede binomický test s použitím první číselné hodnoty nalezené v kategorii po jejím seřazení ve vzestupném pořadí.

Střední bod

Používá číselné hodnoty \geq střed jako případy. Střední hodnota je průměr minimálních a maximálních ukázkových dat.

Bod řezu

Používá číselné hodnoty \geq zadanou hodnotu uzavření objektu jako případy. Nastavení musí být jedna číselná hodnota.

³ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

⁴ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Úroveň

Zachází s uživatelem určenými hodnotami řetězce (mohou být více než 1) jako s případy. Jednotlivé hodnoty oddělte čárkami.

5. Volitelně můžete klepnout na volbu **Kritéria** a zadat nastavení "Bayesian jeden vzorový odvozování: Kritéria" na stránce 103 (procento důvěryhodného intervalu, volby chybějících hodnot a nastavení numerické metody), nebo klepnout na volbu **Priory** a zadat nastavení "Bayesian jeden vzorek Inference: Binomické/Poisson Priors" na stránce 107 (konjugovat nebo vlastní předchozí distribuce).

Bayesian jeden vzorek Inference: Binomické/Poisson Priors

Pro Bayesian One-Sample Inference můžete určit následující předchozí distribuční kritéria:

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahlcen, jak data se zvyšuje. Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Parametr tvaru

Pro binomické priory zadejte parametr shape a_0 pro beta distribuci.

Pro Poisson priors zadejte parametr tvaru a_0 pro distribuci gama.

Musíte zadat jedinou hodnotu, která je větší než 0.

Parametr měřítka

Pro binomické priory zadejte parametr měřítka b_0 pro distribuci beta.

Pro Poisson priors zadejte parametr měřítka b_0 pro distribuci gama.

Musíte zadat jedinou hodnotu, která je větší než 0.

Bayesian jeden vzorek Inference: Poisson

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Bayesian One Sample Inference: Poisson procedura poskytuje možnosti pro provedení Bayesian one-sample inference na Poissonovu distribuci. Poissonova distribuce, užitečný model pro vzácné události, předpokládá, že v malých časových intervalech je pravděpodobnost výskytu události úměrná délce čekací doby. Konjugát před v rámci rodiny distribuce Gamma se používá při kreslení Bayesovské statistické inference na Poissonovo rozdělení.

1. Z nabídek vyberte:

Analyzovat > Bayesovská statistika > Jednovzorkové Poissonovo

2. V seznamu **Dostupné proměnné** vyberte příslušné **Proměnné testu**. Musí být vybrána alespoň jedna proměnná.

Poznámka: Seznam dostupných proměnných poskytuje všechny proměnné kromě proměnných typu Date a String.

3. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtížně a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 4. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H_0 : Hypotéza s hodnotou Null

H_1 : Alternativní hypotéza

5

6

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.
4. Vyberte a/nebo zadejte příslušná nastavení **Hodnoty hypotézy**. Tabulka odráží proměnné, které jsou aktuálně v seznamu **Proměnné testu**. Při přidávání nebo odebírání proměnných z **Proměnné testu** tabulka automaticky přidává nebo odebírá stejné proměnné ze svých sloupců dvojic proměnných.
- Je-li jako volba **Bayesovská analýza** vybrána volba **Znaková distribuce**, nejsou povoleny žádné sloupce.
 - Je-li jako **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo volba **Použit obě metody**, budou povoleny všechny upravitelné sloupce.

Hodnota Null bodu

Povolí a zakáže volbu **Četnost hodnot Null**. Je-li nastavení povoleno, jsou zakázány volby **Předchozí tvar s hodnotou Null** i **Předchozí měřítko s hodnotou Null**.

Předchozí tvar s hodnotou null

Určuje parametr shape a_0 pod nulovou hypotézou Poissonova odvozování.

Předchozí měřítko s hodnotou null

Určuje parametr měřítka b_0 pod nulovou hypotézou Poissonova odvozování.

Četnost hodnot null

Uvádí parametr tvaru a_0 a parametr měřítka b_0 pod nulovou hypotézou pro konjugát před distribucí (pro přízpusobení vztahu Poissonova-gama). Minimální hodnota musí být číselná hodnota větší než 0; maximální hodnota musí být maximální hodnota typu double.

Alternativní předchozí tvar

Povinný parametr pro určení a_1 pod alternativní hypotézou Poissonova odvozování, pokud má být Bayesův faktor odhadnut.

Alternativní předchozí měřítko

Povinný parametr pro uvedení b_1 pod alternativní hypotézou Poissonova odvozování, pokud má být Bayesův faktor odhadnut.

5. Volitelně můžete klepnout na volbu **Kritéria** a zadat nastavení "Bayesian jeden vzorový odvozování: Kritéria" na stránce 103 (procento důvěryhodného intervalu, volby chybějících hodnot a nastavení

⁵ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

⁶ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

numerické metody), nebo klepnout na volbu **Priory** a zadat nastavení “Bayesian jeden vzorek Inference: Binomické/Poisson Priors” na stránce 107 (konjugovat nebo vlastní předchozí distribuce).

Bayesian související ukázkové odvozování: Normální

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Bayesian Related Sample Inference: Normal procedura poskytuje Bayesian one-sample inference options pro spárované vzorky. Názvy proměnných můžete zadat ve dvojicích a spustit bayesovskou analýzu rozdílu středních hodnot.

1. Z nabídek vyberte:

Analyzovat > Bayesovská statistika > Normální související ukázky

2. V seznamu **Dostupné proměnné** vyberte příslušné **Paired Variables**. Musí být vybrána alespoň jedna dvojice zdrojových proměnných a pro danou sadu dvojic nelze vybrat více než dvě zdrojové proměnné.

Poznámka: Seznam dostupných proměnných poskytuje všechny proměnné s výjimkou řetězcových proměnných.

3. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

<i>Tabulka 5. Běžně používané prahové hodnoty pro definování významnosti důkazů</i>					
Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

7

8

- **Použít obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

4. Vyberte a/nebo zadejte příslušná nastavení **Odchylka dat a Hypotéza**. Tabulka odráží dvojice proměnných, které jsou momentálně v seznamu **Paired Variables**. Při přidávání nebo odebrání dvojic

⁷ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

⁸ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

proměnných ze seznamu **Paired Variables** tabulka automaticky přidá nebo odebere stejné dvojice proměnných ze svých sloupců dvojic proměnných.

- Je-li v seznamu **Paired Variables** uvedena jedna nebo více dvojic proměnných, jsou povoleny sloupce **Variance Known** a **Variance Value**.

Známa odchylka

Tuto volbu vyberte pro každou proměnnou, je-li odchylka známa.

Hodnota odchylky

Volitelný parametr, který určuje hodnotu rozptylu, je-li známa, pro pozorovaná data.

- Když je v seznamu **Paired Variables** jedna nebo více dvojic proměnných a není vybrána volba **Characterize Posterior Distribution**, jsou povoleny sloupce **Null Test Value** a **g Value**.

Hodnota testu s hodnotou null

Povinný parametr, který uvádí hodnotu null v odhadu Bayesova faktoru. Je povolena pouze jedna hodnota a výchozí hodnota je 0.

g Hodnota

Určuje hodnotu, která má být definována $ps^2 = g\sigma_x^2$ v odhadu Bayesova faktoru. Je-li zadána volba **Hodnota odchylky**, je výchozí hodnota **g hodnota** 1. Není-li zadána hodnota **Odchylka**, můžete zadat pevnou hodnotu g nebo hodnotu vynechat a integrovat ji.

5. Volitelně můžete klepnout na volbu **Kritéria** a zadat nastavení "Bayesian jeden vzorový odvozování: Kritéria" na stránce 103 (procento důvěryhodného intervalu, volby chybějících hodnot a nastavení numerické metody), nebo klepnout na volbu **Priory** a zadat nastavení "Bayesian jeden vzorek Inference: Binomické/Poisson Priors" na stránce 107 (konjugovat nebo vlastní předchozí distribuce).

Bayesian Independent-Vzorové odvozování

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Procedura Bayesian Independent-Sample Inference poskytuje volby pro použití skupinové proměnné k definování dvou nesouvisajících skupin a k vytvoření Bayesovského odvozování rozdílu dvou středních hodnot skupiny. Můžete odhadnout Bayesovy faktory pomocí různých přístupů a také charakterizovat požadované zadní rozdělení buď za předpokladu, že rozptyly jsou známé, nebo neznámé.

1. Z nabídek vyberte:

Analýza > Bayesovská statistika > Normální nezávislé vzorky

2. Ze seznamu zdrojových proměnných vyberte příslušné **proměnné testu**. Musí být vybrána alespoň jedna zdrojová proměnná.
3. V seznamu **Dostupné proměnné** vyberte příslušnou **proměnnou seskupení**. Proměnná seskupení definuje dvě skupiny pro nepárový *t*-test. Vybraná proměnná seskupení může být buď číselná, nebo řetězcová proměnná.
4. Vyberte požadovanou **Bayesovskou analýzu**:
 - **Charakterizovat následnou distribuci**: Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
 - **Faktor odhadu bayes**: Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 6. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

9

10

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.
5. Pomocí voleb **Definovat skupiny** můžete definovat dvě skupiny pro test t určením dvou hodnot (pro řetězcové proměnné) nebo dvou hodnot, středového bodu nebo bodu řezu (pro číselné proměnné).

Poznámka: Uvedené hodnoty musí v proměnné existovat, jinak se zobrazí chybová zpráva, která označuje, že alespoň jedna ze skupin je prázdná.

Pro číselné proměnné:

- **Použit zadané hodnoty.** Zadejte hodnotu pro skupinu 1 a jinou hodnotu pro skupinu 2. Případy s jinými hodnotami jsou z analýzy vyloučeny. Čísla nemusí být celá čísla (například 6.25 a 12.5 jsou platná).
- **Použit střední hodnotu.** Je-li tato volba vybrána, jsou skupiny rozděleny do $<$ a \geq středových hodnot.
- **Použit bod vyjmutí.**
 - **Cutpoint.** Zadejte číslo, které rozdělí hodnoty proměnné seskupení do dvou sad. Všechny případy s hodnotami, které jsou menší než hodnota cutpoint, tvoří jednu skupinu, a případy s hodnotami, které jsou větší nebo rovny hodnotě cutpoint, tvoří druhou skupinu.

Pro proměnné seskupení řetězců zadejte řetězec pro skupinu 1 a jinou hodnotu pro skupinu 2, například *yes* a *no*. Případy s jinými řetězci jsou z analýzy vyloučeny.

6. Volitelně můžete klepnout na volbu **Kritéria**, chcete-li uvést nastavení “Bayesian Independent-Inference vzorků: kritéria” na stránce 112 (procento věrohodného intervalu, volby chybějících hodnot a nastavení metody adaptivní kvadratury), klepněte na volbu **Priory**, chcete-li uvést nastavení “Bayesian Independent-Inference vzorků: předchozí distribuce” na stránce 112 (odchylka dat, před rozptylem a před střední hodnotou podmíněná rozptylem), nebo klepněte na volbu **Odhadnout faktor bayes**, chcete-li uvést nastavení “Bayesian Independent-Sample Inference: Odhad Bayes Factor” na stránce 113.

Bayesian Independent-Skupiny definice odvozování vzorku (číselné)

V případě číselných proměnných seskupení definujte dvě skupiny pro test t určením dvou hodnot, středového bodu nebo bodu řezu.

⁹ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

¹⁰ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Poznámka: Uvedené hodnoty musí v proměnné existovat, jinak se zobrazí chybová zpráva, která označuje, že alespoň jedna ze skupin je prázdná.

- **Použit zadané hodnoty.** Zadejte hodnotu pro skupinu 1 a jinou hodnotu pro skupinu 2. Případy s jinými hodnotami jsou z analýzy vyloučeny. Čísla nemusí být celá čísla (například 6.25 a 12.5 jsou platná).
- **Použit střední hodnoty.** Je-li tato volba vybrána, jsou skupiny rozděleny do $<$ a \geq středových hodnot.
- **Použit bod vyjmutí.**
 - **Cutpoint.** Zadejte číslo, které rozdělí hodnoty proměnné seskupení do dvou sad. Všechny případy s hodnotami, které jsou menší než hodnota cutpoint, tvoří jednu skupinu, a případy s hodnotami, které jsou větší nebo rovny hodnotě cutpoint, tvoří druhou skupinu.

Bayesian Independent-Skupiny definice odvozování vzorku (řetězec)

Pro proměnné seskupení řetězců zadejte řetězec pro skupinu 1 a jinou hodnotu pro skupinu 2, například *yes* a *no*. Případy s jinými řetězci jsou z analýzy vyloučeny.

Poznámka: Uvedené hodnoty musí v proměnné existovat, jinak se zobrazí chybová zpráva, která označuje, že alespoň jedna ze skupin je prázdná.

Bayesian Independent-Inference vzorků: kritéria

Pro Bayesian Independent-Sample Inference můžete určit následující kritéria analýzy:

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Chybějící hodnoty

Uveďte metodu, ve které se mají řídit chybějící hodnoty.

Vyloučit případy po dvojicích

Toto je výchozí nastavení a vylučuje záznamy s chybějícími hodnotami na základě analýzy.

Záznamy, které obsahují chybějící hodnoty, pro pole, které se používá pro specifický test, jsou z testu vynechány.

Vyloučit případy podle seznamu

Toto nastavení vyloučí záznamy, které obsahují chybějící hodnoty v seznamu. Záznamy, které obsahují chybějící hodnoty pro jakékoli pole, které je pojmenováno v libovolném podpříkazu, jsou vyloučeny ze všech analýz.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použit obě metody**.

Adaptivní kvadrurní metoda

Uveďte hodnoty tolerance a maximální hodnoty iterace pro metodu Adaptive Quadrature Quadrature.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001.

Maximální počet iterací

Zadejte maximální počet iterací metody Adaptive Quadrature. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 500.

Bayesian Independent-Inference vzorků: předchozí distribuce

Pro Bayesian Independent-Sample Inference můžete určit následující kritéria předchozí distribuce:

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahlcen, jak data se zvyšuje. Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Odchylka dat

Poskytuje volby pro definování nastavení odchylky dat.

Známa odchylka

Je-li vybráno, umožňuje vám zadat dvě známé odchylky skupiny. Obě hodnoty musí být > 0 .

Rozptyl skupiny 1

Zadejte první známou hodnotu odchylky skupiny.

Rozptyl skupiny 2

Zadejte druhou známou hodnotu odchylky skupiny.

Předpokládat stejnou odchylku

Řídí, zda se předpokládá, že se obě odchylky skupiny rovnají. Standardně se předpokládá, že odchylky skupiny jsou nestejně. Toto nastavení je ignorováno, když jsou zadány hodnoty pro dvě odchylky skupiny.

Předpokládat nerovnou rozptyl

Řídí, zda se má za to, že dvě rozptyly skupiny jsou nerovné. Standardně se předpokládá, že odchylky skupiny jsou nestejně. Toto nastavení je ignorováno, když jsou zadány hodnoty pro dvě odchylky skupiny.

Před na rozptylu

Uvedte předchozí rozdělení pro dvě stejné odchylky.

Jeffreys

Je-li vybrána tato volba, použije se pro prostor parametrů neinformativní (objektivní) předchozí distribuce.

Inverzní-ChiSquare

Určuje spojitě rozdělení pravděpodobnosti náhodné proměnné s kladnou hodnotou a parametry pro inverzní-selhání²(v_0, σ^2_0), kde $v_0 > 0$ je stupeň volnosti a $\sigma^2_0 > 0$ je parametr měřítka.

Stupně volnosti

Zadejte hodnotu pro počet hodnot v konečném výpočtu, které se mohou lišit.

Parametr měřítka

Zadejte parametr měřítka $\sigma^2_0 > 0$ pro inverzní-práci s volbou- $\chi^2^2(v_0, \sigma^2_0)$. Musíte zadat jedinou hodnotu, která je větší než 0. Čím větší je parametr měřítka, tím více je rozložení rozloženo.

Před střední hodnotou podmíněný rozptyl

Poskytuje volby pro určení předchozí distribuce pro dvě střední hodnoty skupiny.

Poznámka: Volby **Difuzní** a **Normální** jsou k dispozici pouze v případě, že je vybrána volba **Variance known**.

Difuzní

Výchozí nastavení. Uvádí difúzní před.

Běžný

Je-li tato volba vybrána, musíte zadat parametry umístění a měřítka pro definované střední hodnoty skupiny.

Parametr umístění

Zadejte číselné hodnoty, které určují parametry umístění pro skupinové distribuce.

Parametr měřítka

Zadejte parametr měřítka $\sigma^2_0 > 0$ pro inverzní-práci s volbou- $\chi^2^2(v_0, \sigma^2_0)$. Pro každou skupinu musíte zadat jednu hodnotu, která je větší než 0. Čím větší je parametr měřítka, tím více je rozložení rozloženo.

Bayesian Independent-Sample Inference: Odhad Bayes Factor

Můžete určit metodu, která se použije k odhadu Bayesova faktoru.

Metoda směrovače

Je-li tato volba vybrána, vyvolá přístup směrovače. Toto je výchozí nastavení

Gonenova metoda

Je-li tato volba vybrána, vyvolá přístup Gonen a musíte zadat následující nastavení velikosti efektu:

Střední hodnota pro velikost efektu

Zadejte hodnotu, která určuje střední rozdíl mezi dvěma skupinami.

Odchylna velikosti efektu

Zadejte hodnotu, která určuje rozptyl pro dvě skupiny. Hodnota musí být větší než 0.

Hyper-Předchozí metoda

Je-li tato volba vybrána, vyvolá přístup hyper-g, kde musíte zadat jedinou hodnotu. Do pole **Parametr tvaru** zadejte hodnotu mezi -1 a -0.5. Výchozí hodnota je -0.75.

Bayesian Inference o Pearsonově korelaci

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Pearsonův korelační koeficient měří lineární vztah mezi dvěma proměnnými stupnice společně po bivariátním normálním rozdělení. Konvenční statistický závěr o korelačním koeficientu byl široce diskutován a jeho praxe je již dlouho nabízena v IBM SPSS Statistika. Návrh Bayesovského odvozování o Pearsonově korelačním koeficientu umožňuje uživatelům čerpat Bayesovské odvozování odhadem Bayesových faktorů a charakterizací posteriorních distribucí.

1. Z nabídek vyberte:

Analýza > Bayesovská statistika > Pearsonova korelace

2. Ze seznamu **Dostupné proměnné** vyberte příslušné **Proměnné testu**, které se mají použít pro odvozování korelace po dvojicích. Musí být vybrány alespoň dvě zdrojové proměnné. Jsou-li vybrány více než dvě proměnné, spustí se analýza pro všechny kombinace vybraných proměnných po dvojicích.

3. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 7. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

11

12

¹¹ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

- **Použít obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.
4. Zadejte **Maximální počet grafů**, které se mají zobrazit ve výstupu. Sada grafů může obsahovat 3 grafy ve stejném podokně. Grafy jsou generovány v pořadí z první proměnné versus zbývající proměnné, pak druhá proměnná versus zbývající proměnné atd. Definovaná celočíselná hodnota musí být v rozsahu 0 až 50. Standardně je výstup 10 sad grafů, které pojmu pět proměnných. Tato volba není k dispozici, je-li vybrána volba **Odhadovat faktor bayes**.
 5. Volitelně můžete klepnout na volbu **Kritéria**, chcete-li uvést nastavení "Bayesian Pearson korelace: Kritéria" na stránce 115 (procento věrohodného intervalu, volby chybějících hodnot a nastavení numerické metody), klepněte na volbu **Priory** a zadejte "Bayesian Pearson korelace: předchozí distribuce" na stránce 115 nastavení (hodnota c pro předchozí $p(P) (1 - P^2)^c$, nebo klepněte na volbu **Bayes Factor** a zadejte nastavení "Bayesian Independent-Sample Inference: Odhad Bayes Factor" na stránce 113.

Bayesian Pearson korelace: Kritéria

Pro Bayesian Pearson Correlation Inference (párovány) můžete určit následující kritéria analýzy.

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Chybějící hodnoty

Uveďte metodu, ve které se mají řídit chybějící hodnoty.

Vyloučit případy po dvojicích

Toto nastavení vyloučí záznamy, které obsahují chybějící hodnoty po dvojicích.

Vyloučit případy podle seznamu

Toto nastavení vyloučí záznamy, které obsahují chybějící hodnoty v seznamu. Záznamy, které obsahují chybějící hodnoty pro jakékoli pole, které je pojmenováno v libovolném podpříkazu, jsou vyloučeny ze všech analýz.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody**.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Nastavit vlastní počáteční hodnotu

Je-li tato volba vybrána, můžete do pole **Seed** zadat vlastní hodnotu náhodného řetězce.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001.

Maximální počet iterací

Zadejte maximální počet iterací metody. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 2000.

Počet ukázek Monte Carlo

Uveďte počet bodů, které jsou vzorkovány pro aproximaci Monte Carlo. Hodnotou musí být kladné celé číslo. Výchozí hodnota je 10000.

Ukázky simulované pro posteriorní distribuci

Uveďte počet vzorků, které se použijí k nakreslení požadované zadní distribuce. Výchozí hodnota je 10000.

Bayesian Pearson korelace: předchozí distribuce

Můžete zadat hodnotu c pro předchozí $p(P)$ oblast $(1 - P^2)^c$.

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahrncen, jak data se zvyšuje.

¹² Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Jednotná (c = 0)

Je-li vybráno, použije se uniforma před.

Jeffreys (c = -1.5)

Je-li vybráno, použije se neinformativní předchozí distribuce.

Nastavit vlastní hodnotu c

Je-li tato volba vybrána, můžete určit vlastní **hodnotu c**. Je povoleno libovolné reálné číslo.

Bayesian Pearson korelace: Bayes faktor

Můžete určit metodu, která se použije k odhadu Bayesova faktoru. Následující volby jsou k dispozici pouze v případě, že je vybrána volba **Odhadovat bayes faktor** nebo **Obě metody uživatele** Bayesovská analýza.

JZS Bayesova součinitel

Je-li vybráno, vyvolá přístup Zellner-Siow. Toto je výchozí nastavení.

Faktor zlomkových Bayesů

Je-li vybrána tato volba, můžete určit faktor zlomkové hodnoty Bayesovy hodnoty a hodnotu nulové hypotézy. Pro faktor zlomkové hodnoty Bayes musíte zadat hodnotu (0, 1). Výchozí hodnota je 0.5.

Bayesian Inference o lineárních regresních modelech

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Regrese je statistická metoda, která se široce používá v kvantitativním modelování. Lineární regrese je základní a standardní přístup, ve kterém výzkumníci používají hodnoty několika proměnných k vysvětlení nebo predikci hodnot výsledků stupnice. Bayesovská jednorozměrná lineární regrese je přístup k lineární regresi, kde se statistická analýza provádí v kontextu Bayesovského odvozování.

Můžete vyvolat regresní proceduru a definovat úplný model.

1. Z nabídek vyberte:

Analýza > Bayesovská statistika > Lineární regrese

2. Ze seznamu **Dostupné proměnné** vyberte jednu neřetězcovou závislou proměnnou. Musíte vybrat jednu neřetězcovou proměnnou.

3. Ze seznamu **Dostupné proměnné** vyberte jednu nebo více kategorických proměnných faktorů pro model.

4. Ze seznamu **Dostupné proměnné** vyberte jednu nebo více proměnných měřítka bez řetězce.

Poznámka: Seznamy **Faktor (y)** a **Covariate (y)** nemohou být prázdné. Musíte vybrat alespoň jednu proměnnou **Faktor (y)** nebo **Covariate (y)**.

5. Volitelně vyberte jednu neřetězcovou proměnnou, která bude sloužit jako regresní váha, ze seznamu **Dostupné proměnné**.

6. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 8. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

13

14

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

Volitelně můžete:

- Klepněte na volbu **Kritéria**, chcete-li určit nastavení věrohodného intervalu v procentech a numerické metody.
- Klepněte na volbu **Priory**, chcete-li definovat odkazy a konjugovat předchozí nastavení distribuce.
- Klepněte na volbu **Bayesový faktor**, abyste uvedli nastavení Bayesova faktoru.
- Klepnutím na tlačítko **Uložit** identifikujte, které položky se mají uložit, a uložte informace o modelu do souboru XML.
- Chcete-li určit regresory pro Bayesovskou předpověď, klepněte na volbu **Předpověďt**.
- Klepněte na volbu **Zákresy**, chcete-li zakreslit následné distribuce regresních parametrů, rozptyl chybových výrazů a předpovězené hodnoty.
- Klepněte na volbu **F-testy**, chcete-li porovnat statistické modely, abyste identifikovali model, který nejlépe vyhovuje populaci, z níž byl odebrán vzorek.

Bayesovské lineární regresní modely: Kritéria

Pro Bayesovské modely lineární regrese můžete uvést následující kritéria analýzy.

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použit obě metody**.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001.

¹³ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

¹⁴ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Maximální počet iterací

Zadejte maximální počet iterací metody. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 2000.

Bayesovské lineární regresní modely: Priors Distributions

Pro parametry regrese a rozptyl chyb můžete určit následující předchozí nastavení distribuce. Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Znaková distribuce**.

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahrnut, jak data se zvyšuje. Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Odkaz

Je-li vybrána tato volba, vytvoří referenční analýza objektivní Bayesovské odvozování. Inferenční příkazy závisí pouze na předpokládaném modelu a dostupných datech a předchozí distribuce, která se používá k odvozování, je nejméně informativní. Toto je výchozí nastavení.

Konjugát

Poskytuje volby pro definování konjugace předchozích distribucí. Konjugované priory předpokládají rozdělení kloubu Normal-Inverse-Gamma. Ačkoli konjugované priory nejsou při provádění Bayesovských aktualizací vyžadovány, pomáhají výpočetním procesům.

Poznámka: Chcete-li určit konjugované hodnoty pro model lineární regrese, nastavte očekávanou střední hodnotu regresních parametrů v tabulce **Priory pro rozptyl chyb**. Můžete se také rozhodnout použít nastavení **Odchylka matice kovariance** k určení předchozí variance-kovariance.

Převratné odchylky chyb

Parametr tvaru

Zadejte parametr tvaru a_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0.

Parametr měřítka

Zadejte parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0. Čím větší je parametr měřítka, tím více je rozložení rozloženo.

Tabulka uvádí střední hodnotu regresních parametrů (včetně zachycení), která určuje střední hodnotu vektoru θ_0 pro definované regresní parametry. Počet hodnot musí odpovídat počtu regresních parametrů, včetně termínu zachycení.

První název proměnné je vždy INTERCEPT. Ve druhém řádku je sloupec **Proměnné** automaticky naplněn proměnnými určenými pomocí faktorů a proměnných Covariate (s). Sloupec **Střední** neobsahuje žádné výchozí hodnoty.

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Rozptyl matice kovariance: σ^2x

Zadejte V_0 hodnoty v dolním trojúhelníku v matici variance-kovariance pro multivariační normální hodnotu před. Všimněte si, že hodnota V_0 musí být polokladná. Poslední hodnota každého řádku musí být kladná. Další řádek by měl mít o jednu hodnotu více než předchozí řádek. Pro referenční kategorie nejsou uvedeny žádné hodnoty (pokud existují).

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Použití matice identit

Je-li vybráno, použije se škálovaná matice identit. V dolním trojúhelníku v matici variance-kovariance nemůžete zadat hodnoty V_0 pro vícerozměrnou normální hodnotu před.

Bayesovské lineární regresní modely: Bayes faktor

Můžete uvést návrh modelu pro analýzu, včetně přístupu, který se použije k odhadu Bayesova faktoru pro Bayesovské modely lineární regrese. Následující volby jsou k dispozici pouze v případě, že je vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody** Bayesovská analýza.

Model s hodnotou Null

Je-li tato volba vybrána, jsou odhadované Bayesovy faktory založeny na modelu s hodnotou null. Toto je výchozí nastavení.

Úplný model

Je-li vybráno, odhadované Bayesovy faktory jsou založeny na úplném modelu a můžete vybrat proměnné, které se mají použít, a další faktory a kovariáty.

Proměnné

Vypíše všechny proměnné dostupné pro úplný model.

Další činitel (y)

Vyberte proměnné ze seznamu **Proměnné**, které se mají použít jako další faktory.

Další kovariát (y)

Vyberte proměnné ze seznamu **Proměnné**, které se mají použít jako další kovariáty.

Výpočet

Uveďte přístup k odhadu Bayesových faktorů. Výchozí nastavení je metoda JZS.

Metoda JZS

Je-li vybráno, vyvolá přístup Zellner-Siow. Toto je výchozí nastavení.

Zellnerova metoda

Je-li tato volba vybrána, vyvolá Zellnerovu metodu a vy musíte zadat jednu g předchozí hodnotu > 0 (neexistuje žádná výchozí hodnota).

Hyper-Předchozí metoda

Je-li tato volba vybrána, vyvolá přístup hyper- g a vy musíte zadat parametr tvaru a_0 pro distribuci Inverse-Gamma. Musíte zadat jednu hodnotu > 0 (výchozí hodnota je 3).

Metoda směrovače

Je-li tato volba vybrána, vyvolá přístup směrovače a vy musíte zadat parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jednu hodnotu > 0 (výchozí hodnota je 1).

Bayesovské lineární regresní modely: Uložit

Toto dialogové okno vám umožňuje určit, které statistiky mají být vyhodnoceny pro Bayesovskou distribuci předpovědí, a exportovat výsledky modelu do souboru XML.

Posteriorní prediktivní statistika

Můžete skóre následující statistiky, které jsou odvozeny z Bayesovských předpovědí.

Prostředky

Střední hodnota posteriorní prediktivní distribuce.

Odchytky

Rozptyl posteriorní prediktivní distribuce.

Režimy

Režim posteriorní prediktivní distribuce.

Dolní limit důvěryhodného intervalu

Dolní hranice věrohodného intervalu posteriorní prediktivní distribuce.

Horní limit důvěryhodného intervalu

Horní mez věrohodného intervalu posteriorní prediktivní distribuce.

Poznámka: Pro každou statistiku můžete přiřadit odpovídající názvy proměnných.

Exportovat informace o modelu do souboru XML

Zadejte název souboru XML a umístění pro export matice rozptylu parametru s přiděleným skóre-kovariance.

Bayesovské lineární regresní modely: Predikce

Můžete určit regresory pro generování prediktivních distribucí.

Regresory pro Bayesovskou předpověď

V tabulce jsou uvedeny všechny dostupné regresory. Sloupec **Regresory** je automaticky naplněn určitými proměnnými faktoru a Covariate. Zadejte pozorované vektory s hodnotami pro regresory. Každému regresoru lze přiřadit jednu hodnotu nebo řetězec a je povoleno předpovídat pouze jeden případ. Pro faktory jsou povoleny hodnoty i řetězce.

Aby bylo možné spustit předpověď (klepnutím na volbu **Spustit analýzu**), musí být zadána buď všechny, nebo žádná hodnota regrese.

Když je proměnná Factor nebo Covariate odebrána, odpovídající řádek regrese je odebrán z tabulky.

Pro kovariáty lze zadat pouze číselné hodnoty. Pro faktory jsou povoleny číselné hodnoty i řetězce.

Poznámka: Klepnutím na tlačítko **Resetovat** vymažete definované hodnoty.

Bayesovské lineární regresní modely: Plots

Můžete řídit grafy, které jsou výstupem.

Kovariáty

Zobrazí seznam aktuálně definovaných kovariátů.

Zakreslit kovariát (y)

Ze seznamu **Kovariáty** vyberte kovariáty, které mají být zakresleny, a přidejte je do seznamu **Kreslit kovariáty**.

Faktory

Zobrazí seznam aktuálně definovaných faktorů.

Faktor (y) zakreslení

Ze seznamu **Faktory** vyberte faktory, které mají být zakresleny, a přidejte je do seznamu **Faktory výkresu**.

Maximální počet kategorií, které mají být vyneseny

Vyberte maximální počet kategorií, které mají být zakresleny (jedno kladné celé číslo). Nastavení platí pro všechny faktory. Standardně jsou první 2 úrovně zakresleny pro každý faktor.

Zahrnout grafy

Zachytit výraz

Je-li vybráno, je zakreslen výraz zachycení. Ve výchozím nastavení je výběr nastavení zrušen.

Odchylka chybových výrazů

Je-li tato volba vybrána, je zakreslena odchylka chyb. Ve výchozím nastavení je výběr nastavení zrušen.

Bayesovská predikovaná distribuce

Je-li vybráno, zakreslí se prediktivní rozdělení. Ve výchozím nastavení je výběr nastavení zrušen. Nastavení lze vybrat pouze v případě, že jsou vybrány platné hodnoty regresorů.

Bayesovské lineární regresní modely: F-testy

Můžete vytvořit jeden nebo více dílčích F-testů. F-test je jakýkoli statistický test, při kterém má statistika testu rozdělení F podle nulové hypotézy. F-testy se běžně používají při porovnávání statistických modelů, které byly namontovány do datové sady, s cílem identifikovat model, který nejlépe vyhovuje populaci, z níž byly údaje odebrány.

Dostupné proměnné

Vypíše proměnné faktoru a kovariantní proměnné, které jsou vybrány z hlavního dialogového okna Bayesovské lineární regrese. Když jsou proměnné typu faktor a kovariantní proměnné přidány nebo odebrány z hlavního dialogového okna, seznam se odpovídajícím způsobem aktualizuje.

Testovací proměnná (proměnné)

Ze seznamu **Dostupné proměnné** vyberte proměnné faktoru/proměnné, které chcete testovat, a přidejte je do seznamu **Proměnné testování** .

Poznámka: Volba **Zahrnout výraz zachycení** musí být vybrána, když nejsou vybrány žádné testovací faktory nebo kovariáty.

Testovací proměnná (proměnné) a hodnota (hodnoty)

Zadejte hodnoty, které se mají testovat. Počet hodnot musí odpovídat počtu parametrů v původním modelu. Jsou-li zadány hodnoty, musí být uvedena první hodnota pro výraz zachycení (předpokládejme, že všechny hodnoty jsou 0, pokud nejsou explicitně definovány).

Zahrnout výraz zachycení

Je-li tato volba vybrána, jsou do testu zahrnuty výrazy zachycení. Standardně není nastavení vybráno.

Je-li tato volba povolena, zadejte hodnotu do pole **Hodnota testování** .

Testovací štítek (volitelný)

Volitelně můžete určit popisek pro každý test. Můžete zadat hodnotu řetězce s maximální délkou 255 bajtů. Pro každý F-test je povolen pouze jeden štítek.

Bayesovská jednosměrná ANOVA

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Procedura One-Way ANOVA vytváří jednosměrnou analýzu rozptylu pro kvantitativní závislou proměnnou pomocí jednofaktorové (nezávislé) proměnné. Analýza rozptylu se používá k testování hypotézy, že několik střední hodnoty jsou stejné. SPSS Statistika podporuje Bayesovy faktory, konjugované priory a neinformativní priory.

1. Z nabídek vyberte:

Analýza > Bayesovská statistika > Jednosměrná ANOVA

2. Ze seznamu **Dostupné proměnné** vyberte jednu číselnou proměnnou **Závislé** . Musíte vybrat alespoň jednu proměnnou.
3. Vyberte jednu proměnnou **Faktor** pro model ze seznamu **Dostupné proměnné** . Musíte vybrat alespoň jednu proměnnou **Faktor** .
4. Ze seznamu **Dostupné proměnné** vyberte jednu neřetězcovou proměnnou, která bude sloužit jako regrese **Váha** . Pole proměnné **Váha** může být prázdné.
5. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 9. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0

Tabulka 9. Běžně používané prahové hodnoty pro definování významnosti důkazů (pokračování)

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

15

16

- **Použít obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

Volitelně můžete:

- Klepněte na volbu **Kritéria**, chcete-li určit nastavení věrohodného intervalu v procentech a numerické metody.
- Klepněte na volbu **Priory**, chcete-li definovat odkazy a konjugovat předchozí nastavení distribuce.
- Klepněte na volbu **Bayesový faktor**, abyste uvedli nastavení Bayesova faktoru.
- Klepnutím na volbu **Zákresy** můžete řídit grafy, které jsou výstupem.

Bayesovská jednosměrná ANOVA: Kritéria

Můžete uvést následující kritéria analýzy pro Bayesovské jednosměrné modely ANOVA.

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody**.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001.

Maximální počet iterací

Zadejte maximální počet iterací metody. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 2000.

Bayesovská jednosměrná ANOVA: Přívozy

Pro parametry regrese a rozptyl chyb můžete určit následující předchozí nastavení distribuce. Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Znaková distribuce**.

Poznámka: Mnoho aplikovaných výzkumných pracovníků může zpochybnit potřebu specifikovat předchozí. Referenční priors minimalizovat obavy, kde předchozí je obecně zahrčen, jak data se zvyšuje. Jsou-li uvedeny předchozí informativní informace, Bayesovské metody mohou tyto informace efektivně

¹⁵ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

¹⁶ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

využívat. Požadavek na uvedení předchozího by neměl být považován za odstrašující prostředek k použití bayesovské analýzy.

Odkaz

Je-li vybrána tato volba, vytvoří referenční analýza objektivní Bayesovské odvozování. Inferenční příkazy závisí pouze na předpokládaném modelu a dostupných datech a předchozí distribuce, která se používá k odvozování, je nejméně informativní. Toto je výchozí nastavení.

Konjugát

Poskytuje volby pro definování konjugace předchozích distribucí. Konjugované priory předpokládají rozdělení kloubu Normal-Inverse-Gamma. Ačkoli konjugované priory nejsou při provádění Bayesovských aktualizací vyžadovány, pomáhají výpočetním procesům.

Převratné odchylky chyb

Parametr tvaru

Zadejte parametr tvaru α_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0.

Parametr měřítka

Zadejte parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jedinou hodnotu, která je větší než 0. Čím větší je parametr měřítka, tím více je rozložení rozloženo.

Převratné hodnoty regresních parametrů

Uvedte střední vektor β_0 pro střední hodnotu skupiny. Počet hodnot musí odpovídat počtu regresních parametrů, včetně termínu zachycení.

Sloupec **Proměnné** se automaticky naplní úrovněmi faktoru. Sloupec **Střední** neobsahuje žádné výchozí hodnoty.

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Rozptyl matice kovariance: σ^2x

Zadejte V_0 hodnoty v dolním trojúhelníku v matici variance-kovariance pro multivariační normální hodnotu před. Všimněte si, že hodnota V_0 musí být polokladná. Musí být uveden pouze dolní trojúhelník tabulky.

Řádky a sloupce jsou automaticky naplněny úrovněmi faktoru. Všechny diagonální hodnoty jsou 1; všechny hodnoty off-diagonální jsou 0.

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Použití matice identit

Je-li tato volba vybrána, použije se matice identit. V dolním trojúhelníku v matici variance-kovariance nemůžete zadat hodnoty V_0 pro vícerozměrnou normální hodnotu před.

Bayesian One-way ANOVA: Faktor pro bayes

Můžete uvést přístup, který se použije k odhadu Bayesova faktoru pro Bayesovské jednosměrné modely ANOVA. Následující volby jsou k dispozici pouze v případě, že je vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody** Bayesovská analýza.

Výpočet

Uvedte přístup k odhadu Bayesových faktorů. Výchozí nastavení je metoda JZS.

Metoda JZS

Je-li vybráno, vyvolá přístup Zellner-Siow. Toto je výchozí nastavení.

Zellnerova metoda

Je-li tato volba vybrána, vyvolá Zellnerovu metodu a vy musíte zadat jednu g předchozí hodnotu > 0 (neexistuje žádná výchozí hodnota).

Hyper-Předchozí metoda

Je-li tato volba vybrána, vyvolá přístup hyper- g a vy musíte zadat parametr tvaru α_0 pro distribuci Inverse-Gamma. Musíte zadat jednu hodnotu > 0 (výchozí hodnota je 3).

Metoda směrovače

Je-li tato volba vybrána, vyvolá přístup směrovače a vy musíte zadat parametr měřítka b_0 pro distribuci Inverse-Gamma. Musíte zadat jednu hodnotu > 0 (výchozí hodnota je 1).

Bayesian jednosměrná ANOVA: Plots

Můžete řídit grafy, které jsou výstupem.

Skupiny zakreslů

Uveďte podskupiny, které se mají zakreslit. Vykreslit pravděpodobnost, předchozí a zadní pro prostředky uvedených skupin. Seznam **Skupiny** je podmnožinou kategorií faktorové proměnné, takže formát by měl být konzistentní s datovým typem faktoru a skutečnými hodnotami.

Odchylka chybových výrazů

Je-li tato volba vybrána, je zakreslena odchylka chyb. Ve výchozím nastavení je výběr nastavení zrušen. Tato volba není k dispozici, je-li jako Bayesovská analýza vybrána volba **Odhadovat faktor bayes**.

Bayesovské loglineární modely

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

Návrh pro testování nezávislosti dvou faktorů vyžaduje dvě kategorické proměnné pro konstrukci pohotovostní tabulky a provádí Bayesovské odvozování na přidružení řádku-sloupce. Můžete odhadnout Bayesovy faktory předpokládáním různých modelů a charakterizovat požadovanou zadní distribuci simulací současného věrohodného intervalu pro podmínky interakce.

1. Z nabídek vyberte:

Analyzovat > Bayesovská statistika > Loglinear Models

2. Ze seznamu **Dostupné proměnné** vyberte jednu proměnnou řádku bez měřítka. Musíte vybrat alespoň jednu neškálovanou proměnnou.

3. Ze seznamu **Dostupné proměnné** vyberte jednu neškálovanou proměnnou sloupce. Musíte vybrat alespoň jednu neškálovanou proměnnou.

4. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.
- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 10. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

17

18

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

Volitelně můžete:

- Klepněte na volbu **Kritéria**, chcete-li určit nastavení věrohodného intervalu v procentech a numerické metody.
- Klepněte na volbu **Bayesový faktor**, abyste uvedli nastavení Bayesova faktoru.
- Klepněte na tlačítko **Tisk** a určete způsob zobrazení obsahu ve výstupních tabulkách.

Bayesian Loglinear Modely: kritéria

Pro Bayesovské logliniové modely můžete určit následující kritéria analýzy.

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Nastavit vlastní počáteční hodnotu

Je-li tato volba vybrána, můžete do pole **Seed** zadat vlastní hodnotu náhodného řetězce. Zadejte náhodnou hodnotu sady náhodných hodnot. Hodnotou musí být kladné celé číslo. Standardně je přiřazena náhodná počáteční hodnota.

Poznámka: Následující volby jsou k dispozici pouze v případě, že je pro volbu **Bayesovská analýza** vybrána volba **Odhadovat bayes faktor** nebo **Použit obě metody**.

Tolerance

Uveďte hodnotu tolerance pro numerické metody. Výchozí nastavení je 0.000001.

Maximální počet iterací

Zadejte maximální počet iterací metody. Hodnotou musí být kladné celé číslo. Výchozí nastavení je 2000.

Ukázky simulované pro Posterior Distribution

Uveďte počet vzorků, které se použijí k nakreslení požadované zadní distribuce. Výchozí hodnota je 10000.

Formát

Vyberte, zda se kategorie zobrazují v pořadí **Vzestupně** nebo **Sestupně**. Výchozí nastavení je vzestupné.

Bayesian Loglinear Modely: Bayesův faktor

Můžete určit model předpokládaný pro pozorovaná data (Poisson, Multinomial nebo Nonparametric). Multinomiální distribuce je standardně nastavena. Následující volby jsou k dispozici pouze v případě, že je vybrána volba **Odhadovat bayes faktor** nebo **Použit obě metody** Bayesovská analýza.

Poissonův model

Je-li zvolena, předpokládá se Poissonův model pro pozorované údaje.

¹⁷ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

¹⁸ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Víceúčelový model

Je-li tato volba vybrána, předpokládá se pro pozorovaná data model Multinomial. Toto je výchozí nastavení.

Pevné okraje

Vyberte **Celkový součet**, **Součet řádků** nebo **Součet sloupců**, abyste určili pevné mezní součty pro pohotovostní tabulku. Výchozí nastavení je **Celkový součet**.

Předchozí distribuce

Uvedte předchozí typ distribuce při odhadu Bayesova faktoru.

Konjugát

Tuto volbu vyberte, chcete-li určit konjugaci před distribucí. Pomocí tabulky **Parametry tvaru** můžete určit parametry tvaru a_{rs} pro distribuci gama. Je-li jako předchozí typ distribuce vybrána volba **Konjugovat**, musíte zadat parametry tvaru.

Je-li zadána jediná hodnota, předpokládá se, že se všechny a_{rs} rovnají této hodnotě. *příkaz:rs* = 1 je výchozí nastavení. Potřebujete-li zadat více než jednu hodnotu, můžete tyto hodnoty oddělit mezerami.

Počet číselných hodnot, které jsou určeny v každém řádku a každém sloupci, se musí shodovat s dimenzí kontingenční tabulky. Všechny zadané hodnoty musí být > 0 .

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Parametr měřítka

Zadejte parametr měřítka b pro distribuci gama. Musíte zadat jedinou hodnotu > 0 .

Směs Dirichletův

Vyberte, chcete-li uvést směs Dirichletův před distribucí.

Vnitřní

Tuto volbu vyberte, chcete-li určit vnitřní předchozí distribuci.

Neparametrický model

Je-li tato volba vybrána, pro pozorovaná data se předpokládá neparametrický model.

Pevné okraje

Vyberte volbu **Součet řádků** nebo **Součet sloupců**, chcete-li určit pevné mezní součty pro kontingenční tabulku. **Součet řádků** je výchozí nastavení.

Předchozí distribuce

Zadejte parametry pro Dirichletův priors. Je-li vybrána volba **Neparametrický model**, je třeba určit parametry **Předchozí distribuce**. Když je uvedena jediná hodnota, předpokládá se, že se všechny λ_s rovnají této hodnotě. $\lambda_s = 1$ je výchozí nastavení. Potřebujete-li zadat více než jednu hodnotu, můžete tyto hodnoty oddělit mezerami. Všechny zadané hodnoty musí být > 0 . Počet zadaných číselných hodnot se musí shodovat s dimenzí řádku nebo sloupce, která není pro kontingenční tabulku pevná.

Klepnutím na tlačítko **Resetovat** vymažete hodnoty.

Bayesian Loglinear Models: Tisknout

Můžete určit způsob zobrazení obsahu ve výstupních tabulkách.

Návrh tabulky

Potlačit tabulku

Je-li tato volba vybrána, není do výstupu zahrnuta tabulka nepředvídaných událostí. Nastavení není standardně povoleno.

Poznámka: Následující nastavení nemají žádný vliv, je-li povoleno nastavení **Potlačit tabulku**.

Statistika

Zadejte statistiku pro testování nezávislosti.

Chi-kvadrát

Tuto volbu vyberte, chcete-li vypočítat statistiku Pearson Chi-Square, stupně volnosti a oboustrannou asymptotickou významnost. Pro 2 x 2 pohotovostní tabulku toto nastavení také vypočítává Yates kontinuitu korigované statistiky, stupně volnosti a související 2-stranná asymptotická významnost. Pro 2 x 2 pohotovostní tabulku, s alespoň jedním očekávaným počtem buněk < 5, toto nastavení také vypočítá oboustrannou a jednostrannou přesnou významnost Fisherova přesného testu.

Poměr pravděpodobnosti

Tuto volbu vyberte, chcete-li vypočítat statistiku testu poměru pravděpodobnosti, stupně volnosti a přidruženou oboustrannou asymptotickou významnost.

Počty

Určete, které typy počítání jsou zahrnuty v tabulce pro nepředvídané události.

Pozorované

Tuto volbu vyberte, chcete-li do tabulky pro nepředvídané události zahrnout sledované počty buněk.

Očekáváno

Tuto volbu vyberte, chcete-li do tabulky pro nepředvídané události zahrnout očekávané počty buněk.

Procentní části

Určete, které procentní typy jsou zahrnuty do tabulky pro nepředvídané události.

Řádek

Tuto volbu vyberte, chcete-li do tabulky pro nepředvídané události zahrnout procentní části řádků.

Sloupec

Tuto volbu vyberte, chcete-li do tabulky pro nepředvídané události zahrnout procenta sloupců.

Celkem

Tuto volbu vyberte, chcete-li do tabulky pro nepředvídané události zahrnout celkové procentní části.

Bayesian jednosměrná opakovaná opatření AMNOVA modely

Tato funkce vyžaduje SPSS Statistika Standard Edition nebo volbu Rozšířená statistika.

V Bayesovských jednosměrných analýzách rozptylu (ANOVA) se předpokládá, že pro každý subjekt existuje jediné měření. Tento předpoklad však není vždy pravdivý. Není neobvyklé, že cílem návrhu studie je prošetřit průměrné odezvy v průběhu více časových bodů nebo podmínek. Bayesovská jednosměrná opakovaná opatření Postup ANOVA měří jeden faktor ze stejného subjektu v každém jednotlivém časovém bodu nebo podmínce a umožňuje, aby subjekty byly překročeny v rámci úrovně. Předpokládá se, že každý subjekt má jediné pozorování pro každý časový bod nebo stav (jako takový není interakce mezi předmětem a léčbou zohledněna).

1. Z nabídek vyberte:

Analýza > Bayesovská statistika > Jednosměrné opakované ukazatele ANOVA

2. Ze seznamu **Dostupné proměnné** vyberte alespoň dvě proměnné **Opakované ukazatele**.

3. Volitelně vyberte jednu proměnnou, která bude sloužit jako regrese **Váha** ze seznamu **Dostupné proměnné**. Pole proměnné **Váha** může být prázdné.

Poznámka: Seznam dostupných proměnných poskytuje všechny proměnné s výjimkou řetězcových proměnných.

4. Vyberte požadovanou **Bayesovskou analýzu**:

- **Charakterizovat následnou distribuci:** Je-li vybráno, Bayesovské odvozování se provede z perspektivy, ke které se přistupuje charakterizací následných distribucí. Můžete prozkoumat okrajové zadní rozdělení parametrů, které jsou předmětem zájmu, tak, že integrujete ostatní parametry obtíží a dále vytvoříte důvěryhodné intervaly pro přímé odvozování. Toto je výchozí nastavení.

- **Faktor odhadu bayes:** Je-li vybrán, představuje odhad Bayesových faktorů (jedna z pozoruhodných metodologií v Bayesovském odvozování) přirozený poměr pro porovnání okrajových pravděpodobností mezi nulovou a alternativní hypotézou.

Tabulka 11. Běžně používané prahové hodnoty pro definování významnosti důkazů

Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů	Faktor bayes	Kategorie důkazů
> 100	Extrémní důkazy pro H1	1-3 (3)	Neoficiální důkazy pro H1	1/30-1/10	Silné důkazy pro H0
30-100	Velmi silný důkaz pro H1	1	Žádné důkazy	1/100-1/30	Velmi silný důkaz pro H0
10-30	Silné důkazy pro H1	1/3-1	Neoficiální důkazy pro H0	1/100	Extrémní důkazy pro H0
3-10	Střední důkazy pro H1	1/10-1/3	Střední důkazy pro H0		

H0: Hypotéza s hodnotou Null

H1: Alternativní hypotéza

19

20

- **Použit obě metody:** Je-li tato volba vybrána, použijí se obě metody odvozování **Rozložení znaků** a **Odhadnout faktor bayes**.

Volitelně můžete:

- Klepněte na volbu **Kritéria**, chcete-li určit nastavení věrohodného intervalu v procentech a numerické metody.
- Klepněte na volbu **Bayesový faktor**, abyste uvedli nastavení Bayesova faktoru.
- Klepněte na volbu **Zákresy**, chcete-li zakreslit následné distribuce prostředků skupiny.

Bayesovská jednosměrná opakovaná opatření ANOVA: kritéria

Můžete určit následující kritéria analýzy pro Bayesovské jednosměrné opakované ukazatele ANOVA.

Procento důvěryhodného intervalu v%

Zadejte úroveň významnosti pro výpočet věrohodných intervalů. Výchozí úroveň je 95%.

Numerická metoda

Uveďte numerickou metodu, která se použije k odhadu integrálu.

Nastavit vlastní počáteční hodnotu

Je-li tato volba vybrána, můžete do pole **Seed** zadat vlastní hodnotu náhodného řetězce. Výchozí hodnota je 2 000 000. Hodnota musí být kladné celé číslo v rozsahu 1 až 2 147 483 647.

Standardně je přiřazena náhodná počáteční hodnota.

Počet ukázek Monte Carlo

Uveďte počet bodů, které jsou vzorkovány pro aproximaci Monte Carlo. Hodnota musí být kladné celé číslo mezi 10^3 a 10^6 . Výchozí hodnota je 30 000.

¹⁹ Lee, M.D., a Wagenmakers, E.-J. V roce 2013. *Bayesovské modelování pro kognitivní vědu: praktický kurz*. Cambridge University Press.

²⁰ Jeffreys, H. 1961. *Teorie pravděpodobnosti*. Oxford University Press.

Bayesovská jednosměrná opakovaná opatření ANOVA: Bayes faktor

Můžete uvést přístup, který se použije k odhadu Bayesova faktoru pro Bayesovské jednosměrné opakující se ukazatele modelů ANOVA. Následující volby jsou k dispozici pouze v případě, že je vybrána volba **Odhadovat bayes faktor** nebo **Použít obě metody** Bayesovská analýza.

Bayesovská informační kritéria (BIC)

Používá rozšíření aproximace BIC o návrh opakovaných opatření pro odhad Bayesových faktorů. Nastavení odvozuje efektivní účtování velikosti vzorku pro opakovanou korelaci opatření a navrhuje lepší sankční termín při odhadu BIC pro výběr mezi dvěma konkurenčními modely. Toto je výchozí nastavení.

Rouderova smíšená konstrukce

Používá vícerozměrné generalizace Cauchyho distribuce jako předchozí pro standardizovanou velikost efektu a neinformativní předchozí pro rozptyl.

Poznámka: Je-li vybrána tato volba, budou globální nastavení váhy frekvence a regresní váha ignorovány.

Bayesovská jednosměrná opakovaná opatření ANOVA: Plots

Můžete řídit grafy, které jsou výstupem, abyste ilustrovali následné distribuce středních dat skupiny. V tabulce jsou uvedeny všechny proměnné, které jsou vybrány jako opakované ukazatele v dialogovém okně Proměnné. Vyberte proměnné opakovaných ukazatelů, které se mají zakreslit

Regrese hřebenu jádra

Regrese hřebenu jádra je procedura rozšíření, která používá třídu Python **sklearn.kernel_ridge.KernelRidge** k odhadu regresních modelů hřebenu jádra. Regresní modely hřebenu jádra jsou neparаметrické regresní modely, které jsou schopny modelovat lineární a nelineární vztahy mezi proměnnými prediktoru a výsledky. Výsledky mohou být velmi citlivé na volby hyperparametrů modelu. Regrese ridge jádra usnadňuje výběr hodnot hyperparametrů prostřednictvím křížového ověření k-fold na zadaných mřížkách hodnot pomocí třídy **sklearn.model_selection.GridSearchCV**.

Příklad

Statistika

Additive_CHI2, CHI2, Cosine, Laplaceova, Lineární, Polynomial, RBF, Sigmoid, Alpha, Gamma, Coef0, Stupeň, křížová validace, pozorovaná versus předpovězená, zbytky versus předpovězená, koeficienty duální hmotnosti, koeficienty hmotnosti prostoru jádra.

Aspekty dat

Data

- Můžete zadat libovolnou nebo všechny z osmi různých funkcí jádra.
- Vybraná funkce jádra určuje, které hyperparametry jsou aktivní.
- Hyperparametry zahrnují alfa pro regularizaci ridge, které jsou společné pro všechna jádra plus až tři další hyperparametry pro každou specifickou funkci jádra.
- Pokud je zadáno více podpříkazů jádra nebo je zadána více než jedna hodnota pro libovolný parametr, provede se vyhledávání v mřížce s křížovým ověřením pro vyhodnocení modelů a vybere se nejvhodnější model založený na zadržených datech.
- Rozšíření přijímá proměnné rozdělení z procedury Rozdělit soubor a váhy pomocí procedury Případy váhy.
- Jsou-li zahrnuty váhy, jsou použity při vytváření hodnot ve všech analýzách. Kvůli omezením v metodě skóre ve třídě **sklearn.model_selection.GridSearchCV** nejsou vyhodnocení křížového ověření, která se používají pro výběr modelu, vážená.

Předpoklady

Získání regrese hřebenu jádra

1. Z nabídek vyberte:

Analyzovat > regrese > ridge jádra ...

2. Vyberte proměnnou **Závislé**.

3. Vyberte jednu nebo více **nezávislých** proměnných.

4. Výchozí nastavení **Jediný model** se používá, když je pro každý parametr funkce jádra zadána pouze jedna hodnota. Je-li vybráno nastavení **Jediný model**, nemůžete určit další funkce **Jádro** a váhy jsou plně použity v rámci analýzy, vyhodnocení a přidělení skóre výsledků. Můžete také použít ovládací prvky se šipkou nahoru a dolů k přeuspořádání funkcí jádra.

Volitelně vyberte volbu **Výběr modelu** ze seznamu **Režim**.

Je-li v seznamu **Režim** vybrána volba **Výběr modelu**, můžete do seznamu **Kernel (y)** přidat více funkcí jádra.

a. Klepnutím na tlačítko Přidat ovládací prvek (+) zahrnete další funkce jádra.

b. Klepnutím na prázdnou buňku ve sloupci **Kernel** vyberte funkci jádra.

c. Poklepejte na libovolnou buňku řádku funkce jádra a zadejte hodnoty parametrů funkce jádra pro odpovídající sloupec (**Alpha, Gamma, Coef0, Degree**). Více informací naleznete v části [“parametry jádra”](#) na stránce 130. Výchozí parametry ladění funkce jádra jsou uvedeny níže.

Additive_CHI2

ALPHA=1 GAMMA=1

CHI2

ALPHA=1 GAMMA=1

Kosinus

ALPHA=1

Laplacian

ALPHA=1 GAMMA=1/p

Lineární

Výchozí funkce jádra. ALPHA=1

Polynomní

ALPHA=1 GAMMA=1/p COEF0=1 DEGREE=3

RBF

ALPHA=1 GAMMA=1/p

ID podpisu

ALPHA=1 GAMMA=1/p COEF0=1

Poznámka: Je-li pro libovolný parametr funkce jádra zadána více než jedna hodnota, provede se vyhledávání v mřížce s křížovým ověřením pro vyhodnocení modelů a vybere se nejvhodnější model založený na zadržených datech.

5. Volitelně klepněte na volbu **Volby** a zadejte počet záhybů křížového ověřením, volby zobrazení, nastavení výkresu a položky k uložení. Další informace viz téma [“Regrese hřebenu jádra: Volby”](#) na stránce 131.

6. Klepněte na tlačítko **OK**.

parametry jádra

Dialogové okno **Parametry jádra** poskytuje volby pro určení hodnot jednotlivých parametrů funkce jádra a pro určení, že výběr modelu se provádí pomocí vyhledávání v mřížce nad kombinacemi jader a zadanými hodnotami parametrů mřížky.

Zadat jednotlivé parametry

Povolte nastavení pro určení hodnot pro vybraný parametr funkce jádra.

- Zadejte hodnotu a klepněte na tlačítko **Přidat** , abyste zahrnuli hodnotu do parametru funkce jádra.
- Chcete-li aktualizovat hodnotu, vyberte hodnotu parametru a klepněte na tlačítko **Změnit** .
- Vyberte hodnotu parametru a klepněte na tlačítko **Odebrat** , abyste hodnotu odstranili.

Určit parametry mřížky

Povolte nastavení, chcete-li určit, že výběr modelu se provádí pomocí vyhledávání v mřížce nad kombinacemi jader a zadanými hodnotami parametrů mřížky.

Regrese hřebenu jádra: Volby

Dialogové okno **Zákresy** poskytuje volby pro určení počtu záhybů křížového ověření, voleb zobrazení, nastavení výkresu a položek k uložení.

Počet záhybů křížového ověření

Počet rozdělení nebo záhybů v křížovém ověření s hledáním mřížky pro výběr modelu. Zadejte celočíselnou hodnotu větší než 1. Výchozí hodnota je 5. Nastavení je k dispozici pouze v případě, že je jako **Režim** v primárním dialogovém okně **Regrese hřebenu jádra** vybrána volba **Výběr modelu** .

Zobrazit

Poskytuje volby pro určení, který výstup se má zobrazit, když je v platnosti křížové ověření.

Nejlepší

Výchozí nastavení zobrazuje pouze základní výsledky pro vybraný nejlepší model.

Porovnat

Zobrazí základní výsledky pro všechny vyhodnocené modely.

Porovnání modelů a záhybů

Zobrazí úplné výsledky pro každé rozdělení nebo přeložení pro každý vyhodnocený model.

Graf

Poskytuje volby pro určení grafů pozorovaných nebo zbytkových hodnot oproti předpovězených hodnot.

Pozorované vs. předpokládané

Zobrazí bodový graf pozorovaných versus předpokládaných hodnot pro určený nebo nejlepší model.

Rezidua vs. předpovězeno

Zobrazí bodový graf zbytkových chyb oproti předpovězených hodnot pro uvedený nebo nejlepší model.

Uložit

Tabulka poskytuje volby pro určení proměnných, které se mají uložit do aktivní datové sady.

Předpokládané hodnoty

Uloží předpovězené hodnoty ze zadaného nebo nejlepšího modelu do aktivní datové sady. Lze zahrnout volitelný název proměnné.

Zbytkové chyby

Uloží rezidua z uvedených nebo nejlepších předpovědí modelu do aktivní datové sady. Lze zahrnout volitelný název proměnné.

Duální koeficienty

Uloží duální koeficienty hmotnosti nebo koeficienty váhy prostoru jádra ze zadaného modelu do aktivní datové sady. Lze zahrnout volitelný název proměnné. Nastavení není k dispozici, je-li jako **Režim** v primárním dialogovém okně **Regrese hřebenu jádra** vybrána volba **Výběr modelu** .

Poznámky

Uvedené informace se týkají produktů a služeb nabízených v USA. Společnost IBM může tento materiál poskytovat v jiných jazycích. K povolení přístupu však může být vyžadováno vlastnictví kopie produktu nebo verze produktu v tomto jazyce.

Společnost IBM nemusí nabízet produkty, služby nebo funkce uvedené v tomto dokumentu v jiných zemích. Informace o produktech a službách, které jsou aktuálně k dispozici ve vaší oblasti, získáte od lokálního zástupce společnosti IBM. Odkazy na produkty, programy nebo služby společnosti IBM neuvádí ani neimplikují, že lze použít pouze daný produkt, program nebo službu společnosti IBM. Lze použít libovolný funkčně ekvivalentní produkt, program nebo službu neporušující práva duševního vlastnictví společnosti IBM. Vyhodnocení a ověření funkčnosti produktů, programů nebo služeb, které nepatří společnosti IBM, je však zodpovědností uživatele.

Společnost IBM může vlastnit patenty nebo nevyřízené žádosti o patenty zahrnující předměty popsané v tomto dokumentu. Vlastnictví tohoto dokumentu neposkytuje licenci k těmto patentům. Dotazy na licence můžete písemně odeslat na následující adresu:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
nám*

S dotazy na licence týkající se dvoubajtových informací (DBCS) se obraťte na oddělení intelektuálního vlastnictví společnosti IBM v dané zemi, nebo je odešlete písemně na následující adresu:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

SPOLEČNOST INTERNATIONAL BUSINESS MACHINES CORPORATION TUTO PUBLIKACI POSKYTUJE TAK, JAK JE (AS-IS), BEZ JAKÝCHKOLIV ZÁRUK, VYJÁDŘENÝCH VÝSLOVNĚ NEBO VYPLÝVAJÍCÍCH Z OKOLNOSTÍ, VČETNĚ, A TO ZEJMÉNA, ZÁRUK NEPORUŠENÍ PRÁV TŘETÍCH STRAN, PRODEJNOSTI NEBO VHODNOSTI PRO URČITÝ ÚČEL VYPLÝVAJÍCÍCH Z OKOLNOSTÍ. Některé jurisdikce u určitých transakcí nepřipouštějí vyloučení záruk výslovně vyjádřených nebo vyplývajících z okolností, a proto se na vás výše uvedené omezení nemusí vztahovat.

Uvedené údaje mohou obsahovat technické nepřesnosti nebo typografické chyby. Údaje zde uvedené jsou pravidelně upravovány a tyto změny budou zahrnuty v nových vydáních této publikace. Společnost IBM může kdykoli bez upozornění provádět vylepšení nebo změny v produktech či programech popsaných v této publikaci.

Jakékoli odkazy na jiný web než web společnosti IBM jsou v této příručce uvedeny pouze pro doplnění a v žádném případě neslouží jako oficiální uznání tohoto webu. Materiály uvedené na těchto webových stránkách nejsou součástí materiálů pro tento produkt IBM a použití uvedených stránek je pouze na vlastní nebezpečí.

Společnost IBM může použít nebo distribuovat jakékoli informace, které jí poskytnete, libovolným způsobem, který společnost považuje za odpovídající, bez vyžádání vašeho svolení.

Vlastníci licence k tomuto programu, kteří chtějí získat informace o možnostech (i) výměny informací s nezávisle vytvořenými programy a jinými programy (včetně tohoto) a (ii) oboustranného využití vyměňovaných informací, mohou kontaktovat informační středisko na adrese:

*IBM Director of Licensing
IBM Corporation*

North Castle Drive, MD-NC119
Armonk, NY 10504-1785
nám

Poskytnutí takových informací může být podmíněno dodržáním určitých podmínek a požadavků zahrnujících v některých případech uhrazení stanoveného poplatku.

Licencovaný program popsáný v tomto dokumentu a veškerý licencovaný materiál k němu dostupný jsou společností IBM poskytovány na základě podmínek uvedených ve smlouvách IBM Customer Agreement, IBM International Program License Agreement nebo v jiné ekvivalentní smlouvě.

Data o výkonu a příklady klienta jsou uváděny pouze pro účely ilustrace. Skutečné výsledky výkonu se mohou lišit v závislosti na specifických konfiguracích a provozních podmínkách.

Informace ohledně produktů jiných společností než IBM byly získány od dodavatelů těchto produktů, z jimi publikovaných prohlášení či jiných veřejně přístupných zdrojů. IBM všechny tyto produkty netestovala a nemůže potvrdit věrohodnost údajů o výkonu, kompatibilitě či jiných tvrzení týkajících se těchto produktů nepocházejících od IBM. Další případné otázky ve věci produktů, které nejsou dodávány společností IBM, směřujte přímo na dodavatele těchto produktů.

Tvrzení týkající se budoucího směru vývoje nebo záměrů společnosti IBM se mohou bez upozornění změnit nebo mohou být zrušena a reprezentují pouze cíle a plány společnosti.

Tyto údaje obsahují příklady dat a sestav používaných v běžných obchodních operacích. Aby byla představa úplná, používají se v příkladech jména osob, společností, značek a produktů. Všechna tato jména jsou fiktivní a jejich podobnost se jmény skutečných osob nebo organizací je zcela náhodná.

COPYRIGHT - LICENCE:

Tyto informace obsahují ukázkové aplikační programy ve zdrojovém jazyku a ilustrují různé programovací techniky na různých operačních platformách. Tyto ukázkové programy můžete bez závazků vůči společnosti IBM jakýmkoli způsobem kopírovat, měnit a distribuovat za účelem vývoje, používání, odbytu či distribuce aplikačních programů odpovídajících rozhraní API pro operační platformu, pro kterou byly ukázkové programy napsány. Tyto příklady nebyly důkladně testovány ve všech podmínkách. Společnost IBM proto nemůže zaručit spolehlivost, upotřebitelnost nebo funkčnost těchto programů. Ukázkové programy jsou poskytovány "JAK JSOU", bez záruky jakéhokoli druhu. IBM nenes odpovědnost za žádné škody vzniklé ve spojení s Vaším užíváním ukázkových programů.

Jakákoli kopie nebo část těchto ukázkových programů nebo jakékoli odvozené dílo musí obsahovat následující poznámku o autorských právech:

© Copyright IBM Corp. 2021. Části tohoto kódu jsou odvozeny ze vzorových programů společnosti IBM Corp. Vzorové programy.

© Copyright IBM Corp. 1989-2021. Všechna práva vyhrazena.

Ochranné známky

IBM, logo IBM a ibm.com jsou ochranné známky nebo registrované ochranné známky společnosti International Business Machines Corp., registrované v mnoha jurisdikcích po celém světě. Ostatní názvy produktů a služeb mohou být ochrannými známkami společnosti IBM nebo jiných společností. Aktuální seznam ochranných známek IBM je k dispozici na webu na stránce "Copyright and trademark information" (Copyright a ochranné známky) na adrese www.ibm.com/legal/copytrade.shtml.

Adobe, logo Adobe, PostScript a logo PostScript jsou buď registrované ochranné známky, nebo ochranné známky společnosti Adobe Systems Incorporated ve Spojených státech anebo v dalších zemích.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium a Pentium jsou ochranné známky nebo registrované ochranné známky společnosti Intel Corporation nebo jejich dceřiných společností ve Spojených státech a případně v dalších jiných zemích.

Linux je registrovaná ochranná známka Linuse Torvaldse ve Spojených státech a případně v dalších jiných zemích.

Microsoft, Windows, Windows NT a logo Windows jsou ochranné známky společnosti Microsoft Corporation ve Spojených státech a případně v dalších jiných zemích.

UNIX je registrovaná ochranná známka společnosti The Open Group ve Spojených státech a případně v dalších jiných zemích.

Java a všechny ochranné známky a loga založená na termínu Java jsou ochranné známky nebo registrované ochranné známky společnosti Oracle a případně jejích příbuzných společností.

Rejstřík

A

- analýza kovariance
 - v GLM Multivariate [1](#)
- analýza přežití
 - Časově závislá regrese Cox [92](#)
 - v Cox Regression [89](#)
 - v Kaplan-Meieru [87](#)
 - v regresi hřebenu jádra [129](#)
 - v životních tabulkách [68](#)
- analýza rozptylu
 - v komponentách Variance [19](#)
 - v zobecněných lineárních smíšených modelech [47](#)
- ANOVA
 - v GLM Multivariate [1](#)
 - v GLM opakovaných ukazatelích [9](#)

B

- binomická distribuce
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- Bodování Fisher
 - v lineárních smíšených modelech [24](#)
- Bonferroni
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)

C

- cenzurované případy
 - v Cox Regression [89](#)
 - v Kaplan-Meieru [87](#)
 - v životních tabulkách [68](#)
- Coxova regrese
 - časově závislé kovariáty [92](#)
 - definovat událost [92](#)
 - DfBeta(s) [91](#)
 - dílčí zbytkové chyby [91](#)
 - dodatečné funkce příkazu [92](#)
 - funkce přežití [91](#)
 - funkce rizika [91](#)
 - funkce úrovně baseline [91](#)
 - Iterace [91](#)
 - kategorické kovariáty [90](#)
 - kontrasty [90](#)
 - kovariáty [89](#)
 - postupné zadávání a odstraňování [91](#)
 - proměnná stavu přežití [92](#)
 - příklad [89](#)
 - řetězcové kovariáty [90](#)
 - statistiky [89](#), [91](#)
 - ukládání nových proměnných [91](#)
 - výkresy [90](#)

D

- Deskriptivní statistika
 - v lineárních smíšených modelech [25](#)
 - v zobecněných lineárních modelech [33](#)
 - ve zobecněných odhadovacích rovnicích [44](#)
- distribuce gama
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- Distribuce tweedie
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- Dobrá kondice
 - v zobecněných lineárních modelech [33](#)
 - ve zobecněných odhadovacích rovnicích [44](#)
- doplňková funkce log-log link
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- Duncanův test s více rozsahy
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)
- Dunnettovo C
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)
- Dunnettův t test
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)

F

- faktory
 - v GLM opakovaných ukazatelích [11](#)
- frekvence
 - v Logineární analýza výběru modelu [62](#)
- funkce elektrického propojení kurzů
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce logit link
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce odkazu na doplněk protokolu
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce probit link
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce propojení
 - zobecněné lineární smíšené modely [50](#)
- funkce propojení identity
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce propojení napájení
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)
- funkce propojení protokolu
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadových rovnicích [39](#)

funkce přežití
v životních tabulkách [68](#)

G

Gabrielův test porovnávání po dvojicích
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

Gehanův test
v životních tabulkách [69](#)

generování třídy
v Logineární analýza výběru modelu [61](#)

GLM
ukládání matic [6](#)
ukládání proměnných [6](#)

GLOR
v obecné logliniové analýze [62](#)

H

Hessovská konvergence
v zobecněných lineárních modelech [32](#)
ve zobecněných odhadovacích rovnicích [42](#)

hierarchická dekompozice
v komponentách Variance [19](#)

hierarchické logliniové modely [60](#)

hierarchické modely
zobecněné lineární smíšené modely [47](#)

historie iterací
v lineárních smíšených modelech [24](#)
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)

hodnoty pákového efektu
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
v zobecněných lineárních modelech [35](#)

Hochbergův GT2
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

Hry a Howelův porovnávací test po dvojicích
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

Hřeben jádra
Alfa [129](#)
coef0 [129](#)
DEGREE [129](#)
Gama [129](#)
jeden model [129](#)
Výběr modelu [129](#)

I

informace na úrovni faktoru
v lineárních smíšených modelech [25](#)

Informace o modelu
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)

intervaly spolehlivosti
v lineárních smíšených modelech [25](#)
v Logit Loglinear Analysis (loglineární analýza) [67](#)
v obecné logliniové analýze [64](#)

inverzní Gaussovo rozdělení
v zobecněných lineárních modelech [27](#)

inverzní Gaussovo rozdělení (*pokračování*)
v zobecněných odhadovacích rovnicích [39](#)

Iterace

v Logineární analýza výběru modelu [62](#)
v zobecněných lineárních modelech [32](#)
ve zobecněných odhadovacích rovnicích [42](#)

K

Kaplan-Meier
definování událostí [88](#)
dodatečné funkce příkazu [89](#)
kvartily [88](#)
lineární trend pro úroveň faktoru [88](#)
porovnání úrovní faktorů [88](#)
proměnné stavu přežití [88](#)
příklad [87](#)
statistiky [87](#), [88](#)
střední a střední doba přežití [88](#)
tabulky přežití [88](#)
ukládání nových proměnných [88](#)
výkresy [88](#)

kompletní faktoriální modely
v GLM opakovaných ukazatelích [11](#)
v komponentách Variance [18](#)

Komponenty odchylky
dodatečné funkce příkazu [20](#)
model [18](#)
Ukládání výsledků [20](#)
volby [19](#)

kontrasty
v Cox Regression [90](#)
v Logit Loglinear Analysis (loglineární analýza) [65](#)
v obecné logliniové analýze [62](#)

konvergence parametrů
v lineárních smíšených modelech [24](#)
v zobecněných lineárních modelech [32](#)
ve zobecněných odhadovacích rovnicích [42](#)

korelační matice
v lineárních smíšených modelech [25](#)
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)

kovariance matice
v GLM [6](#)
v lineárních smíšených modelech [25](#)
v zobecněných lineárních modelech [32](#), [33](#)
ve zobecněných odhadovacích rovnicích [42](#), [44](#)

kovariance struktury
v lineárních smíšených modelech [96](#)

kovariáty
v Cox Regression [90](#)

křížové abulace
v Logineární analýza výběru modelu [60](#)

kumulativní doplňková funkce log-log link
v zobecněných lineárních modelech [27](#)
v zobecněných odhadovacích rovnicích [39](#)

kumulativní funkce logit link
v zobecněných lineárních modelech [27](#)
v zobecněných odhadovacích rovnicích [39](#)

kumulativní funkce odkazu Cauchit
v zobecněných lineárních modelech [27](#)
v zobecněných odhadovacích rovnicích [39](#)

kumulativní funkce probit link
v zobecněných lineárních modelech [27](#)

kumulativní funkce probit link (*pokračování*)
v zobecněných odhadových rovnicích [39](#)
kumulativní negativní funkce log-log link
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)

L

Lagrangeova multiplikační zkouška
v zobecněných lineárních modelech [33](#)
Lineární smíšené modely
dodatečné funkce příkazu [27](#)
kovariance struktura [96](#)
kritéria odhadu [24](#)
model [25](#)
náhodné účinky [24](#), [27](#)
odhadovaný mezní střední počet [26](#)
pevné účinky [23](#)
podmínky interakce [23](#)
podmínky sestavení [23](#)
uložit proměnné [26](#)
log-pravděpodobnost konvergence
v lineárních smíšených modelech [24](#)
v zobecněných lineárních modelech [32](#)
ve zobecněných odhadovacích rovnicích [42](#)
logistická regrese
zobecněné lineární smíšené modely [47](#)
Logit Loglineární analýza
faktory [65](#)
intervaly spolehlivosti [67](#)
kontrasty [65](#)
kovariáty buněk [65](#)
Podmínky [67](#)
předpovězené hodnoty [67](#)
rezidentní chyby [67](#)
rozdělení počtu buněk [65](#)
specifikace modelu [66](#)
struktury buněk [65](#)
ukládání proměnných [67](#)
volby zobrazení [67](#)
výkresy [67](#)
loglineární analýza
Logit Loglineární analýza [65](#)
Obecná logliniová analýza [62](#)
v zobecněných lineárních smíšených modelech [47](#)

M

matice kontrastních koeficientů
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)
matice kovariance parametrů
v lineárních smíšených modelech [25](#)
Matice L
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)
matrice kovariance s náhodným efektem
v lineárních smíšených modelech [25](#)
MINQUE (nečinnost)
v komponentách Variance [19](#)
míra rizika
v životních tabulkách [68](#)
model proporcionálních rizik

model proporcionálních rizik (*pokračování*)
v Cox Regression [89](#)
multinomiální distribuce
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)
multinomiální logistická regrese
zobecněné lineární smíšené modely [47](#)
multinomiální logitové modely [65](#)
multivariační ANOVA [1](#)

N

náhodné účinky
v lineárních smíšených modelech [24](#), [27](#)
náhodný efekt priors
v komponentách Variance [19](#)
nasyčené modely
v Logineární analýza výběru modelu [61](#)
negativní binomická distribuce
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)
negativní binomická funkce propojení
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)
negativní funkce log-log link
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)
nejméně významný rozdíl
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
nestandardizované zbytkové chyby
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
Newton-Raphsonova metoda
v Logit Loglinear Analysis (loglineární analýza) [65](#)
v obecné logliniové analýze [62](#)
normální pravděpodobnostní grafy
v Logineární analýza výběru modelu [62](#)
normální rozdělení
v zobecněných lineárních modelech [27](#)
v zobecněných odhadových rovnicích [39](#)
Nový muž-Keuls
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

O

Obecná logliniová analýza
dodatečné funkce příkazu [65](#)
faktory [62](#)
intervaly spolehlivosti [64](#)
kontrasty [62](#)
kovariáty buněk [62](#)
Podmínky [64](#)
rezidentní chyby [64](#)
rozdělení počtu buněk [62](#)
specifikace modelu [63](#)
struktury buněk [62](#)
ukládání proměnných [64](#)
uložení předpovězených hodnot [64](#)
volby zobrazení [64](#)
výkresy [64](#)
obecná odhadnutelná funkce

- obecná odhadnutelná funkce (*pokračování*)
 - ve zobecněných odhadovacích rovnicích [44](#)
- obecné odhadnutelné funkce
 - v zobecněných lineárních modelech [33](#)
- obecný lineární model
 - zobecněné lineární smíšené modely [47](#)
- oddělení
 - v zobecněných lineárních modelech [32](#)
 - ve zobecněných odhadovacích rovnicích [42](#)
- odhad maximální pravděpodobnosti
 - v komponentách Variance [19](#)
- odhad omezené maximální pravděpodobnosti
 - v komponentách Variance [19](#)
- odhadovaný mezní střední počet
 - v lineárních smíšených modelech [26](#)
 - v zobecněných lineárních modelech [34](#)
 - ve zobecněných odhadovacích rovnicích [45](#)
- odhady parametrů
 - v lineárních smíšených modelech [25](#)
 - v Logineární analýza výběru modelu [62](#)
 - v Logit Loglinear Analysis (loglineární analýza) [65](#)
 - v obecné logliniové analýze [62](#)
 - v zobecněných lineárních modelech [33](#)
 - ve zobecněných odhadovacích rovnicích [44](#)
- odchylky zbytkových hodnot
 - v zobecněných lineárních modelech [35](#)
- odstraněné rezidua
 - v GLM [6](#)
 - v GLM opakovaných ukazatelích [15](#)
- Opakovaná opatření GLM
 - definovat faktory [11](#)
 - dodatečné funkce příkazu [16](#)
 - model [11](#)
 - post hoc testy [13](#)
 - profilové grafy [13](#)
 - ukládání proměnných [15](#)

P

- parametr měřítka
 - v zobecněných lineárních modelech [32](#)
 - ve zobecněných odhadovacích rovnicích [42](#)
- Parametrické časové modely urychleného selhání
 - analýza [70](#)
 - exportovat [75](#)
 - model [71](#)
 - odhad [72](#)
 - Podmínky [71](#)
 - předvídat [74](#)
 - tisk [73](#)
 - zakreslit [74](#)
- Parametrické modely křehkosti
 - proměnné stavu přežití [82](#)
- Parametrické sdílené modely křehkosti
 - analýza [76](#)
 - exportovat [82](#)
 - model [78](#)
 - odhad [79](#)
 - Podmínky [77](#)
 - předvídat [80](#)
 - tisk [80](#)
 - zakreslit [81](#)
- Pearsonovy pozůstalí
 - v zobecněných lineárních modelech [35](#)

- Pearsonovy pozůstalí (*pokračování*)
 - ve zobecněných odhadovacích rovnicích [46](#)
- pevné predikované hodnoty
 - v lineárních smíšených modelech [26](#)
- pevné účinky
 - v lineárních smíšených modelech [23](#)
- podélné modely
 - zobecněné lineární smíšené modely [47](#)
- podmínky interakce
 - v lineárních smíšených modelech [23](#)
- podmínky sestavení [3](#), [11](#), [18](#), [62](#), [64](#), [67](#)
- Poissonova regrese
 - v obecné logliniové analýze [62](#)
 - zobecněné lineární smíšené modely [47](#)
- Poissonovo rozdělení
 - v zobecněných lineárních modelech [27](#)
 - v zobecněných odhadovacích rovnicích [39](#)
- poměr kurzů
 - v obecné logliniové analýze [62](#)
- pravděpodobná rezidua
 - v zobecněných lineárních modelech [35](#)
- probační analýza
 - zobecněné lineární smíšené modely [47](#)
- profilové grafy
 - v GLM Multivariate [4](#)
 - v GLM opakovaných ukazatelích [13](#)
- proměnné opakovaných ukazatelů
 - v lineárních smíšených modelech [21](#)
- proměnné předmětů
 - v lineárních smíšených modelech [21](#)
- předpovězené hodnoty
 - v lineárních smíšených modelech [26](#)
 - v Logit Loglinear Analysis (loglineární analýza) [67](#)
 - v obecné logliniové analýze [64](#)
- Přežívání AFT
 - Dialogové okno přežití-proměnné kategorie [75](#)
- přídělení skóre
 - v lineárních smíšených modelech [24](#)

R

- R-E-G-W F
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)
- R-E-G-W Q
 - v GLM Multivariate [5](#)
 - v GLM opakovaných ukazatelích [13](#)
- referenční kategorie
 - v zobecněných lineárních modelech [30](#)
 - ve zobecněných odhadovacích rovnicích [41](#), [42](#)
- Regrese hřebenu jádra
 - obrazovka [131](#)
 - parametry [130](#)
 - parametry mřížky [130](#)
 - uložit [131](#)
 - výkresy [131](#)
 - Záhyby křížové validace [131](#)
- rezidentní chyby
 - v lineárních smíšených modelech [26](#)
 - v Logineární analýza výběru modelu [62](#)
 - v Logit Loglinear Analysis (loglineární analýza) [67](#)
 - v obecné logliniové analýze [64](#)
 - v zobecněných lineárních modelech [35](#)
 - ve zobecněných odhadovacích rovnicích [46](#)

Ryan-Einot-Gabriel-Welsch s více rozsahy
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Ryan-Einot-Gabriel-Welsch více F
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Rybářský LSD
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

Ř

řetězcové kovariáty
v Cox Regression [90](#)

S

segmentované kovariáty závislé na čase
v Cox Regression [92](#)
Scheffé zkouška
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Sidakův t test
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
smíšené modely
lineární [20](#)
zobecněné lineární smíšené modely [47](#)
součet čtverců [3](#), [12](#)
součty čtverců
v komponentách Variance [19](#)
v lineárních smíšených modelech [23](#)
souhrn zpracování případů
v zobecněných lineárních modelech [33](#)
ve zobecněných odhadovacích rovnicích [44](#)
standardizované zbytkové chyby
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
standardní chyba
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
Statistika wald
v Logit Loglinear Analysis (loglineární analýza) [65](#)
v obecné logliniové analýze [62](#)
Studentka-Newman-Keuls
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
stupňové-půlení
v lineárních smíšených modelech [24](#)
v zobecněných lineárních modelech [32](#)
ve zobecněných odhadovacích rovnicích [42](#)

T

T3 společnosti Dunnett
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
tabulky pro nepředvídané události
v obecné logliniové analýze [62](#)
Tamhanův T2
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Tarone-Ware test

Tarone-Ware test (*pokračování*)
v Kaplan-Meieru [88](#)
test ohodnocení důležitosti protokolu
v Kaplan-Meieru [88](#)
test parametrů kovariance
v lineárních smíšených modelech [25](#)
tolerance singularity
v lineárních smíšených modelech [24](#)
Tukey je b test
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Tukey je upřímně významný rozdíl
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)

V

vážené predikované hodnoty
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
vícerozměrná regrese [1](#)
víceúrovňové modely
zobecněné lineární smíšené modely [47](#)
vícevariantní GLM [1](#)
Vícevariantní GLM
faktory [1](#)
kovariáty [1](#)
post hoc testy [5](#)
profilové grafy [4](#)
Závislá proměnná [1](#)
vlastní modely
v GLM opakovaných ukazatelích [11](#)
v komponentách Variance [18](#)
v Loglineární analýza výběru modelu [61](#)
vnořené výrazy
v lineárních smíšených modelech [23](#)
v zobecněných lineárních modelech [31](#)
ve zobecněných odhadovacích rovnicích [42](#)
Výběr modelu Loglinear Analysis
definování rozsahů faktorů [61](#)
dodatečné funkce příkazu [62](#)
modely [61](#)
volby [62](#)
výkresy
v Logit Loglinear Analysis (loglineární analýza) [67](#)
v obecné logliniové analýze [64](#)
Vzdálenost kuchaře
v GLM [6](#)
v GLM opakovaných ukazatelích [15](#)
v zobecněných lineárních modelech [35](#)

W

Waller-Duncanův test
v GLM Multivariate [5](#)
v GLM opakovaných ukazatelích [13](#)
Wilcoxonův test
v životních tabulkách [69](#)

Z

zbytková kovarianční matice
v lineárních smíšených modelech [25](#)

Zkouška Breslow
v Kaplan-Meieru [88](#)

Zobecněné lineární modely
export modelu [36](#)
funkce propojení [27](#)
kritéria odhadu [32](#)
odezva [30](#)
odhadovaný mezní střední počet [34](#)
Počáteční hodnoty [33](#)
prediktory [30](#)
referenční kategorie pro binární odezvu [30](#)
Rozdělení [27](#)
specifikace modelu [31](#)
statistiky [33](#)
typy modelů [27](#)
uložit proměnné do aktivní datové sady [35](#)
volby pro kategoriální faktory [31](#)

zobecněné lineární smíšené modely
cílová distribuce [50](#)
datová struktura [57](#)
export modelu [56](#), [57](#)
funkce propojení [50](#)
náhodné účinky [53](#)
náhodný blok efektů [53](#)
náhodný efekt kovariance [59](#)
odhadnuté prostředky [60](#)
odhadovaný mezní střední počet [55](#)
parametry kovariance [59](#)
pevné koeficienty [58](#)
pevné účinky [52](#), [58](#)
posunutí [54](#)
předpovězeno podle pozorovaného [58](#)
souhrn modelu [57](#)
tabulka klasifikací [58](#)
Uložit pole [56](#), [57](#)
váha analýzy [54](#)
vlastní výrazy [52](#)
Zobrazení modelu [57](#)

Zobecněné odhadovací rovnice
export modelu [46](#)
kritéria odhadu [42](#)
odezva [40](#)
odhadovaný mezní střední počet [45](#)
Počáteční hodnoty [43](#)
prediktory [41](#)
referenční kategorie pro binární odezvu [41](#)
specifikace modelu [42](#)
statistiky [44](#)
typ modelu [39](#)
uložit proměnné do aktivní datové sady [46](#)
volby pro kategoriální faktory [42](#)

zobecněný lineární model
v zobecněných lineárních smíšených modelech [47](#)

zobecněný poměr log-odds
v obecné logliniové analýze [62](#)

Zobrazení modelu
v zobecněných lineárních smíšených modelech [57](#)

zpětná eliminace
v Logineární analýza výběru modelu [60](#)

Ž

Životní tabulky
dodatečné funkce příkazu [70](#)

Životní tabulky (*pokračování*)
funkce přežití [68](#)
míra rizika [68](#)
porovnání úrovní faktorů [69](#)
potlačení zobrazení tabulky [69](#)
proměnné faktoru [69](#)
proměnné stavu přežití [75](#)
příklad [68](#)
statistiky [68](#)
výkresy [69](#)
Zkouška Wilcoxon (Gehan) [69](#)

