

# IBM SPSS REGRESSION

## Zpřesněte své predikce solistikovanými regresními modely

Využijte široké spektrum nelineárních modelovacích procedur modulu IBM SPSS Regression a aplikujte tak vhodné sofistikované modely na svá data, která překračují model lineární regrese založené na běžné OLS metodě (ta je součástí Statistics Base) a to ať již pracujete v komerční, akademické nebo státní sféře.

## Nejdůležitější:

- predikovat kategorizované proměnné s více než dvěma kategoriemi,
- snadno klasifikovat data do dvou kategorií,
- odhadovat parametry nelineárních modelů,
- přiřadit větší váhy závažnějším pozorováním,
- kontrolovat závislosti prediktorů a chybových členů,
- naleznout intenzity stimulů.

## Predikce kategorizované proměnné s více než dvěma kategoriemi

Díky proceduře mnohorozměrné logistické regrese (MLR) nejste vázáni na modelování dichotomických proměnných jako např. odpovědí typu ano/ne. Můžete například vytvořit model, který najde klíčové faktory pro rozhodnutí, zda si zákazník koupí výrobek A, B nebo C.

## Klasifikace dat do dvou kategorií

Využijte binární logistickou regresi pro predikci dichotomických proměnných typu koupí/nekoupí, zvolil/nezvolil. Procedura obsahuje metody pro postupný výběr nejvhodnějších spojitých i kategorizovaných prediktorů.

## Práce s modelem

S procedurami pro tvorbu nelineárních regresních modelů s okrajovými podmínkami i bez nich získáte plnou kontrolu nad svými modely. Procedury nabízí dva algoritmy pro výpočet odhadu parametrů v nelineárních modelech. Pro modely bez okrajových podmínek je k dispozici Levenberg-Marquardtův algoritmus. Sekvenční kvadratické programování otevírá možnost zadat okrajové podmínky pro hodnoty parametrů, definovat vlastní ztrátovou funkci či získat odhady standardních chyb parametrů metodou bootstrap.

## Alternativní přístup pro nesplněné předpoklady

Nesplňují-li Vaše data předpoklady pro použití standardní metody nejmenších čtverců (OLS), můžete využít jejího zobecnění – metodu vážených nejmenších čtverců (WLS) nebo dvoustupňovou metodu nejmenších čtverců (2SLS). V metodě WLS lze závažnějším pozorováním přiřadit větší váhu. Metoda 2SLS zajistí vyšší kontrolu nad závislostí prediktorů a chybových členů, která se často objevuje v časových řadách.

## Nejúčinnější stimuly

Probitové a logitové modely poskytují analýzu účinnosti různých podnětů, jako jsou např. dávkování léků, ceny či nabídky, na pozitivní výsledek akce. Procedura Probit zkoumá významnost podnětu na základě logitové nebo probitové transformace podílu responze k úrovni nebo intenzitě stimulu.

## Nejúčinnější stimuly

Probitové a logitové modely poskytují analýzu účinnosti různých podnětů, jako jsou např. dávkování léků, ceny či nabídky, na pozitivní výsledek akce. Procedura Probit zkoumá významnost podnětu na základě logitové nebo probitové transformace podílu responze k úrovni nebo intenzitě stimulu.

## Spoluprací dosáhnete hodnotnějších výsledků

Chcete-li soubory a výsledky efektivně sdílet, mít je připravené pro opětovné použití, chránit je dle Vašich interních zásad nebo externích požadavků a publikovat je většímu počtu uživatelů tak, aby je mohli prohlížet a pracovat s nimi, zvažte rozšíření softwaru SPSS Statistics o IBM SPSS Collaboration and Deployment Services. Více informací o těchto užitečných funkcích naleznete na adrese [www.ibm.com/spss/cds](http://www.ibm.com/spss/cds).

## SPECIFIKACE

### Binární logistická regrese

- Závislost dichotomické proměnné na množině nezávislých proměnných:
- postupné přidávání a ubírání proměnných, model s pevně danými nezávislými proměnnými,
- transformace kategorizovaných proměnných: indikační proměnné, odchylkové kontrasty, jednoduché kontrasty, diferenční kontrasty, Helmertovy kontrasty, porovnání sousedních kategorií, polynomiální kontrasty, uživatelem definované kontrasty,
- kritéria pro definování modelu: pravděpodobnost skóre pro vstupní hodnoty, ubírání proměnných na základě pravděpodobnosti Waldovy statistiky nebo poměru věrohodností,
- ukládání výsledných statistik do datové matice: odhadovaná pravděpodobnost a skupina, rezidua, odchylky, logit, studentizovaná a standardizovaná rezidua, leverage, analogie Cookovy statistiky vlivu, vliv na rozdíly v parametrech beta,
- export modelu v jazyce XML.

## Multinomická logistická regrese (MLR)

Závislost kategorizované proměnné s více než dvěma kategoriemi na několika nezávislých proměnných.

- volba hodnot parametrů pro odladění algoritmů pomocí specifikace CRITERIA,
- možnost do modelu zahrnout interakční členy,
- přímé stanovení nulové hypotézy jako lineární kombinace parametrů pomocí příkazu TEST,
- zadání disperzní škálovací hodnoty příkazem SCALE,
- model s konstantou nebo bez ní,
- interval spolehlivosti pro poměr šancí (odds ratio),
- možnost uložení následujících statistik: predikční (odhadovaná, očekávaná) pravděpodobnost kategorií, predikovaná kategorie závislé proměnné, pravděpodobnost této predikované kategorie a pravděpodobnost skutečné kategorie responze,
- specifikace referenční kategorie pro závislé proměnné,
- nalezení nejlepších prediktorů z desítek možných,
  - » prediktory určíte pomocí jedné ze čtyř metod: forward entry, backward elimination, forward stepwise nebo backward stepwise,
  - » volba pravidla pro vstup nebo výstup proměnných z analýzy,
  - » vstupy nebo výstupy jsou založeny na splnění hierarchického požadavku pro všechny efekty, nebo pro efekty obsahující pouze faktory, nebo je požadavek na zahrnutí aplikován na všechny efekty faktorů,
  - » vstup nebo výstup vybraného faktoru je volitelně povolen i bez splnění podmínek,
- metody Score a Wald napomáhají rychlému dosažení výsledků v případě velkého množství prediktorů,
- shoda modelu a dat je posuzována Akaikeho informačním kritériem (AIC) a Bayesovým informačním kritériem (BIC, jinak také Schwarz-Bayesovo kritérium – SBC),
- zpracování i rozsáhlých množství dat,
- volba ověřovacích statistik pro klasifikační tabulky,
- procento konkordance,
- procento vazeb,
- procento diskordance,
- C-hodnota pro logistický mód.

## Nelineární regrese s omezením (CNLR)

Využívá lineární i nelineární omezení na jakoukoli kombinaci parametrů.

- ukládání odhadů, reziduí a derivací,
- číselné nebo uživatelem definované derivace.

## Nelineární regrese (NLR)

Odhaduje modely s libovolným vztahem mezi závislými a nezávislými proměnnými s využitím iteračních algoritmů:

- specifikace ztrátové funkce,
- bootstrapové odhady standardních chyb.

## Metoda vážených nejmenších čtverců (WLS)

Závažnější pozorování mají větší váhu.

- váhy počítané na základě zdrojových proměnných a delta hodnot nebo převzaté z existujících řad,
- výstupy pro vypočtené váhy: logaritmická věrohodnostní funkce pro každou hodnotu parametru delta; R, R<sup>2</sup>, adjustované R<sup>2</sup>, standardní chyby, analýza rozptylu a t testy jednotlivých koeficientů pro hodnoty delta s maximalizovanou logaritmickou věrohodnostní funkcí,
- zobrazení výstupů v pivotní tabulce.

## Dvoustupňové nejmenší čtverce (2SLS)

Pomáhá při odstraňování korelací mezi prediktory a chybovou složkou.

- strukturální rovnice a instrumentální proměnné
- sledování korelace mezi prediktory a chybovými členy
- zobrazení výstupů v pivotní tabulce

## PROBIT

Analýza dávkových dat a příbuzné metody; logitové nebo probitové modely responsí.

- transformace prediktorů: dekadické logaritmy, přirozené logaritmy, logaritmy o libovolném základu,
- přirozený nebo specifikovaný poměr responsí,
- parametry algoritmu: konvergence, maximální počet iterací, pravděpodobnosti kritéria heterogenity,

- statistiky: četnosti, fiduciální intervaly spolehlivosti, relativní mediánová potence, test parallelismu, grafy pozorovaných PROBITů nebo LOGITů,
- zobrazení výstupů v pivotní tabulce,
- kvantilová regrese - modeluje vztah mezi množinou prediktorů (nezávislé proměnné) a určeným percentilem (kvantilem) cílové (závislé) proměnné, nejčastěji medián, - výhody - není předpoklad o rozdělení cílové proměnné, robustní vůči odlehlým pozorováním.

## Systémové požadavky

Požadavky se liší v závislosti na použité platformě. Více informací lze nalézt na [www.ibm.com/spss/requirements](http://www.ibm.com/spss/requirements).

## IBM Business Analytics

IBM Business Analytics software poskytuje akční vhled, který manažeři potřebují k dosažení lepších obchodních výsledků. IBM nabízí komplexní, jednotné portfolio složeno z business intelligence, prediktivních a sofistikovaných analýz, finančních výpočtů a strategického managementu, kontroly rizik a podmínek a analytických aplikací.

Pomocí IBM software dokážou společnosti odhalit trendy, vzory chování a odchylky od nich, porovnat scénáře „co by bylo kdyby“, predikovat potencionální hrozby a příležitosti, identifikovat a řídit klíčová obchodní rizika a plány, rozvrhnout a odhadovat zdroje. Díky těmto analytickým možnostem dokážou naši zákazníci po celém světě lépe pochopit, předvídat a formovat obchodní výsledky.