

LOGISTICKÁ REGRESE

ŠANCE A PRAVDĚPODOBŇNOST

Šance a pravděpodobnosti

- Citujme klasika:
 - šance (O^*) je poměr pravděpodobnosti jedné varianty (události) ke druhé variantě (událost nenastala)
- (Tomáš Katrňák, osobní web)
- Jak na výpočet šancí (4 polní tabulky)
- Interpretace (důležitost hodnoty 1, asymetričnost škály a hodnoty menší než 1), plus exkurz o závislosti
- * O je zkratkou anglického výrazu ODD=ŠANCE

Poměr šancí

- Podělíme-li dvě šance (uděláme jejich poměr) získáme poměr šancí – OR*
- Jak na výpočet poměru šancí (4 polní tabulky)
- Interpretace (důležitost hodnoty 1, asymetričnost škály a hodnoty menší než 1)
- * OR je zkratkou anglického výrazu ODDS RATIO

Důvody pro užívání šancí resp. poměrů šancí

- Podobně jako jiné popisné statistiky zjednodušují popis dat
- Statisticky umožňují zmenšit počet odhadovaných parametrů či vůbec řešit úlohu
- Existují poměry šancí vyšších řádů (vhodné zejména pro loglineární modelování)

LOGIT

Logit

- Zlogaritmovanou šanci či poměr šancí nazýváme logit
- Jakých hodnot může teoreticky nabýt?
- Shrnutí:
- **pravděpodobnost \neq šance \neq logit**

BINÁRNÍ LOGISTICKÁ REGRESE

Logistická regrese a možné zdroje informací

- Klasické texty: Applied logistic regression /Hosmer, Lemeshow (1989,2000,375 s.), Regression models for Categorical and Limited Dependent Data, /Scott Long(1997,296 s.), Logistic regression :a primer /Fred C. Pampel (2000,85s.).
Categorical Data Analysis/Agresti (2002,710s.)
- Český: **Hebák** a kol.(3.díl), **Hendl a hlavně Řeháková-Nebojte se logistické regrese (Soc. časopis 4:475-492)**
- Aplikace v SPSS-opět Discovering statistics using SPSS for Windows :advanced techniques for the beginner /Andy **Field** (2000,2005,2009) a opět **Norušis**

Předpoklady a proměnné

- připomenutí lineární regrese a jejích předpokladů včetně požadavků na proměnné
- v logistické regresi závislá proměnná trojí charakter
 - a) dichotomická (binární logistická regrese-binary) - dnešní lekce
 - b) ordinální(ordinal)- okrajově projdeme princip
 - c) nominální s více než dvěmi kategoriemi (polytomická logistická regrese-multinomial) – okrajově projdeme princip
- nezávislé proměnné mohou být nominální (včetně dichotomické) i spojité
- logistická regrese jako obdoba diskriminační analýzy (ale s méně striktními požadavky)

Binární logistická regrese

- závislá proměnná-dvě varianty - úspěch a neúspěch, muž žena atd.
- nezávislé proměnné - nominální i spojité

Model logistické regrese

- **pravděpodobnost** (úspěchu či neúspěchu)-tu nelze korektně regresně modelovat-Proč?
- Ukázka předpovědi pravděpodobnosti z lineárního regresního modelu
- Ukázka nejjednodušší dvourozměrné kontingenční tabulky 2x2
- **Šance**-podíl pravděpodobnosti úspěchu a neúspěchu
- **Podíl šancí**-stále kladné číslo
- **Logit**-přirozený logaritmus podílu šancí (resp. někdy i šance) - hodnoty od minus do plus nekonečna - s tímto už lze v rámci regrese pracovat (Připomenutí logaritmické funkce)

Rovnice, nezávislé proměnné a interakce

- rovnice logistické regrese
- nominální nezávislé proměnné-bud' dichotomie (pak s nimi lze pracovat obdobně jako v lineární regresi), nebo více kategorií (pak převádíme na umělé (dummy) proměnné- SPSS toto udělá za nás - poslední kategorie je pak vztahová=referenční kategorie)
- interakce-vzájemné spolupůsobení dvou či více nezávislých proměnných na proměnnou závislou (viz dříve u lineární regrese)

Základní otázky v binární logistické regresi

- Jak je celkově vhodný náš model pro data která máme (lineární regrese F-test a index determinace) a jak silné jsou vazby mezi proměnnými? (obdoby indexu determinace a celkové testy vhodnosti LR testy, Hosmer-Lemeshow test)
- Která z nezávislých proměnných důležitější než ostatní, která do modelu nepatří (lineární regrese t-testy a beta koeficienty, logistická Wald test)
- Jsou splněny předpoklady modelu (lineární regrese existuje lineární vztah mezi závislou a nezávislými proměnnými, v logistické regresi linearita vztahu mezi logitem a nezávislými proměnnými)

!!!Logistická regrese předpoklady!!!

- Nekorelovanost nezávislých proměnných (opak multikolinearita)
- Závislá proměnná dichotomická (ordinální, nominální)
- Nezávislé v zásadě jakékoliv (v diskriminační analýze musí mít nezávislé proměnné vícerozměrné normální rozdělení)

Metody odhadu parametrů

- **V logistické regresi nelze užít metodu nejmenších čtverců MNČ (resp. OLS)**
- Užívaná metoda: **ML** viz (i loglineární modely, faktorová analýza a mnohé jiné metody)
- Řešení logistické regrese nelze získat v jednom kroku výpočtu, užívají se tzv. iterativní postupy

Binární logistická regrese - příklad

- Matějů, P., Vlachová, K. 2000. Nerovnost, spravedlnost, politika. SLON, str. 284 an. – vysvětlení volby určité strany pomocí sociodemografických charakteristik a škál levice-pravice, liberárně-autoritářské atd.
- binární logistická regrese-hledání faktorů ovlivňujících volbu levicové či pravicové strany
- Náš příklad – predikce užívání internetu z dat WIP 2006

Binární logistická regrese v SPSS

- výsledkem procedury v SPSS je logistická regresní rovnice, otestování významnosti regresního modelu a jednotlivých parametrů včetně signalizace jednotlivých problémů
- zadání pomocí nabídky-jednotlivé důležité volby
- menu-Analyze-Regression-Binary Logistic

Metody pro zahrnutí proměnných do modelu (2)

- forward (dopředná metoda)-postup kdy od modelu konstanty (bez nezávislých proměnných) přecházíme k "optimálnímu" modelu postupným zařazováním proměnných
- backward (zpětná metoda)-vycházíme od modelu zahrnujícího všechny nezávislé proměnné a postupně se vyřazují proměnné

3 různá kritéria pro zařazení proměnných do modelu

- LR(likelihood-ratio)-vychází z kritéria maximální věrohodnosti (na metodě maximální věrohodnosti je založen odhad parametrů v logistické regresi- v lineární se zpravidla užívá metody nejmenších čtverců-viz přednáška o regresní analýze)-zařazují se tedy proměnné, které nejvíce zlepší (zvětší) maximální věrohodnost (resp. zmenší minus dvojnásobek záporného logaritmu maximální věrohodnosti, tzv. $-2LL$ nebo deviance)
- Wald-kritérium méně vhodné, využívá se i pro testování vhodnosti zařazení jednotlivých proměnných do modelu (tedy obdoba t-testů v lineární regresi-viz dále)
- conditional-obdoba LR, výpočetně jednodušší
- kombinací dvou uvedených metod a tří kritérií získáváme šest metod výběru "optimální" sady nezávislých proměnných

Zadání v SPSS-další volby

Volba **Categorical**-definování proměnných a určení nominálních proměnných včetně definování kontrastů a vztahové kategorie

Volba **Save**

- různé typy reziduí-slouží k diagnostikování regresních problémů, dále identifikace vlivných pozorování a také pravděpodobnost a skupinovou příslušnost jednotlivých pozorování

Zadání v SPSS-další volby - Options

- Hosmer-Lemeshow test dobré shody-Rozdělí soubor dle odhadnutých pravděpodobností do deseti skupin a poté zjišťuje zda odhadovaná a napozorovaný úspěch (neúspěch) v těchto skupinách je "rozumně" rozdělen (tedy odhadnuté četnosti a napozorované četnosti se neliší)- požadavek testu je aby byl poměrně velký počet případů a očekávané četnosti ve většině skupin přesáhly 5 (a žádná ze skupin nemůže mít četnost menší než 1) Nulová hypotéza testu: odhadnuté a napozorované četnosti úspěchů a neúspěchů se neliší
- Klasifikační charakteristiky (Classification Plots)
- ukazuje zařazení jednotlivých případů na základě modelu logistické regrese v klasifikační tabulce a graficky (viz komentář k výstupům)

Binární logistická regrese v SPSS: syntax

zadání pomocí příkazu (pro metodu ENTER a
vybrané výstupy)

```
LOGISTIC REGRESSION VAR=volbalp  
/METHOD=ENTER age edyrs polor  
/CLASSPLOT  
/PRINT=GOODFIT CI(95)  
/CRITERIA PIN(.05) POUT(.10) ITERATE(20)  
CUT(.5) .
```

Logistická regrese v SPSS- výstupy

- Odhad parametrů modelu (Variables in the Equation)
- Koeficient a jeho interpretace-udává změnu logitu při změně závislé proměnné o jednotku (u spojitých proměnných) či ve vztahu k referenční kategorii (u umělých proměnných) pomocí $\exp(B)$ -lepší pro interpretaci než koeficient, značí změnu šance úspěchu (varianty závislé proměnné kódované jako 1)
- při změně nezávislé proměnné o jednotku resp. ve vztahu k referenční kategorii, je-li větší než 1 šance se zmenší, je-li větší než 1 šance vzroste (šance je pro připomenutí podíl pravděpodobností úspěchu a neúspěchu)

Logistická regrese v SPSS- výstupy

- **Waldova statistika** - obdoba dílčích t-testů v lineární regresi, problém je v případě velkých koeficientů, pak vychází testy (sloupec Sig.) jako nevýznamné a nutí nás nezamítat nulovou hypotézu o nulovosti příslušného regresního koeficientu (a tedy uvažovat o vyřazení proměnné z modelu)-je tedy nutno při velkých hodnotách regresních koeficientů brát tyto testy s rezervou-lze alternativně projít všechny modely pomocí backward LR metody s nastavenou hodnotou pro vyřazení proměnné 1 - removal v Options (tedy postupně vyřadí všechny proměnné) a pomocí změny logaritmu věrohodnosti lze usoudit, zda má proměnná v modelu smysl či nikoliv
- **Interval spolehlivosti pro $\exp(B)$** -95 % intervalový odhad, obsahuje-li jednotku pak zařazení proměnné do modelu není namístě (týž postup jako u Waldovy statistiky a hodnoty Sig.)

Logistická regrese v SPSS- výstupy

Shoda modelu s daty (Improvement of Goodness-of-Fit)

- test založený na logaritmu maximální věrohodnosti-ta má rozdělení Chí-kvadrát (počet stupňů volnosti je roven počtu odhadovaných parametrů bez konstanty)- v tabulce Changes of Goodness-of-Fit tři statistiky odvozené od této
- **Model**-testuje zda náš model je lepší než model konstanty
- **Block**-v případě, že v zadání modelu zadáme proměnné blokově (aby vstupovaly po několika v blocích) testuje změnu mezi modely obsahujícími postupně jednotlivé bloky proměnných
- **Step**-Testuje rozdíl oproti předchozími modelu, má význam pokud budujeme postupně více modelů
- Obecně jsou tyto testy obdobou celkového F-testu v regresní analýze

Logistická regrese v SPSS- výstupy

- Obdoba R^2 (indexu determinace) v logistické regresi (viz Model summary)
- nelze definovat variabilitu jako v lineární regresi
- existuje více způsobů jak obdobu R^2 spočítat, všechny vychází z porovnávání věrohodnosti našeho modelu s modelem obsahujícím pouze konstantu
- Cox a Snell R^2 -hodnoty mezi 0 a 1 (této hodnoty nikdy nedosáhne)
- Nagelkerke R^2 -modifikace předchozího, hodnoty 1 může dosáhnout
- Mc Fadden-jen u proměnné s více variantami (polytomické logistické regrese)
- nejvhodnější obdobou R^2 z lineární regrese je zřejmě Nagelkerkovo R^2

Logistická regrese v SPSS- výstupy

- **Klasifikační tabulka** (Classification table)-udává kolik procent objektů se mi podařilo správně zařadit z hlediska odhadu hodnoty závislé proměnné (obdoba procesu resubstituce u diskriminační analýzy-viz 7. přednáška)
- **Histogram odhadnutých pravděpodobností** (Histogram of estimated probabilities)-optimálně by jednotky se závislou proměnnou nabývající hodnoty 1 (úspěch) měly být nad úrovní 0,5 na ose x, jednotky s hodnotou závislé proměnné 0 (neúspěch) pod touto úrovní
- Tabulka pro **Hosmer-Lemeshow test** a interpretace výsledků-viz výše (zadání v SPSS)

Logistická regrese jiné výstupy

- Pro lepší interpretaci významu jednotlivých proměnných je dobré z modelu odhadnout pravděpodobnosti pro jednotlivé úrovně nezávislé proměnné a tyto nakreslit do grafu, například pro 5 věkových skupin (na ose x) na osu y vynášíme vypočtenou pravděpodobnost
- Těchto křivek lze vykreslit více například pro různé skupiny-dle pohlaví, vzdělání apod.
- A co když jsou vztahy v různých skupinách diametrálně odlišné?

Logistická regrese-připomínka problému skupin

- **Jak tuto situaci řešit?**
- Analýzy provést zvlášť (TEMP nebo SPLIT FILE) a teprve poté dohromady, je-li pro toto důvod (test o shodě hodnot regresních koeficientů- v SPSS není obsažen)
- Lze užít i **víceúrovňové/hierarchické modely**
- Obecně je problém často složitější a skupin může být více a i zde platí: **Nejdříve zkoumejme, jak vypadá vztah v jednotlivých skupinách** a je-li podobný, lze analyzovat dohromady

Logistická regrese a její problémy

Diagnostika problémů a jejich řešení

- Analýza reziduí opět lze diagnostikovat vlivná (influentials) a odlehlá pozorování (outliers)
- multikolinearita - závislost mezi vysvětlujícími proměnnými, je téměř vždy přítomná, problémem je škodlivá multikolinearita

Logistická regrese a její problémy

- měření škodlivé multikolinearity - orientační kritérium alespoň jeden párový korelační koeficient mezi vysvětlujícími proměnnými ve výši 0,8 (existují i rozličné exaktní testy),
- Měření v SPSS-viz minulá lekce v proceduře pro lineární regresi (nutno zadat všechny nezávislé poměnné a jakoukoliv závislou a sledovat jen tabulku s diagnostikou kolinearity)

Připomínka důležité zásady

- Vždy před použitím analýzy by měla předcházet úvaha o souvislostech, tedy budujeme jen model, který má nějaké teoretické opodstatnění!!!

Poučky k regresi

- **Publikační náležitosti:** Při publikaci výsledků by měl autor dbát, zda mluví o logitech či šancích, které kategorie závislé proměnné se porovnávají, měly by být zmíněny výsledky Waldova testu pro jednotlivé parametry, celkový test pro model, Hosmer Lemeshow test, období R^2 (mělo by být uvedeno o jakou jde- Nagelkerke, Cox a Snell či Mc Fadden), lze uvést i úspěšnost klasifikace pomocí modelu

POLYTOMICKÁ LOGISTICKÁ REGRESE

Předpoklady a proměnné

- v polytomické logistické regresí závislá proměnná nominální s více kategoriemi (ale ne ordinální)
- nezávislé proměnné mohou být nominální (včetně dichotomické) i spojité
- Polytomická regrese jako obdoba diskriminační analýzy (viz konec lekce)

Polytomická logistická regrese

Model polytomické regrese

- **pravděpodobnost zařazení do kategorie m a srovnávací-tu nelze korektně regresně modelovat- Proč?**
- **Modelujeme obdobu šance-podíl pravděpodobnosti zařazení do kategorie m a srovnávací kategorie (první nebo poslední)**
- **A opět logit-přirozený logaritmus podílu pravděpodobností - hodnoty od minus do plus nekonečna - s tímto už lze v rámci regrese pracovat**

Rovnice, nezávislé proměnné a interakce

- rovnice polytomické regrese - není jedna ale je jich více, konkrétně o jednu méně než je kategorií závislé proměnné
- nominální nezávislé proměnné - buď dichotomie (pak s nimi lze pracovat obdobně jako v lineární regresi), nebo více kategorií (pak převádíme na umělé (dummy) proměnné - SPSS toto udělá za nás - poslední kategorie je pak vztahová = referenční kategorie)
- interakce - vzájemné spolupůsobení dvou či více nezávislých proměnných na proměnnou závislou (viz dříve lineární regrese)

Polytomická logistická regrese - příklad

- V zásadě několik binárních regresí, kde srovnáváme oproti posledním kategoriím
- Příklad-opět viz Matějů,P., Vlachová,K. 2000. Nerovnost, spravedlnost, politika. SLON , str. 284 an. –vysvětlení volby určité strany pomocí sociodemografických charakteristik a škál levice-pravice, liberárně-autoritářské atd.
- hledání faktorů ovlivňujících volbu konkrétní politické strany

Polytomická logistická regrese v SPSS

- výsledkem procedury v SPSS jsou rovnice (o jednu méně než je původních kategorií závislé proměnné, otestování významnosti regresního modelu a jednotlivých parametrů včetně signalizace jednotlivých problémů
- zadání pomocí nabídky-jednotlivé důležité volby
- menu-Analyze-Regression-Multinomial Logistic

Zadání v SPSS- další volby

Categorical vs. Covariate aneb nominální a ordinální vs. spojitá proměnná, u nominálních nemůžeme rozhodnout o kontrastech, vždy se srovnává proti poslední kategorii

Volba Statistics

- - Clasificaion Table

Volba **Model**

- -zadání proměnných a interakcí

Volba **Model**

- **Predicted category**
- **Volba Options – různé transformační funkce**

Zadání v SPSS-syntax

NOMREG

volba BY muz WITH rightor libaut astat intaln
extaln anomie age edyrs polor

/CRITERIA = CIN(95) DELTA(0) MXITER(100)
MXSTEP(5) LCONVERGE(0)

PCONVERGE(1.0E-6) SINGULAR(1.0E-8)

/MODEL

/INTERCEPT = INCLUDE

/PRINT = CLASSTABLE FIT PARAMETER
SUMMARY LRT .

Polytomická regrese v SPSS- výstupy

- Odhad parametrů modelu (Variables in the Equation)
- koeficient a jeho interpretace-udává změnu logitu při změně nezávislé proměnné o jednotku (u spojitých proměnných) ve vztahu k referenční kategorii

Polytomická regrese v SPSS- výstupy

- interpretace koeficientů - změny logitu (příslušné kategorie závislé proměnné k referenční kategorii) při změně spojité nezávislé proměnné o jednotku (resp. u nominální nezávislé proměnné ve vztahu k referenční kategorii této nezávislé proměnné)
- Exp (B)- opět lépe interpretovatelné- jde o změnu šance příslušné kategorie závislé proměnné k referenční kategorii při změně spojité nezávislé proměnné o jednotku (resp. u nominální nezávislé proměnné ve vztahu k referenční kategorii této nezávislé proměnné)

Polytomická regrese v SPSS- výstupy

- LR testy pro jednotlivé proměnné a Waldovy testy pro jednotlivé proměnné (rozdílnost testovaných hypotéz)
- někdy může být otázkou i zda nemá smysl sloučit některé kategorie závislé proměnné, protože mezi nimi model příliš nerozlišuje - poznáme z klasifikační tabulky (existuje i statistický test, ale SPSS ho neumí) a tento krok děláme i pokud je některá kategorie málo zastoupena (obdoba slučování/kolapsování v kontingenčních tabulkách)

Polytomická regrese v SPSS- výstupy

Pseudo R-Square - viz výklad o obdobách R^2 v binární logistické regresi

Goodness-of fit statistics

- testování shody modelu s daty, obdoba Hosmer Lemeshow testu z binární logistické regrese
- Pearson a Deviance- nulová hypotéza: rozdělení očekávaných (odhadů závislé proměnné) a napozorovaných (hodnot závislé proměnné) četností si odpovídají

Polytomická regrese jiné výstupy

Co dělat pokud chci srovnávat jednu kategorii závislé proměnné s jinou než poslední?

- 1) Model spočítat s rekódovanou závislou proměnnou, kde překóduji jako poslední kategorii tu se kterou chci srovnávat
- 2) lze vyjít i ze vzorce, kterým lze koeficienty dopočítávat:

$$\beta_{p/r} = \beta_{p/k} - \beta_{r/k}$$

A poté převést na koeficient u šance

Polytomická regrese jiné výstupy

- Pro lepší interpretaci významu jednotlivých proměnných je dobré z modelu odhadnout pravděpodobnosti pro jednotlivé úrovně nezávislé proměnné a tyto nakreslit do grafu, například pro 5 věkových skupin (na ose x) na osu y vynášíme vypočtenou pravděpodobnost
- Těchto křivek lze vykreslit více například pro různé skupiny-dle pohlaví, vzdělání apod.
- Obecně v polytomické regresi máme mnoho různých srovnání a interpretace závěrů jsou složité (možnosti řešení)
- Doporučení: V polytomické regresi používejte málo nezávislých proměnných, pokud možno s málo kategoriemi

ORDINÁLNÍ LOGISTICKÁ REGRESE

Předpoklady a proměnné

- v ordinální regresí závislá proměnná ordinální(ordinal)-viz ověření testem rovnoběžnosti křivek (parallel lines)
- nezávislé proměnné mohou být nominální (včetně dichotomické) i spojité
- ordinální regrese jako obdoba loglineárních modelů s ordinální proměnnou

Ordinální logistická regrese

Model ordinální regrese

- **pravděpodobnost zařazení do kategorie m a nižší)**- tu nelze korektně regresně modelovat-Proč?
- **Modelujeme obdobu šance**-podíl pravděpodobnosti zařazení do kategorie m a nižší vs. pravděpodobnosti zařazení do kategorie vyšší než m
- **A opět logit**-přirozený logaritmus podílu pravděpodobností - hodnoty od minus do plus nekonečna - s tímto už lze v rámci regrese pracovat

Rovnice, nezávislé proměnné a interakce

- rovnice ordinální regrese
- nominální nezávislé proměnné-bud' dichotomie (pak s nimi lze pracovat obdobně jako v lineární regresi), nebo více kategorií (pak převádíme na umělé (dummy) proměnné-SPSS toto udělá za nás - poslední kategorie je pak vztahová=referenční kategorie)
- interakce-vzájemné spolupůsobení dvou či více nezávislých proměnných na proměnnou závislou (viz dříve lineární regrese)

Ordinální logistická regrese - příklad

- Nejdříve srovnání ordinální a binární logistické regrese k určení prediktorů členství v profesních organizacích
- Ordinální logistická regrese-hledání faktorů ovlivňujících zájem o politiku

Ordinální logistická regrese v SPSS

- výsledkem procedury v SPSS jsou rovnice (o jednu méně než je původních kategorií závislé proměnné, otestování významnosti regresního modelu a jednotlivých parametrů včetně signalizace jednotlivých problémů
- zadání pomocí nabídky-jednotlivé důležité volby
- menu-Analyze-Regression-Ordinal

Zadání v SPSS-další volby

Volba Output

- - Cell Information, Test of Parallel Lines

Volba **Save**

- -různé typy reziduí-slouží k diagnostikování regresních problémů, dále identifikace vlivných pozorování a také pravděpodobnost a skupinovou příslušnost jednotlivých pozorování
- -estimated probabilities

Volba **Options** – různé transformací funkce

Zadání v SPSS-další volby

Volba **Options** – různé transformační funkce

- **Logit, Probit** – pro rovnoměrně rozdělené kategorie závislé proměnné
- **Complementary log-log** $\ln(-\ln(1-P))$ -jsou-li častější **vyšší** hodnoty závislé proměnné
- **Negative log-log** $-\ln(-\ln(P))$ -jsou-li častější **nižší** hodnoty závislé proměnné
- **Cauchy** $\text{tg}(\pi(p-0,5))$ – při výskytu mnoha extrémních hodnot závislé proměnné
- V praxi nejužívanější logit či probit (u latentní proměnné)- logit jednodušší interpretace

Ordinální logistická regrese v SPSS-syntax

PLUM

q30crec BY q36rec

/CRITERIA = CIN(95) DELTA(0)

LCONVERGE(0) MXITER(100) MXSTEP(5)

PCONVERGE(1.0E-6) SINGULAR(1.0E-8)

/LINK = LOGIT

/PRINT = CELLINFO FIT PARAMETER

SUMMARY TPARALLEL .

*PLUM-**P**o**L**ytomous **U**niversal **M**odel (SPSS 10)

Ordinální regrese v SPSS-výstupy

- Odhad parametrů modelu (Variables in the Equation)
- koeficient a jeho interpretace-udává změnu logitu při změně nezávislé proměnné o jednotku (u spojitých proměnných) či ve vztahu k referenční kategorii (u umělých proměnných) pomocí $\exp(B)$ -lepší pro interpretaci než koeficient, značí změnu šance vyšší kategorie oproti nižší

Ordinální regrese v SPSS-výstupy

- Obdoba R^2 (indexu determinace) Mc Fadden, Nagelkerke, Cox Snell
- Cox a Snell R^2 -hodnoty mezi 0 a 1 (této hodnoty nikdy nedosáhne)
- Nagelkerke R^2 -modifikace předchozího, hodnoty 1 může dosáhnout
- Mc Fadden- R^2 -hodnoty mezi 0 a 1 (této hodnoty nikdy nedosáhne)
- nejvhodnější obdobou R^2 z lineární regrese je zřejmě Nagelkerkovo R^2

Ordinální regrese - jiné výstupy

- Pro lepší interpretaci významu jednotlivých proměnných je dobré z modelu odhadnout pravděpodobnosti pro jednotlivé úrovně nezávislé proměnné a tyto nakreslit do grafu, například pro 5 věkových skupin (na ose x) na osu y vynášíme vypočtenou pravděpodobnost
- Těchto křivek lze vykreslit více například pro různé skupiny-dle pohlaví, vzdělání apod.

PŘÍBUZNÍ LOGISTICKÉ REGRESE

Podobné techniky k logistické regresi

- Diskriminační analýza
- Loglineární analýza
- Logitová analýza

- Podobnosti a odlišnosti
- Kde lze najít v SPSS
- Demontrace diskriminační analýzy (7.lekce)

- Návazné procedury – ROC křivky, neuronové sítě