

IBM SPSS Categories 29



Poznámka

Před použitím těchto informací a produktu, který podporují, si přečtěte informace v tématu [“Upozornění” na stránce 43](#).

Informace o produktu

Toto vydání se vztahuje k verzi 29, vydání 0, modifikaci 1 produktu IBM® SPSS Statistics a ke všem následujícím vydáním a modifikacím, dokud nebude v nových vydáních uvedeno jinak.

© Copyright International Business Machines Corporation .

Kapitola 1. Kategorie.....	1
Úvod do optimálního škálování postupů pro kategorická data.....	1
Co je optimální škálování?.....	1
Proč používat optimální škálování?.....	1
Optimální úroveň rozšiřitelnosti a úroveň měření.....	1
Který postup je nejlepší pro vaši aplikaci?.....	4
Poměr stran v optimálních grafech škálování.....	9
Kategorická regrese (CATREG).....	9
Definovat měřítko v kategorickém regresí.....	10
Diskreční diskretnost regresní oblasti.....	10
Kategorické regrese chybějících hodnot.....	11
Kategorické volby regrese.....	11
Kategorická regularizace regresních regresí.....	12
Kategorický regresní výstup.....	12
Kategorické ukládání regresí.....	13
Kategorické transformace regresní transformace.....	13
Další funkce příkazu CATREG.....	13
Kategorická analýza hlavních komponent (CATPCA).....	14
Definování měřítka a váhy v produktu CATPCA.....	15
Diskreční diskretizace v kategorii Hlavní komponenty.....	15
Analýza kategorických komponent hlavní komponenty-chybějící hodnoty.....	16
Volby analýzy Kategorických komponent hlavní komponenty.....	16
Výstup analýzy hlavních komponent hlavních komponent.....	18
Uložení analýzy hlavních komponent hlavních komponent.....	18
Kategorické objekty analýzy objektů hlavních komponent.....	18
Kategorie analýzy kategorických komponent hlavní komponenty.....	19
Kategorické analýzy hlavních komponent-naložení šarží.....	19
Zaváděcí program analýzy hlavních komponent hlavních komponent.....	19
Další funkce příkazu CATPCA.....	20
Analýza nelineárních kanonických korelací (OVERALS).....	20
Definovat rozsah a měřítko.....	21
Definovat rozsah.....	22
Volby analýzy nelineárních kanonických korelací korelace.....	22
Další funkce příkazu OVERALS.....	23
Analýza korespondence.....	23
Definovat rozsah řádků v analýze korespondencí.....	24
Definovat rozsah sloupců v analýze Korespondence.....	24
Model analýzy korespondence.....	25
Statistika analýzy korespondence.....	26
Ploty analýzy korespondence.....	26
Další funkce příkazu CORRESPONDENCE.....	27
Analýza více korespondencí.....	27
Definování váhy proměnné ve více korespondenční analýze.....	28
Diskretizace s více korespondenční analýzou.....	28
Chybějící hodnoty analýzy více korespondence.....	28
Volby analýzy více korespondencí.....	29
Výstup analýzy více korespondencí.....	30
Hromadná analýza korespondenční analýzy.....	30
Hromadné objekty analýzy korespondenční analýzy.....	30
Hromadné položky analýzy s více korespondenční analýzou.....	30
Další funkce příkazu MULTIPLE CORRESPONDENCE.....	31

Vícerozměrné škálování (PROXSCAL).....	31
Proximities v Matrices ve sloupcích.....	32
Proximaty ve sloupcích.....	32
Proximaty v jednom sloupci.....	33
Vytvořit proximaty z dat.....	33
Vytvořit ukazatel z dat.....	33
Definice modelu Multidimensional Scaling Model.....	34
Omezení vícerozměrného škálování.....	34
Vícerozměrné volby škálování.....	35
Multidimensional Scaling Plots, verze 1.....	35
Multidimensional Scaling Plots, verze 2.....	36
Vícerozměrný výstup škálování.....	36
Další funkce příkazu PROXSCAL.....	36
Multidimenzionální rozkládání (PREFSCAL).....	37
Definovat vícerozměrný model rozkládání.....	37
Omezení vícenásobného rozkládání.....	38
Volby vícenásobného rozkládání.....	38
Vícerozměrné skládací paty.....	39
Vícerozměrný výstup pro rozkládání.....	40
Další funkce příkazu PREFSCAL.....	41
Upozornění.....	43
Ochranné známky.....	44
Rejstřík.....	47

Kapitola 1. Kategorie

Následující kategorie funkcí jsou obsaženy v části SPSS Statistics Professional Edition nebo Kategorie.

Úvod do optimálního škálování postupů pro kategorická data

Procedury kategorií používají optimální škálování k analýze dat, která jsou obtížná nebo neproveditelná pro standardní statistické postupy pro analýzu. Tato kapitola popisuje, co každý postup dělá, situace, kdy je každý postup nejvhodnější, vztahy mezi postupy a vztahy těchto postupů k jejich standardním statistickým protějškům.

Poznámka: Tyto postupy a jejich implementace v produktu IBM SPSS Statistics byly vyvinuty skupinou DTSS (Data Theory Scaling System Group), která se skládá z členů oddělení pro vzdělávání a psychologii, Fakulty sociálních a behaviorálních věd, Leidenské univerzity.

Co je optimální škálování?

Myšlenka optimální škálování spočívá v přidělení numerických kvantit do kategorií každé proměnné, čímž umožňuje použití standardních postupů k získání roztoku na kvantifikovaných proměnných.

Optimální hodnoty stupnice jsou přiřazeny kategoriím každé proměnné na základě optimalizujícího kritéria používané procedury. Na rozdíl od původních jmenovek nominálních nebo ordinálních proměnných v analýze mají tyto hodnoty měřítka vlastnosti metriky.

Ve většině procedur kategorií je optimální kvantifikace pro každou škálovanou proměnnou získána prostřednictvím iterační metody zvané **střídající se nejmenších čtverců**, v níž jsou po použití aktuálního kvantitu k nalezení řešení aktualizovány kvantity tímto řešením. Aktualizované kvantity se pak použijí k nalezení nového řešení, které se používá k aktualizaci kvantitů, a tak dále, dokud se nedosáhne určitého kritéria, které signalizuje zastavení procesu.

Proč používat optimální škálování?

Categorical data are often found in marketing research, survey research, and research in the social and behavioral sciences. Ve skutečnosti řada výzkumníků řeší téměř výhradně kategorická data.

Přestože existují úpravy většiny standardních modelů specificky pro analýzu kategorických dat, často neprovádějí dobře pro datové sady, které jsou součástí:

- Příliš málo pozorování
- Příliš mnoho proměnných
- Příliš mnoho hodnot na proměnnou

Pomocí kvantifikace kategorií se optimálním metodám škálování vyhýbá problémům v těchto situacích. Kromě toho jsou užitečné i v případech, že jsou vhodné specializované techniky.

Místo interpretace odhadů parametrů je interpretace optimálního výstupu měřítka často založena na grafických displejích. Optimální techniky škálování nabízejí vynikající průzkumné analýzy, které doplňují ostatní modely IBM SPSS Statistics dobře. Díky zúžení fokusu vašeho vyšetřování může vizualizace dat díky optimálnímu škálování tvořit základ analýzy, která se zaměřuje na interpretaci modelových parametrů.

Optimální úroveň rozšiřitelnosti a úroveň měření

To může být velmi matoucí koncept při prvním použití procedur kategorií. Při zadávání úrovně určujete nikoli úroveň, na které jsou proměnné *měřeny*, ale úroveň, na které jsou proměnné *zmenšeny*. Myšlenka spočívá v tom, že proměnné, které mají být kvantifikovány, mohou mít nelineární vztahy bez ohledu na to, jak se měří.

Pro účely kategorií existují tři základní úrovně měření:

- Úroveň **nominal** znamená, že hodnoty proměnné reprezentují neřazené kategorie. Příklady proměnných, které mohou být nominální, jsou oblasti, oblast PSČ, náboženská příslušnost a více kategorií voleb.
- Úroveň **ordinal** znamená, že hodnoty proměnné reprezentují seřazené kategorie. Mezi příklady patří klasifikační stupnice představující stupeň spokojenosti nebo sebedůvěra a skóre hodnocení preferencí.
- Úroveň **numerical** znamená, že hodnoty proměnné reprezentují seřazené kategorie se smysluplnou metrikou, takže porovnání vzdálenosti mezi kategoriemi je vhodné. Příklady zahrnují věk v rocích a příjem v tisících dolarů.

Předpokládejme například, že proměnné *region*, *joba age* jsou kódovány způsobem uvedeným v následující tabulce.

Tabulka 1. Schéma kódování pro region, pracovní místo a věk

Kód regionu	Hodnota oblasti	Kód úlohy	Hodnota úlohy	Věk
1	Sever	1	internitka	20
2	Jih	2	PRODEJNÍ ZÁSTUPCE	22
3	Východ	3	Správce	25
4	Západ			27

Zobrazené hodnoty představují kategorie každé proměnné. *Region* by byla nominální proměnná. Existují čtyři kategorie *oblasti* bez vnitřního uspořádání. Hodnoty 1 až 4 jednoduše reprezentují čtyři kategorie; schéma kódování je zcela libovolné. *Úlohy* na druhou stranu předpokládá jako ordinální proměnnou. Původní kategorie tvoří postup od internisty manažerovi. Větší kódy představují vyšší pozici ve firemním žebříčku. Nicméně pouze informace o pořadí je známo -- nic nelze říci o vzdálenosti mezi sousedními kategoriemi. Na rozdíl od *stáří* lze předpokládat, že má být číselná proměnná. V případě *stáří* jsou vzdálenosti mezi hodnotami věcně smysluplné. Vzdálenost mezi 20 a 22 je stejná jako vzdálenost mezi 25 a 27, zatímco vzdálenost mezi 22 a 25 je větší než jedna z těchto.

Výběr optimální úrovně měřítka

Je důležité pochopit, že neexistují žádné vnitřní vlastnosti proměnné, které by automaticky předdefinovaly optimální úroveň měřítka, která byste pro ni měli zadávat. Můžete zkoumat vaše data jakýmkoli způsobem, který by dávalo smysl a usnadnil interpretaci. Analýzou číselné proměnné na ordinální úrovni může například použití nelineární transformace umožnit řešení v menším počtu rozměrů.

Následující dva příklady ilustrují, jak úroveň "zřejmých" měření nemusí být nejlepší optimální úroveň škálování. Předpokládejme, že proměnná řadí objekty do věkových skupin. Ačkoli věk může být škálován jako číselná proměnná, to může být pravda, že pro lidi mladší než 25 bezpečnosti má pozitivní vztah s věkem, zatímco pro lidi starší než 60 bezpečnost má negativní vztah s věkem. V tomto případě by mohlo být lepší léčit věk jako nominální proměnná.

Jako další příklad se zdá, že proměnná, která seřadí osoby podle politických preferencí, je v zásadě nominální. Pořádáte-li však politické levici politické levici, možná budete chtít, aby byla kvantifikace stran respektována použitím ordinální úrovně analýzy.

I když neexistují žádné předdefinované vlastnosti proměnné, které činí výhradně jednu úroveň nebo jinou, existují obecné pokyny, které pomáhají uživateli nováčků. S jednoduchou-nominálním kvantifikace obvykle neznáte pořadí kategorií, ale chcete, aby analýza vnucl jeden. Je-li známo pořadí kategorií, měli byste zkusit pořadové kvantifikaci. Pokud jsou kategorie neobjednatelné, můžete zkusit více nominálního kvantifikace.

Plosy transformace

Různé úrovně, při kterých může být každá proměnná škálována, ukládají různá omezení na kvantitu. Transformační grafy znázorňují vztah mezi kvantitami a původními kategoriemi, které jsou výsledkem vybrané optimální úrovně škálování. Např. výsledky lineárního transformačního grafu, když je proměnná

považována za numerickou. Proměnné považované za ordinální výsledky v neklesajícím transformačním grafu. Transformační grafy pro proměnné považované nominálně, které jsou ve tvaru U (nebo inverzní funkce), zobrazují kvadratický vztah. Nominální proměnné by mohly také přinést transformační zkusné plochy bez zjevných trendů tím, že se změní pořadí kategorií zcela. Na následujícím obrázku je zobrazen ukázkový diagram transformace.

Transformační grafy se hodí zejména k určení, jak dobře se vybraná optimální úroveň škálování provádí. Pokud několik kategorií obdrží podobné kvantily, může být záruka těchto kategorií do jedné kategorie oprávněná. Jinak, pokud se proměnná považovaná za nominální přijme kvanta, která vykazuje rostoucí trend, může být výsledná transformace výsledkem podobného přizpůsobení. Je-li tento trend lineární, může být vhodná numerická léčba. Je-li však zaručeno sbalení kategorií nebo změna úrovně škálování, analýza se výrazně nezmění.

Kódy kategorií

Při kódování kategoriálních proměnných je třeba věnovat pozornost kódování kategoriálních proměnných, protože některé režimy kódování mohou přinést nežádoucí výstup nebo neúplné analýzy. Možné režimy kódování pro *úlohu* jsou zobrazeny v následující tabulce.

Tabulka 2. Alternativní režimy kódování pro <i>úlohu</i>				
Kategorie	A	B	C	D
internitka	1	1	5	1
PRODEJNÍ ZÁSTUPCE	2	2	6	5
Správce	3	7	7	3

Některé postupy kategorie vyžadují, aby byl definován rozsah každé použité proměnné. Jakákoli hodnota mimo tento rozsah je považována za chybějící hodnotu. Minimální hodnota kategorie je vždy 1. Hodnota maximální kategorie je dodána uživatelem. Tato hodnota není *počet* kategorií pro proměnnou-je to hodnota kategorie *největší*. Např. v tabulce A má schéma A maximální hodnotu kategorie 3 a schéma B má maximální hodnotu kategorie 7, ale obě schémata obsahují stejné tři kategorie.

Rozsah proměnných určuje, které kategorie budou z analýzy vynechány. Libovolné kategorie s kódy mimo definovaný rozsah se vynechají z analýzy. Jedná se o jednoduchou metodu pro vynechání kategorií, ale může vést k nechtěným analýzám. Nesprávně definovaná maximální kategorie může z analýzy vynechat *platné* kategorie. Např. pro schéma B definující maximální hodnotu kategorie 3 označuje, že *úloha* má kategorie kódované od 1 do 3; kategorie *manager* je považována za chybějící. Protože žádná kategorie nebyla ve skutečnosti kódována 3, třetí kategorie v analýze neobsahuje žádné případy. Pokud jste chtěli vynechat všechny kategorie správců, tato analýza by byla vhodná. Pokud však mají být zahrnuti manažeři, musí být maximální kategorie definována jako 7 a chybějící hodnoty musí být kódovány hodnotami nad 7 nebo nižší než 1.

Pro proměnné považované za nominální nebo ordinální, rozsah kategorií nemá vliv na výsledky. U nominálních proměnných je důležitá pouze jmenovka a ne hodnota přidružená k tomuto popisku. Pro ordinální proměnné je pořadí kategorií zachováno v souladu s kvantitami; hodnoty kategorie samotné nejsou důležité. Všechny systémy kódování, které mají za následek stejné řazení kategorií, budou mít stejné výsledky. První tři schémata v tabulce jsou například funkčně ekvivalentní, pokud je *úloha* analyzována na ordinální úrovni. Pořadí kategorií je v těchto režimech totožné. Režim D na druhé straně invertuje druhou a třetí kategorii a přinese odlišné výsledky než ostatní programy.

Ačkoli řada schémat kódování pro proměnnou je funkčně ekvivalentní, jsou preferovány systémy s malými rozdíly mezi kódy, protože kódy mají vliv na objem výstupu vytvořeného procedurou. Všechny kategorie kódované s hodnotami mezi 1 a uživatelem definované maximum jsou platné. Je-li kterákoli z těchto kategorií prázdná, odpovídající kvantily budou buď systémem chybět, nebo 0, v závislosti na postupu. Přestože ani jedno z těchto přiřazení neovlivní analýzy, je pro tyto kategorie vytvořen výstup. Tudíž pro schéma B má *úloha* čtyři kategorie, které přijímají systémově chybějící hodnoty. U schématu C existují také čtyři kategorie, které přijímají systémově chybějící ukazatele. Naproti tomu u režimu A neexistují žádné systémové kvantily chybějící v systému. Použití po sobě jdoucích celých čísel jako kódů pro

proměnné považované za nominální nebo ordinální výsledky za mnohem menší výstup bez dopadu na výsledky.

Schémata kódu pro proměnné považované za číselné proměnné jsou omezenější než ordinální případ. U těchto proměnných jsou důležité rozdíly mezi následnými kategoriemi. V následující tabulce jsou zobrazena tři schémata kódování pro *age*.

Tabulka 3. Alternativní režimy kódování pro stáří

Kategorie	A	B	C
20	20	1	1
22	22	3	2
25	25	6	3
27	27	8	4

Jakýkoli zpětný kód číselných proměnných musí zachovat rozdíly mezi kategoriemi. Použití původních hodnot je jedna z metod pro zajištění zachování rozdílů. To však může vést k mnoha kategoriím, které mají systémem chybějící indikátory. Např. schéma A používá původní pozorované hodnoty. Pro všechny procedury kategorie s výjimkou Korespondenční analýzy je maximální hodnota kategorie 27 a hodnota minimální kategorie je nastavena na 1. Prvních 19 kategorií je prázdné a systém obdrží indikátory chybějících systémů. Výstup se může rychle stát poměrně těžkopádným, pokud maximální kategorie je mnohem větší než 1 a mezi 1 a maximem je mnoho prázdných kategorií.

Chcete-li snížit objem výstupu, je možné provést zpětný odraz. V číselném případě by však nemělo být použito automatické zařízení Recde. Kódování po sobě jdoucí celá čísla vede k rozdílům mezi 1 mezi všechny po sobě jdoucí kategorií, a, jako výsledek, všechny kvantace bude rovnoměrně rozložený. Charakteristiky metriky, které jsou považovány za důležité při zpracování proměnné jako numerické, jsou zničeny rekódováním po sobě jdoucí celá čísla. Například schéma C v tabulce odpovídá automaticky překódování *stáří*. Rozdíl mezi kategoriemi 22 a 25 se změnil ze tří na jeden, a kvanta bude odrážet posledně zmíněný rozdíl.

Alternativní schéma kódování, které zachovává rozdíly mezi kategoriemi, je odčítání nejmenší hodnoty kategorie ze všech kategorií a přičítají se k jednotlivým odlišnostem. Schéma B má za následek tuto transformaci. Do každého výsledku byla přidána nejmenší hodnota kategorie, 20, odečtena od každé kategorie a 1. Transformované kódy mají minimálně 1 a všechny rozdíly jsou identické s původními daty. Maximální hodnota kategorie je nyní 8 a před první nenulové kvantifikace jsou odstraněny všechny kvantity, které byly odstraněny. Nenulová kvantita odpovídající jednotlivým kategoriím, které jsou výsledkem režimu B, jsou však shodná s kvantitami ze systému A.

Který postup je nejlepší pro vaši aplikaci?

Techniky obsažené ve čtyřech z těchto postupů (Korespondenční analýza, vícenásobná korespondence Analýza, Analýza kategorických komponent a nelineární kanonická analýza korelace) spadají do obecné oblasti vícerozměrné analýzy dat známé jako **snížení dimenze**. To znamená, že vztahy mezi proměnnými jsou reprezentovány v několika dimenzích-řekněme dva nebo tři-tak často, jak je to jen možné. To vám umožní popsat struktury nebo vzory ve vztazích, které by byly příliš obtížné na fathom v jejich originálním bohatství a složitosti. V aplikacích průzkumu trhu mohou být tyto techniky způsobem **percepční mapování**. Hlavní výhodou těchto postupů je to, že se přizpůsobují datům s různou úrovní optimálního škálování.

Kategorická regrese popisuje vztah mezi kategorickou proměnnou odezvy a kombinací proměnných kategorického prediktoru. Vliv každé proměnné prediktoru na proměnnou odezvy je popsán pomocí příslušné regresní váhy. Stejně jako v jiných procedurách lze data analyzovat s různými úrovněmi optimálního škálování.

Vícerozměrné škálování a vícerozměrné rozkládání popisují vztahy mezi objekty v nízkém dimenzionálním prostoru za použití proximity mezi objekty.

Níže jsou uvedeny stručné pokyny pro každý z postupů:

- Kategorickou regresí použijte k předpovídání hodnot kategoriální závislé proměnné z kombinace kategoriálních nezávislých proměnných.
- Kategorické analýzy hlavních komponent použijte k zohlednění vzorců variace v jedné sadě proměnných se smíšenými optimálními úrovněmi měřítka.
- Použijte nelineární kanonickou analýzu korelace k posouzení rozsahu, do kterého jsou korelovány dvě nebo více sad proměnných smíšených optimálních úrovní měřítka.
- Korespondenční analýza se používá k analýze dvousměrné tabulky pro nepředvídané události nebo dat, které lze vyjádřit jako dvoucestné tabulky, jako jsou preference značky nebo sociometric choice data.
- Použití více korespondenčních analýz k analýze kategorické vícerozměrné datové matice, když jste ochotni udělat žádný silnější předpoklad, že všechny proměnné jsou analyzovány na nominální úrovni.
- Použijte Multidimensional Scaling k analýze dat blízkosti k nalezení znázornění jednotlivých sad objektů v nízkém dimenzionálním prostoru.
- Vícedimenzionální rozkládání použít k analýze dat blízkých slov k vyhledání dvou sad objektů v nízkém dimenzionálním prostoru.

Kategorická regrese

Použití Categorical Regression je nejvhodnější, když cílem analýzy je předpovědět závislou proměnnou (odpověď) ze sady proměnných nezávislých (prediktorů). Stejně jako u všech optimálních postupů škálování jsou hodnoty stupnice přiřazeny ke každé kategorii každé proměnné tak, aby tyto hodnoty byly optimální s ohledem na regresi. Řešením kategorické regrese maximalizuje druhou korelaci mezi transformovanou odezvou a váženým kombinací transformovaných prediktorů.

Vztah k jiným postupům kategorie. Kategorická regrese s optimálním škálováním je srovnatelná s optimální analýzou kanonické korelace škálování se dvěma sadami, z nichž jedna obsahuje pouze závislou proměnnou. V posledně uvedené technice se podobnost sad odvozuje porovnáním každé sady s neznámou proměnnou, která leží někde mezi všemi sadami. V kategorické regresí se podobnost s transformovanou odezvou a lineární kombinací transformovaných prediktorů hodnotí přímo.

Vztah ke standardním technikám. V standardní lineární regresí lze kategorické proměnné buď rekódovat jako proměnné indikátoru nebo být považovány za proměnné úrovně intervalu stejně jako proměnné úrovně intervalu. V prvním přístupu model obsahuje samostatné zachycení a sklon pro každou kombinaci úrovní kategoriálních proměnných. To vede k interpretaci velkého počtu parametrů. Ve druhém přístupu se odhaduje pouze jeden parametr pro každou proměnnou. Nicméně, svévolná povaha kóje kategorie znemožňuje generalizace.

Nejsou-li některé proměnné souvislé, jsou k dispozici alternativní analýzy. Je-li odezva souvislá a prediktory jsou kategorické, často se používá analýza rozptylu. Je-li odpověď kategorická a prediktory jsou souvislé, logistická regrese nebo analýza diskriminantů může být vhodná. Jsou-li odezva a prediktory kategorické, často se používají lineární modely loglinear.

Regrese s optimálním škálováním nabízí tři úrovně měřítka pro každou proměnnou. Kombinace těchto úrovní může brát v úvahu celou řadu nelineárních vztahů, pro které je jakákoliv jedna "standardní" metoda špatně vhodná. Optimální rozšiřitelnost proto nabízí větší flexibilitu než standardní přístupy s minimální přidanou složitostí.

Kromě toho nelineární transformace prediktorů obvykle redukuje závislosti mezi prediktory. Pokud porovnáte vlastní hodnoty matice korelace pro prediktory s hodnotami matice korelace pro optimálně škálované prediktory, bude tato druhá sada obvykle méně proměnlivá než hodnota dřívější. Jinými slovy, v případě kategorické regrese činí optimální škálování větší vlastní hodnoty korelační matice prediktorů menší a menší vlastní hodnoty jsou větší.

Analýza kategorických komponent hlavních komponenty

Použití analýzy Categorical Principal Components je nejvhodnější, když chcete vzít v úvahu vzorce rozptylu v jedné sadě proměnných smíšených optimálních úrovní měřítka. Tato technika se snaží snížit dimenzionalitu sady proměnných, zatímco je co nejvíce započítávání variace. Hodnoty stupnice jsou přiřazeny ke každé kategorii každé proměnné, aby tyto hodnoty byly optimální s ohledem na řešení

hlavních komponent. Objekty v analýze přijímají skóre komponenty založené na kvantifikovaných datech. Vykreslí skóre komponent odhalují vzory mezi objekty v analýze a mohou v datech odhalit neobvyklé objekty. Řešení analýzy kategoriálních komponent činitele maximalizuje korelace skóre objektu s každou z kvantifikovaných proměnných pro určený počet komponent (rozměrů).

Důležitým použitím kategoriálních komponent činitelů je zkoumání údajů o preferencích, v nichž respondenti ohodnocovali nebo ohodnocovali řadu položek s ohledem na preference. V běžné konfiguraci dat produktu IBM SPSS Statistics jsou řádky jednotlivci, sloupce jsou měřeními položek a skóre mezi řádky jsou skóre předvoleb (například v měřítku 0 až 10), čímž se datový řádek podmíní. Pro data předvoleb můžete chtít zacházet s jednotlivci jako s proměnnými. Pomocí procedury Transpose můžete data transponovat. Rattery se stávají proměnnými a všechny proměnné jsou deklarovány pořadovým číslem. Neexistují žádné námitky proti použití více proměnných než objektů v produktu CATPCA.

Vztah k jiným postupům kategorie. Pokud jsou všechny proměnné deklarovány vícenásobnými nominálními komponentami, analýza kategoriálních hlavních komponent vytvoří analýzu ekvivalentní ke spuštění analýzy více korespondence na stejných proměnných. Analýzu kategoriálních činitelů lze proto považovat za typ vícenásobné analýzy korespondence, ve které jsou některé proměnné deklarovány jako ordinální nebo numerické.

Vztah ke standardním technikám. Pokud jsou všechny proměnné zmenšeny na číselné úrovni, analýza základních komponent činitele je ekvivalentní standardní analýze hlavních komponent.

Obecněji řečeno, analýza kategoriálních komponent představuje alternativu ke zpracování korelace mezi nečíselnými stupnicemi a jejich analýzou pomocí standardní hlavní komponenty nebo metody analýzy faktoru. Naivní použití obvyklého Pearsonova korelačního koeficientu jako měřítka přidružení pro ordinální data může vést k netriviální zaujatosti při odhadu korelací.

Analýza nelineárních kanonických korelací

Nelineární kanonická analýza korelace je velmi obecná procedura s mnoha různými aplikacemi. Cílem nelineárních kanonických analytických analýz je analýza vztahů mezi dvěma nebo více sadami proměnných místo mezi samotnými proměnnými, jako v analýze hlavních komponent. Například můžete mít dvě sady proměnných, kde jedna sada proměnných může být demografickými položkami pozadí na sadě respondentů a druhá sada může být odpovědí na sadu položek přístupu. Úrovně škálování v analýze mohou být různé kombinace jmenovitých, ordinálních a numerických. Optimální škálování kanonické korelační analýzy určuje podobnost mezi sadami současně porovnáním kanonických proměnných z každé sady do kompromisního souboru skóre přiřazených k objektům.

Vztah k jiným postupům kategorie. Pokud existují dvě nebo více sad proměnných s pouze jednou proměnnou na sadu, optimální analýza kanonické korelace škálování je ekvivalentní k analýze komponent optimálního škálování škálování. Jsou-li všechny proměnné v jedné analýze jedné proměnné na sadu více nominálních, optimální analýza kanonické korelace škálování je ekvivalentní vícenásobné analýze korespondence. Pokud existují dvě sady proměnných, z nichž jedna obsahuje pouze jednu proměnnou, je optimální analýza kanonické korelace škálování ekvivalentní kategoriální regresi s optimálním škálováním.

Vztah ke standardním technikám. Standardní kanonická korelační analýza je statistická technika, která vyhledá lineární kombinaci jedné sady proměnných a lineární kombinaci druhé sady proměnných, které jsou maximálně korelované. Vzhledem k této sadě lineárních kombinací, kanonická analýza korelace může najít následné nezávislé sady lineárních kombinací, které se označují jako kanonické proměnné, až do maximálního počtu rovnající se počtu proměnných v menší sadě.

Pokud existují dvě sady proměnných v analýze a všechny proměnné jsou definovány jako numerické, optimální škálování kanonické korelační analýzy je ekvivalentní standardní kanonické analýze korelace. Ačkoli produkt IBM SPSS Statistics nemá proceduru kanonické korelace korelace, mnoho relevantních statistik lze získat od vícevariantní analýzy rozptylu.

Optimální kanonická korelační analýza škálování má různé jiné aplikace. Pokud máte dvě sady proměnných a jedna sada obsahuje nominální proměnnou deklarovanou jako jednu nominální hodnotu, lze výsledky analýzy kanonické korelace škálování optimálně interpretovat podobně jako regresní analýzu. Pokud proměnnou považujete za vícenásobnou nominální hodnotu, je optimální analýza škálování alternativou k analýze diskriminačních úloh. Seskupení proměnných do více než dvou sad poskytuje různé způsoby analýzy vašich dat.

Analýza korespondence

Cílem analýzy korespondence je dělat bifáty pro korespondenční tabulky. V tabulce korespondenční tabulky se předpokládá, že řádková a sloupcová proměnná představuje neřazené kategorie; proto je vždy použita nominální optimální úroveň měřítka. Obě proměnné jsou kontrolovány pouze pro své nominální informace. To znamená, že jediným hlediskem je skutečnost, že některé objekty jsou ve stejné kategorii, zatímco jiné nikoli. Nepředpokládá se nic o vzdálenosti nebo pořadí mezi kategoriemi stejné proměnné.

Jednou z konkrétních použití analýzy korespondence je analýza dvoucestných pohotovostních tabulek. Má-li tabulka r aktivních řádků a c aktivních sloupců, je počet dimenzí v řešení korespondenční analýzy minimální hodnotou r minus 1 nebo c minus 1, podle toho, která hodnota je nižší. Jinými slovy, můžete dokonale reprezentovat kategorie řádků nebo kategorie sloupců rezervní tabulky v prostoru dimenzí. Prakticky řečeno, vy byste chtěli reprezentovat řádek a sloupec kategorií dvoucestné tabulky v low-dimenzionálním prostoru, řekněme dvou rozměrů, z důvodu, že dvojrozměrné grafy jsou snadněji srozumitelné než vícerozměrné prostorové reprezentace.

Je-li použito méně než maximálního počtu možných dimenzí, statistika vytvořená v analýze popisuje, jak dobře jsou v nízkorozměrné reprezentaci reprezentovány kategorie řádků a sloupců. Za předpokladu, že kvalita reprezentace dvoudimenzionálního řešení je dobrá, můžete zkontrolovat grafy řádků a body sloupce, aby se dozvěděli, které kategorie řádkové proměnné jsou podobné, které kategorie proměnné sloupce jsou podobné a které kategorie řádků a sloupců jsou podobné sobě navzájem.

Vztah k jiným postupům kategorie. Korespondenční analýza je omezena na dvoucestné tabulky. Pokud existuje více než dvě proměnné zájmu, můžete kombinovat proměnné a vytvořit proměnné interakce. Například pro proměnné *region*, *úkola* věkmůžete kombinovat *region* a *úkol* a vytvořit tak novou proměnnou *přeúloha* s 12 kategoriemi zobrazenými v následující tabulce. Tato nová proměnná tvoří dvoucestnou tabulku s *věkem* (12 řádků, 4 sloupci), které lze analyzovat v rámci analýzy korespondence.

Kód kategorie	Definice kategorie	Kód kategorie	Definice kategorie
1	sever, stáž	7	východ, stáž
2	Severní, obchodní zástupce	8	Východní, obchodní zástupce
3	Severní, vedoucí	9	Východ, manažer
4	Jih, inter	10	západ, stáž
5	Jih, obchodní zástupce	11	Západ, obchodní zástupce
6	Jižní, vedoucí	12	Západ, manažer

Jeden nedostatek tohoto přístupu je, že jakýkoli pár proměnných může být kombinován. Můžeme kombinovat *úlohu* a *stáří*, což vede k další 12-kategorii proměnné. Nebo můžeme kombinovat *region* a *age*, což vede k nové proměnné 16. kategorie. Každá z těchto proměnných interakce vytváří dvoucestnou tabulku se zbývající proměnnou. Korespondenční analýzy těchto tří tabulek nepřinesou stejné výsledky, přesto je každý z nich platný přístup. Kromě toho, pokud existují čtyři nebo více proměnných, lze konstruovat dvoucestné tabulky porovnávající proměnnou interakce s jinou proměnnou interakce. Počet možných tabulek pro analýzu může získat poměrně velký, dokonce i pro několik proměnných. Můžete vybrat jednu z těchto tabulek, které chcete analyzovat, nebo můžete všechny tyto tabulky analyzovat. Alternativně lze použít proceduru analýzy více korespondence k prozkoumání všech proměnných najednou bez nutnosti vytvářet proměnné interakce.

Vztah ke standardním technikám. Postup kontingenční tabulky lze také použít k analýze pohotovostních tabulek s nezávislostí jako se společným zaměřením v analýzách. Avšak i v malých tabulkách může být odhalení příčiny odklonění od nezávislosti obtížné. Obslužný program korespondence analýzy spočívá v zobrazení takových vzorů pro dvousměrné tabulky jakékoli velikosti. Existuje-li přidružení mezi řádkem a sloupci sloupců -- to znamená, že je-li hodnota chí kvadrát významná -- analýza korespondence může pomoci odhalit povahu vztahu.

Analýza více korespondencí

Analýza více korespondencí se pokouší o vytvoření řešení, v němž jsou objekty v rámci stejné kategorie zakresleny blízko sebe a objekty v různých kategoriích jsou vykresleny daleko od sebe. Každý objekt je co nejbližší ke kategoriím kategorií, které se vztahují na daný objekt. Tímto způsobem rozdělují kategorie objekty do homogenních podskupin. Proměnné jsou považovány za homogenní, když klasifikují objekty ve stejných kategoriích do stejných podskupin.

Pro jednorozměrné řešení přiřazuje vícenásobné korespondenční analýzy optimální hodnoty stupnice (kategorie kvantitativní kategorie) do každé kategorie každé proměnné tak, aby se v průměru v průměru rozšířilo maximální rozložené kategorie. Pro dvourozměrné řešení vyhledá vícenásobná korespondence druhou sadu kvantitativní z kategorií každé proměnné nesouvisející s první sadou, pokouší se znovu maximalizovat šíření, a tak dále. Vzhledem k tomu, že kategorie proměnné obdrží tolik výstřelů, kolik je rozměrů, předpokládá se, že proměnné v analýze jsou více nominálně na optimální úroveň škálování.

Analýza více korespondence také přiřazuje skóre objektům v analýze takovým způsobem, aby byly kvantitativní kategorií průměrnými nebo centroidy objektů v dané kategorii.

Vztah k jiným postupům kategorie. Analýza více korespondence je také známá jako analýza homogenity nebo duální škálování. To dává srovnatelné, ale ne identické, výsledky k korespondenční analýze, když existují pouze dvě proměnné. Korespondenční analýza vytváří jedinečný výstup shrnující vhodnost a kvalitu znázornění řešení, včetně informací o stabilitě. Analýza korespondence je tedy obvykle vhodnější než vícenásobné korespondenční analýzy v obou případech. Dalším rozdílem mezi oběma procedurami je, že vstup do vícenásobné korespondenční analýzy je datová matice, kde řádky jsou objekty a sloupce jsou proměnné, zatímco vstup do analýzy korespondence může být stejná datová matice, obecné matice blízkosti, nebo společná pohotovostní tabulka, která je agregovanou maticí, v níž jak řádky a sloupce představují kategorie proměnných. Analýza více korespondence může být také pomyšlena jako hlavní komponenty analýzy dat, jejichž rozsah je změněn na více nominálních úrovní.

Vztah ke standardním technikám. Analýza více korespondence může být považována za analýzu vícecestné tabulky pro nepředvídané události. Vícecestné tabulky pro nepředvídané události lze také analyzovat pomocí procedury kontingenční tabulky, ale kontingenční tabulky poskytují samostatné souhrnné statistické údaje pro každou kategorii každé řídicí proměnné. Pomocí analýzy více korespondence je často možné shrnout vztah mezi všemi proměnnými s jedním dvojrozměrným zákresem. Rozšířené použití analýzy více korespondence má nahradit původní hodnoty kategorií hodnotami optimálního rozsahu z první dimenze a provádět sekundární vícerozměrnou analýzu. Vzhledem k tomu, že analýza více korespondence nahrazuje popisky kategorií číselnými hodnotami desetinných míst, mnoho různých procedur, které vyžadují číselná data, lze použít po provedení vícenásobné analýzy korespondence. Například procedura Analýza faktorů vygeneruje první hlavní komponentu, která je ekvivalentem první dimenze analýzy více korespondence. Skóre komponenty v první dimenzi se rovnají skóre objektu a druhé naložení komponenty se rovnají diskriminačním ukazávám. Druhá dimenze analýzy více korespondence se však neshoduje s druhou dimenzí analýzy faktorů.

Multidimenzionální škálování

Použití Multidimensional Scaling je nejvhodnější, když je cílem vaší analýzy najít strukturu v sadě opatření na vzdálenost mezi jednou sadou objektů nebo případů. Toho lze dosáhnout přiřazením pozorování ke specifickým umístěním v konceptuálním nízdídimenzionálním prostoru, takže vzdálenosti mezi body v prostoru odpovídají danému (dis) podobnostem co nejbližší. Výsledkem je nejmenších čtverců znázornění objektů v tomto málo dimenzionálním prostoru, což vám v mnoha případech pomůže blíže porozumět vašim datům.

Vztah k jiným postupům kategorie. Máte-li více dat, z nichž vytváříte vzdálenosti a které pak analyzujete s vícerozměrným škálováním, výsledky se podobají analýze dat pomocí analýzy kategoriálních komponent činitele s normalizací činitele objektu. Tento druh PCA je také znám jako hlavní souřadnice souřadnic.

Vztah ke standardním technikám. Kategorie Multidimensional Scaling procedure (PROXSCAL) nabízí několik vylepšení v závislosti na proceduře škálování, která je k dispozici ve volbě Statistika základu statistiky (ALSCAL). PROXSCAL nabízí akcelerovaný algoritmus pro určité modely a umožňuje vám klást omezení na společný prostor. Navíc PROJSCAL se snaží minimalizovat normalizované hrubé zatížení spíše

než S-stres (označovaný také jako **kmen**). Normalizovaná prvotní zátěž je obecně preferována, protože se jedná o opatření založené na vzdálenostech, zatímco S-stres je založen na čtvercové vzdálenosti.

Vícziónální rozkládání

Použití vícerozměrného skládání je nejvhodnější, když je cílem vaší analýzy najít strukturu v sadě opatření na vzdálenost mezi dvěma sadami objektů (označované jako objekty řádků a sloupců). Toho lze dosáhnout přiřazením pozorování ke specifickým umístěním v konceptuálním nízdídimenzionálním prostoru, takže vzdálenosti mezi body v prostoru odpovídají danému (dis) podobnostem co nejblíže. Výsledkem je znázornění nejmenších čtverců řádku a sloupce objektů v malém dimenzionálním prostoru, což vám v mnoha případech pomůže blíže porozumět vašim datům.

Vztah k jiným postupům kategorie. Pokud se vaše data skládají ze vzdáleností mezi jednoduchou sadou objektů (čtverec, symetrická matice), použijte Multidimensional Scaling (Multidimensional Scaling).

Vztah ke standardním technikám. Kategorie Multidimensional Unfolding procedure (PREFSCAL) nabízí několik vylepšení v případě neskládací funkčnosti dostupné ve volbě Základní statistika (prostřednictvím ALSCAL). PREFSCAL vám umožňuje klást omezení na společný prostor; navíc, PREFSCAL se snaží minimalizovat penalizující stresové opatření, které mu pomáhá vyhnout se degenerování řešení (ke kterým jsou náchylné starší algoritmy náchylné).

Poměr stran v optimálních grafech škálování

Poměr stran v optimálních parních plochách je izotropní. V dvourozměrném výkresu se vzdálenost reprezentující jednu jednotku v dimenzi 1 rovná vzdálenosti představující jednu jednotku v dimenzi 2. Změníte-li rozsah dimenze v dvourozměrném výkresu, systém změní velikost druhé dimenze, aby se udržela fyzická vzdálenost. Izotropický poměr stran nemůže být potlačen u optimálních postupů škálování.

Kategorická regrese (CATREG)

Kategorická regrese kvantifikuje kategoriální data přiřazením číselných hodnot do kategorií, což má za následek optimální lineární regresní rovnici pro transformované proměnné. Kategorická regrese je také známa zkratkou CATREG, pro *categorical reg*regrese.

Standardní lineární regresní analýza zahrnuje minimalizaci součtu čtvercových rozdílů mezi proměnnou odezvy (závislá) a váženou kombinací prediktorů (nezávislých) proměnných. Proměnné jsou obvykle kvantitativní, s (nominálními) kategoriálními daty převrácenými do binárních nebo kontrastních proměnných. V důsledku toho kategoričné proměnné slouží k oddělení skupin případů a technika odhaduje samostatné sady parametrů pro každou skupinu. Odhadnuté koeficienty odrážejí, jak změny v prediktorů ovlivňují odezvu. Předpověď odezvy je možné pro jakoukoliv kombinaci hodnot prediktoru.

Alternativním přístupem je regresi odezvy na kategoriální prediktory hodnot samotné. V důsledku toho se pro každou proměnnou odhaduje jeden koeficient. Nicméně pro kategoriální proměnné jsou hodnoty kategorie libovolné. Kódování kategorií různými způsoby přináší různé koeficienty, takže srovnání v rámci analýz stejných proměnných je obtížné.

Produkt CATREG rozšiřuje standardní přístup tím, že současně škáluje nominální, ordinální a číselné proměnné. Procedura kvantifikuje kategoriální proměnné tak, aby kvanta odrážel charakteristiky původních kategorií. Procedura považuje kvantifikované kategoričné proměnné stejným způsobem jako číselné proměnné. Použití nelineárních transformací umožňuje analýzu proměnných na různých úrovních, aby bylo nalezeno nejvhodnější model.

Příklad. Kategorická regrese by mohla být použita k popisu, jak uspokojování úloh závisí na kategorii úloh, geografické oblasti a rozsahu cesty. Možná zjistíte, že vysoké úrovně spokojenosti odpovídají manažerům a nízkým cestovům. Výsledná regresní rovnice by mohla být použita k předpovědi spokojenosti pracovních míst pro jakoukoli kombinaci těchto tří nezávislých proměnných.

Statistika a grafy. Frekvence, regresní koeficienty, tabulka ANOVA, historie iterace, vyčíslení kategorií, korelace mezi netransformovanými prediktory, korelace mezi transformovanými prediktory, reziduální grafy a transformační grafy.

Aspekty kategoriálního regresního data

Data. Produkt CATREG pracuje s proměnnými indikátoru kategorie. Indikátory kategorií by měly být kladná celá čísla. Dialogové okno Diskretizace můžete použít k převodu proměnných zlomkové proměnné a řetězcových proměnných na kladná celá čísla.

Předpoklady. Je povolena pouze jedna proměnná odezvy, ale maximální počet proměnných prediktoru je 200. Data musí obsahovat alespoň tři platné případy a počet platných případů musí být vyšší než počet proměnných prediktoru plus jedna.

Související postupy. CATREG je ekvivalentní kategorické kanonické analýze korelace s optimálním škálováním (OVERALS) se dvěma sadami, z nichž jedna obsahuje pouze jednu proměnnou. Scaling all variables at the numerical level cords to standard multiple regression analysis

Jak získat kategorickou regresi

1. Z nabídky vyberte:

Analyzovat > Regrese > Optimální škálování (CATREG) ...

2. Vyberte závislou proměnnou a nezávislou proměnnou (s).

3. Klepněte na tlačítko **OK**.

Volitelně můžete změnit úroveň měřítka pro každou proměnnou.

Definovat měřítko v kategorickém regresí

Můžete nastavit optimální úroveň měřítka pro závislé a nezávislé proměnné. Ve výchozím nastavení jsou tyto monotónní spline (ordinální) se dvěma vnitřními uzly škálované. Kromě toho můžete nastavit váhu proměnných analýzy.

Optimální úroveň měřítka. Můžete také vybrat úroveň měřítka pro kvantifikaci každé proměnné.

- **Ordinal.** Pořadí kategorií sledované proměnné je zachováno v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace je hladký monotónní polynom na zvoleném stupni. Kusy jsou specifikovány uživatelem-zadaným číslem a procedurou-stanovené umístění vnitřních uzlíků.
- **Spline Nominal.** Jediné informace v pozorované proměnné, které jsou zachovány v optimálně proměnlivé proměnné, je seskupení objektů v kategoriích. Pořadí kategorií sledované proměnné není zachováno. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace je hladký, možná nonmonotónní, po kousku polynomiální polynomiální části zvoleného stupně. Kusy jsou specifikovány uživatelem-zadaným číslem a procedurou-stanovené umístění vnitřních uzlíků.
- **Pořadový.** Pořadí kategorií pozorovaných proměnných je zachováno v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace se hodí lépe než transformace spline, ale je méně hladká.
- **Jmenovitá hodnota.** Jediná informace v pozorované proměnné, která je zachována v optimálně proměnlivé proměnné, je seskupení objektů v kategoriích. Pořadí kategorií sledované proměnné není zachováno. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace se hodí lépe než nominální transformace sleziny, ale je méně hladká.
- **Číslo.** Kategorie se zpracovávají jako seřazené a rovnoměrně rozložené (úroveň intervalu). Pořadí kategorií a rovné vzdálenosti mezi čísly kategorií pozorovaných proměnných jsou zachovány v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Jsou-li všechny proměnné na číselné úrovni, je analýza analogická ke standardní analýze hlavních komponent.

Diskreční diskretnost regresní oblasti

Dialogové okno Diskretnost vám umožňuje vybrat způsob, jak se překódovat proměnné. Proměnné zlomkové hodnoty jsou seskupeny do sedmi kategorií (nebo do počtu odlišených hodnot proměnné, pokud je toto číslo menší než sedm) s přibližně normálním rozdělením, není-li uvedeno jinak. Řetězcové proměnné jsou vždy převáděny na kladná celá čísla přiřazováním indikátorů kategorií podle vzestupného

alfanumerického pořadí. Diskretizace pro řetězcové proměnné platí pro tato celá čísla. Ostatní proměnné jsou při výchozím nastavení ponechány samostatně. V analýze se pak použijí diskretizované proměnné.

. Vyberte si mezi skupinami, očíslování pořadí a vynásobením.

- **Seskupení.** Překóduje se do zadaného počtu kategorií nebo se přečísluje intervalem.
- **Hodnocení.** Proměnná je diskretizována podle důležitosti případů.
- **Vynásobením.** Aktuální hodnoty proměnné jsou standardizovány, násobené 10, zaokrouhleny a mají konstantní přidanou hodnotu, takže nejnižší diskretizovaná hodnota je 1.

Seskupení. Následující volby jsou k dispozici při diskretizaci proměnných podle seskupení:

- **Počet kategorií.** Zadejte počet kategorií a zda mají být hodnoty proměnné v rámci těchto kategorií přibližně normální nebo rovnoměrné rozdělení.
- **Rovnoměrné intervaly.** Proměnné se překotnou do kategorií definovaných v těchto stejně velkých intervalech. Musíte uvést délku intervalů.

Kategorická regrese chybějících hodnot

Dialogové okno Chybějící hodnoty umožňuje zvolit strategii pro zpracování chybějících hodnot v proměnných analýzy a doplňujících proměnných.

Strategie. Chcete-li vyloučit objekty s chybějícími hodnotami (deletwise deletion) nebo impute chybějící hodnoty (aktivní léčba).

- **Vyloučit objekty s chybějícími hodnotami u této proměnné.** Objekty s chybějícími hodnotami na vybrané proměnné jsou vyloučeny z analýzy. Tato strategie není k dispozici pro doplňkové proměnné.
- **Chybějící hodnoty Impute.** Objekty s chybějícími hodnotami na zvolené proměnné mají imputované hodnoty. Můžete zvolit metodu imputace. Vyberte **Režim**, chcete-li nahradit chybějící hodnoty nejčetnější kategorií. Je-li více režimů, použije se ten s nejmenším indikátorem kategorie. Vyberte **Přebytečná kategorie**, chcete-li nahradit chybějící hodnoty se shodným kvantifikací zvláštní kategorie. To znamená, že objekty s chybějící hodnotou na této proměnné jsou považovány za objekty patřící do stejné kategorie (extra).

Kategorické volby regrese

Dialogové okno Volby vám umožňuje vybrat počáteční styl konfigurace, uvést iteraci a konvergenční kritéria, vybrat doplňkové objekty a nastavit popisky grafů.

Doplňkové objekty. To vám umožňuje uvést objekty, které chcete považovat za doplňkové. Jednoduše zadejte číslo doplňkového objektu (nebo uveďte rozsah případů) a klepněte na **Přidat**. Doplňkové objekty typu weight (určená váha jsou ignorovány).

Počáteční konfigurace. Pokud se s žádnými proměnnými zachází jako s nominální hodnotou, vyberte konfiguraci **Číselná**. Je-li alespoň jedna proměnná považována za nominální hodnotu, vyberte konfiguraci **Náhodná**.

Případně, pokud má alespoň jedna proměnná úroveň ordinální nebo spline ordinální velikosti, může běžný model vybavování modelů způsobit suboptimální řešení. Při výběru volby **Vícenásobné systematické spuštění** se všemi možnými vzory podpisu se vždy najde optimální řešení, ale doba potřebná pro zpracování se rychle zvyšuje, když se zvýší počet pořadových a třířadných proměnných v množném množině dat. Počet vzorků testu můžete snížit tak, že uvedete procentní část prahové hodnoty ztráty odchylky, kdy čím vyšší je prahová hodnota, tím více vzorků podpisu bude vyloučeno. S touto volbou, získání optimálního řešení není garantueed, ale šance na získání suboptimální řešení je snížena. Také v případě, že optimální řešení není nalezeno, je šance, že suboptimální řešení je velmi odlišné od optimálního řešení se sníží. Je-li požadováno více systematických spuštění, jsou příznaky regresních koeficientů pro každý začátek zapsány do externího datového souboru IBM SPSS Statistics nebo datové sady v aktuální relaci. Další informace naleznete v tématu [“Kategorické ukládání regrese”](#) na stránce 13.

Výsledky předchozího spuštění s vícenásobným systematickým spuštěním vám umožní **Použít pevné příznaky pro regresní koeficienty**. Znamky (označené jako 1 a - 1) musí být v řádku zadané datové

sady nebo souboru. Počáteční číslo s hodnotou typu integer je číslo případu řádku v tomto souboru, který obsahuje znaky, které se mají použít.

Kritéria. Můžete určit maximální počet iterací, které může regrese projít ve svých výpočtech. Můžete také vybrat hodnotu kritéria konvergence. Regresní se zastaví iterování, pokud je rozdíl mezi posledními dvěma iteracemi menší než hodnota konvergence, nebo pokud je dosažen maximální počet iterací.

Popisek se kreslí podle. Umožňuje určit, zda se v grafech použijí proměnné a názvy proměnných nebo názvy proměnných, a hodnoty. Můžete také uvést maximální délku pro popisky.

Kategorická regularizace regresních regresí

. Regulační metody mohou zlepšit prediktivní chybu modelu tím, že snižují variabilitu v odhadech regresního koeficientu tím, že se odhady sníží na 0. Lasso a Elastic Net bude zmenšovat některé koeficienty odhad na přesně 0, čímž se vytvoří tvar proměnné výběru. Je-li požadována metoda regulizace, regulizovaný model a koeficienty pro každou hodnotu koeficientu pokuty se zapíše do externího datového souboru IBM SPSS Statistics nebo datové sady v aktuální relaci. Další informace naleznete v tématu "[Kategorické ukládání regresí](#)" na stránce 13 .

- **regrese Ridge.** Ridge regression shrnlinks koeficienty by introducing a penalty term equal to the sum of squared coefficients times a **penalizační koeficient** Tento koeficient se může pohybovat v rozsahu od 0 (žádná sankce) do 1; procedura vyhledá "nejlepší" hodnotu penále, pokud uvedete rozsah a přírůstek.
- **Lasso.** Lasův trest je založen na součtu absolutních koeficientů, a specifikace penále koeficientu je podobný tomu, že Ridge regresi, nicméně, Lasso je více výpočetně intenzivní.
- **Elastická síť.** Elastic Net prostě kombinuje Lasso a Ridge regresní sankce, a bude hledat nad mířížkou hodnot stanovených najít "nejlepší" Lasso a Ridge regresní koeficienty penaly. Pro daný pár lasso a Ridge regresní pokuty, Elastická síť není mnohem více výpočetně drahé než lasso.

Zobrazit regularizační grafy. Jedná se o grafy regresních koeficientů versus regularizační trest. Při vyhledávání rozsahu hodnot pro "nejlepší" koeficient penále poskytuje pohled na to, jak se regresní koeficienty mění v daném rozsahu.

Elastické čisté paty. Pro metodu Elastic Net jsou oddělené zkusné plochy produkovány hodnotami sankce Ridge Regrese penaly. **Všechny možné grafy** používá každou hodnotu v rozsahu určeném hodnotami minima a maximální regresní regrese Ridge uvedené. **V případě některých penále** vám umožňuje uvést podmnožinu hodnot v rozsahu určeném minimem a maximem. Jednoduše zadejte hodnotu penále (nebo uveďte rozsah hodnot) a klepněte na **Přidat**.

Kategorický regresní výstup

Dialogové okno Výstup vám umožňuje vybrat statistiky, které se mají zobrazit ve výstupu.

Tabulky. Vytváří tabulky pro:

- **Multiple R.** zahrnuje P^2 , upravené P^2 a upravené P^2 s optimálním škálováním v úvahu.
- **ANOVA.** Tato volba zahrnuje regresní a zbytkové součty čtverců, středních čtverců a F . Jsou zobrazeny dvě tabulky ANOVA: jeden se stupněm volnosti pro regresi se rovná počtu proměnných prediktoru a jeden se stupněm volnosti pro regresi při použití optimálního škálování v úvahu.
- **Koeficienty.** Tato volba dává tři tabulky: tabulku koeficientů, která obsahuje beta, standardní chyba hodnot beta, t hodnot a významnost; koeficienty-optimální škálování s standardní chybou beta při použití optimálního škálování stupně volnosti; a tabulky s nulovým pořadím, částečným a částečným korelací, relativní důležitostí Pratt pro transformované prediktory a tolerance před transformací a po transformaci.
- **Historie iterací.** Pro každou iteraci, včetně počátečních hodnot algoritmu, se zobrazí více chyb R a regrese. Zvýšení více R je vypsáno od první iterace.
- **Korelace původních proměnných.** Zobrazí se matice zobrazující korelace mezi netransformovanými proměnnými.
- **Korelace transformovaných proměnných.** Zobrazí se matice zobrazující korelace mezi transformovanými proměnnými.

- **Regularizované modely a koeficienty.** Zobrazí hodnoty pokut, R-čtvereček a regresní koeficienty pro každý zlegalizovaný model. Je-li zadána metoda resampling nebo jsou-li zadány doplňkové objekty (testovací případy), zobrazí se také chyba předpovědi nebo testování MSE.

Odbavení. Metody opětovného zpracování vám poskytují odhad predikce chyby modelu.

- **Křížové ověření.** Crossvalidation rozděljuje vzorek na několik dílčích vzorků nebo záhyby. Následně jsou generovány kategoricky regresní modely s vyloučením dat z jednotlivých dílčích vzorků postupně. První model je založen na všech případech s výjimkou těch, které se nacházejí v prvním vzorku, druhý model je založen na všech případech s výjimkou těch, které jsou uvedeny v druhém vzorku a tak dále. Pro každý model se odhaduje chyba předpovědi použitím modelu na dílčí ukázkou, která má být generována při generování.
- **Zaváděcí program .632.** Při použití samozaváděcího programu se pozorování kreslí náhodně z dat s náhradou, opakování tohoto procesu za účelem získání počtu vzorků pro samozavedení. Model se hodí pro každý vzorek zaváděcího programu. Chyba predikce pro každý model se odhaduje použitím namontovaného modelu na případy, které nejsou v ukázce zaváděcího programu.

Kategorie Quantifications. Zobrazí se **tabulky** zobrazující transformované hodnoty vybraných proměnných.

Deskriptivní statistika. Zobrazí se **tabulky** zobrazující frekvence, chybějící hodnoty a režimy vybraných proměnných.

Kategorické ukládání regresí

Dialogové okno Uložení vám umožňuje uložit předpokládané hodnoty, zbytkové chyby a transformované hodnoty do aktivní datové sady a/nebo uložit diskretizovaná data, transformované hodnoty, regularizované modely a koeficienty a příznaky regresních koeficientů na externí datový soubor nebo datovou sadu IBM SPSS Statistics v aktuální relaci.

- Datové sady jsou k dispozici během aktuální relace, ale nejsou k dispozici v následných relacích, pokud je výslovně neuložíte jako datové soubory. Názvy datových sad musí dodržovat pravidla pojmenování proměnných.
- Názvy názvů souborů nebo názvů datových sad se musí lišit pro každý typ uložených dat.

Regularizované modely a koeficienty jsou ukládány, kdykoli je v dialogovém okně **Regularizace** vybrána metoda regulování. Při výchozím nastavení procedura vytvoří novou datovou sadu s jedinečným názvem, ale vy můžete samozřejmě zadat název vlastního výběru nebo zápisu do externího souboru.

Příznaky regresních koeficientů se ukládají při každém použití vícenásobného systematického spuštění jako počáteční konfigurace v dialogovém okně **Volby**. Při výchozím nastavení procedura vytvoří novou datovou sadu s jedinečným názvem, ale vy můžete samozřejmě zadat název vlastního výběru nebo zápisu do externího souboru.

Kategorické transformace regresní transformace

Dialogové okno **Plot** vám umožňuje uvést proměnné, které budou produkovat transformaci a zbytkové grafy.

Vykreslované prvky transformace. Pro každou z těchto proměnných jsou vyčíslení kategorií vykreslena proti původním hodnotám kategorie. Prázdné kategorie se objevují na vodorovné ose, ale nemají vliv na výpočty. Tyto kategorie jsou identifikovány zalomením v řádku, který spojuje kvantily.

Zbývající počet částí. Pro každou z těchto proměnných jsou zbytkové hodnoty (vypočítány pro závislou proměnnou predikované ze všech proměnných prediktorů kromě proměnné prediktoru) vykresleny proti indikátorům kategorií a optimální kvantita kategorie násobená beta proti indikátorům kategorie.

Další funkce příkazu CATREG

Tuto kategorickou regresí můžete upravit, pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu CATREG. Jazyk syntaxe příkazu vám také umožňuje:

- Při ukládání do aktivní datové sady (s dílčím příkazem SAVE) určete názvy rootnames pro transformované proměnné.

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

Kategorická analýza hlavních komponent (CATPCA)

Tento postup současně kvantifikuje kategoriální proměnné a zároveň snižuje dimenzionalitu dat. Kategorické analýzy komponent činitele jsou známy také akronym CATPCA pro analýzu komponent *categorical principal*.

Cílem analýzy hlavních komponent je snížit původní sadu proměnných do menší sady nekorelovaných komponent, které představují většinu informací nalezených v původních proměnných. Tato technika je nejužitečnější, pokud velký počet proměnných zakazuje efektivní interpretaci vztahů mezi objekty (předměty a jednotky). Snížením dimenzionalitu interpretujete spíše několik komponent než velké množství proměnných.

Analýza standardních komponent hlavního činitele předpokládá lineární vztahy mezi numerickými proměnnými. Optimální škálování umožňuje na druhé straně proměnné, které mají být škálovány na různých úrovních. Kategorické proměnné jsou optimálně kvantifikovány v zadané dimenzionalitě. V důsledku toho lze modelovat nelineární vztahy mezi proměnnými.

Příklad. Kategorická analýza hlavních komponent by mohla být použita ke grafickému zobrazení vztahu mezi kategorií úloh, oddělením úloh, oblastí, množstvím cest (vysoká, střední a nízká) a uspokojením pracovních míst. Možná zjistíte, že dva rozměry představují velké množství rozptylu. První rozměr může oddělit pracovní kategorii od regionu, zatímco druhá dimenze může oddělení pracovních míst oddělit od rozsahu cesty. Také můžete zjistit, že vysoká spokojenost s uspokojením souvisí se střední částkou cesty.

Statistika a grafy. Frekvence, chybějící hodnoty, optimální úroveň škálování, režim, odchylka zaúčtovaná centroid souřadnice, vektorové souřadnice, celková hodnota na proměnné a na rozměr, složka zatížení pro vektor-kvantifikované proměnné, kategorie kvantity a souřadnice, iterace historie, korelace transformovaných proměnných a eigenvalues of the correlation matrix, correlations of the object scores, category plots, joint caparts, transformation plots, residual plots, projected centroid parts, object plots, residual parts, conaded centroid plots, object plots, residual plots, concented centroid plots, object plots, Parcely, triploty a lomeniny komponenty.

Posouzení dat analýzy hlavních komponent hlavních komponent

Data. Hodnoty řetězcové proměnné jsou vždy převedeny na kladná celá čísla ve vzestupném alfanumerickém pořadí. Uživatelem definované chybějící hodnoty, systémově chybějící hodnoty a hodnoty menší než 1 se považují za chybějící; můžete překódovat nebo přidat konstantu do proměnných s hodnotami nižšími než 1, aby se z nich nechyběly.

Předpoklady. Data musí obsahovat alespoň tři platné případy. Analýza je založena na kladných celočíselných datech. Volba diskretizace automaticky kategorizuje zlomkovou proměnnou tak, že seskupí její hodnoty do kategorií s blízkým "normálním" distribucí a automaticky převede hodnoty řetězcových proměnných na kladná celá čísla. Můžete uvést jiné systémy diskretizace.

Související postupy. Rozšiřitelnost všech proměnných na číselné úrovni odpovídá standardní analýze komponent činitele. Alternativní mapovací funkce jsou k dispozici za použití transformovaných proměnných ve standardní lineární analýze hlavních komponent. Pokud mají všechny proměnné více nominálních úrovní škálování, analýza komponent kategorických činitelů je identická s více korespondenční analýzou. Pokud sady proměnných jsou předmětem zájmu, měla by být použita metoda kanonické korelace (nelineární).

Chcete-li získat analýzu kategorických komponent hlavní komponenty

1. Z nabídky vyberte:

Analyzovat > Redukce dimenzí > Optimální škálování ...

2. Vyberte volbu **Některá proměnná (proměnné) není vícenásobná.**
3. Vyberte volbu **Jedna sada.**

4. Klepněte na tlačítko **Definovat**.
5. Vyberte alespoň dvě proměnné analýzy a určete počet dimenzí v řešení.
6. Klepněte na tlačítko **OK**.

Volitelně můžete uvést doplňkové proměnné, které jsou umístěny do nalezeného řešení, nebo popisky pro grafy.

Definování měřítka a váhy v produktu CATPCA

Můžete nastavit optimální úroveň měřítka pro proměnné analýzy a doplňkové proměnné. Ve výchozím nastavení jsou tyto monotónní spline (ordinální) se dvěma vnitřními uzly škálované. Kromě toho můžete nastavit váhu proměnných analýzy.

Proměnná váha. Můžete zvolit, zda má být pro každou proměnnou definována váha. Uvedená hodnota musí být kladné celé číslo. Výchozí hodnota je 1.

Optimální úroveň měřítka. Můžete také vybrat úroveň měřítka, která má být použita ke kvantifikaci jednotlivých proměnných.

- **Křivka spline.** Pořadí kategorií sledované proměnné je zachováno v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace je hladký monotónní polynom na zvoleném stupni. Kusy jsou specifikovány uživatelem-zadaným číslem a procedurou-stanovené umístění vnitřních uzlíků.
- **Spline nominal.** Jediné informace v pozorované proměnné, které jsou zachovány v optimálně proměnlivé proměnné, je seskupení objektů v kategoriích. Pořadí kategorií sledované proměnné není zachováno. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace je hladký, možná nonmonotónní, po kousku polynomiální polynomiální části zvoleného stupně. Kusy jsou specifikovány uživatelem-zadaným číslem a procedurou-stanovené umístění vnitřních uzlíků.
- **Vícenásobný nominální.** Jediné informace v pozorované proměnné, které jsou zachovány v optimálně proměnlivé proměnné, je seskupení objektů v kategoriích. Pořadí kategorií sledované proměnné není zachováno. Kategorie bodů budou v centroidu objektů v jednotlivých kategoriích. *Vícenásobný* označuje, že pro každou dimenzi jsou získány různé sady kvantitů.
- **Pořadový.** Pořadí kategorií pozorovaných proměnných je zachováno v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace se hodí lépe než transformace spline, ale je méně hladká.
- **Jmenovitá hodnota.** Jediná informace v pozorované proměnné, která je zachována v optimálně proměnlivé proměnné, je seskupení objektů v kategoriích. Pořadí kategorií sledované proměnné není zachováno. Body kategorií budou na přímce (vektor) prostřednictvím původu. Výsledná transformace se hodí lépe než nominální transformace sleziny, ale je méně hladká.
- **Číslo.** Kategorie se zpracovávají jako seřazené a rovnoměrně rozložené (úroveň intervalu). Pořadí kategorií a rovné vzdálenosti mezi čísly kategorií pozorovaných proměnných jsou zachovány v optimálně proměnlivé proměnné. Body kategorií budou na přímce (vektor) prostřednictvím původu. Jsou-li všechny proměnné na číselné úrovni, je analýza analogická ke standardní analýze hlavních komponent.

Diskreční diskretizace v kategorii Hlavní komponenty

Dialogové okno Diskrétnost vám umožňuje vybrat způsob, jak se překódovat proměnné. Proměnné s zlomkovou hodnotou jsou seskupeny do sedmi kategorií (nebo do počtu odlišených hodnot proměnné, pokud je tento počet menší než sedm) s přibližně normálním rozdělením, pokud není uvedeno jinak. Řetězcové proměnné jsou vždy převáděny na kladná celá čísla přiřazením indikátorů kategorií podle vzestupného alfanumerického pořadí. Diskretizace pro řetězcové proměnné platí pro tato celá čísla. Ostatní proměnné jsou při výchozím nastavení ponechány samostatně. V analýze se pak použijí diskretizované proměnné.

- Vyberte si mezi skupinami, očíslování pořadí a vynásobením.
- **Seskupení.** Překóduje se do zadaného počtu kategorií nebo se přečísluje intervalem.
- **Hodnocení.** Proměnná je diskretizována podle důležitosti případů.

- **Vynásobení.** Aktuální hodnoty proměnné jsou standardizovány, násobené 10, zaokrouhleny a mají konstantu přidanou tak, že nejnižší diskretizovaná hodnota je 1.

Seskupení. Následující volby jsou k dispozici, když provádíte diskretizace proměnných podle seskupení:

- **Počet kategorií.** Zadejte počet kategorií a zda mají být hodnoty proměnné v rámci těchto kategorií přibližně normální nebo rovnoměrné rozdělení.
- **Rovnoměrné intervaly.** Proměnné se překotnou do kategorií definovaných v těchto stejně velkých intervalech. Musíte uvést délku intervalů.

Analýza kategorických komponent hlavní komponenty-chybějící hodnoty

Použijte dialogové okno Chybějící hodnoty k výběru strategie pro zpracování chybějících hodnot v proměnných analýzy a doplňujících proměnných.

Strategie. Vyberte vyloučení chybějících hodnot (pasivní léčba), impute chybějící hodnoty (aktivní léčba) nebo vynechte objekty s chybějícími hodnotami (odstranění lišových vláken).

- **Vyloučit chybějící hodnoty; pro korelace impute po kvantifikaci.** Objekty s chybějícími hodnotami u vybrané proměnné nepříspějí k analýze této proměnné. Jsou-li použity všechny proměnné pasivní léčba, pak s objekty s chybějícími hodnotami na všech proměnných se zachází jako s doplňkovými hodnotami. Jsou-li korelace zadány v dialogovém okně Výstup, pak (po analýze) chybějící hodnoty jsou imputovány s nejčastějším kategoriím nebo režimem proměnné pro korelace původních proměnných. Pro korelace optimálně škálovaných proměnných si můžete zvolit metodu imputace.
 - **Režim.** Nahradte chybějící hodnoty v režimu optimálně proměnlivé proměnné.
 - **Další kategorie.** Nahradte chybějící hodnoty kvantifikací přebytečné kategorie. Toto nastavení znamená, že objekty s chybějící hodnotou na této proměnné jsou považovány za objekty náležející do stejné kategorie (extra).
 - **Náhodná kategorie.** Impute každou chybějící hodnotu na proměnné s kvantifikovanou hodnotou jiného náhodného čísla kategorie na základě mezních frekvencí kategorií dané proměnné.
- **Chybějící hodnoty Impute.** Objekty s chybějícími hodnotami na zvolené proměnné mají imputované hodnoty. Můžete zvolit metodu imputace.
 - **Režim.** Nahradte chybějící hodnoty nejčastější kategorií. Je-li více režimů, použije se ten s nejmenším indikátorem kategorie.
 - **Další kategorie.** Nahradte chybějící hodnoty shodnými hodnotami kvantifikace zvláštní kategorie. Toto nastavení znamená, že objekty s chybějící hodnotou na této proměnné jsou považovány za objekty náležející do stejné kategorie (extra).
 - **Náhodná kategorie.** Každá chybějící hodnota na proměnné nahradte jinou číselnou hodnotou kategorie založenou na okrajových frekvencích kategorií.
- **Vyloučit objekty s chybějícími hodnotami u této proměnné.** Objekty s chybějícími hodnotami na vybrané proměnné jsou vyloučeny z analýzy. Tato strategie není k dispozici pro doplňkové proměnné.

Volby analýzy Kategorických komponent hlavní komponenty

Dialogové okno Volby poskytuje ovládací prvky pro výběr počáteční konfigurace, uvedení iterací a konvergenčních kritérií, vyberte metodu normalizace, zvolte metodu pro označování grafů a uveďte doplňkové objekty.

Doplňkové objekty. Určete číslo případu objektu nebo první a poslední číslo případu rozsahu objektů, které chcete provést jako doplňkové, a poté klepněte na tlačítko **Přidat**. Je-li objekt zadán jako doplňkový, pak jsou pro tento objekt ignorována váha případu.

Normalizační metoda. Můžete určit jednu z pěti voleb pro normalizaci skóre objektu a proměnných. V každé analýze může být použita pouze jedna normalizační metoda.

- **Činitel proměnné.** Tato volba optimalizuje přidružení mezi proměnnými. Souřadnice proměnných v prostoru objektů jsou načtenými komponenty (korelace s hlavními komponentami, jako např. skóre dimenzí a objektů). Tato metoda je užitečná, když se primárně zajímáte o korelaci mezi proměnnými.

- **Objekt Principal.** Tato volba optimalizuje vzdálenosti mezi objekty. Tato metoda je užitečná, když se zajímáte především o rozdíly nebo podobnosti mezi objekty.
- **Symmetrický.** Tuto normalizační volbu použijte v případě, že se zajímáte především o vztah mezi objekty a proměnnými.
- **Nezávislá.** Tuto normalizační volbu použijte, chcete-li zkoumat vzdálenosti mezi objekty a korelacemi mezi proměnnými samostatně.
- **Vlastní.** V uzavřeném intervalu $[-1, 1]$ můžete zadat libovolnou reálnou hodnotu. Hodnota 1 je rovna metodě Principal objektu. Hodnota 0 odpovídá metodě Symmetric. Hodnota -1 se rovná metodě Variable Principal. Zadáte-li hodnotu větší než -1 a menší než 1, můžete rozložit hodnotu eigenvalue nad objekty i proměnné. Tato metoda je užitečná pro výrobu fázového a fázového souboru s bitem nebo tripartem.

Kritéria. Můžete určit maximální počet iterací, které mohou procedury projít ve svých výpočtech. Můžete také vybrat hodnotu kritéria konvergence. Algoritmus přestane opakovat, pokud je rozdíl mezi posledními dvěma iteracemi menší než hodnota konvergence, nebo pokud je dosažen maximální počet iterací.

Popisek se kreslí podle. Můžete určit, zda se v grafech budou používat proměnné a popisky hodnot nebo názvy proměnných a hodnoty. Můžete také uvést maximální délku pro popisky.

Zákresové dimenze. Dimenze zobrazené ve výstupu můžete řídit.

- **Zobrazit všechny dimenze v řešení.** Všechny dimenze v řešení se zobrazí v matici bodového grafu.
- **Omezte počet dimenzí.** Zobrazené dimenze jsou omezeny na vykreslené dvojice. Pokud omezíte rozměry, musíte vybrat nejnižší a nejvyšší dimenze, které mají být zakresleny. Nejnižší dimenze může být v rozsahu od 1 do počtu dimenzí v řešení minus 1 a je vykreslena proti vyšším rozměrům. Nejvyšší rozměrová hodnota může být v rozsahu od 2 do počtu dimenzí v řešení a označuje nejvyšší dimenzi, která má být použita při vykreslování dvojic dimenze. Tato specifikace se vztahuje na všechny požadované vícerozměrné grafy.

Rotace. Chcete-li dosáhnout rotovaných výsledků, můžete vybrat metodu rotace.

Poznámka: Tyto metody rotace nejsou dostupné, pokud v dialogovém okně **Zaváděcí program** vyberete volbu **Provést bootstrapping**.

- **Varimax.** Metoda rotace ortogonální rotace, která minimalizuje počet proměnných, které mají vysoké zatížení na každé komponentě. Zjednodušuje interpretaci komponent.
- **Kvartimax.** Metoda rotace, která minimalizuje počet komponent potřebných k vysvětlení každé proměnné. Zjednodušuje interpretaci sledovaných proměnných.
- **Equamax.** Metoda rotace, která je kombinací metody Varimax, která zjednodušuje komponenty, a metoda Quartimax, která zjednodušuje proměnné. Počet proměnných, které načítají vysokou hodnotu komponenty a počet komponent potřebných k vysvětlení proměnné jsou minimalizovány.
- **Oblimová.** Metoda rotace šikmé (neortogonální). Je-li rozdílová data rovna 0, komponenty jsou nejvíce šikmé. Když se Δ stane více negativní, složky se stanou méně nakloněné. Kladné hodnoty povolují další korelaci komponent. Hodnota parametru **Delta** musí být menší nebo rovna 0,8.
- **Promax.** Rotace šikmé (neortogonální), která umožňuje korelaci komponent. Může být vypočítán rychleji než přímá Oblimonová rotace, takže je užitečné pro velké datové sady. Množství korelace (šikšivost), které je povoleno, je řízeno parametrem kappa. Hodnota **Kappa** musí být větší než nebo rovna 1 a méně 10 000.

Konfigurace. Můžete číst data ze souboru, který obsahuje souřadnice konfigurace. První proměnná v souboru obsahuje souřadnice pro první dimenzi. Druhá proměnná obsahuje souřadnice pro druhou dimenzi a tak dále.

- **Počáteční.** Jako počáteční bod analýzy se použije konfigurace v souboru, který je zadán.
- **Opraveno.** Konfigurace v souboru, který je zadán, se použije pro přizpůsobení proměnných. Proměnné, které jsou instalovány, musí být vybrány jako proměnné analýzy, ale protože je konfigurace opravena, se s nimi zachází jako s doplňujícími proměnnými (takže nemusí být jako doplňkové proměnné vybrány).

Výstup analýzy hlavních komponent hlavních komponent

Zobrazí se dialogové okno Výstup, které bude určovat zobrazení výsledků.

Skóre objektu. Zobrazí skóre objektu a má následující volby:

- **Zahrnout kategorie.** Zobrazí indikátory kategorie vybraných proměnných analýzy.
- **Popisek objektu Skóre podle.** Ze seznamu proměnných, které jsou určeny jako proměnné popisky, můžete vybrat jeden objekt, který bude označovat objekty jako popisky.

Načtení komponenty. Zobrazí zavedení komponenty pro všechny proměnné, kterým nebylo poskytnuto více nominálních úrovní škálování. Založení komponenty můžete řadit podle velikosti.

Historie iterací. Pro každou iteraci se zobrazí odchylka, která se zaúčtovala, ztráta a zvýšení rozptylu.

Korelace původních proměnných. Zobrazí korelační matici původních proměnných a eigenvalues této matrice.

Korelace transformovaných proměnných. Zobrazí korelační matici transformovaných (optimálně změněných) proměnných a ideových hodnot této matrice.

Odchylka účtována. Zobrazí množství rozptylu zachycený centroid souřadnice, vektorové souřadnice a celkový součet (souřadnice centroid a vektorové souřadnice) na proměnnou a na rozměr.

Kategorií kategorií. Poskytuje kategorie kvantování a souřadnice pro každou dimenzi proměnných, které jsou vybrány.

Deskriptivní statistika. Zobrazí frekvence, počet chybějících hodnot a režim vybraných proměnných.

Uložení analýzy hlavních komponent hlavních komponent

Ovládací prvky v dialogovém okně Uložení ukládají diskrétní data, skóre objektů, transformované hodnoty a jiné výsledky do aktivní datové sady, nové datové sady v aktuální relaci nebo externího souboru.

- Datové sady jsou k dispozici během aktuální relace, ale nejsou k dispozici v následných relacích, pokud je výslovně neuložíte jako datové soubory. Názvy datových sad musí dodržovat pravidla pojmenování proměnných.
- Názvy názvů souborů nebo názvů datových sad se musí lišit pro každý typ uložených dat.
- Uložíte-li skóre objektu nebo transformované hodnoty do aktivní datové sady, můžete zadat počet více nominálních rozměrů.
- Volby ve skupině **Elipsa důvěry pro zaváděcí program** jsou dostupné pouze v případě, že v dialogovém okně Zaváděcí program vyberete volbu **Provést bootstrappování**.

Kategorické objekty analýzy objektů hlavních komponent

Dialogové okno Objekt a Proměnné dávky vám umožňuje určit typy grafů, které chcete, a proměnné, pro které budou vykreslené grafy vytvořeny.

Objektové body. Zobrazí se zakres bodů objektu.

Objekty a proměnné (biflot). Objektové body jsou vykresleny podle vašeho výběru proměnných souřadnic -- component loadings nebo variable centroids.

Objekty, loadings a centroids (triplot). Objektové body jsou vyneseny s centroidy více proměnných nominálního škálování a zavedení komponent z jiných proměnných.

Proměnné oblasti Biplot a Triplot. Můžete zvolit použití všech proměnných pro dvousložky a triploty nebo vybrat dílčí sadu.

Objekty štítku. Můžete zvolit, zda mají být objekty označené kategoriemi vybraných proměnných (můžete zvolit hodnoty indikátoru kategorie nebo popisek hodnot v dialogovém okně Volby) nebo s jejich čísly případů. Je-li vybrána volba **Proměnná**, vytvoří se jeden zakres za každou proměnnou.

Kategorie analýzy kategorických komponent hlavní komponenty

Dialogové okno Plot kategorie vám umožňuje uvést typy grafů, které chcete, a proměnné, pro které budou vykreslené grafy vytvořeny.

Ploty kategorie. Pro každou vybranou proměnnou je vykreslen graf centroid a vektorové souřadnice. Pro proměnné s více nominálními úrovněmi měřítka jsou kategorie v centrojích objektů v jednotlivých kategoriích. Pro všechny ostatní úrovně měřítka jsou kategorie na vektoru přes výchozí bod.

Společné partie kategorie. Toto je jeden zákres hodnot těžiště a vektorové souřadnice každé zvolené proměnné.

Vykreslované prvky transformace. Zobrazí zákresový graf optimálních kvantit kategorie ve srovnání s indikátory kategorie. Můžete určit počet dimenzí pro proměnné s více nominálními úrovněmi škálování; pro každou dimenzi bude vygenerován jeden zákres. Můžete také zvolit zobrazení zbytkových grafů pro každou vybranou proměnnou.

Centroidy projektu. Pro vybrané proměnné můžete zvolit proměnnou a projekt jako centroidy. Proměnné s více nominálními úrovněmi měřítka nelze vybrat k projektu. Když je tento plot žádán, zobrazí se také tabulka se souřadnicemi projektovaných centroidů.

Kategorické analýzy hlavních komponent-naložení šarží

Dialogové okno Započtení režijních nákladů řídí proměnné, které jsou zahrnuty v zákresu, zobrazení centroidů v zákresovém grafu a zobrazení zakreslených ploch rozptylu.

Odchylka účtována. Pro každou dimenzi se zobrazí zákres rozptylu, který je k dispozici.

Zobrazte načtení komponent. Je-li vybrána tato volba, zobrazí se zákres zatížení komponenty.

Načítání proměnných. Můžete zvolit, zda chcete použít všechny proměnné pro vykreslení komponenty nebo vybrat dílčí sadu.

Zahrnout centroidy. Proměnné s více nominálními úrovněmi měřítka nemají naložení komponenty, ale můžete zvolit zahrnutí centroidů těchto proměnných do zákresu. Můžete zvolit použití všech více nominálních proměnných nebo výběr dílčí sady.

Zaváděcí program analýzy hlavních komponent hlavních komponent

Dialogové okno programu pro samozavedení určuje parametry pro analýzu zaváděcího programu.

Provést bootstrapping. Provádí opětné připojení zaváděcího programu.

- Jsou-li požadovány grafy načtení, kategorie nebo skóre komponent, zobrazí se další grafy. V těchto grafech jsou zobrazeny body pro ukázkou dat a odhady pro zaváděcí program. Transformační grafy zahrnují regiony důvěry. Zobrazí se také zákres pro hodnoty eigenvalues.
- Je-li uvedeno dvourozměrné řešení, jsou zobrazeny eliptické grafy s elipsou důvěry pro vlastní hodnoty, načtení komponent, body kategorie a body objektu.
- Pokud zadáte metodu rotace v dialogovém okně Volby, nebude zaváděcí program pro samozavedení k dispozici.

Můžete vybrat volbu **Vyvážený** nebo **Nevyvážený** bootstrapping.

Počet ukázek. Počet vzorků zaváděcího programu, které se používají k výpočtu odhadů zaváděcího programu. Hodnotou musí být kladné celé číslo.

Úroveň důvěry. Úroveň důvěry odhadů zaváděcího programu, vyjádřená jako procentní část. Hodnota musí být kladné číslo menší než 100.

Odpovídající metoda. Dvě alternativy jsou **Prokřížky** a **Reflektování**.

Sebehy s důvěrou. Řídí oblast prahové hodnoty pro elipsu důvěry v zakreslených plochách. Pro oblasti větší než (> operátor) uvedená hodnota se sníží počet výpustek, jak se zvyšuje hodnota prahové hodnoty.

- Nastavení v této skupině jsou k dispozici pouze v případě, že počet dimenzí, které jsou uvedeny v hlavním dialogovém okně, je 2.
- Tři tečky pro načítání grafů jsou k dispozici pouze v případě, že je vybrána volba **Zobrazit načtení komponenty** v dialogovém okně Zavádí se loje.
- Výpustky pro vykreslení objektů jsou k dispozici pouze v případě, že je v dialogovém okně Objekt a v políčkách proměnných vybráno **Objektové body**.
- Tři tečky pro vykreslení kategorie jsou k dispozici pouze v případě, že je v seznamu **Plot kategorie** v dialogovém okně Přepočítání kategorií určena jedna nebo více proměnných.

Počet obrysových bodů elipsy. Počet bodů výkresu použitých k vykreslení každé elipsy důvěry. Větší hodnoty produkují hladší elipsu. Hodnota musí být kladné celé číslo menší než nebo rovné 100.

Další funkce příkazu CATPCA

Analýzu základních komponent činitele můžete upravit, pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu CATPCA. Jazyk syntaxe příkazu vám také umožňuje:

- Uvedte rootnames pro transformované proměnné, skóre objektů a aproximace, když je ukládáte do aktivní datové sady (s dílčím příkazem SAVE).
- Určete maximální délku pro popisky pro každý výkres zvlášť (s dílčím příkazem PLOT).
- Uvedte samostatný seznam proměnných pro zbytkové grafy (s dílčím příkazem PLOT).

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

Analýza nelineárních kanonických korelací (OVERALS)

Nelineární kanonická korelační analýza odpovídá kategorické kanonické analýze korelace s optimálním škálováním. Účelem této procedury je určit, jak podobné sady kategorických proměnných mají být jedna. Nelineární kanonická korelační analýza je známa také zkratkou OVERALS.

Standardní kanonická korelační analýza je rozšířením vícenásobné regrese, kdy druhá sada neobsahuje jedinou proměnnou odezvy, ale místo toho obsahuje více proměnných odezvy. Cílem je co nejvíce vysvětlit odchylku ve vztazích mezi dvěma sadami numerických veličin v nízkém rozměrovém prostoru. Zpočátku, proměnné v každé sadě jsou lineárně kombinované tak, že lineární kombinace mají maximální korelace. Vzhledem k těmto kombinacím dochází k následné lineární kombinaci, které jsou nekorelované s předchozími kombinacemi a které mají největší možnou korelaci.

Optimální přístup škálování rozšiřuje standardní analýzu o tři klíčové způsoby. Za prvé, PŘEHLED povoluje více než dvě skupiny proměnných. Zadruhé, proměnné lze nastavit jako nominální, ordinální, nebo numerickou. V důsledku toho lze analyzovat nelineární vztahy mezi proměnnými. Nakonec místo maximalizace korelací mezi sadami proměnných se sady porovnávají s neznámou sadou kompromisů, která je definována skóre objektu.

Příklad. Kategorická analýza kanonické korelace s optimálním škálováním by mohla být použita ke grafickému zobrazení vztahu mezi jednou sadou proměnných obsahujících kategorii úlohy a roky vzdělávání a jinou sadou proměnných, které obsahují region bydliště a pohlaví. Možná zjistíte, že roky vzdělávání a region pobytu diskriminují lépe než zbývající proměnné. Můžete také zjistit, že roky vzdělávání diskriminují nejlepší na první rozměr.

Statistika a grafy. Frekvence, centroidy, historie iterace, skóre objektů, kategorie kvantity, váhy, načtení komponent, jednoduché a víceřadé, objektové skóre grafů, kategorie souřadnic plots, loadings komponenty, kategorie centroidů pozemků, transformační grafy.

Podmínky pro analýzu dat v nelineární kanonické analýze

Data. Použijte celá čísla, abyste kódlí kategoriální proměnné (nominální nebo ordinální úroveň měřítka). Chcete-li minimalizovat výstup, použijte po sobě jdoucí celá čísla začínající 1 pro kód každé proměnné. Proměnné, které jsou škálovány na numerické úrovni, by neměly být převráceny na následná celá čísla. Chcete-li minimalizovat výstup, pro každou proměnnou, která je zmenšena na číselné úrovni, odečtěte

nejmenší pozorovanou hodnotu od každé hodnoty a přidejte 1. Zlomkové hodnoty jsou za desetinnou čárkou oříznuty.

Předpoklady. Proměnné lze klasifikovat do dvou nebo více sad. Proměnné v analýze jsou škálované jako násobek nominálního, jednoho nominálního, ordinálního nebo numerického. Maximální počet dimenzí použitých v proceduře závisí na optimální úrovni měřítka proměnných. Jsou-li všechny proměnné zadány jako ordinální, jedna nominální nebo číselná, maximální počet dimenzí je menší z následujících dvou hodnot: počet pozorování minus 1 nebo celkový počet proměnných. Jsou-li však definovány pouze dvě sady proměnných, je maximální počet dimenzí počtem proměnných v menší sadě. Jsou-li některé proměnné více nominálně, je maximální počet rozměrů celkový počet vícenásobných nominálních kategorií plus počet nevícenásobných nominálních proměnných minus počet vícenásobných nominálních proměnných. Například, pokud analýza zahrnuje pět proměnných, z nichž jedna je vícenásobná nominální se čtyřmi kategoriemi, maximální počet rozměrů je $(4 + 4 - 1)$ nebo 7. Pokud uvedete číslo, které je větší než maximum, použije se maximální hodnota.

Související postupy. Pokud každá sada obsahuje jednu proměnnou, nelineární kanonická korelační analýza je ekvivalentní analýze základních komponent s optimálním škálováním. Je-li každá z těchto proměnných vícenominální, analýza odpovídá vícenásobné analýze korespondence. Pokud jsou zahrnuty dvě sady proměnných a jedna z těchto sad obsahuje pouze jednu proměnnou, je analýza identická s kategorickou regresí s optimálním škálováním.

Chcete-li získat nelineární kanonickou analýzu korelace

1. Z nabídky vyberte:

Analyzovat > Redukce dimenzí > Optimální škálování ...

2. Vyberte buď **Všechny proměnné více nominálně**, nebo **Některé proměnné nejsou vícenásobné nominální hodnoty**.

3. Vyberte volbu **Více sad**.

4. Klepněte na tlačítko **Definovat**.

5. Definujte alespoň dvě sady proměnných. Vyberte proměnnou (proměnné), kterou chcete zahrnout do první sady. Chcete-li přejít k další sadě, klepněte na tlačítko **Další** vyberte proměnné, které chcete zahrnout do druhé sady. Můžete přidat další sady. Klepnutím na tlačítko **Předchozí** se vrátíte do dříve definované sady proměnných.

6. Definujte rozsah hodnot a měřítko měření (optimální úroveň měřítka) pro každou vybranou proměnnou.

7. Klepněte na tlačítko **OK**.

8. Volitelně:

- Vyberte jednu nebo více proměnných pro poskytnutí bodových popisků pro grafy s bodovým skóre objektu. Každá proměnná vytvoří oddělený výkres, s body označenými hodnotami této proměnné. Pro každou z těchto proměnných návěští výkresu musíte definovat rozsah. Používáte-li dialogové okno, nelze v analýze použít jedinou proměnnou a jako proměnnou značení. Chcete-li označit objekt skóre skóre pomocí proměnné, která se používá v analýze, použijte pro vytvoření kopie této proměnné prostředek Compute (dostupný z nabídky Transform). Použijte novou proměnnou k označení zákresu. Případně lze použít syntaxi příkazu.
- Uvedte počet rozměrů, které chcete v řešení. Obecně platí, že je třeba zvolit si co nejméně rozměrů, aby bylo možné většinu variace vysvětlit. Pokud analýza zahrnuje více než dvě dimenze, vytvoří se tři rozměrové grafy prvních tří rozměrů. Ostatní dimenze lze zobrazit úpravou grafu.

Definovat rozsah a měřítko

Pro každou proměnnou musíte definovat rozsah. Uvedená maximální hodnota musí být celé číslo. Zlomkové datové hodnoty jsou v analýze zkráceny. Hodnota kategorie, která je mimo určený rozsah, je v analýze ignorována. Chcete-li minimalizovat výstup, použijte k vytvoření následných kategorií začínajících na 1 pro proměnné, které jsou považovány za nominální nebo ordinální, funkci Automatic Recode facility (dostupné z nabídky Transform). Přecenění na následná celá čísla se nedoporučuje pro proměnné, které jsou zmenšeny na numerické úrovni. Chcete-li minimalizovat výstup pro proměnné, které

jsou považovány za numerické, pro každou proměnnou, odečtete minimální hodnotu od každé hodnoty a přidejte 1.

Musíte také vybrat měřítko, které má být použito pro kvantifikaci každé proměnné.

- **Pořadový.** Pořadí kategorií pozorovaných proměnných je zachováno v kvantifikované proměnné.
- **Jeden nominální.** Ve kvantifikované proměnné obdrží objekty ve stejné kategorii stejné skóre.
- **Více nominálních hodnot.** Kvantifikace se mohou lišit pro každou dimenzi.
- **Diskrétní číslo.** Kategorie se zpracovávají jako seřazené a rovnoměrně rozložené. Rozdíly mezi čísly kategorií a pořadím kategorií pozorované proměnné jsou v kvantifikované proměnné uchovány.

Definovat rozsah

Pro každou proměnnou musíte definovat rozsah. Uvedená maximální hodnota musí být celé číslo. Zlomkové datové hodnoty jsou v analýze zkráceny. Hodnota kategorie, která je mimo určený rozsah, je v analýze ignorována. Chcete-li minimalizovat výstup, použijte funkci Automatic Recode (dostupná z nabídky Transform) k vytvoření následných kategorií začínajících na 1.

Musíte také definovat rozsah pro každou proměnnou, která se používá k označení grafů skóre objektů. Na plochách se však zobrazí popisky pro kategorie s hodnotami dat, které jsou mimo definovaný rozsah pro danou proměnnou.

Volby analýzy nelineárních kanonických korelací korelace

Dialogové okno Volby vám umožňuje vybrat si volitelné statistiky a grafy, uložit skóre objektů jako nové proměnné v aktivní datové sadě, uvést iterace a konvergenční kritéria a uvést počáteční konfiguraci pro analýzu.

Obrazovka. Dostupné statistiky zahrnují mezní frekvence (počty), centros, historii iterací, váhy a zatížení komponent, kvantily kategorií, skóre objektů a jednotlivé statistiky a jejich vícenásobná přizpůsobení.

- *Centroidy.* Ohodnocení kategorií a projektované a skutečné průměry skóre objektů pro objekty (případy) zahrnuté do každé sady pro tyto případy náležící do stejné kategorie proměnné.
- *Váhy a naložení komponenty.* regresní koeficienty v každé dimenzi pro každou kvantifikovanou proměnnou v sadě, kde jsou skóre objektů regresí na kvantifikovaných proměnných a projekce kvantifikované proměnné v prostoru objektů. Poskytuje informace o příspěvcích jednotlivých proměnných do dimenze v rámci každé sady.
- *Jednoduché a vícenásobné přizpůsobení.* Měřítko dobroty pro jedno-a více-kategorie souřadnic/kategorizační kvanta s ohledem na objekty.
- *Zdůvodnění kategorií.* Optimální hodnoty měřítka přiřazené kategoriím proměnné.
- *Skóre objektu.* Optimální skóre přiřazené k objektu (případu) v konkrétní dimenzi.

Zakreslit. Můžete vytvářet grafy se souřadnicemi kategorií, skóre objektů, loadingy komponent, centroidy kategorie a transformací.

Ukládat skóre objektů. Skóre objektu můžete uložit jako nové proměnné v aktivní datové sadě. Skóre objektu se ukládají pro počet dimenzí, které jsou uvedeny v hlavním dialogovém okně.

Použit náhodnou počáteční konfiguraci. Měla by být použita náhodná počáteční konfigurace, pokud jsou některé nebo všechny proměnné jednotlivé nominální hodnoty. Není-li tato volba vybrána, bude použita vnořená počáteční konfigurace.

Kritéria. Můžete určit maximální počet iterací, které nelineární kanonická korelační analýza může provádět ve svých výpočtech. Můžete také vybrat hodnotu kritéria konvergence. Analýza se zastaví iterováním, pokud je rozdíl mezi posledními dvěma iteracemi menší než hodnota konvergence nebo pokud je dosažen maximální počet iterací.

Další funkce příkazu OVERALS

Pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu OVERALS, můžete upravit nelineární kanonickou korelační analýzu. Jazyk syntaxe příkazu vám také umožňuje:

- Určete páry dimenzí, které mají být vykresleny, namísto vykreslování všech extrahovaných dimenzí (pomocí klíčového slova `NDIM` na dílčím příkazu `PLOT`).
- Uvedte počet znaků návěští hodnoty, které se používají k označení bodů na plochách (s dílčím příkazem `PLOT`).
- Určit více než pět proměnných jako popisování proměnných pro parcely objektů (s dílčím příkazem `PLOT`).
- Vyberte proměnné, které se použijí v analýze jako popisky proměnných pro parcely objektů (s dílčím příkazem `PLOT`).
- Vyberte proměnné, které mají poskytovat bodové popisky pro skóre kvantifikace skóre (s dílčím příkazem `PLOT`).
- Určete počet případů, které mají být zahrnuty do analýzy, pokud nechcete použít všechny případy v aktivní datové sadě (s dílčím příkazem `NOBSERVATIONS`).
- Určete názvy `rootnames` pro proměnné vytvořené uložením skóre objektů (s dílčím příkazem `SAVE`).
- Zadejte počet dimenzí, které mají být uloženy, místo ukládání všech extrahovaných dimenzí (s dílčím příkazem `SAVE`).
- Zapište do matricového souboru matice kategorií (pomocí dílčího příkazu `MATRIX`).
- Vytvořit grafy s nízkým rozlišením, které mohou být snadněji čitelné než obvyklé grafy s vysokým rozlišením (pomocí příkazu `SET`).
- Produkce centroid a transformačních grafů pouze pro zadané proměnné (s dílčím příkazem `PLOT`).

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

Analýza korespondence

Jedním z cílů analýzy korespondence je popsat vztahy mezi dvěma nominálními proměnnými v tabulce korespondenční tabulky v nedimenzionálním prostoru a současně popisovat vztahy mezi kategoriemi pro každou proměnnou. Pro každou proměnnou odrazí vzdálenost mezi body kategorie v grafu vztahy mezi kategoriemi s podobnými kategoriemi vykreslenými blízko sebe navzájem. Vysunutí bodů pro jednu proměnnou na vektoru z výchozího bodu do bodu kategorie pro druhou proměnnou popisují vztah mezi proměnnými.

Analýza pohotovostních tabulek často zahrnuje prověřování profilů řádků a sloupců a testování nezávislosti prostřednictvím statistiky chí kvadrát. Počet profilů však může být poměrně velký a chí kvadrát test neodhalí strukturu závislosti. Procedura kontingenčních tabulek nabízí několik ukazatelů přidružení a testů přidružení, které však nemohou graficky znázorňovat vztahy mezi proměnnými.

Analýza faktoru je standardní metoda pro popis vztahů mezi proměnnými v nízkém dimenzionálním prostoru. Analýza faktoru však vyžaduje intervalová data a počet pozorování by měl být pětikrát vyšší než počet proměnných. Korespondenční analýza, na druhé straně, předpokládá nominální proměnné a může popisovat vztahy mezi kategoriemi každé proměnné, stejně jako vztah mezi proměnnými. Korespondenční analýza může být navíc použita k analýze jakékoli tabulky pozitivních korespondenčních opatření.

Příklad. Korespondenční analýza by mohla být použita ke grafickému zobrazení vztahu mezi kategorií personálu a stravovacími návyky. Možná zjistíte, že s ohledem na kouření se mladí manažeři liší od sekretářky, ale sekretářky se neliší od vyšších manažerů. Může se také stát, že je těžké kouření spojeno s mladými manažery, zatímco světlo kouření je spojeno s tajemníky.

Statistika a grafy. Korespondenční opatření, profily řádků a sloupců, singulární hodnoty, skóre řádků a sloupců, setrvačná hmotnost, hmotnost, počet řádků a sloupců, statistiky důvěry, singulární hodnota statistiky důvěry, transformační grafy, parcely na řádcích, plochy sloupcových bodů a bisety.

Aspekty dat analýzy korespondence

Data. Kategorické proměnné, které mají být analyzovány, jsou škálovány nominálně. Pro agregovaná data nebo pro jiné korespondenci než frekvence použijte váhový proměnnou s kladnými hodnotami podobnosti. Alternativně pro data tabulky použijte syntaxi pro čtení tabulky.

Předpoklady. Maximální počet rozměrů použitých v proceduře závisí na počtu aktivních řádků a kategorií sloupců a na počtu omezení rovnosti. Nejsou-li použity žádné podmínky rovnosti a všechny kategorie jsou aktivní, maximální dimenzionalita je o jednu nižší než počet kategorií pro proměnnou s nejmenším počtem kategorií. Má-li například jedna proměnná pět kategorií a druhá má čtyři, je maximální počet dimenzí tři. Doplnkové kategorie nejsou aktivní. Má-li například jedna proměnná pět kategorií, z nichž dvě jsou doplnkové, a druhá proměnná má čtyři kategorie, maximální počet dimenzí je dva. Zacházet se všemi sadami kategorií, které jsou omezeny tak, aby byly stejné jako jedna kategorie. Pokud má například proměnná pět kategorií, z nichž tři jsou omezeny na stejnou hodnotu, měla by se s touto proměnnou při určování maximální dimenzionality zacházet jako se třemi kategoriemi. Dvě z kategorií jsou neomezeně a třetí kategorie odpovídá třem omezeným kategoriím. Určíte-li počet dimenzí, které jsou větší než maximum, použije se maximální hodnota.

Související postupy. Pokud jsou zahrnuty více než dvě proměnné, použijte analýzu více korespondence. Pokud by měly být proměnné odstupňovány koordinovaně, použijte analýzu kategoriálních komponent činitele.

Chcete-li získat Korespondenční analýzu

1. Z nabídky vyberte:

Analyzovat > Redukce dimenzí > Analýza korespondence ...

2. Vyberte řádkovou proměnnou.

3. Vyberte proměnnou sloupce.

4. Definujte rozsahy pro proměnné.

5. Klepněte na tlačítko **OK**.

Definovat rozsah řádků v analýze korespondencí

Pro řádkovou proměnnou je třeba definovat rozsah. Uvedené minimální a maximální hodnoty musí být celá čísla. Zlomkové datové hodnoty jsou v analýze zkráceny. Hodnota kategorie, která je mimo určený rozsah, je v analýze ignorována.

Všechny kategorie jsou zpočátku bez omezení a jsou aktivní. Můžete omezit kategorie řádků tak, aby se shodovaly s ostatními kategoriemi řádků, nebo můžete definovat kategorii řádku jako doplňující.

- **Kategorie musí být stejné.** Kategorie musí mít stejné skóre. Použijte omezení rovnosti, pokud je obdržena objednávka pro kategorie nežádoucí nebo kontrain intuitivní. Maximální počet kategorií řádků, které lze omezit na stejný počet, je celkový počet aktivních kategorií řádků minus 1. Chcete-li pro sady kategorií uložit různá omezení rovnosti, použijte syntaxi. Například použijte syntaxi k omezení kategorií 1 a 2, aby byla stejná a kategorie 3 a 4, aby byla stejná.
- **Kategorie je doplňková.** Doplnkové kategorie nemají vliv na analýzu, ale jsou reprezentovány v prostoru definovaném aktivními kategoriemi. Doplnkové kategorie nehrají žádnou roli při definování dimenzí. Maximální počet doplňkových kategorií řádků je celkový počet kategorií řádků minus 2.

Definovat rozsah sloupců v analýze Korespondence

Pro proměnnou sloupce je třeba definovat rozsah. Uvedené minimální a maximální hodnoty musí být celá čísla. Zlomkové datové hodnoty jsou v analýze zkráceny. Hodnota kategorie, která je mimo určený rozsah, je v analýze ignorována.

Všechny kategorie jsou zpočátku bez omezení a jsou aktivní. Kategorie sloupců můžete omezit tak, aby se shodovaly s ostatními kategoriemi sloupců, nebo můžete definovat kategorii sloupců jako doplňující.

- **Kategorie musí být stejné.** Kategorie musí mít stejné skóre. Použijte omezení rovnosti, pokud je obdržena objednávka pro kategorie nežádoucí nebo kontrain intuitivní. Maximální počet kategorií sloupců, které lze omezit na stejný počet, je celkový počet aktivních kategorií sloupců minus 1. Chcete-li pro sady

kategorií uložit různá omezení rovnosti, použijte syntaxi. Například použijte syntaxi k omezení kategorií 1 a 2, aby byla stejná a kategorie 3 a 4, aby byla stejná.

- **Kategorie je doplňková.** Doplňkové kategorie nemají vliv na analýzu, ale jsou reprezentovány v prostoru definovaném aktivními kategoriemi. Doplňkové kategorie nehrají žádnou roli při definování dimenzí. Maximální počet doplňkových kategorií sloupců je celkový počet kategorií sloupců minus 2.

Model analýzy korespondence

Dialogové okno Model vám umožňuje uvést počet dimenzí, měřítko vzdálenosti, normalizační metodu a normalizační metodu.

Dimenze v řešení. Zadejte počet dimenzí. Obecně platí, že je třeba zvolit si co nejméně rozměrů, aby bylo možné většinu variace vysvětlit. Maximální počet dimenzí závisí na počtu aktivních kategorií použitých v analýze a na omezeních rovnosti. Maximální počet dimenzí je menší z následujících hodnot:

- Počet kategorií aktivních řádků minus počet kategorií řádků, které mají být stejné, plus počet omezených sad kategorie řádků.
- Počet kategorií aktivních sloupců minus počet kategorií sloupců, které mají být stejné, plus počet omezených sad kategorie sloupců.

Měřítko vzdálenosti. Můžete vybrat měřítko vzdálenosti mezi řádky a sloupci tabulky korespondenční tabulky. Vyberte jednu z následujících možností:

- **Chi-náměstí.** Použijte váženou vzdálenost profilu, kde váha je hmotnost řádků nebo sloupců. Toto opatření je nezbytné pro standardní analýzu korespondence.
- **euklidovský.** Použijte druhou odmocninu součtu čtvercových rozdílů mezi dvojicemi řádků a dvojicemi sloupců.

Standardizační metoda. Vyberte jednu z následujících možností:

- **Pořadí řádků a sloupců je odebráno.** Oba řádky a sloupce se vystředí. Tato metoda je vyžadována pro standardní analýzu korespondence.
- **Řádkové prostředky jsou odebrány.** Pouze řádky se vystředí.
- **Hodnota sloupce znamená, že jsou odebrány.** Vystředí se pouze sloupce.
- **Celkové součty řádků jsou vyrovnány a prostředky byly odebrány.** Před centrování řádků jsou okraje řádků vyrovnány.
- **Sloupcové souhrny jsou vyrovnány a prostředky jsou odebrány.** Před cenzurování sloupců jsou okraje sloupců rovnávané.

Normalizační metoda. Vyberte jednu z následujících možností:

- **Symetrický.** Pro každou dimenzi jsou skóre řádku váženým průměrem skóre sloupců vydělených odpovídající hodnotou v jednotném čísle a skóre sloupce jsou váženým průměrem skóre řádku děleného odpovídající hodnotou v jednotném čísle. Tuto metodu použijte, pokud chcete zkoumat rozdíly nebo podobnosti mezi kategoriemi těchto dvou proměnných.
- **Činitel.** Vzdálenosti mezi řadami a sloupními body jsou aproximace vzdáleností v korespondenční tabulce v souladu s vybraným ukazatelem vzdálenosti. Tuto metodu použijte v případě, že chcete zkoumat rozdíly mezi kategoriemi obou nebo obou proměnných namísto rozdílů mezi dvěma proměnnými.
- **Činitel řádku.** Vzdálenosti mezi řadami řádku jsou aproximace vzdáleností v korespondenční tabulce podle vybraného ukazatele vzdálenosti. Skóre řádku je vážený průměr skóre sloupce. Tuto metodu použijte v případě, že chcete prozkoumat rozdíly nebo podobnosti mezi kategoriemi proměnné řádku.
- **Činitel sloupců.** Vzdálenosti mezi sloupci bodů jsou aproximace vzdáleností v korespondenční tabulce v souladu s vybraným ukazatelem vzdálenosti. Skóre sloupce je vážený průměr skóre řádku. Tuto metodu použijte v případě, že chcete prozkoumat rozdíly nebo podobnosti mezi kategoriemi proměnné sloupce.
- **Vlastní.** Musíte uvést hodnotu mezi -1 a 1. Hodnota -1 odpovídá činiteli sloupce. Hodnota 1 odpovídá činiteli řádku. Hodnota 0 odpovídá symetrickým hodnotám. Všechny ostatní hodnoty rozptýlí

setrvačnost přes skóre řádku i sloupce na různé stupně. Tato metoda je užitečná pro výrobu bifátů z krejčoviny.

Statistika analýzy korespondence

Dialogové okno Statistika vám umožňuje uvést numerický výstup vytvořený.

Korespondenční tabulka. Scrostable vstupních proměnných s okrajovými součty řádků a sloupců.

Přehled řádkových bodů. Pro každou kategorii řádku se skóre, hmotnost, setrvačnost, příspěvek k setrvačné hmotnosti rozměru a příspěvek dimenze k setrvačnosti daného bodu.

Přehled sloupcových bodů. Pro každou kategorii sloupců skóre, hmotnost, setrvačnost, přínos k setrvačnosti dimenze a přínos dimenze k netečnosti daného bodu.

Profily řádků. Pro každou kategorii řádků rozdělení mezi kategoriemi proměnné sloupce.

Profily sloupců. Pro každou kategorii sloupců rozdělení v rámci kategorií řádků proměnné.

Permutace v korespondenční tabulce. Korespondenční tabulka reorganizovala tak, že řádky a sloupce jsou ve vzestupném pořadí podle skóre na první dimenzi. Volitelně můžete určit maximální číslo dimenze, pro které budou vytvořeny zmutované tabulky. Je vytvořena permutovaná tabulka pro každou dimenzi od 1 do zadaného čísla.

Statistika důvěry pro řádkové body. Zahrnuje směrodatnou odchylku a korelace pro všechny nedoplňkové řádkové body.

Statistika důvěry pro sloupcové body. Zahrnuje směrodatnou odchylku a korelace pro všechny nedoplňkové sloupce.

Ploty analýzy korespondence

Dialogové okno Plot umožňuje určit, které grafy se mají vytvořit.

Bodové grafy. Vytvoří matici všech párových grafů dimenzí. Dostupné rozptylové grafy zahrnují:

- **Biplot.** Vytvoří matici společných grafů řádku a sloupců. Je-li vybrána normalizace činitele, není k dispozici dvoufázová fáze.
- **Řádkové body.** Vytvoří matici grafů na řádcích.
- **Sloupcové body.** Vytvoří matici grafů ve sloupcích.

Volitelně můžete uvést, kolik znaků popisku hodnoty se má použít při označování bodů. Tato hodnota musí být nezáporné celé číslo menší než nebo rovné 20.

Ploty čar. Vytvoří zakres pro každou dimenzi vybrané proměnné. Dostupné spojnicové grafy zahrnují:

- **Transformované kategorie řádků.** Vytvoří zakresový graf původních hodnot kategorie řádku vzhledem k jejich odpovídajícím skóre řádků.
- **Transformované kategorie sloupců.** Vytvoří zakres původních hodnot kategorií sloupců vzhledem k jejich odpovídajícím skóre sloupců.

Volitelně můžete určit, kolik znaků popisku hodnoty se má použít při označování osy kategorie. Tato hodnota musí být nezáporné celé číslo menší než nebo rovné 20.

Záksesové dimenze. Umožňuje vám řídit dimenze zobrazené ve výstupu.

- **Zobrazit všechny dimenze v řešení.** Všechny dimenze v řešení se zobrazí v matici bodového grafu.
- **Omezte počet dimenzí.** Zobrazené dimenze jsou omezeny na vykreslené dvojice. Pokud omezíte rozměry, musíte vybrat nejnižší a nejvyšší dimenze, které mají být zakresleny. Nejnižší dimenze může být v rozsahu od 1 do počtu dimenzí v řešení minus 1 a je vykreslena proti vyšším rozměrům. Nejvyšší rozměrová hodnota může být v rozsahu od 2 do počtu dimenzí v řešení a udává nejvyšší dimenzi, která má být použita při vykreslování dvojic dimenze. Tato specifikace se vztahuje na všechny požadované vícerozměrné grafy.

Další funkce příkazu CORRESPONDENCE

Analýzu korespondence můžete přizpůsobit, pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu CORRESPONDENCE . Jazyk syntaxe příkazu vám také umožňuje:

- Určit data tabulky jako vstup namísto použití dat z případů (pomocí dílčího příkazu TABLE = ALL).
- Uvedte počet znaků pro popisek hodnoty použitých pro popisky bodů pro každý typ matice bodového grafu nebo dvousložkové matice (s dílčím příkazem PLOT).
- Uvedte počet znaků popisku hodnoty použitých pro popisky bodů pro každý typ zakreslení řádku (s dílčím příkazem PLOT).
- Zapište matici řádků a sloupců skóre do souboru matice dat (s dílčím příkazem OUTFILE).
- Zapište matici statistiky důvěry (rozptyly a kovariance) pro jednotlivé hodnoty a skóre do maticového datového souboru (s dílčím příkazem OUTFILE).
- Určete více sad kategorií, které mají být shodné (s dílčím příkazem EQUAL).

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

Analýza více korespondencí

Analýza více korespondence vyčísluje nominální (kategoriální) data přiřazením číselných hodnot k případům (objektům) a kategoriím tak, aby objekty v rámci stejné kategorie byly blízko sebe a objekty v různých kategoriích jsou daleko od sebe. Každý objekt je co nejbližší ke kategoriím kategorií, které se vztahují na daný objekt. Tímto způsobem rozdělují kategorie objekty do homogenních podskupin. Proměnné jsou považovány za homogenní, když klasifikují objekty ve stejných kategoriích do stejných podskupin.

Příklad. Analýzu více korespondence lze použít ke grafickému zobrazení vztahu mezi kategorií úloh, menšinovou klasifikací a pohlavím. Možná zjistíte, že menšinová klasifikace a pohlaví rozlišují mezi lidmi, ale že tato kategorie pracovní pozice ne. Můžete také zjistit, že latinskoamerické a africko-americké kategorie jsou podobné sobě navzájem.

Statistika a grafy. Objektová skóre, diskriminační opatření, historie iterace, korelace původních a transformovaných proměnných, kvantita ke kategoriím, popisná statistika, body objektů pozemků, biosložky, kategorie pozemků, společné kategorie pozemků, transformační grafy a diskriminační opatření.

Aspekty dat analýzy více korespondenční analýzy

Data. Hodnoty řetězcové proměnné jsou vždy převedeny na kladná celá čísla ve vzestupném alfanumerickém pořadí. Uživatelem definované chybějící hodnoty, systémově chybějící hodnoty a hodnoty menší než 1 se považují za chybějící; můžete překódovat nebo přidat konstantu do proměnných s hodnotami nižšími než 1, aby se z nich nechyběly.

Předpoklady. Všechny proměnné mají vícenásobnou nominální úroveň škálování. Data musí obsahovat alespoň tři platné případy. Analýza je založena na kladných celočíselných datech. Volba diskretizace automaticky kategorizuje zlomkovou proměnnou seskupením svých hodnot do kategorií pomocí funkce close-to-normal a automaticky převede hodnoty řetězcových proměnných na kladná celá čísla. Můžete uvést jiné systémy diskretizace.

Související postupy. Pro dvě proměnné je analýza více korespondence obdobná analýze Korespondence. Pokud se domníváte, že proměnné mají pořadové nebo číselné vlastnosti, měla by se použít Categorical Principal Components Analysis. Pokud sady proměnných jsou zajímavé, měla by být použita nelineární kanonická analýza korelace.

Chcete-li získat analýzu více korespondence,

1. Z nabídky vyberte:

Analyzovat > Redukce dimenzí > Optimální škálování ...

2. Vyberte volbu **Všechny proměnné s více nominálními hodnotami**.
3. Vyberte volbu **Jedna sada**.

4. Klepněte na tlačítko **Definovat**.
5. Vyberte alespoň dvě proměnné analýzy a určete počet dimenzí v řešení.
6. Klepněte na tlačítko **OK**.

Volitelně můžete uvést doplňkové proměnné, které jsou umístěny do nalezeného řešení, nebo popisky pro grafy.

Definování váhy proměnné ve více korespondenční analýze

Váhu můžete nastavit pro proměnné analýzy.

Proměnná váha. Můžete zvolit, zda má být pro každou proměnnou definována váha. Uvedená hodnota musí být kladné celé číslo. Výchozí hodnota je 1.

Diskretizace s více korespondenční analýzou

Dialogové okno Diskrétnost vám umožňuje vybrat způsob, jak se překódovat proměnné. Zlomkové proměnné jsou seskupeny do sedmi kategorií (nebo do počtu odlišených hodnot proměnné, pokud je toto číslo menší než sedm) s přibližně normálním rozdělením, pokud není uvedeno jinak. Řetězcové proměnné jsou vždy převáděny na kladná celá čísla přiřazováním indikátorů kategorií podle vzestupného alfanumerického pořadí. Diskretizace pro řetězcové proměnné platí pro tato celá čísla. Ostatní proměnné jsou při výchozím nastavení ponechány samostatně. V analýze se pak použijí diskretizované proměnné.

. Vyberte si mezi skupinami, očíslování pořadí a vynásobením.

- **Seskupení.** Překóduje se do zadaného počtu kategorií nebo se přečísluje intervalem.
- **Hodnocení.** Proměnná je diskretizována podle důležitosti případů.
- **Vynásobení.** Aktuální hodnoty proměnné jsou standardizovány, násobené 10, zaokrouhleny a mají konstantní přidanou hodnotu, takže nejnižší diskretizovaná hodnota je 1.

Seskupení. Následující volby jsou k dispozici při diskretizaci proměnných podle seskupení:

- **Počet kategorií.** Zadejte počet kategorií a zda mají být hodnoty proměnné v rámci těchto kategorií přibližně normální nebo rovnoměrné rozdělení.
- **Rovnoměrné intervaly.** Proměnné se překotnou do kategorií definovaných v těchto stejně velkých intervalech. Musíte uvést délku intervalů.

Chybějící hodnoty analýzy více korespondence

Dialogové okno Chybějící hodnoty umožňuje zvolit strategii pro zpracování chybějících hodnot v proměnných analýzy a doplňujících proměnných.

Strategie chybějící hodnoty. Vyberte vyloučení chybějících hodnot (pasivní léčba), impute chybějící hodnoty (aktivní léčba) nebo vynechte objekty s chybějícími hodnotami (odstranění lišových vláken).

- **Vyloučit chybějící hodnoty; pro korelace impute po kvantifikaci.** Objekty s chybějícími hodnotami u vybrané proměnné nepřispívají k analýze této proměnné. Jsou-li použity všechny proměnné pasivní léčba, pak s objekty s chybějícími hodnotami na všech proměnných se zachází jako s doplňkovými hodnotami. Jsou-li korelace zadány v dialogovém okně Výstup, pak (po analýze) chybějící hodnoty jsou imputovány s nejčastějším kategoriím nebo režimem proměnné pro korelace původních proměnných. Pro korelace optimálně škálovaných proměnných si můžete zvolit metodu imputace. Vyberte **Režim**, chcete-li nahradit chybějící hodnoty v režimu optimálně škálované proměnné. Vyberte **Přebytečná kategorie**, chcete-li nahradit chybějící hodnoty s kvantifikací přebytečné kategorie. To znamená, že objekty s chybějící hodnotou na této proměnné jsou považovány za objekty patřící do stejné kategorie (extra).
- **Chybějící hodnoty Impute.** Objekty s chybějícími hodnotami na zvolené proměnné mají imputované hodnoty. Můžete zvolit metodu imputace. Vyberte **Režim**, chcete-li nahradit chybějící hodnoty nejčetnější kategorií. Je-li více režimů, použije se ten s nejmenším indikátorem kategorie. Vyberte **Přebytečná kategorie**, chcete-li nahradit chybějící hodnoty se shodným kvantifikací zvláštní kategorie.

To znamená, že objekty s chybějící hodnotou na této proměnné jsou považovány za objekty patřící do stejné kategorie (extra).

- **Vyloučit objekty s chybějícími hodnotami u této proměnné.** Objekty s chybějícími hodnotami na vybrané proměnné jsou vyloučeny z analýzy. Tato strategie není k dispozici pro doplňkové proměnné.

Volby analýzy více korespondencí

Dialogové okno Volby vám umožňuje vybrat počáteční konfiguraci, uvést iteraci a konvergenční kritéria, vybrat normalizační metodu, zvolit metodu pro označení grafů a uvádět doplňkové objekty.

Doplňkové objekty. Zadejte číslo případu objektu (nebo čísla prvního a posledního případu rozsahu objektů), které chcete přidat, a poté klepněte na tlačítko **Přidat**. Pokračujte, dokud jste neuvedli všechny své doplňkové objekty. Je-li objekt zadán jako doplňkový, pak jsou pro tento objekt ignorována váha případu.

Normalizační metoda. Můžete určit jednu z pěti voleb pro normalizaci skóre objektu a proměnných. V dané analýze může být použita pouze jedna metoda normalizace.

- **Činitel proměnné.** Tato volba optimalizuje přidružení mezi proměnnými. Souřadnice proměnných v prostoru objektů jsou načtenými komponenty (korelace s hlavními komponentami, jako např. skóre dimenzí a objektů). To je užitečné, když se zajímáte především o korelaci mezi proměnnými.
- **Objekt Principal.** Tato volba optimalizuje vzdálenosti mezi objekty. To je užitečné, když se zajímáte především o rozdíly nebo podobnosti mezi objekty.
- **Symmetrický.** Tuto normalizační volbu použijte, pokud se zajímáte především o vztah mezi objekty a proměnnými.
- **Nezávislá.** Tuto normalizační volbu použijte, chcete-li zkoumat vzdálenosti mezi objekty a korelacemi mezi proměnnými samostatně.
- **Vlastní.** V uzavřeném intervalu $[-1, 1]$ můžete zadat libovolnou reálnou hodnotu. Hodnota 1 se rovná metodě Principal objektu, hodnota 0 se rovná symetrické metodě Symetrický a hodnota -1 se rovná metodě Variable Principal. Zadáte-li hodnotu větší než -1 a menší než 1, můžete rozložit hodnotu eigenvalue nad objekty i proměnné. Tato metoda je užitečná pro výrobu fázového a fázového souboru s bitem nebo tripartem.

Kritéria. Můžete určit maximální počet iterací, které mohou procedury projít ve svých výpočtech. Můžete také vybrat hodnotu kritéria konvergence. Algoritmus přestane opakovat, pokud je rozdíl mezi posledními dvěma iteracemi menší než hodnota konvergence, nebo pokud je dosažen maximální počet iterací.

Popisek se kreslí podle. Umožňuje určit, zda se v grafech použijí proměnné a názvy proměnných nebo názvy proměnných, a hodnoty. Můžete také uvést maximální délku pro popisky.

Zákresové dimenze. Umožňuje vám řídit dimenze zobrazené ve výstupu.

- **Zobrazit všechny dimenze v řešení.** Všechny dimenze v řešení se zobrazí v matici bodového grafu.
- **Omezte počet dimenzí.** Zobrazené dimenze jsou omezeny na vykreslené dvojice. Pokud omezíte rozměry, musíte vybrat nejnižší a nejvyšší dimenze, které mají být zakresleny. Nejnižší dimenze může být v rozsahu od 1 do počtu dimenzí v řešení minus 1 a je vykreslena proti vyšším rozměrům. Nejvyšší rozměrová hodnota může být v rozsahu od 2 do počtu dimenzí v řešení a označuje nejvyšší dimenzi, která má být použita při vykreslování dvojic dimenze. Tato specifikace se vztahuje na všechny požadované vícerozměrné grafy.

Konfigurace. Můžete číst data ze souboru obsahujícího souřadnice konfigurace. První proměnná v souboru by měla obsahovat souřadnice pro první dimenzi, druhá proměnná by měla obsahovat souřadnice pro druhou dimenzi atd.

- **Počáteční.** Jako výchozí bod analýzy bude použita konfigurace uvedená v uvedeném souboru.
- **Opraveno.** Konfigurace v uvedeném souboru bude použita k uložení do proměnných. Proměnné, které jsou instalovány, musí být vybrány jako proměnné analýzy, ale protože je konfigurace opravena, se s nimi zachází jako s doplňkovými proměnnými (takže nemusí být jako doplňkové proměnné vybrány).

Výstup analýzy více korespondencí

Výstupní dialogové okno vám umožňuje vytvářet tabulky pro skóre objektů, diskriminačních opatření, historie iterace, korelace původních a transformovaných proměnných, kvantily kategorií pro vybrané proměnné a deskriptivní statistiky pro vybrané proměnné.

Skóre objektu. Zobrazí skóre objektů, včetně hmotnosti, setrvačnosti a příspěvků a má následující volby:

- **Zahrnout kategorie.** Zobrazí indikátory kategorie vybraných proměnných analýzy.
- **Popisek objektu Skóre podle.** Ze seznamu proměnných, které jsou uvedeny jako proměnné popisků, můžete vybrat jeden objekt, který bude označovat objekty jako popisky.

Diskriminovaná opatření. Zobrazí diskriminační ukazatele na proměnnou a dimenzi.

Historie iterací. Pro každou iteraci se zobrazí odchylka, která se zaúčtovala, ztráta a zvýšení rozptylu.

Korelace původních proměnných. Zobrazí korelační matici původních proměnných a eigenvalues této matrice.

Korelace transformovaných proměnných. Zobrazí korelační matici transformovaných (optimálně změněných) proměnných a ideových hodnot této matrice.

Kategorie a příspěvky kategorie. Poskytuje kategorie (souřadnice) kategorie (souřadnice), včetně hmotnosti, setrvačnosti a příspěvků, pro každou velikost vybraných proměnných.

Poznámka: souřadnice a příspěvky (včetně hmotnosti a setrvačné hmotnosti) se zobrazují v samostatných vrstvách výstupu kontingenční tabulky se souřadnicemi zobrazenými při výchozím nastavení. Chcete-li zobrazit příspěvky, aktivujte ji (poklepejte) na tabulku a vyberte příspěvky z rozevíracího seznamu Vrstva.

Deskriptivní statistika. Zobrazí frekvence, počet chybějících hodnot a režim zvolené proměnné (proměnných).

Hromadná analýza korespondenční analýzy

Dialogové okno Uložení vám umožňuje uložit diskretizovaná data, skóre objektů a transformované hodnoty do externího datového souboru IBM SPSS Statistics nebo datové sady v aktuální relaci. Také můžete uložit transformované hodnoty a skóre objektů do aktivní datové sady.

- Datové sady jsou k dispozici během aktuální relace, ale nejsou k dispozici v následných relacích, pokud je výslovně neuložíte jako datové soubory. Názvy datových sad musí dodržovat pravidla pojmenování proměnných.
- Názvy názvů souborů nebo názvů datových sad se musí lišit pro každý typ uložených dat.
- Uložíte-li skóre objektu nebo transformované hodnoty do aktivní datové sady, můžete zadat počet více nominálních rozměrů.

Hromadné objekty analýzy korespondenční analýzy

Dialogové okno Vykreslování objektu vám umožňuje uvést typy grafů, které chcete, a proměnné, které mají být zakresleny

Objektové body. Zobrazí se zakres bodů objektu.

Objekty a centros (biflot). Body objektu jsou vyneseny s proměnnou centroids.

Proměnné Biplot. Můžete zvolit použití všech proměnných pro dvousložky nebo výběr dílčí sady.

Objekty štítku. Můžete zvolit, zda mají být objekty označené kategoriemi vybraných proměnných (můžete zvolit hodnoty indikátoru kategorie nebo popisek hodnot v dialogovém okně Volby) nebo s jejich čísly případů. Je-li vybrána volba **Proměnná**, vytvoří se jeden zakres za každou proměnnou.

Hromadné položky analýzy s více korespondenční analýzou

Dialogové okno Proměnné dávky vám umožňuje uvést typy grafů, které chcete, a proměnné, které mají být zakresleny.

Ploty kategorie. Pro každou vybranou proměnnou se zakreslí zakres centroidních souřadnic. Kategorie jsou v centrojích objektů v jednotlivých kategoriích.

Společné partie kategorie. Toto je jeden zakres centroid souřadnice každé zvolené proměnné.

Vykreslované prvky transformace. Zobrazí zakresový graf optimálních kvantitu kategorie ve srovnání s indikátory kategorie. Můžete určit počet dimenzí; pro každou dimenzi bude vygenerován jeden zakres. Můžete také zvolit zobrazení zbytkových grafů pro každou vybranou proměnnou.

Rozlišovací opatření. Vytvoří jeden zakresový graf diskriminačních opatření pro vybrané proměnné.

Další funkce příkazu MULTIPLE CORRESPONDENCE

Pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu MULTIPLE CORRESPONDENCE, můžete upravit analýzu více korespondence. Jazyk syntaxe příkazu vám také umožňuje:

- Uvedte rootnames pro transformované proměnné, skóre objektů a aproximace, když je ukládáte do aktivní datové sady (s dílčím příkazem SAVE).
- Určete maximální délku pro popisky pro každý výkres zvlášť (s dílčím příkazem PLOT).
- Uvedte samostatný seznam proměnných pro zbytkové grafy (s dílčím příkazem PLOT).

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

Vícerozměrné škálování (PROXSCAL)

Vícerozměrné škálování se pokouší najít strukturu v sadě přibližných ukazatelů mezi objekty. Tento proces se provádí přiřazením pozorování ke specifickým umístěním v konceptuálním nízkodimenzionálním prostoru tak, že vzdálenosti mezi body v prostoru odpovídají dané (dis) podobnostem co nejvěrněji. Výsledkem je nejmenších čtverců znázornění objektů v tomto low-dimenzionálním prostoru, který vám v mnoha případech pomůže blíže porozumět vašim datům.

Příklad. Vícerozměrné škálování může být velmi užitečné při určování percepčních vztahů. Například při zvažování obrazu vašeho produktu můžete provést průzkum za účelem získání datové sady, která popisuje vnímaná podobnost (nebo blízkost) vašeho produktu k těm, kteří jsou soutěžiteli. Pomocí těchto proximity a nezávislých proměnných (jako např. cena) se můžete pokusit určit, které proměnné jsou důležité pro to, jak lidé tyto produkty vidí, a můžete odpovídajícím způsobem upravit obrázek.

Statistika a grafy. Historie iterací, zátěžová opatření, rozklad napětí, souřadnice společného prostoru, vzdálenosti objektů v rámci konečné konfigurace, individuální váhy prostoru, jednotlivé prostory, transformované proximity, transformované nezávislé proměnné, zátěžové grafy, rozptylové grafy se společnými prostory, jednotlivé prostory bodových grafů, jednotlivé prostory bodových grafů, transformační grafy, zbytkové grafy Sheppard a nezávislé proměnné transformace proměnných.

Aspekty použití více rozměrových dat

Data. Data lze dodat ve formě bezdotykové matrice nebo proměnných, které se převádějí na bezdotykové matrice. Matice mohou být formátovány ve sloupcích nebo ve řádcích. S proximity lze zacházet podle poměru, intervalu, ordinální hodnoty nebo úrovně škálování spline.

Předpoklady. Musí být uvedeny alespoň tři proměnné. Počet dimenzí nemůže překročit počet objektů minus jedna. Snížení stavu prostorového uspořádání je vynecháno, pokud je kombinováno s vícenásobným náhodnou začátkem. Je-li zadán pouze jeden zdroj, jsou všechny modely ekvivalentní modelu identity; proto je pro model identity použita výchozí hodnota analýzy.

Související postupy. Rozšiřitelnost všech proměnných na číselné úrovni odpovídá standardní vícerozměrné analýze škálování.

Získání vícerozměrného škálování

1. Z nabídky vyberte:

Analyzovat > Měřítka > Vícerozměrné škálování (PROXSCAL) ...

Otevře se dialogové okno Formát dat.

2. Určete formát dat:

Formát dat. Uvedte, zda se data skládají z blízkých ukazatelů, nebo chcete vytvořit proximaty z dat.

Počet zdrojů. Pokud jsou vaše data proximaty, uvedte, zda máte jeden zdroj nebo více zdrojů blízkých opatření.

Jeden zdroj. Pokud existuje jeden zdroj proximities, uvedte, zda je vaše datová sada formátována s proximitou v matice přes sloupce nebo v jednom sloupci se dvěma samostatnými proměnnými, abyste označili řádek a sloupec každé blízkosti.

- *Proximaty jsou ve sloupcích ve více sloupcích.* Blízkost matice se rozprostírá přes počet sloupců, které se rovnají počtu objektů. To vede k dialogovému oknu Proximit v Matrices v dialogovém okně Sloupec.
- *Proximities jsou v jednom sloupci.* Matice blízkosti je sbalena do jednoho sloupce nebo proměnné. Nezbytné jsou dvě další proměnné identifikující řádek a sloupec pro každou buňku. To vede k dialogovému oknu Proximity v jednom sloupci.

Více zdrojů. Pokud existuje více zdrojů proximit, uvedte, zda je datová sada formátována s proximitou ve skládaných maticích přes sloupce, ve více sloupcích s jedním zdrojem na sloupec, nebo v jednom sloupci.

- *Proximities jsou ve skládaných maticích přes sloupce.* Bezdotykové matice se rozloží na několik sloupců rovnými počtu objektů a jsou skládány nad sebe, přes počet řádků, které se rovnají počtu objektů, kolikrát je počet zdrojů. To vede k dialogovému oknu Proximit v Matrices v dialogovém okně Sloupec.
- *Proximities jsou ve sloupcích, jeden zdroj na sloupec.* Matice blízkosti jsou sbaleny do více sloupců nebo proměnných. Nezbytné jsou dvě další proměnné identifikující řádek a sloupec pro každou buňku. To vede k proximatosti v dialogovém okně Sloupec.
- *Proximity jsou skládány v jednom sloupci.* Teplotní matice jsou sbaleny do jednoho sloupce nebo proměnné. Nezbytné jsou tři další proměnné, identifikující řádek, sloupec a zdroj pro každou buňku. To vede k dialogovému oknu Proximity v jednom sloupci.

3. Klepněte na tlačítko **Definovat**.

Proximities v Matrices ve sloupcích

Pokud vyberete proximální hodnoty v datovém modelu matice pro jeden zdroj nebo více zdrojů v dialogovém okně Formát dat, pak postupujte takto:

1. Vyberte tři nebo více proměnných proximit. (Ujistěte se, že pořadí proměnných v seznamu se shoduje s pořadím sloupců proximaty.)
2. Volitelně můžete vybrat počet proměnných váhy rovnající se počtu proměnných proximit. (Ujistěte se, že se pořadí vah shoduje s pořadím proximit, které váží.)
3. Volitelně, pokud existuje více zdrojů, vyberte proměnnou zdrojů. (Počet případů v každé proměnné proximit by měl být roven počtu proximit proměnných krát počet zdrojů.)

Dále můžete definovat model pro vícerozměrné škálování, umístit omezení na společný prostor, nastavit kritéria konvergence, určit počáteční konfiguraci, která má být použita, a zvolit vykreslení a výstup.

Proximaty ve sloupcích

Vyberete-li v dialogovém okně Formát dat model více sloupců pro více zdrojů, postupujte takto:

1. Vyberte dvě nebo více proměnných proximit. (Každá proměnná se předpokládá jako matice proximit ze samostatného zdroje.)
2. Vyberte proměnnou řádků, chcete-li definovat umístění řádků pro proximitu v každé proměnné proximit.

3. Vyberte proměnnou sloupce, chcete-li definovat umístění sloupců pro proximity v každé proměnné proximit. (Buňky bezdotykové matrice, které nemají označení sloupce/sloupce, jsou považovány za chybějící.)
4. Volitelně můžete vybrat počet proměnných váhy rovnající se počtu proměnných proximit.

Dále můžete definovat model pro vícerozměrné škálování, umístit omezení na společný prostor, nastavit kritéria konvergence, určit počáteční konfiguraci, která má být použita, a zvolit vykreslení a výstup.

Proximity v jednom sloupci

Pokud vyberete jeden zdrojový nebo více zdrojů v dialogovém okně Formát dat, postupujte takto:

1. Vyberte proměnnou proximities. (t se předpokládá, že je jedna nebo více matic proximities.)
2. Vyberte proměnnou řádků, chcete-li definovat umístění řádků pro proximity v proměnné proximit.
3. Vyberte proměnnou sloupce, chcete-li definovat umístění sloupců pro proximity v proměnné proximities.
4. Existuje-li více zdrojů, vyberte proměnnou zdrojů. (U každého zdroje se za chybějící označení řádku či sloupce považují buňky v blízkosti matice) za chybějící.)
5. Volitelně můžete vybrat proměnnou vah.

Dále můžete definovat model pro vícerozměrné škálování, umístit omezení na společný prostor, nastavit kritéria konvergence, určit počáteční konfiguraci, která má být použita, a zvolit vykreslení a výstup.

Vytvořit proximity z dat

Pokud se rozhodnete vytvořit proximaty z dat v dialogovém okně Formát dat, pak postupujte takto:

1. Pokud vytvoříte vzdálenosti mezi proměnnými (viz dialogové okno Vytvořit ukazatel z dat), vyberte alespoň tři proměnné. Tyto proměnné budou použity k vytvoření bezdotykové matice (nebo matic, pokud existuje více zdrojů). Pokud vytváříte vzdálenosti mezi případy, je zapotřebí pouze jedna proměnná.
2. Existuje-li více zdrojů, vyberte proměnnou zdrojů.
3. Volitelně můžete vybrat ukazatel pro vytvoření proximit.

Dále můžete definovat model pro vícerozměrné škálování, umístit omezení na společný prostor, nastavit kritéria konvergence, určit počáteční konfiguraci, která má být použita, a zvolit vykreslení a výstup.

Vytvořit ukazatel z dat

Vícedimenzionální škálování používá data nesoupodobnost k vytvoření řešení škálování. Pokud jsou vaše data vícerozměrná data (hodnoty naměřených proměnných), musíte pro výpočet vícerozměrného řešení škálování vytvořit data dispodobnosti. Můžete uvést podrobnosti o vytváření dispodobnosti ukazatelů z vašich dat.

Ukazatel. Umožňuje určit míru nepodobnosti pro analýzu. Vyberte jednu alternativu ze skupiny ukazatelů odpovídající vašemu typu dat, a pak vyberte jedno z ukazatelů z rozevřacího seznamu odpovídajícího typu ukazatele. Dostupné alternativy jsou:

- **Interval.** euklidovský odstup, Squared euklidovský odstup, Chebychev, Block, Minkowského, nebo Upravené.
- **Počty.** Chi-kvadrát ukazatel nebo Phi-square opatření.
- **Binární.** euklidovský odstup, Squared euklidovský vzdálenost, Velikost rozdíl, Vzor rozdíl, Odchylka, nebo Lance a Williams.

Vytvořit matici vzdálenosti. Umožňuje zvolit jednotku analýzy. Alternativy jsou Mezi proměnnými nebo Mezi případy.

Transformace hodnot. V některých případech, například při měření proměnných na velmi odlišných měřících, chcete standardizovat hodnoty před výpočtem proximities (nepoužitelné pro binární data).

Vyberte standardizační metodu z rozevíracího seznamu Standardizace (pokud není požadována žádná standardizace, vyberte volbu **Žádná**).

Definice modelu Multidimensional Scaling Model

V dialogovém okně Model můžete určit model škálování, jeho minimální a maximální počet dimenzí, strukturu matice blízkosti, transformaci, která má být použita na proximitách, a informace o tom, zda jsou proximity transformovány v rámci každého zdroje samostatně nebo nepodmíněně ve zdroji.

Model škálování. Vyberte si z následujících možností:

- **Identita.** Všechny zdroje mají stejnou konfiguraci.
- **Vážené euklidovský.** Tento model je individuálním modelem rozdílů. Každý zdroj má individuální prostor, v němž je každá dimenze společného prostoru exponenciálně vážená.
- **Generalizovaný euklidovský.** Tento model je individuálním modelem rozdílů. Každý zdroj má určitý prostor, který se rovná rotaci prostoru společného prostoru, za nímž následuje rozdílová váha dimenze.
- **Nížší úroveň hodnocení.** Tento model je zobecněný euklidovský model, pro který můžete uvést hodnotu jednotlivých prostor. Je třeba určit ohodnocení důležitosti, které je větší nebo rovno 1 a menší než maximální počet dimenzí.

Tvar. Uvedte, zda by měly být proximity odebrány z dolní trojúhelníkové části nebo z horní trojúhelníkové části matice blízkosti. Můžete uvést, že se použije úplná matice, v tom případě bude analyzován vážený součet horní trojúhelníkové části a dolní trojúhelníkové části. V každém případě by měla být uvedena úplná matice, včetně diagonála, ačkoli se použijí pouze uvedené části.

Blížkost. Určete, zda matice blízkosti obsahuje ukazatele podobnosti nebo nepodobnosti.

Proximity transformace. Vyberte si z následujících možností:

- **Poměr.** Přetransformované proximity jsou úměrné původním proximitám. To je přípustné pouze pro vysoce hodnotné proximaty.
- **Interval.** Přetransformované proximity jsou proporcionální k původní proximity plus intersecí období. Zachycení zajišťuje, že všechny transformované proximity mají být pozitivní.
- **Pořadový.** Přetransformované proximity mají stejné uspořádání jako původní proximaty. Určíte, zda mají být svázané proximity drženy svázané nebo se mohou stát nevázanými.
- **Spline.** Přetransformované proximity jsou hladké neklesající polynomiální transformace původních proximací. Určíte stupeň polynomu a počet vnitřních uzlů.

Použit transformace. Uvedte, zda jsou porovnávány pouze proximity v rámci každého zdroje, nebo zda jsou porovnání nepodmíněná na zdroji.

Dimenze. Standardně je řešení vypočítáno ve dvou dimenzích (minimum = 2, maximum = 2). Vyberete celočíselné minimum a maximum od 1 do počtu objektů minus 1 (pokud je minimum menší nebo rovno maximum). Procedura vypočítá řešení v maximálních dimenzích a poté sníží dimenzionalitu v krocích, dokud nebude dosaženo nejnižší hodnoty.

Omezení vícerozměrného škálování

Dialogové okno Omezení vám umožňuje umístit omezení na společný prostor.

Omezení ve společném prostoru. Uvedte typ omezení, které chcete.

- **Bez omezení.** Na společném prostoru nejsou umístěna žádná omezení.
- **Některé souřadnice jsou pevné.** První vybraná proměnná obsahuje souřadnice objektů v první dimenzi, druhá proměnná odpovídá souřadnicím na druhé dimenzi a tak dále. Chybějící hodnota označuje, že souřadnice na dimenzi je volná. Počet vybraných proměnných musí být roven maximálnímu počtu požadovaných rozměrů.
- **Lineární kombinace nezávislých proměnných.** Společný prostor je omezen na lineární kombinaci vybraných proměnných.

Proměnné omezení. Vyberte proměnné, které definují omezení na společném prostoru. Pokud jste uvedli lineární kombinaci, uveďte interval, nominální, ordinální, nebo spline transformaci pro proměnné omezení. V obou případech se musí počet případů pro každou proměnnou rovnat počtu objektů.

Vícerozměrné volby škálování

Dialogové okno Volby vám umožňuje vybrat počáteční styl konfigurace, zadat iteraci a konvergenční kritéria a vybrat standardní nebo uvolněné aktualizace.

Počáteční konfigurace. Vyberte jednu z následujících možností:

- **Simplexní.** Objekty jsou umístěny ve stejné vzdálenosti od sebe v maximální dimenzi. Je provedena jedna iterace, která má zlepšit tuto vysokorozměrnou konfiguraci, po níž následuje operace redukce dimenze za účelem získání počáteční konfigurace, která má maximální počet rozměrů, které jste uvedli v dialogovém okně Model.
- **Torgerson.** Klasické řešení škálování se používá jako počáteční konfigurace.
- **Jeden náhodný začátek.** Konfigurace je vybrána náhodně.
- **Spustí se více náhodných hodnot.** Několik konfigurací se volí náhodně a jako počáteční konfigurace se použije konfigurace s nejnižším normalizovaným prvotním napětím.
- **Vlastní.** Vyberete proměnné, které obsahují souřadnice vaší vlastní počáteční konfigurace. Počet vybraných proměnných by měl být roven maximálnímu počtu uvedených veličin, přičemž první proměnná odpovídající souřadnicím na dimenzi 1, druhá proměnná odpovídající souřadnicím na dimenzi 2 atd. Počet případů v každé proměnné by měl být roven počtu objektů.

Kritéria iterace. Určete hodnoty iteračních kritérií.

- **Sbližování napětí.** Algoritmus přestane opakovat, je-li rozdíl v následných normalizovaných prvotních zátěžových hodnotách nižší než zde uvedené číslo, které musí ležet v rozsahu 0,0 až 1,0.
- **Minimální zatížení.** Algoritmus se zastaví, když se normalizované zatížení dostane pod číslo, které je zde uvedeno, které musí ležet v rozsahu 0,0 až 1,0.
- **Maximum iterací.** Algoritmus bude provádět počet zadaných iterací, pokud není nejprve splněn některý z výše uvedených kritérií.
- **Použijte uvolněné aktualizace.** Relaxované aktualizace urychlí algoritmus; tyto aktualizace nelze použít s jinými modely, než je model identity nebo používají omezení.

Multidimensional Scaling Plots, verze 1

Dialogové okno Plot vám umožňuje uvést, které výkresy se budou produkovat. Toto téma popisuje dialogové okno Vykreslování, pokud máte formát Proximities ve sloupcích Sloupců dat. Pro **Hmotnosti jednotlivých prostor, Původní versus transformované proximitya Transformované proximity na vzdálenostech** určete zdroje, pro které mají být vytvořeny grafy. Seznam dostupných zdrojů je seznamem proměnných proximit v hlavním dialogovém okně.

Stres. Vykreslení je vyrobeno z normalizovaného zatížení ve srovnání s rozměry. Tento zákres se vyrábí pouze tehdy, je-li maximální počet rozměrů větší než minimální počet rozměrů.

Společný prostor. Zobrazí se bodová matice souřadnic společného prostoru.

Jednotlivé prostory. Pro každý zdroj se souřadnice jednotlivých prostorů zobrazují v maticích scatterplot. Toto je možné pouze v případě, že je v dialogovém okně Model zadán jeden z jednotlivých modelů rozdílů.

Jednotlivá váha prostoru. Bodový graf se vytváří z hmotností jednotlivých prostor. Toto je možné pouze v případě, že je v dialogovém okně Model zadán jeden z jednotlivých modelů rozdílů. Pro vážený euklidovský model se váhy tisknou na zkusných plochách, přičemž jedna dimenze na každé ose je. Pro zevšeobecněný euklidovský model se vytváří jeden zápletka na rozměr, což svědčí o otočení a vážení uvedené dimenze. Omezeným modelem hodnocení je stejný výkres jako generalizovaný euklidovský model, ale snižuje počet rozměrů pro jednotlivé prostory.

Původní versus transformované proximity. Plot se vyrábí z původních proximatů oproti transformovaným proximitám.

Transformované proximity na vzdálenostech. Vykreslují se transformované proximity ve srovnání se vzdálenostmi.

Transformované nezávislé proměnné. Transformační grafy se vytvářejí pro nezávislé proměnné.

Korelace proměnných a dimenzí. Zobrazí se zakres korelací mezi nezávislými proměnnými a rozměry prostoru společného prostoru.

Multidimensional Scaling Plots, verze 2

Dialogové okno Plot vám umožňuje uvést, které výkresy se budou produkovat. Toto téma popisuje dialogové okno Vykreslování, je-li váš formát dat cokoli jiného než Proximities ve sloupcích.

Pro **Hmotnosti jednotlivých prostor, Původní versus transformované proximity a Transformované proximity na vzdálenostech** určete zdroje, pro které mají být vytvořeny grafy. Zadaná zdrojová čísla musí být hodnoty proměnné zdroje, která je uvedena v hlavním dialogovém okně a musí být v rozsahu od 1 do počtu zdrojů.

Vícerozměrný výstup škálování

Dialogové okno Výstup vám umožňuje řídit množství zobrazeného výstupu a uložit některé z nich do samostatných souborů.

Obrazovka. Vyberte jednu nebo více následujících položek pro zobrazení:

- **Společné souřadnice prostoru.** Zobrazí souřadnice společného prostoru.
- **Jednotlivé souřadnice prostoru.** Souřadnice jednotlivých prostorů jsou zobrazeny pouze v případě, že model není model identity.
- **Jednotlivá váha prostoru.** Zobrazí se váhy jednotlivých prostorů pouze v případě, že je zadán jeden z jednotlivých modelů rozdílů. V závislosti na modelu se váhy prostoru rozloží vahami a váhami rozměru, které se také zobrazí.
- **Vzdálenosti.** Zobrazí vzdálenost mezi objekty v konfiguraci.
- **Transformované proximity.** Zobrazí transformované proximity mezi objekty v konfiguraci.
- **Vstupní data.** Zahrnuje původní proximaty a, jsou-li přítomny, údaje o hmotnosti, počáteční uspořádání a pevné souřadnice nezávislých proměnných.
- **Stres pro náhodné spuštění.** Zobrazí náhodný počet náhodných čísel a normalizovanou hodnotu prvotního namáhání každého náhodného začátku.
- **Historie iterací.** Zobrazí historii iterací hlavního algoritmu.
- **Více zátěžových ukazatelů.** Zobrazí různé hodnoty napětí. Tabulka obsahuje hodnoty pro normalizované prvotní napětí, Stress-I, Stress-II, S-Stress, Dispersion Accounted For (DAF) a Tucker's Coefficient of Congruence.
- **Rozklad Stres.** Zobrazí rozklad objektů a zdrojů pro konečnou normalizovanou prvotní zátěž, včetně průměru na objekt a průměru na jeden zdroj.
- **Transformované nezávislé proměnné.** Pokud bylo vybráno omezení lineární kombinace, jsou zobrazeny transformované nezávislé proměnné a odpovídající regresní váhy.
- **Korelace proměnných a dimenzí.** Pokud bylo vybráno lineární kombinované omezení, zobrazí se korelace mezi nezávislými proměnnými a dimenzemi společného prostoru.

Uložit do nového souboru. Můžete uložit společné souřadnice prostoru, jednotlivé váhy prostoru, vzdálenosti, transformované proximity a transformované nezávislé proměnné pro oddělení datových souborů produktu IBM SPSS Statistics .

Další funkce příkazu PROXSCAL

Můžete přizpůsobit své vícerozměrné škálování analýzy proximities, pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu PROXSCAL . Jazyk syntaxe příkazu vám také umožňuje:

- Uvedte oddělené seznamy proměnných pro transformace a grafy zbytkových chyb (s dílčím příkazem PLOT).
- Určete samostatné zdrojové seznamy pro plošné váhy prostoru, transformace a zbytkové chyby (pomocí dílčího příkazu PLOT).
- Určete podmnožinu grafů transformace nezávislých proměnných, které mají být zobrazeny (spolu s dílčím příkazem PLOT).

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

Multidimenzionální rozkládání (PREFSCAL)

Procedura Multidimensional Unfolding se pokusí najít společné kvantitativní měřítko, které umožňuje vizuálně zkoumat vztahy mezi dvěma sadami objektů.

Příklady. Požádali jste 21 osob, aby zařadila 15 snídaňových položek v pořadí podle preference, 1 až 15. Pomocí Multidimensional Unfolding můžete určit, že jednotlivci diskriminují položky mezi snídaní ve dvou hlavních směrech: mezi měkkými a tvrdými babs, a mezi výkrmny a non-výkrmnými předměty.

Případně jste se zeptal skupiny řidičů na sazbu 26 modelů aut na 10 atributů na 6-bodové stupnici od 1 = "není pravda vůbec" až 6 = "velmi pravda." Hodnoty jsou ve srovnání s jednotlivci brány jako podobnosti. Pomocí Vícedimenzionálního rozkládání najdete klastrování podobných modelů a atributů, se kterými jsou nejvíce asociovány.

Statistika a grafy. Procedura Multidimensional Unfolding může vytvořit historii iterace, zátěžová opatření, rozklad napětí, souřadnice společného prostoru, vzdálenosti objektů v rámci konečné konfigurace, jednotlivé váhy prostoru, jednotlivé prostory, transformované proximity, zátěžové grafy, společné prostory bodových grafů, rozptylové grafy jednotlivých prostor, jednotlivé prostory bodových grafů, transformační grafy a zbytkové grafy Shepard.

Aspekty multidimenzionálního rozkládání dat

Data. Data jsou dodávána ve formě obdélníkových teplotních matic. Každý sloupec je považován za samostatný sloupcový objekt. Každý řádek bezdotykové matice je považován za samostatný objekt řádku. Když existuje více zdrojů proximit, jsou matice skládány.

Předpoklady. Musí být uvedeny alespoň dvě proměnné. Počet rozměrů v řešení nesmí překročit počet objektů minus jedna. Je-li zadán pouze jeden zdroj, jsou všechny modely ekvivalentní modelu identity; proto je pro model identity použita výchozí hodnota analýzy.

Získání vícedimenzionálního rozkládání

1. Z nabídky vyberte:

Analyzovat > Měřítko > Vícerozměrné rozkládání (PREFSCAL) ...

2. Vyberte dvě nebo více proměnných, které identifikují sloupce v obdélníkové matici blízkosti. Každá proměnná reprezentuje samostatný sloupcový objekt.
3. Volitelně můžete vybrat počet proměnných váhy rovnající se počtu proměnných objektů sloupce. Pořadí proměnných vah by se mělo shodovat s pořadím objektů ve sloupci, které váží.
4. Volitelně můžete vybrat proměnnou řádků. Hodnoty (nebo popisky hodnot) této proměnné se používají k označení řádkových objektů ve výstupu.
5. Existuje-li více zdrojů, můžete volitelně vybrat proměnnou zdrojů. Počet případů v datovém souboru by měl být roven počtu řádků s počtem řádků, kolikrát je počet zdrojů.

Navíc můžete definovat model pro vícerozměrný rozkládání, umístit omezení na společný prostor, nastavit kritéria konvergence, určit počáteční konfiguraci, která má být použita, a zvolit vykreslení a výstup.

Definovat vícerozměrný model rozkládání

V dialogovém okně Model můžete určit model škálování, jeho minimální a maximální počet dimenzí, strukturu matice blízkosti, transformaci, která má být použita na proximitách, a zda jsou proximity transformovány v řádku, podmíněny zdrojem nebo bezpodmínečně na zdroji.

Model škálování. Vyberte si z následujících možností:

- **Identita.** Všechny zdroje mají stejnou konfiguraci.
- **Vážené euklidovský.** Tento model je individuálním modelem rozdílů. Každý zdroj má individuální prostor, v němž je každá dimenze společného prostoru exponenciálně vážená.
- **Generalizovaný euklidovský.** Tento model je individuálním modelem rozdílů. Každý zdroj má určitý prostor, který se rovná rotaci prostoru společného prostoru, za nímž následuje rozdílová váha dimenze.

Blížkost. Určete, zda matice blízkosti obsahuje ukazatele podobnosti nebo nepodobnosti.

Dimenze. Standardně je řešení vypočítáno ve dvou dimenzích (minimum = 2, maximum = 2). Můžete zvolit celočíselné minimum a maximum od 1 k počtu objektů minus 1, dokud je minimum menší nebo rovno maximu. Procedura vypočítá řešení v maximální dimenzionalitě a poté sníží dimenzionalitu v krocích, dokud nebude dosaženo nejnižší hodnoty.

Proximity transformace. Vyberte si z následujících možností:

- **Není nutno vykonat žádnou akci.** Proximity se netransformují. Volitelně můžete vybrat volbu **Zahrnout zachycení**, v tom případě mohou být proximity posunuty konstantním termínem.
- **Lineární.** Transformované proximity jsou úměrné původním proximitám; to znamená, že transformační funkce odhaduje svah a průsečík se stanoví na 0. Tomu se také říká transformace poměru. Volitelně můžete vybrat volbu **Zahrnout zachycení**, v tom případě mohou být proximity také posunuty konstantním termínem. Tomu se také říká transformace intervalu.
- **Spline.** Přetransformované proximity jsou hladké neklesající polynomiální transformace původních proximací. Můžete uvést stupeň polynomu a počet vnitřních uzlů. Volitelně můžete vybrat volbu **Zahrnout zachycení**, v tom případě mohou být proximity také posunuty konstantním termínem.
- **Vyhladit.** Přetransformované proximity mají stejné pořadí jako původní proximaty, včetně omezení, které bere rozdíly mezi následujícími hodnotami v úvahu. Výsledkem je transformace "hladké pořadové číslo". Můžete určit, zda mají být svázané proximity drženy drženy nebo mít možnost stát se nevázanými.
- **Pořadový.** Přetransformované proximity mají stejné uspořádání jako původní proximaty. Můžete určit, zda mají být svázané proximity drženy drženy nebo mít možnost stát se nevázanými.

Použit transformace. Uvedte, zda se v každém řádku navzájem porovnají pouze proximální hodnoty, nebo se navzájem porovnají pouze proximity v rámci každého zdroje, nebo jsou porovnání nepodmíněná na řádku nebo zdroji; to znamená, zda jsou transformace provedeny za řádek, na zdroj, nebo přes všechny proximity najednou.

Omezení vícenásobného rozkládání

Dialogové okno Omezení vám umožňuje umístit omezení na společný prostor.

Omezení ve společném prostoru. Můžete zvolit opravu souřadnic řádků a/nebo sloupců objektů ve společném prostoru.

Proměnné omezení řádek/sloupce. Vyberte soubor obsahující omezení a vyberte proměnné, které definují omezení na společném prostoru. První vybraná proměnná obsahuje souřadnice objektů v první dimenzi, druhá proměnná odpovídá souřadnicím na druhé dimenzi a tak dále. Chybějící hodnota označuje, že souřadnice na dimenzi je volná. Počet vybraných proměnných musí být roven maximálnímu počtu požadovaných rozměrů. Počet případů pro každou proměnnou se musí rovnat počtu objektů.

Volby vícenásobného rozkládání

Dialogové okno Volby vám umožňuje vybrat počáteční styl konfigurace, zadat iterační a konvergenční kritéria a nastavit podmínky trestu pro stres.

Počáteční konfigurace. Vyberte jednu z následujících možností:

- **Klasika.** Pravoúhlé matrice se používá k doplnění vnitřních bloků (hodnot mezi řádky a mezi sloupci) kompletní symetrické matrice MDS. Jakmile je vytvořena úplná matice, použije se jako počáteční konfigurace klasický řešení škálování. Tyto vnitřní bloky mohou být naplněny prostřednictvím imputace pomocí trojúhelníku nerovnosti nebo Spearman vzdálenosti.

- **Ross-Cliff.** The Ross-Cliff start uses the results of a singular value decomposition on the double centered and squared axity matrix as the initial values for the row and column objects.
- **Korespondence.** Korespondenční start používá výsledky analýzy korespondence na obrácené data (podobnosti místo nepodobnosti) se symetrickou normalizací skóre řádku a sloupce.
- **Centroidy.** Procedura začne umístěním objektů řádků v konfiguraci pomocí dekompozice eigenvalue. Potom jsou objekty sloupce umístěny na centroid uvedených voleb. Pro počet voleb uveďte kladné celé číslo mezi 1 a počet proměnných proximit.
- **Spustí se více náhodných hodnot.** Řešení se počítají pro několik počátečních konfigurací vybraných náhodně, a jeden s nejnižším penalizovaným stresem se zobrazí jako nejlepší řešení.
- **Vlastní.** Můžete vybrat proměnné, které obsahují souřadnice vaší vlastní počáteční konfigurace. Počet vybraných proměnných by měl být roven maximálnímu počtu uvedených veličin, přičemž první proměnná odpovídající souřadnicím na dimenzi 1, druhá proměnná odpovídající souřadnicím na dimenzi 2 atd. Počet případů v každé proměnné by měl být roven kombinovanému počtu řádků řádků a sloupců. Souřadnice řádku a sloupce by měly být skládány, se sloupcovou souřadnicemi za souřadnicemi řádku.

Kritéria iterace. Určete hodnoty iteračních kritérií.

- **Sbližování napětí.** Algoritmus přestane opakovat, je-li relativní rozdíl v po sobě jdoucích penalizovaných zátěžových hodnotách nižší než zde zadané číslo, které musí být nezáporné.
- **Minimální zatížení.** Algoritmus se zastaví, když penalizovaný stres klesne pod zde uvedené číslo, které musí být nezáporné.
- **Maximum iterací.** Algoritmus provede zadaný počet iterací, pokud není nejprve splněn jeden z výše uvedených kritérií.

Podmínka trestu. Algoritmus se snaží minimalizovat penalizovaný stres, dobrota-propadá míru rovnající se produktu Kruskal Stres-I a trest založený na koeficientu variace na transformované proximities. Tyto ovládací prvky umožňují nastavit sílu a rozsah trestu.

- **Síla.** Čím nižší je hodnota pevnostní parametru, tím silnější je sankce. Zadejte hodnotu v rozsahu 0,0 až 1,0.
- **Rozsah.** Tento parametr nastavuje okamžik, kdy bude pokuta aktivní. Je-li hodnota nastavena na 0.0, tato sankce je neaktivní. Zvýšení hodnoty způsobí, že algoritmus vyhledá řešení s většími rozdíly mezi transformovanou proximatou. Uveďte nezápornou hodnotu.

Vícerozměrné skládací paty

Dialogové okno Plot vám umožňuje uvést, které výkresy se budou produkovat.

Ploty. K dispozici jsou následující grafy:

- **Spustí se více spuštění.** Zobrazí skládaný histogram penalizovaného stresu zobrazujícího jak zátěž, tak sankci.
- **Počáteční společný prostor.** Zobrazí matici bodového grafu souřadnic počátečního společného prostoru.
- **Stres na dimenzi.** Vytváří linepitu penalizovaného stresu oproti dimenzionalitě. Tento zakres se vyrábí pouze tehdy, je-li maximální počet rozměrů větší než minimální počet rozměrů.
- **Konečný společný prostor.** Zobrazí se bodová matice souřadnic společného prostoru.
- **Váhy prostoru.** Bodový graf se vytváří z hmotností jednotlivých prostor. Toto je možné pouze v případě, že je v dialogovém okně Model zadán jeden z jednotlivých modelů rozdílů. Pro vážený euklidovský model se váhy pro všechny zdroje zobrazují na zkusné ploše s jednou dimenzí na každé ose. Pro zevšeobecněný euklidovský model se vytváří jeden zakres na rozměr, který označuje rotaci a váhu této dimenze pro každý zdroj.
- **Jednotlivé prostory.** Zobrazí se bodová matice souřadnic individuálního prostoru každého zdroje. Toto je možné pouze v případě, že je v dialogovém okně Model zadán jeden z jednotlivých modelů rozdílů.

- **Transformační grafy.** Scatterplot se vyrábí z původních proximatů proti transformovaným proximitám. V závislosti na způsobu použití transformací je ke každému řádku nebo zdroji přiřazena samostatná barva. Nepodmíněná transformace vytvoří jednu barvu.
- **Shepard parcel.** Původní proximity proti oběma transformovaným proximitám a vzdálenosti. Vzdálenosti jsou signalizovány body a transformované proximity jsou označeny řádkem. V závislosti na způsobu použití transformací se pro každý řádek nebo zdroj vytváří samostatný řádek. Nepodmíněná transformace vytvoří jeden řádek.
- **Scatterplot se hodí.** Zobrazí se bodový graf transformovaných proximatů ve vztahu ke vzdálenostem. Ke každému zdroji je přiřazena samostatná barva, pokud je určeno více zdrojů.
- **Grafy Residuals.** Zobrazí se bodový graf transformovaných proximatů oproti zbytkům (transformované proximity minus vzdálenosti). Ke každému zdroji je přiřazena samostatná barva, pokud je určeno více zdrojů.

Styly řádkových objektů. To vám dává další kontrolu nad zobrazením objektů řádků v grafech. Hodnoty volitelné proměnné barvy se používají k cyklicku ve všech barvách. Hodnoty volitelných proměnných značek se používají k cyklicku mezi všemi možnými značkami.

Vlastnosti zdroje. Pro **Jednotlivé prostory**, **Bodový graf rozptylu** a **Vykreslení reziduí**-a jsou-li transformace použity zdrojem, pro **Transformační grafy** a **Shepard parcel**-můžete zadat zdroje, pro které by měly být pozemky vyprodukovány. Zadaná zdrojová čísla musí být hodnoty proměnné zdroje uvedené v hlavním dialogovém okně a rozsah od 1 do počtu zdrojů.

Plot řádku. Pokud jsou transformace použity řádkem, pro **Transformační grafy** a **Shepard parcel**, můžete určit řádek, pro který se mají vytvořit zakreslené grafy. Zadaná čísla řádků musí být v rozsahu od 1 do počtu řádků.

Vícerozměrný výstup pro rozkládání

Dialogové okno Výstup vám umožňuje řídit množství zobrazeného výstupu a uložit některé z nich do samostatných souborů.

Obrazovka. Vyberte jednu nebo více z následujících možností pro zobrazení:

- **Vstupní data.** Zahrnuje původní proximity a, jsou-li přítomny, údaje o váhách dat, počáteční konfiguraci a pevné souřadnice.
- **Spustí se více spuštění.** Zobrazí náhodný počet náhodných čísel a penalizovanou hodnotu namáhání každého náhodného začátku.
- **Počáteční data.** Zobrazí souřadnice počátečního společného prostoru.
- **Historie iterací.** Zobrazí historii iterací hlavního algoritmu.
- **Přízpusobit ukazatele.** Zobrazuje různá měřítka. Tabulka obsahuje řadu dobročinnosti, drsností, korelací, variací a nondegeneracy opatření.
- **Rozklad Stres.** Zobrazí objekty, řádky a zdroje dekompozice penalizovaného namáhání, včetně řádku, sloupce a zdrojových středních hodnot a směrodatných odchylek.
- **Transformované proximity.** Zobrazí transformované proximity.
- **Konečný společný prostor.** Zobrazí souřadnice společného prostoru.
- **Váhy prostoru.** Zobrazí váhu jednotlivých prostorů. Tato volba je k dispozici pouze v případě, že je uveden jeden z jednotlivých modelů rozdílů. V závislosti na modelu se váhy prostoru rozloží vahami a vahami rozměru, které se také zobrazí.
- **Jednotlivé prostory.** Zobrazí se souřadnice jednotlivých prostorů. Tato volba je k dispozici pouze v případě, že je uveden jeden z jednotlivých modelů rozdílů.
- **Uhrané vzdálenosti.** Zobrazí vzdálenost mezi objekty v konfiguraci.

Uložit do nového souboru. Můžete uložit společné souřadnice prostoru, jednotlivé váhy prostoru, vzdálenosti a transformované proximity do samostatných datových souborů produktu IBM SPSS Statistics .

Další funkce příkazu PREFSCAL

Pokud vložíte své výběry do okna syntaxe a upravíte výslednou syntaxi příkazu PREFSCAL , můžete upravit vícerozměrné rozkládání proximity analýzy. Jazyk syntaxe příkazu vám také umožňuje:

- Určete více zdrojových seznamů pro jednotlivé prostory, grafy Scatterparts a grafy Residuals-a v případě maticových podmíněných transformací, pro diagramy transformace a grafy Shepard-je-li k dispozici více zdrojů (s dílčím příkazem PLOT).
- Zadejte více seznamů řádků pro diagramy transformace a vykreslení Shepard v případě podmíněných transformací řádku (s dílčím příkazem PLOT).
- Uveďte počet řádků místo proměnné ID řádku (s dílčím příkazem INPUT).
- Uveďte počet zdrojů místo proměnné ID zdroje (s dílčím příkazem INPUT).

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

Upozornění

Tyto informace byly vytvořeny pro produkty a služby poskytované v USA. Tento materiál může být dostupný od IBM 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
USA*

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 POSKYTUJE TUTO PUBLIKACI "TAKOVOU, JAKÁ JE", BEZ JAKÝCHKOLIV ZÁRUK, VYJÁDŘENÝCH NEBO ODVOZENÝCH VČETNĚ, MIMO JINÉ, ODVOZENÝCH ZÁRUK NEPORUŠENÍ PRÁV TŘETÍCH STRAN, ZÁRUKY PRODEJNOSTI NEBO VHODNOSTI PRO URČITÝ ÚČEL. Některé právní řády u určitých transakcí nepřipouštějí vyloučení záruk výslovně vyjádřených nebo vyplývajících z okolností, a proto se na vás výše uvedené omezení nemusí vztahovat, a proto se vás toto prohlášení nemusí týkat.

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.

Jakékoliv odkazy v této publikaci na webové stránky jiných společností než IBM jsou poskytovány pouze pro pohodlí uživatele a nemohou být žádným způsobem vykládány jako doporučení těchto webových stránek. Materiály uvedené na těchto webových stránkách nejsou součástí materiálů pro tento produkt IBM a použití uvedených stránek je pouze na vlastní nebezpečí.

IBM může použít nebo distribuovat jakékoli informace, které jí poskytnete, libovolným způsobem, který společnost považuje za odpovídající, bez vzniku jakýchkoliv závazků vůči vám.

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
USA

Poskytnutí takových informací může být podmíněno dodržením 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.

Citovaná data o výkonu a příklady klienta jsou uvedeny pouze pro názornost. Skutečné výsledky výkonu se mohou lišit v závislosti na specifických konfiguracích a provozních podmínkách.

Informace týkající se produktů jiných společností než IBM byly získány od dodavatelů těchto produktů, z jejich publikovaných sdělení, nebo z jiných veřejně dostupných zdrojů. IBM tyto produkty netestovala a nemůže potvrdit přesnost údajů o výkonu, kompatibilitě nebo jiná tvrzení týkající se produktů jiných společností než IBM. Otázky týkající se možností produktů jiných společností než IBM by měly být adresovány dodavatelům těchto produktů.

Prohlášení týkající se budoucího směru vývoje nebo záměrů společnosti IBM se mohou změnit nebo mohou být zrušena bez předchozího upozornění a představují pouze cíle a záměry.

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 jakákoliv podobnost se skutečnými lidmi nebo obchodními podniky je čistě 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 společnosti IBM je k dispozici na webu na stránce "Copyright and trademark information" 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 založené na jazyce Java jsou ochranné známky nebo registrované ochranné známky společnosti Oracle anebo příbuzných společností.

Rejstřík

A

- Analýza kategorických komponent hlavní komponenty
 - bootstrapping [19](#)
 - optimální úroveň měřítka [15](#)
 - příkaz additional features [20](#)
 - ukládání proměnných [18](#)
- Analýza korespondence
 - příkaz additional features [27](#)
 - statistika [23](#)
 - výkresy [23](#)
- Analýza nelineárních kanonických korelací
 - příkaz additional features [23](#)
 - statistika [20](#)
 - výkresy [20](#)
- Analýza více korespondencí
 - optimální úroveň měřítka [28](#)
 - příkaz additional features [31](#)
 - ukládání proměnných [30](#)
- ANOVA
 - v Categorical Regrese [12](#)

B

- bipartity
 - v analýze Korespondence [26](#)
 - v kategorii Categorical Principal Components Analysis [18](#)
 - v rámci vícenásobné analýzy korespondence [30](#)
- bodový graf pro přizpůsobení
 - v multidimenzionálním rozkládání [39](#)
- bootstrapping
 - Analýza kategorických komponent hlavní komponenty [19](#)

C

- centroidy
 - v analýze nelineárních kanonických korelace [22](#)

D

- Deskriptivní statistika
 - v Categorical Regrese [12](#)
- dimenze
 - v analýze Korespondence [25](#)
- diskretizace
 - v Categorical Regrese [10](#)
 - v kategorii Categorical Principal Components Analysis [15](#)
 - v rámci vícenásobné analýzy korespondence [28](#)
- diskriminační opatření
 - v rámci vícenásobné analýzy korespondence [30](#)
- doplňkové objekty
 - v Categorical Regrese [11](#)

E

- elastická síť
 - v Categorical Regrese [12](#)

G

- generalized euklidovský model
 - v multidimenzionálním rozkládání [37](#)
- grafy jednotlivých prostorů
 - v multidimenzionálním rozkládání [39](#)
 - ve vícerozměrném škálování [35](#)
- grafy kategorie
 - v kategorii Categorical Principal Components Analysis [19](#)
 - v rámci vícenásobné analýzy korespondence [30](#)
- grafy načtení komponent
 - v kategorii Categorical Principal Components Analysis [19](#)
- grafy s diskriminací
 - v rámci vícenásobné analýzy korespondence [30](#)
- grafy s hmotností jednotlivých prostor
 - v multidimenzionálním rozkládání [39](#)
 - ve vícerozměrném škálování [35](#)
- grafy s hmotností prostoru
 - v multidimenzionálním rozkládání [39](#)
- grafy zbytkových chyb
 - v multidimenzionálním rozkládání [39](#)

H

- historie iterace
 - v kategorii Categorical Principal Components Analysis [18](#)
 - v multidimenzionálním rozkládání [40](#)
 - v rámci vícenásobné analýzy korespondence [30](#)
 - ve vícerozměrném škálování [36](#)
- hmotnost jednotlivých prostor
 - v multidimenzionálním rozkládání [40](#)
 - ve vícerozměrném škálování [36](#)

CH

- chybějící hodnoty
 - v Categorical Regrese [11](#)
 - v kategorii Categorical Principal Components Analysis [16](#)
 - v rámci vícenásobné analýzy korespondence [28](#)

I

- iterační kritéria
 - v multidimenzionálním rozkládání [38](#)
 - ve vícerozměrném škálování [35](#)

J

jednotlivé souřadnice prostoru
v multidimenzionálním rozkládání [40](#)

K

Kategorická regrese

Legalizace [12](#)
optimální úroveň měřítka [10](#)
příkaz additional features [13](#)
statistika [9](#)
uložit [13](#)
výkresy [9](#)

konečné společné prostorové grafy
v multidimenzionálním rozkládání [39](#)

Korelace

ve vícerozměrném škálování [36](#)

Korelace grafů

ve vícerozměrném škálování [35](#)

korelační matice

v kategorii Categorical Principal Components Analysis
[18](#)

v rámci vícenásobné analýzy korespondence [30](#)

L

laso

v Categorical Regrese [12](#)

loadings komponenty

v analýze nelineárních kanonických korelace [22](#)

v kategorii Categorical Principal Components Analysis
[18](#)

M

model identity

v multidimenzionálním rozkládání [37](#)

model škálování

v multidimenzionálním rozkládání [37](#)

Multidimenzionální škálování

model [34](#)

omezení [34](#)

příkaz additional features [36](#)

statistika [31](#)

volby [35](#)

výkresy [31](#), [35](#), [36](#)

výstup [36](#)

N

netečnost

v analýze Korespondence [26](#)

normalizace

v analýze Korespondence [25](#)

O

odchylka záúčtovaná

v kategorii Categorical Principal Components Analysis
[18](#)

omezení

ve vícerozměrném škálování [34](#)

omezení na společný prostor

v multidimenzionálním rozkládání [38](#)

opatření na vzdálenost

v analýze Korespondence [25](#)

Opevnění kategorie

v analýze nelineárních kanonických korelace [22](#)

v Categorical Regrese [12](#)

v kategorii Categorical Principal Components Analysis
[18](#)

v rámci vícenásobné analýzy korespondence [30](#)

optimální úroveň měřítka

v kategorii Categorical Principal Components Analysis
[15](#)

v rámci vícenásobné analýzy korespondence [28](#)

P

Počáteční konfigurace

v analýze nelineárních kanonických korelace [22](#)

v Categorical Regrese [11](#)

v multidimenzionálním rozkládání [38](#)

ve vícerozměrném škálování [35](#)

počáteční společné prostorové grafy

v multidimenzionálním rozkládání [39](#)

PREFEKT [37](#)

projektované centroidy

v kategorii Categorical Principal Components Analysis
[19](#)

proměnlivá váha

v kategorii Categorical Principal Components Analysis
[15](#)

v rámci vícenásobné analýzy korespondence [28](#)

proximity

v multidimenzionálním rozkládání [40](#)

ve vícerozměrném škálování [36](#)

proxity

v multidimenzionálním rozkládání [37](#)

Přízpůsobit

v analýze nelineárních kanonických korelace [22](#)

R

regresní koeficienty

v Categorical Regrese [12](#)

regresní regrese

v Categorical Regrese [12](#)

S

Sheppard

v multidimenzionálním rozkládání [39](#)

skóre objektů

v analýze nelineárních kanonických korelace [22](#)

v kategorii Categorical Principal Components Analysis
[18](#)

v rámci vícenásobné analýzy korespondence [30](#)

společné prostorové grafy

v multidimenzionálním rozkládání [39](#)

ve vícerozměrném škálování [35](#)

společné souřadnice prostoru

v multidimenzionálním rozkládání [40](#)

ve vícerozměrném škálování [36](#)

standardizace

standardizace (*pokračování*)
v analýze Korespondence [25](#)
statistika důvěry
v analýze Korespondence [26](#)

T

transformační grafy
v kategorii Categorical Principal Components Analysis
[19](#)
v multidimenzionálním rozkládání [39](#)
v rámci vícenásobné analýzy korespondence [30](#)
ve vícerozměrném škálování [35](#)
transformované nezávislé proměnné
ve vícerozměrném škálování [36](#)
trest
v multidimenzionálním rozkládání [38](#)
trigrafy
v kategorii Categorical Principal Components Analysis
[18](#)

U

uvolněné aktualizace
ve vícerozměrném škálování [35](#)

V

váhy
v analýze nelineárních kanonických korelace [22](#)
váženou euklidovský model
v multidimenzionálním rozkládání [37](#)
více R
v Categorical Regrese [12](#)
více spusticích grafů
v multidimenzionálním rozkládání [39](#)
Vícizionální rozkládání
model [37](#)
omezení na společný prostor [38](#)
příkaz additional features [41](#)
statistika [37](#)
volby [38](#)
výkresy [37](#), [39](#)
výstup [40](#)
výkresy
v analýze Korespondence [26](#)
v analýze nelineárních kanonických korelace [22](#)
v Categorical Regrese [13](#)
ve vícerozměrném škálování [35](#), [36](#)
výkresy společných kategorií
v kategorii Categorical Principal Components Analysis
[19](#)
v rámci vícenásobné analýzy korespondence [30](#)
vzdálenosti
v multidimenzionálním rozkládání [40](#)
ve vícerozměrném škálování [36](#)

Z

zakreslených bodů objektů
v kategorii Categorical Principal Components Analysis
[18](#)
v rámci vícenásobné analýzy korespondence [30](#)

zátěžová opatření
v multidimenzionálním rozkládání [40](#)
ve vícerozměrném škálování [36](#)
zátěžové grafy
v multidimenzionálním rozkládání [39](#)
ve vícerozměrném škálování [35](#)

