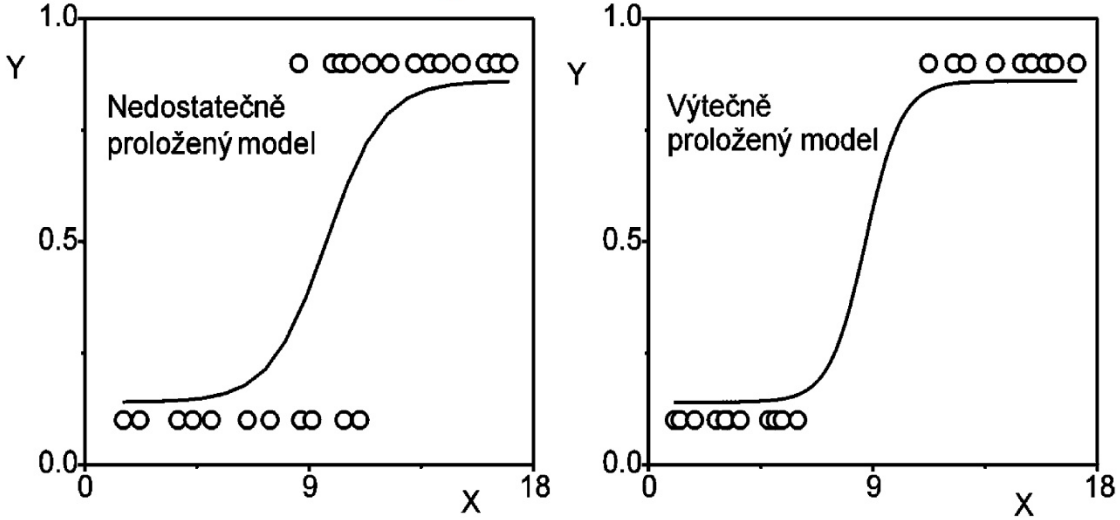
Logistická regrese LR

1. Zaměření metody LR

Navržena v 60tých letech jako alternativní postup k MNČ: závisle proměnná *y* je binární (medicina) značí přítomnost (1) nebo nepřítomnost (0) choroby.

Jde o klasifikaci, když není splněna normalita vícerozměrného modelu.



Krajní případy užití logistického regresního modelu za rozličného proložení.

Rozdíl od lineární regrese: predikuje pravděpodobnost události, která se buď stala (1) nebo nestala (0).

Logitová transformace vede na sigmoidální vztah mezi závisle proměnnou y a nezávisle proměnnými $\boldsymbol{x}$.

Při nízkých x se pravděpodobnost proměnné y blíží k nule, při vysokých x se blíží k jedné.

Logistická regrese používá ***kategorickou závisle proměnnou***
zatímco **lineární regrese** užívá pouze ***spojitou vysvětlovanou proměnnou***.
Logitová transformace **vychází** z ***poměru šancí*** či ***naděje***.

Dle typu *y* se rozlišují:

Binární logistická regrese: *binární závisle proměnná* nabývá pouze dvou hodnot, například přítomnost-absence, muž-žena.
Vektor $\boldsymbol{x}$ obsahuje jednu či více spojitých proměnných (***prediktory***) nebo diskrétních, kategorických (***faktory***).

Ordinální logistická regrese: *ordinální závisle proměnná* nabývá tři a více možných stavů, např. silný nesouhlas, nesouhlas, souhlas, silný souhlas.

Vektor $\boldsymbol{x}$ nezávisle proměnných obsahuje jak ***prediktory*** tak i ***faktory***.

Nominální logistická regrese: *nominální závisle proměnná* o více než třech úrovních, např. mezi kterými je definována pouze odlišnost.

Vektor $\boldsymbol{x}$ může obsahovat jak ***prediktory***, tak i ***faktory***.

2. Logistický regresní model

Potřebujeme vědět, zda se událost stala (1) nebo nestala (0).

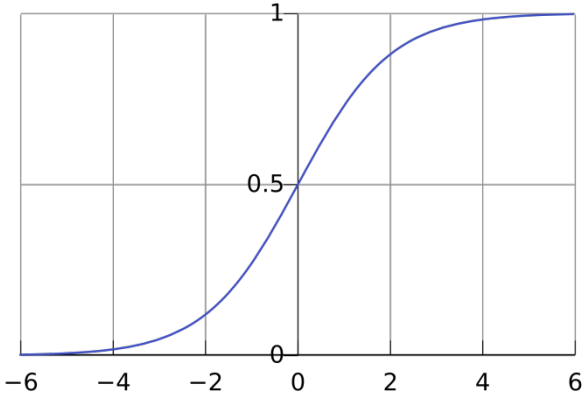
Dichotomická hodnota 0 - 1 závisle proměnné *y* predikuje odhad pravděpodobnosti, že se událost stala (1) či nestala (0).

Je-li predikovaná pravděpodobnost větší než 0.50, pak se událost stala (1),
je-li menší než 0.50, pak se nestala (0).

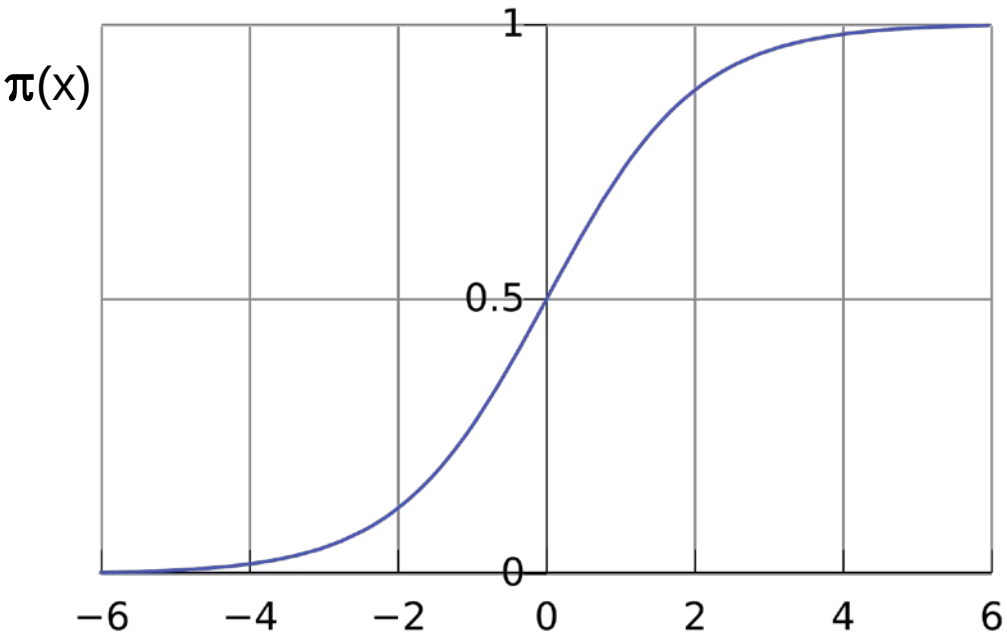
LR porovnává pravděpodobnost události odehrané $L_{(1)}$ vůči pravděpodobnosti události neodehrané $L_{(0)} = 1 - L_{(1)}$.

Využijeme ***pravděpodobnostní poměr*** $\boldsymbol{L_{(1)}/L_{(0)}}$, ve kterém pravděpodobnost $L_{(1)}$ je vyjádřena logistickou funkcí

$$L_{(1)} = \frac{1}{1 + e^{C - Z}}$$



Graf logistické funkce



$$\frac{\pi(x)}{1-\pi(x)} = e^{\beta_0+\beta_1 x}.$$

$$g(x) = \ln \frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x,$$

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{e^{\beta_0+\beta_1 x}+1} = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}.$$

Pravděpodobnostní poměr (zvaný "*poměr šancí*") je vyjádřen

$$\frac{L(1)}{L(0)} = e^{a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_p x_p}$$

kde odhadované koeficienty a_0, a_1, a_2, ..., a_p jsou **míry změny** poměru pravděpodobností $L_{(1)}/L_{(0)}$.

Poměr je lineární funkcí diskriminační funkce o ***p* nezávisle proměnných**

$$Z = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_p x_p$$

Po zlogaritmování a úpravě vyjde $c - z = \ln\left(\frac{L(0)}{L(1)}\right)$, kde C je absolutní člen a_0.

Dle klasifikačního postupu je

$L_{(0)} = P(G = 1 \mid x)$ a

$L_{(1)} = P(G = 1 \mid x) = 1 - P(G = 1 \mid x)$.

a po úpravách bude $\ln\left(\frac{L(1)}{L(0)}\right) = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_p x_p$,

kde $b_0 = -C + a_0$, $b_i = a_i$ pro $i = 1, \ldots, p$.

Například: ve sportu řekneme, že tým má šanci 3:1. Tvrzení říká, že favorizovaný tým má pravděpodobnost vítězství $\frac{3}{3+1} = \frac{3}{4} = 0.75$. Platí tedy pravděpodobnostní poměr $\frac{L(1)}{L(0)} = \frac{0.75}{1 - 0.75} = \frac{3}{1}$.

Aposteriorní pravděpodobnost $P(G = j \mid x)$ zařazení do j-té kategorie:

logistický model lze rozšířit na případ K tříd, a předpokládat, že aposteriorní pravděpodobnost $P(G = j \mid x)$ zařazení do j-té kategorie bude

$$\ln \frac{P(G = 1 \mid \boldsymbol{x})}{P(G = K \mid \boldsymbol{x})} = b_{1,\,0} + \boldsymbol{b}_1^T \boldsymbol{x}$$

$$\ln \frac{P(G = 2 \mid \boldsymbol{x})}{P(G = K \mid \boldsymbol{x})} = b_{2,\,0} + \boldsymbol{b}_2^T \boldsymbol{x}$$

$$\dots$$

$$\ln \frac{P(G = K-1 \mid \boldsymbol{x})}{P(G = K \mid \boldsymbol{x})} = b_{K-1,\,0} + \boldsymbol{b}_K^T \boldsymbol{x}$$

Po zpětné transformaci vychází

$$P(G = j \mid \boldsymbol{x}) = \frac{\exp\,(b_{j,\,0} + \boldsymbol{b}_j^T \boldsymbol{x})}{1 + \sum_{l=1}^{K-1} \exp\,(b_{l,\,0} + \boldsymbol{b}_l^T \boldsymbol{x})}$$

a

$$P(G = K \mid \boldsymbol{x}) = \frac{1}{1 + \sum_{l=1}^{K-1} \exp\,(b_{l,\,0} + \boldsymbol{b}_l^T \boldsymbol{x})}$$

Označíme pravděpodobnost $P(G = K \mid \boldsymbol{x}) = p_k(\mathbf{x}, \mathbf{b})$, aby se zvýraznilo, že jde o funkci regresních parametrů $\boldsymbol{b} = [b_{1,0}, \boldsymbol{b}_1, b_{2,0}, \boldsymbol{b}_2, \dots, b_{K-1,0}, \boldsymbol{b}_{K-1}]$.
(Pro $K = 2$ přechází model na logistický model pro binární proměnnou $y = \mathrm{G}$).

Odhady parametrů:

Pro odhad parametrů logistických modelů se používá metoda maximální věrohodnosti.

Přítomnost v první třídě $y = 1$ je pro $G = 1$.
Nepřítomnost v první třídě $y = 0$ je pro $G = 2$ čili přítomnost ve druhé třídě.

Výchozí data: vektor $\boldsymbol{y}$ rozměru $n \times 1$ a matice $\boldsymbol{X}$ rozměru $n \times m$.
Pro i-tý objekt má y_i hodnotu buď 0, nebo 1 a $\boldsymbol{x}_i^T$ je i-tý řádek matice $\boldsymbol{X}$.

Označme $p(\boldsymbol{x}, \boldsymbol{b}) = p_1(\boldsymbol{x}, \boldsymbol{b})$ a $1 - p(\boldsymbol{x}, \boldsymbol{b}) = p_2(\boldsymbol{x}, \boldsymbol{b})$

a za předpokladu binomického rozdělení $\boldsymbol{y}$ lze zapsat logaritmus věrohodnostní funkce ve tvaru

$$\ln L(\boldsymbol{b}) = \sum_{i=1}^{n} \{y_i \ln p(\boldsymbol{x}_i, \boldsymbol{b}) + (1 - y_i) \ln(1 - p(\boldsymbol{x}_i, \boldsymbol{b})\}$$
$$= \sum_{i=1}^{n} \{y_i \boldsymbol{b}^T \boldsymbol{x}_i - \ln(1 + \exp(\boldsymbol{b}^T \boldsymbol{x}_i\}$$

kde $\boldsymbol{b}^T = \{b_0, \boldsymbol{b}_1\}$ a předpokládá se, že první sloupec matice $\boldsymbol{X}$ obsahuje pouze jedničky (absolutní člen).

Metoda odhadu parametrů:

Pro maximalizaci ln $L(\boldsymbol{b})$ se využívá nulity prvních derivací

$$\boldsymbol{J} = \frac{\mathrm{d}\ln L(\boldsymbol{b})}{\mathrm{d}\boldsymbol{b}} = \sum_{i=1}^{n} x_i \, (y_i - p(\boldsymbol{x}_i, \boldsymbol{b})) = 0$$

Jde o soustavu $m + 1$ nelineárních rovnic vzhledem k $\boldsymbol{b}$.

Řešení soustavy nelineárních rovnic využívá Newtonův-Raphsonovův algoritmus, který vyžaduje matici druhých derivací (hessiánu)

$$\boldsymbol{H} = \frac{\mathrm{d}^2 L(\boldsymbol{b})}{\mathrm{d}\boldsymbol{b}\,\mathrm{d}\boldsymbol{b}^T} = -\sum_{i=1}^{n} \boldsymbol{x}_i \, \boldsymbol{x}_i^T \, p(\boldsymbol{x}_i, \boldsymbol{b}) \, (1 - p(\boldsymbol{x}_i, \boldsymbol{b}))$$

Newtonova-Raphsonova metoda je iterativní, takže výsledkem j-té iterace je zpřesněný odhad

$$\boldsymbol{b}_{(j+1)} = \boldsymbol{b}_{(j+1)} - \boldsymbol{H}_{(j)}^{-1} \, \boldsymbol{J}_j$$

kde pro vektor pravděpodobnosti $\boldsymbol{p}$ rozměru $n \times 1$ s prvky $\boldsymbol{p}(\boldsymbol{x}_i, \boldsymbol{b}_{(j)})$, a diagonální matici vah $\boldsymbol{W}$ rozměru $n \times m$ s prvky lze psát

$$\boldsymbol{J}_{(j)} = \boldsymbol{X}^T (\boldsymbol{y} - \boldsymbol{p}) \quad \text{a} \quad \boldsymbol{H} = -\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}$$

Interpretace regresních koeficientů

Předpoklady o x nejsou a x mohou být **diskrétní** (*faktory*) a **spojité** veličiny (*prediktory*).

$\ln(L_{(1)}/L_{(0)})$ je lineární funkcí nezávisle proměnných a $\ln(L_{(1)}/L_{(0)})$ je nazván **logit** nebo-li **logit transformace pravděpodobnosti.**

Model se nazývá ***vícenásobný logistický regresní model*** (krátce ***logit***) a koeficienty b_i jsou interpretovány jako regresní parametry b_i.

Logit lze dále upravit: dosazením za $L_{(1)} = (1 - L_{(0)})$ dostaneme

$$L_{(0)} = \frac{1}{1 + \exp[-(b_0 + b_1x_1 + b_2x_2 + \ldots + b_px_p)]}$$

Obecně: *Kladné znaménko* koeficientu b_i zvyšuje pravděpodobnost $L_{(0)}$ a *záporné znaménko* tuto pravděpodobnost snižuje.

Diskuse koeficientu b_i:

1) Je-li b_i kladné, funkce exp je větší než 1 a pravděpodobnostní poměr ($L_{(1)}/L_{(0)}$) se bude zvyšovat.

Zvýšení se objeví, když predikovaná pravděpodobnost odehrané události $L_{(1)}$ se zvýší a predikovaná pravděpodobnost neodehrané události $L_{(0)}$ se sníží. Proto má model vyšší predikovanou pravděpodobnost odehrané události $L_{(1)}$.

2) Je-li b_i záporné, je funkce exp menší než 1 a pravděpodobnostní poměr ($L_{(1)}/L_{(0)}$) se bude snižovat.

3) Pro b_i roven 0, vede funkce exp k hodnotě 1 čili k žádné změně pravděpodobnosti.

$L_{(1)}$ se mění od 0 do 1 a pravděpodobnostní poměr $L_{(1)}/L_{(0)}$ se mění od 0 do $\propto$.

Je-li $L_{(1)} = 0.5$, je poměr $L_{(1)}/L_{(0)}$ roven 1.

Ve stupnicí pravděpodobnostního poměru odpovídají hodnoty od 0 do 1 hodnotám pravděpodobnosti $L_{(1)}$ od 0 do 0.5.

Na druhé straně hodnoty pravděpodobnosti $L_{(1)}$ od 0.5 do 1 vedou k poměru $L_{(1)}/L_{(0)}$ od 1 do $\propto$.

Po zlogaritmování poměru $L_{(1)}/L_{(0)}$ bude tato asymetrie odstraněna, neboť pro

pro $L_{(1)} = 0$ je $\ln (L_{(1)}/L_{(0)}) = \propto$, pro $L_{(1)} = 0.5$ je $\ln (L_{(1)}/L_{(0)}) = 0.0$,

pro $L_{(1)} = 1.0$ je $\ln (L_{(1)}/L_{(0)}) = -\propto$.

Test významnosti regresních koeficientů

Logistická regrese umožňuje **testovat významnost** koeficientů čili ověřit, že regresní koeficient se liší od nuly.

Nula zde značí, že pravděpodobnostní poměr $L_{(1)}/L_{(0)}$ se nemění a pravděpodobnost tím pádem není ovlivněna.

Studentův *t*-test k vyšetření statistické významnosti jednotlivých regresních koeficientů.

Waldovo testační kritérium $\boldsymbol{W_{a,i} = (b_i/s(b_i))^2}$ vyčísluje statistickou významnost pro odhady regresních koeficientů stejně jako ve vícenásobné regresi.

Pro kategorické proměnné má $W_{a,i}$ počet stupňů volnosti roven o 1 méně než je počet kategorií.

Waldova statistika W_a má nežádoucí vlastnost: pro velikou hodnotu regresního koeficientu b_i a veliký odhad jeho směrodatné odchylky $s(b_i)$ je výsledkem příliš malá hodnota testačního kritéria $W_{a,i}$, která vede k selhání zamítnutí nulové hypotézy, že regresní koeficient je nulový. Proto, je -li regresní koeficient veliký, neužijeme Waldova kritéria.

Parciální korelace

Je obtížné určit příspěvek jednotlivých proměnných.

Příspěvek každé proměnné závisí také na ostatních proměnných v logistickém modelu.
K vyšetření parciální korelace mezi závisle proměnnou a každou nezávisle proměnnou se užívá **korelační koeficient** R_i, (v intervalu od -1 do +1).

1) **Kladné hodnoty** R_i: když roste hodnota R_i, zvyšuje se pravděpodobnost objektu "v události" $L_{(1)}$.
2) **Záporné hodnoty** R_i: naopak snižuje se pravděpodobnost objektů "v události" $L_{(1)}$.
3) **Malé hodnoty** R_i: proměnná má malý vliv na model.

Korelační koeficient R_i se vyčíslí $R_i = \pm\sqrt{\frac{W_{a,i} - 2df}{-2\ln L_{(0)}}}$, kde $2df$ značí počet stupňů volnosti a týká se počtu odhadovaných parametrů.

$-2 \ln L_{(0)}$ je záporný dvojnásobek logaritmu pravděpodobnosti základního logistického modelu, který neobsahuje žádné proměnné kromě absolutního členu (úseku) b_0.
Je-li Waldova statistika $W_{a,i}$ menší než $2df$, je R_i položeno definitoricky 0.

Kategorické proměnné

Výhodou logistického modelu je možnost užívat i kategorické nezávisle proměnné *x,* zvané **faktory**.

Za **faktor** lze použít numerickou, textovou nebo datumovou hodnotu, zvanou *úroveň* nebo *referenční hladina*.

Nejjednodušší situace je **jediný faktor** *x* se dvěma možnými hodnotami.

Pojmu šance se hodně využívá v biomedikálních aplikacích. Je mírou spojení binární proměnné, jako je faktor risku výskytu dané události, například nemoci.

Kategorická proměnná čili **faktor** má dvě úrovně, tj. $x = 0$ značící muže a $x = 1$ značící ženy a logistickou rovnici pak bude $L(1) = \frac{1}{1 + e^{-a - b x}}$, odhad parametru *a* a odhad *b*.

Odhad *b* představuje přirozený logaritmus pravděpodobnostního poměru žen a mužů.

Odhad *a* je přirozený logaritmus pravděpodobnostního poměru mužů ($x = 0$). Existuje -li pouze jedna dichotomní proměnná, není potřebné provádět logistickou regresní analýzu.

Přibližný interval spolehlivosti pro pravděpodobnostní poměr jako pro binární proměnnou se vypočte užitím odhadu směrnice *b* a odhadu je jí směrodatné odchylky.

3. Volba proměnných

Například úloha logistické předpovědi infarktu: data jsou z dlouhodobého sledování z počátku zdravých pacientů, u kterých byla dlouhodobě provedena opakovaná měření. Několik jedinců bylo postiženo infarktem. Byl sledován **výběr nezávisle proměnných**, které by mohly odhalit blížící se infarkt.

Výběr **účinných** nezávisle proměnných byl předem lékaři vytypován. Častěji však uživatel předem neví nic o nezávisle proměnných.

Proměnné ***x*** jsou nejprve vyšetřovány, která je nejvíce spjata z dichotomní závisle proměnnou.

Studentův *t*-test významnosti jednotlivých parametrů: užívá se dostatečně vysoká hladina významnosti, například $\alpha = 0.15$, aby užitečná nezávisle proměnná nemohla být odstraněna.

Vyšetření zredukuje počet nezávisle proměnných na 10 či ještě méně.

Pak nastoupí *kroková logistická regresní analýza*: jde o test, zda proměnná x_i zlepší predikční schopnost modelu. Postupy a jejich kritéria jsou užita k rozhodování, kolik proměnných x_i a které je třeba užít.

Testy v dopředné krokové analýze jsou postaveny na χ^2-statistice: velká hodnota χ^2 nebo malá spočtená hladina významnosti P ukazují, že nezávisle proměnná by měla být zařazena do proměnných. Nalezená velká hodnota χ^2 ukazuje , že proměnné jsou užitečné.

4. Těsnost proložení logistickým modelem

Před analýzou je třeba posoudit, zda nejsou odlehlé hodnoty. Rozptylové diagramy snadno odhalí **odlehlé body**. Proměnné nemusí být **normálně** rozděleny. **Regresní diagnostika** s analýzou vlivných bodů odhalí *O* a *E*.

Logistická křivka má esovitý tvar a vystihuje logistický model, který je vzhledem ke koeficientům ***b*** nelineární.

Mírou těsnosti proložení navrženého modelu daty je hodnota pravděpodobnosti $L_{(1)}$, že se událost uskuteční.

Místo veličiny $L_{(1)}$ se používá tzv. ***odchylka*, deviance** $D = -2 \ln L_{(1)}$ čili
D* = -2*LL, když ***D* představuje** *míru těsnosti proložení dat logistickým modelem:*

1) ***Dobrý model*** vede k vysoké pravděpodobnosti objektů v události $L_{(1)}$, což přetransformáno do veličiny $-2 \ln L_{(1)}$ poskytne malou hodnotu *blízkou nule*.

Je-li L(1) = 0,9999, pak je D = -2LL = 0,0002. Jeli 0,9, pak je D = 0,210. Je-li 0,8, pak je D = 0,446.

2) ***Minimální hodnotou*** pro $-2 \ln L_{(1)}$ je *nula*, při které je dosaženo naprosto perfektní těsnosti proložení.

Rozdíl v odchylce je definován G = *D*(model bez proměnné) - *D*(model s proměnnou)

čili $G = -2 \ln \frac{\text{pravděpodobnost modelu bez proměnné}}{\text{pravděpodobnost modelu s proměnnou}}$. Veličina *G* proto odpovídá věrohodnostnímu poměru.

Těsnost proložení: spočívá porovnání experimentálních hodnot E s vypočtenými V :

Pearsonův test dobré shody χ^2 se užije, když model platí:

-Velká hodnota χ^2 indikuje špatné proložení modelu.

-Malé hodnoty vypočtené hladiny významnosti P indikují špatné proložení modelu.

Nejužívanější způsoby posouzení těsnosti proložení:

a) **Klasický Pearsonův přístup** začíná s identifikováním různých kombinací hodnot proměnných v regresním modelu, tj. *vzorů*.

Například **dvě dichotomní proměnné**,(pohlaví a zaměstnán) vedou na 4 kombinace: *muž zaměstnán, muž nezaměstnán, žena zaměstnána, žena nezaměstnána.*

- Pro každou kombinaci vyčíslíme počet E experimentálních hodnot jednotlivců (objektů) ve třídě I a II.

- Podobně pro každého jednotlivce vypočteme pravděpodobnost, že se nachází ve třídě I a ve třídě II logistickou regresní analýzou.

- Suma těchto pravděpodobností pro daný vzor se označí V.

- Testační statistika testu dobré shody χ^2 se vyčíslí jako $\chi^2_{\exp} = \sum_{i=1}^{\cdot} 2E\left(\ln \frac{E}{V}\right)$, kde suma se provede přes všechny odlišné vzory. Rezidua se sledují právě pro tyto odlišné vzory.

b) **Hosmerův-Lemeshowův test dobré shody** byl navržen v 1982. Pearsonův χ^2-test dobré shody k redukci v logaritmech hodnoty pravděpodobnosti je mírou sledování zlepšení těsnosti zavedením jedné či více nezávisle proměnných.

Základní model, který je podobný výpočtu sumy čtverců při použití pouze průměrů, poskytuje nulovou linii k porovnání.

Vedle χ^2-testu existuje několik R^2-podobných měr k posouzení těsnosti proložení, obdoba koeficientu determinace ve vícenásobné regresi.

"Pseudo R^2" v logistické regresi pro logitový model se vypočte dle

$$R^2_{logit} = \frac{2\ln L_{nul} - (-2\ln L_{model})}{-2\ln L_{nul}} = -\frac{D_{model} + D_{nul}}{D_{nul}}$$

c) **Metoda klasifikačních matic**, vyvinutých v diskriminační analýze slouží k vyhodnocení predikční schopnosti v pojmech zařazení do třídy.

Pravděpodobnost zařazení do třídy I je vypočtena pro každého jednotlivce (objekt) ve výběru a výsledný počet je uspořádán vzestupně.
Pravděpodobnosti jsou pak rozděleny do 10 skupin (decily).

Pro každý naměřený počet jednotlivců ve třídě I je vyčíslen počet *E*. Užitím logické regrese jsou pro jedince v každém decilu vypočteny počty *V*. Pak se vyčíslí ***Pearsonova χ^2 - statistika testu dobré shody***

$$\chi^2_{\exp} = \sum_{i=1}^{n} \frac{(E - V)^2}{V}$$

kde sumace se provede přes obě třídy a 10 decilů.

Velká hodnota χ^2 nebo malá hodnota *P* indikují **dobré proložení modelu.**

5. Kvalita vyhodnocení logistickou regresí

Třídíme objekty do tříd, musíme nalézt *prahový bod pravděpodobnosti* P_C: objekt je "v události", když pravděpodobnost události větší nebo rovna hodnotě P_C.

Graf prahové operační charakteristiky ROC k detekci signálu, když signál nebylo vždy možné správně přijmout.

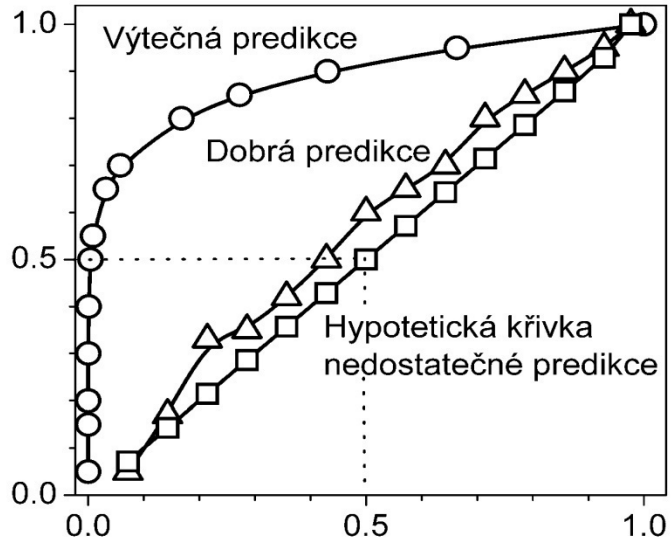
Na ose *y* je procento správně zařazených objektů "v události" nazvané **pozitivní podíl** (a v lékařském výzkumu nazývané **citlivost**).

Na ose *x* je procento nesprávně zařazených objektů nazvané **falešný podíl** nebo v lékařském výzkumu **"1 minus specificita"** (v lékařském výzkumu nazývané **senzitivita** zařazených krys pro správně zařazené krysy a **specificita** krys pro falešně zařazené krysy).

1) **Horní křivka v grafu ROC představu je výtečnou predikci:** i pro malé podíly nesprávně zařazených objektů se získá vysoké procento správně zařazených objektů, které **skutečně jsou "v události"**.

2) **Střední křivka je skutečná křivka při uvažování malého počtu nezávisle proměnných,** třeba dvou. Vysoké procento (80 %) objektů správně zařazených v události je v poměru k 65 % chybně zařazených v události na nepřijatelné hladině.

3) **Dolní hypotetická křivka, (přímka) odpovídá nahodilým výsledkům,** například házení mincí. Blízkost střední křivky k dolní ukazuje , že je potřeba buď volit jinou, anebo přidat ještě další nezávisle proměnnou, abychom získali lepší model, i když je ale tento model statisticky významný na spočtené hladině $P = 0.009$.



Graf prahové operační charakteristiky ROC k detekci signálu, když signál není možné vždy správně přijmout.

(a) **Vybereme prahový bod na dolní části křivky grafu ROC** a nechceme mít příliš mnoho objektů, zařazených jako "v události", bude se nazývat **přísný práh**.
Nevýhoda: je ztráta mnoha objektů, které jsou "v události".

(b) **Vybereme prahový bod na horní části křivky grafu ROC** a chceme mít hodně objektů zařazených jako "v události", bude se nazývat **nedbalý práh.**
Nevýhoda: sice velmi málo objektů "v události" bude ztraceno ale mnoho objektů "v neudálosti" bude chybně označeno jako "v události".

Křivky v grafu ROC **musí procházet body (0, 0) a (1, 1).**

Maximální plocha pod křivkou je je dna čili 100%.
Numerická hodnota velikosti plochy bude blízká 1, když predikce modelu bude výtečná.
Když bude **plocha blízká hodnotě 0.5**, bude predikce modelu špatná.

Křivka ROC je proto užitečná při rozhodování, který ze dvou logistických modelů vybrat: lepší model dosáhne větší plochy pod křivkou ROC ale také větší výšky prahového bodu na křivce ROC.

Většina programů vybírá logistický model podle kritéria největší plochy pod křivkou ROC.

6. Aplikace logistické regrese

Modelu vícenásobné logistické regrese se často užívá k odhadu pravděpodobnosti jisté události, která se přihodí danému objektu.

Výběr dat může být uskutečněn dvojím způsobem:

1. Výběr cross-validation: je získán náhodným způsobem a pozorování provedeno v uvedeném časovém období. Z tohoto výběru se vyčlení dva podvýběry: **první podvýběr**, který obsahuje hodně zkušenosti o události, a **druhý podvýběr**, který obsahuje zbylé údaje.

Na datech prvního podvýběru se vyčíslí logistický regresní model, který pak může být aplikován na člena druhého podvýběru.

2. Případ řídícího výběru: spočívá v získání dvou náhodných výběrů: **první výběr**, ve kterém se událost objeví, a **druhý výběr**, ve kterém se událost neobjeví.

Hodnoty predikovaných proměnných se musí získat retrospektivním způsobem, z minulých záznamů nebo ze vzpomínek. Konstanta *a* musí být nastavena tak, aby vyjadřovala pravý poměr objektu v události.

Existují důležité požadavky:

1. Model předpokládá, že logaritmus pravděpodobnostního poměru je lineárně závislý na nezávislých proměnných. Nesplnění by mělo být předem prověřeno buď užitím měr těsnosti proložení, nebo jinými způsoby. To může vyžadovat transformaci dat.

2. Výpočty jsou často časově náročné, a proto by měl uživatel rozumně redukovat počet proměnných.

3. Logistická regrese by se neměla užívat k vyhodnocení faktorů risku v dlouhodobých studiích, ve kterých jsou je dnotlivé studie rozličné délky.

4. Regresní koeficienty pro nezávisle proměnnou v logistickém regresním modelu závisí na ostatních proměnných, zařazených do logistického modelu. Koeficienty pro stejnou nezávisle proměnnou, když se použijí různé výběry proměnných, mohou být zcela odlišné.

5. Je -li užita sehraná analýza, kterákoliv proměnná pro sehrání nemůže být použita jako nezávisle proměnná.

6. Jsou okolnosti, kde metoda maximální věrohodnosti odhadovaných regresních koeficientů neposkytne odhady, tj. nekonverguje .

Příklad 4.26 *Volba proměnných k popisu leukemie*

Lee (1980) publikoval data o leukemii pacientů. Závisle proměnnou je binární proměnná *REMISS*, zda se objeví *ústup leukemie y* (1) či neobjeví (0). Nezávisle proměnnými ***x*** jsou:

CELL celulirita, buněčnost sraženiny kostní dřeně,
SMEAR skvrna diferenčního procenta napadení,
INFIL procento infiltrátu kostní dřeně buňkou leukemie,
LI procento označeného indexu leukemických buněk kostní dřeně,
BLAST absolutní počet napadení v periferní krvi,
TEMP nejvyšší teplota před začátkem léčby.

Otázkou je, které nezávisle proměnné jsou statisticky významné v navrženém logistickém regresním modelu.

Data: $n = 27$, $p = 6$,

Index	*REMISS*	*CELL*	*SMEAR*	*INFIL*	*LI*	*BLAST*	*TEMP*
1	1	0.8	0.83	0.66	1.9	1.1	0.996
2	1	0.9	0.36	0.32	1.4	0.74	0.992
3	0	0.8	0.88	0.7	0.8	0.176	0.982
4	0	1	0.87	0.87	0.7	1.053	0.986
5	1	0.9	0.75	0.68	1.3	0.519	0.98
6	0	1	0.65	0.65	0.6	0.519	0.982
7	1	0.95	0.97	0.92	1	1.23	0.992
8	0	0.95	0.87	0.83	1.9	1.354	1.02
9	0	1	0.45	0.45	0.8	0.322	0.999
10	0	0.95	0.36	0.34	0.5	0	1.038
11	0	0.85	0.39	0.33	0.7	0.279	0.988
12	0	0.7	0.76	0.53	1.2	0.146	0.982
13	0	0.8	0.46	0.37	0.4	0.38	1.006

Index	*REMISS*	*CELL*	*SMEAR*	*INFIL*	*LI*	*BLAST*	*TEMP*
14	0	0.2	0.39	0.08	0.8	0.114	0.99
15	0	1	0.9	0.9	1.1	1.037	0.99
16	1	1	0.84	0.84	1.9	2.064	1.02
17	0	0.65	0.42	0.27	0.5	0.114	1.014
18	0	1	0.75	0.75	1	1.322	1.004
19	0	0.5	0.44	0.22	0.6	0.114	0.99
20	1	1	0.63	0.63	1.1	1.072	0.986
21	0	1	0.33	0.33	0.4	0.176	1.01
22	0	0.9	0.93	0.84	0.6	1.591	1.02
23	1	1	0.58	0.58	1	0.531	1.002
24	0	0.95	0.32	0.3	1.6	0.886	0.988
25	1	1	0.6	0.6	1.7	0.964	0.99
26	1	1	0.69	0.69	0.9	0.398	0.986
27	0	1	0.73	0.73	0.7	0.398	0.986

Řešení: Byl užit program NCSS2000.

1. Odhad regresních koeficientů.

Proměnná	**Regresní koefficient**	**Směrodatná odchylka**	**χ^2 pro $\beta = 0$**	**Spočtená hladina *P***	**Poslední R^2**
Úsek	58.0387	71.23627	0.66	0.415224	0.032124
CELL	24.66053	47.83722	0.27	0.606197	0.013113
SMEAR	19.29247	57.94952	0.11	0.739196	0.005511
INFIL	-19.60012	61.68098	0.10	0.750662	0.005023
LI	3.895928	2.3371	2.78	0.095516	0.121993
BLASTS	0.1510942	2.278567	0.00	0.947130	0.000220
TEMP	-87.43308	67.57322	1.67	0.195699	0.077243

χ^2 udává Pearsonovo testační kritérium χ^2 pro 1 stupeň volnosti k testu H_0: $\beta_i = 0$ vs. H_A: $\beta_i \neq 0$. Vyčíslí se Waldovo kritérium $W^2_{a,i} = [b_i / s(b_i)]^2$.

Test významnosti b_i: je-li spočtená hladina *P* menší než předvolená $\alpha = 0.05$, je parametr b_i statisticky významný. Všechny prediktory se jeví jako statisticky nevýznamné.

Poslední R^2 udává hodnotu, která se přičte k celkové R^2, když se tato nezávisle proměnná přidá do logistického regresního modelu. Vypočte se dle $R^2 = \chi^2(df)/[\chi^2(df) + n - p - 1]$.

2. Nalezený model v transformované formě.

Nalezený logistický regresní model: 58.0387 + 24.66053**CELL* + 19.29247**SMEAR* - 19.60012**INFIL* + 3.895928**LI* + 0.1510942**BLAST* - 87.43308**TEMP*.

3. Přehled modelu.

R^2 modelu	*df*	Odchylka *D*	Spočtená hladina významnosti *P*
0.386900	6	12.62	0.049463

Odchylka *D* testuje, zda všechny regresní koeficienty β_i kromě úseku β_0 jsou rovny nule. Protože je spočtená *P* menší než $\alpha = 0.05$, je regresní model statisticky významný.

4. Klasifikační tabulka.

		Nalezeno predikcí logistickým modelem		
Dáno závisle proměnnou		Ne	Ano	Celkově
Ne	**Četnost**	15	3	18
	Řádkové procento	83.33	16.67	100.00
	Sloupcové procento	78.95	37.50	66.67
Ano	**Četnost**	4	5	9
	Řádkové procento	44.44	55.56	100.00
	Sloupcové procento	21.05	62.50	33.33
Celkově	**Četnost**	19	8	27
	Řádkové procento	70.37	29.63	
	Sloupcové procento	100.00	100.00	

Procento správně klasifikovaných = 74.07

Tabulka přináší četnosti, řádková procenta a sloupcová procenta predikovaných objektů a nakonec je procento správně klasifikovaných objektů. Jde o procento z celkového počtu, které padne na diagonálu tabulky.

5. Predikovaná klasifikace.

Řádek	Daná třída	Nalezená třída	Logistické skóre	Reziduum
1	Ano (1)	Ano (1)	0.790659	0.209341
2	Ano (1)	Ne (0)	0.435096	0.564904
3	Ne (0)	Ne (0)	0.155573	-0.155573
4	Ne (0)	Ne (0)	0.290862	-0.290862
5	Ano (1)	Ano (1)	0.697019	0.302981
6	Ne (0)	Ne (0)	0.280081	-0.280081
7	Ano (1)	Ne (0)	0.376587	0.623413
8	Ne (0)	Ano (1)	0.600523	-0.600523
9	Ne (0)	Ne (0)	0.165277	-0.165277
10	Ne (0)	Ne (0)	0.000858	-0.000858
11	Ne (0)	Ne (0)	0.027695	-0.027695
12	Ne (0)	Ne (0)	0.169829	-0.169829
13	Ne (0)	Ne (0)	0.000956	-0.000956
14	Ne (0)	Ne (0)	0.000001	-0.000001
15	Ne (0)	Ano (1)	0.575858	-0.575858
16	Ano (1)	Ano (1)	0.725752	0.274248
17	Ne (0)	Ne (0)	0.000055	-0.000055
18	Ne (0)	Ne (0)	0.228173	-0.228173
19	Ne (0)	Ne (0)	0.000064	-0.000064
20	Ano (1)	Ano (1)	0.677843	0.322157
21	Ne (0)	Ne (0)	0.015910	-0.015910
22	Ne (0)	Ne (0)	0.007445	-0.007445
23	Ano (1)	Ne (0)	0.247684	0.752316
24	Ne (0)	Ano (1)	0.851096	-0.851096

25	Ano (1)	Ano (1)	0.938464	0.061536
26	Ano (1)	Ne (0)	0.461177	0.538823
27	Ne (0)	Ne (0)	0.279469	-0.279469

Daná třída určuje zadanou skutečnou třídu. ***Nalezená třída*** představuje nalezenou třídu na základě logistického regresního modelu. ***Logistické skóre*** je odhad pravděpodobnosti, že objekt patří do třídy *Ne*. Reziduum představu je to rozdíl mezi Logistickým skóre a indexem skutečné třídy. Index třídy *Ne* je 0 a index třídy *Ano* je 1.

6. Chybně klasifikované objekty.

Řádek	Daná třída	Nalezená třída	Skóre	Reziduum
2	Ano (1)	Ne (0)	0.435096	0.564904
7	Ano (1)	Ne (0)	0.376587	0.623413
8	Ne (0)	Ano (1)	0.600523	-0.600523
15	Ne (0)	Ano (1)	0.575858	-0.575858
23	Ano (1)	Ne (0)	0.247684	0.752316
24	Ne (0)	Ano (1)	0.851096	-0.851096
26	Ano (1)	Ne (0)	0.461177	0.538823

Jsou zde zobrazeny pouze chybně zařazené řádky.

Cvičení v programu STATISTICA

Úloha B29 *Rozlišení vrozené vady metabolismu u 8 až 9 letých dětí TNT testem ITPA*

Výběr 54 dětí stáří 8 až 9 let s vrozenou vadu metabolizmu Transient Neonatal Tyrosinemia (TNT) byl porovnán s kontrolním výběrem zdravých dětí Illinoiskými testy psycholingualní schopnosti ITPA. Testujte vztah kódu nemocnosti GP (1 nemocný a 0 zdravý) vůči 11 nezávislým proměnným, a to vypouštěním některých nezávisle proměnných které vrozené vady v metabolismu dětí lze posuzovat testem ITPA.

◯ ***Data:*** Výběr dat TNT obsahuje **54 dětí v řádcích pro 10 znaků ve sloupcích**:
i (ID) značící identifikační číslo dítěte,
***j* (GP) kód nemocnosti je závisle proměnná: 1 je nemocný, 0 je zdravý**,
a nezávisle proměnné:
jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (AUDITORY) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (VISMEM) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (VISCLOS) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (MANUAL) značí manuální zručnost.

i	j	jj	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
1	1	1	26	38	44	26	37	38	22	28	30	33
...	...	...	...	...	...	...	...	...	...	...	...	...
54	0	0	38	40	40	32	36	41	36	43	41	28

Efekt	GP - Odhady parametrů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 95, %	Horní LS 95, %	p
Abs.člen		1	-6,46735	3,145394	4,227685	-12,6322	-0,302491	0,039770
SEX		2	0,81050	0,742746	1,190770	-0,6453	2,266256	0,275174
AUDITORY		3	0,11115	0,070430	2,490416	-0,0269	0,249187	0,114541
VISUAL		4	-0,01189	0,050823	0,054716	-0,1115	0,087723	0,815052
VISMEM		5	0,07274	0,051966	1,959094	-0,0291	0,174588	0,161610
AUDASSOC		6	-0,02514	0,045743	0,301975	-0,1148	0,064517	0,582647
AUDMEMO		7	0,02674	0,047616	0,315328	-0,0666	0,120065	0,574429
VISASSOC		8	-0,00693	0,056582	0,015007	-0,1178	0,103967	0,902500
VISCLOS		9	0,16040	0,089228	3,231441	-0,0145	0,335281	0,072237
VISEXPR		10	-0,01088	0,041853	0,067547	-0,0929	0,071153	0,794942
GRAMCLOS		11	-0,02802	0,061647	0,206569	-0,1488	0,092808	0,649470
MANUAL		12	-0,09979	0,061550	2,628478	-0,2204	0,020847	0,104962
Měřítko			1,00000	0,000000		1,0000	1,000000	

Odhady parametrů	QCEXPERT		
Parametr	**Odhad**	**Sm. odch**	**p-hodnota**
Abs	6,467351017	2,104989489	0,0037194
SEX	-0,810501741	0,49706694	0,110459008
AUDITORY	-0,111146088	0,047133869	0,023100532
VISUAL	0,01188821	0,034012202	0,728439655
VISMEM	-0,072735842	0,034777283	0,042571755
AUDASSOC	0,025136564	0,030612266	0,416210482
AUDMEMO	-0,026738505	0,031866209	0,406171438
VISASSOC	0,006931486	0,037866294	0,855638323
VISCLOS	-0,160397609	0,059713777	0,010309523
VISEXPR	0,010877566	0,028009299	0,699714598
GRAMCLOS	0,028018574	0,04125615	0,500776803
MANUAL	0,099788434	0,041191042	0,019803483

Odhady parametrů:

jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (**AUDITORY**) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (**VISMEM**) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

Průběh iteračního procesu zjemňování odhadů parametrů *b*

Efekt	GP - Historie iterací (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0						
	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen		1	0,0000	-5,1009	-6,2696	-6,4624	-6,4673
SEX		2	0,0000	0,5787	0,7741	0,8096	0,8105
AUDITORY		3	0,0000	0,0785	0,1063	0,1110	0,1111
VISUAL		4	0,0000	-0,0081	-0,0112	-0,0119	-0,0119
VISMEM		5	0,0000	0,0569	0,0704	0,0727	0,0727
AUDASSOC		6	0,0000	-0,0252	-0,0256	-0,0252	-0,0251
AUDMEMO		7	0,0000	0,0209	0,0258	0,0267	0,0267
VISASSOC		8	0,0000	-0,0078	-0,0071	-0,0069	-0,0069
VISCLOS		9	0,0000	0,1331	0,1573	0,1603	0,1604
VISEXPR		10	0,0000	-0,0061	-0,0100	-0,0109	-0,0109
GRAMCLOS		11	0,0000	-0,0167	-0,0264	-0,0280	-0,0280
MANUAL		12	0,0000	-0,0763	-0,0967	-0,0997	-0,0998
Věrohodnost			-37,4299	-30,8522	-30,5934	-30,5885	-30,5885

Posouzení statistické významnosti odhadů parametrů dle jejich intervalu spolehlivosti

Efekt	GP - Intervaly spolehlivost odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Úroveň Efekt	Sloupec	Dolní LS 95, %	Horní LS 95, %
Abs.člen		1	-12,6322	-0,302491
SEX		2	-0,6453	2,266256
AUDITORY		3	-0,0269	0,249187
VISUAL		4	-0,1115	0,087723
VISMEM		5	-0,0291	0,174588
AUDASSOC		6	-0,1148	0,064517
AUDMEMO		7	-0,0666	0,120065
VISASSOC		8	-0,1178	0,103967
VISCLOS		9	-0,0145	0,335281
VISEXPR		10	-0,0929	0,071153
GRAMCLOS		11	-0,1488	0,092808
MANUAL		12	-0,2204	0,020847

Efekt	GP - Odhadnutá korelační matice odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0						
	Abs.člen	SEX	AUDITOR	VISUAL	VISMEM	AUDASSOC	AUDMEMO
Abs.člen	1,000000	-0,492832	-0,316816	0,011010	-0,390001	0,072154	-0,180429
SEX	-0,492832	1,000000	0,309219	-0,090876	0,297540	-0,001727	-0,143760
AUDITORY	-0,316816	0,309219	1,000000	-0,172352	0,277184	-0,164041	0,155892
VISUAL	0,011010	-0,090876	-0,172352	1,000000	-0,149863	0,135106	-0,152214
VISMEM	-0,390001	0,297540	0,277184	-0,149863	1,000000	-0,071590	0,077652
AUDASSOC	0,072154	-0,001727	-0,164041	0,135106	-0,071590	1,000000	-0,038633
AUDMEMO	-0,180429	-0,143760	0,155892	-0,152214	0,077652	-0,038633	1,000000
VISASSOC	-0,113032	-0,060806	-0,171600	-0,103774	-0,071437	-0,155525	-0,143360
VISCLOS	-0,348970	0,144857	0,006998	-0,079119	-0,152896	-0,146624	-0,160934
VISEXPR	0,195515	-0,096102	-0,389055	-0,085330	0,058670	-0,180347	-0,080294
GRAMCLOS	-0,089562	0,060292	-0,163300	-0,188959	-0,222943	-0,009959	-0,064025
MANUAL	-0,110185	0,057594	-0,201454	0,002873	-0,113120	-0,155879	-0,248147

Posouzení korelace prediktorů

jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (**AUDITORY**) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (**VISMEM**) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

pokračování Efekt	GP - Odhadnutá korelační matice odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0				
	VISASSOC	VISCLOS	VISEXPR	GRAMCLOS	MANUAL
Abs.člen	-0,113032	-0,348970	0,195515	-0,089562	-0,110185
SEX	-0,060806	0,144857	-0,096102	0,060292	0,057594
AUDITORY	-0,171600	0,006998	-0,389055	-0,163300	-0,201454
VISUAL	-0,103774	-0,079119	-0,085330	-0,188959	0,002873
VISMEM	-0,071437	-0,152896	0,058670	-0,222943	-0,113120
AUDASSOC	-0,155525	-0,146624	-0,180347	-0,009959	-0,155879
AUDMEMO	-0,143360	-0,160934	-0,080294	-0,064025	-0,248147
VISASSOC	1,000000	0,068065	-0,000418	-0,184095	-0,066991
VISCLOS	0,068065	1,000000	-0,149529	-0,191816	-0,138085
VISEXPR	-0,000418	-0,149529	1,000000	-0,046606	-0,086142
GRAMCLOS	-0,184095	-0,191816	-0,046606	1,000000	0,146760
MANUAL	-0,066991	-0,138085	-0,086142	0,146760	1,000000

Posouzení kvality těsnosti proložení experimentálních dat logistickým modelem

Odezva	GP - Kvalita proložení: Hosmer-Lemeshow Test (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 6.5670, p hodn. = 0.5839							
	Skupi1a	Skupi2a	Skupi3a	Skupi4a	Skupi5a	Skupi6a	Skupi7a	Skupi8a
0: Pozorov.	5,00	4,00	3,00	3,00	2,00	2,00	3,00	2,00
Očekáv.	4,54	4,03	3,56	3,07	2,72	2,43	2,05	1,77
1: Pozorov.	0,00	1,00	2,00	2,00	3,00	3,00	2,00	3,00
Očekáv.	0,46	0,97	1,44	1,93	2,28	2,57	2,95	3,23
Vš. skup.	5,00	5,00	5,00	5,00	5,00	5,00	5,00	5,00

Odezva	GP - Kvalita proložení: Hosmer-Lemeshow Test (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 6.5670, p hodn. = 0.5839		
	Skupi9a	Skupi10	Row Tot.
0: Pozorov.	3,00	0,00	27,0
Očekáv.	1,37	1,45	
1: Pozorov.	2,00	9,00	27,0
Očekáv.	3,63	7,55	
Vš. skup.	5,00	9,00	54,0

Stat.	GP - Statistiky kvality proložení (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0		
	SV	Stat.	Stat/sv
Odchylka	42	61,1770	1,456594
Deviance v měřítku	42	61,1770	1,456594
Pearsonovo Chi2	42	50,8459	1,210617
Scaled P. Chi2	42	50,8459	1,210617
AIC		85,1770	
BIC		109,0448	
Cox-Snell R2		0,2238	
Nagelkerke R2		0,2984	
Log-věrohodnost		-30,5885	

Testy statistické významnosti jednotlivých parametrů pro alfa = 0,20:

jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (**AUDITORY**) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (**VISMEM**) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

Efekt	GP - Test všech efektů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0		
	Stupně volnosti	Wald. Stat.	p
Abs. člen	1	4,227685	0,039770
SEX	1	1,190770	0,275174
AUDITORY	1	2,490416	0,114541
VISUAL	1	0,054716	0,815052
VISMEM	1	1,959094	0,161610
AUDASSOC	1	0,301975	0,582647
AUDMEMO	1	0,315328	0,574429
VISASSOC	1	0,015007	0,902500
VISCLOS	1	3,231441	0,072237
VISEXPR	1	0,067547	0,794942
GRAMCLOS	1	0,206569	0,649470
MANUAL	1	2,628478	0,104962

Efekt	GP - Věrohodn. test typ 1 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-37,4299		
SEX	1	-37,3927	0,074500	0,784894
AUDITORY	1	-35,6439	3,497555	0,061459
VISUAL	1	-35,5665	0,154770	0,694018
VISMEM	1	-34,0424	3,048217	0,080826
AUDASSOC	1	-33,8976	0,289663	0,590437
AUDMEMO	1	-33,7563	0,282585	0,595012
VISASSOC	1	-33,6946	0,123414	0,725361
VISCLOS	1	-32,1339	3,121384	0,077271
VISEXPR	1	-32,0116	0,244634	0,620879
GRAMCLOS	1	-31,9782	0,066815	0,796033
MANUAL	1	-30,5885	2,779401	0,095484

Efekt	GP - Věrohodn. test typ 3 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
SEX	1	-31,2186	1,260237	0,261606
AUDITORY	1	-31,9934	2,809776	0,093691
VISUAL	1	-30,6158	0,054544	0,815337
VISMEM	1	-31,6370	2,097071	0,147582
AUDASSOC	1	-30,7412	0,305398	0,580518
AUDMEMO	1	-30,7459	0,314911	0,574682
VISASSOC	1	-30,5960	0,015033	0,902416
VISCLOS	1	-32,3726	3,568183	0,058897
VISEXPR	1	-30,6223	0,067716	0,794692
GRAMCLOS	1	-30,6922	0,207407	0,648807
MANUAL	1	-31,9782	2,779401	0,095484

Statistická analýza reziduí

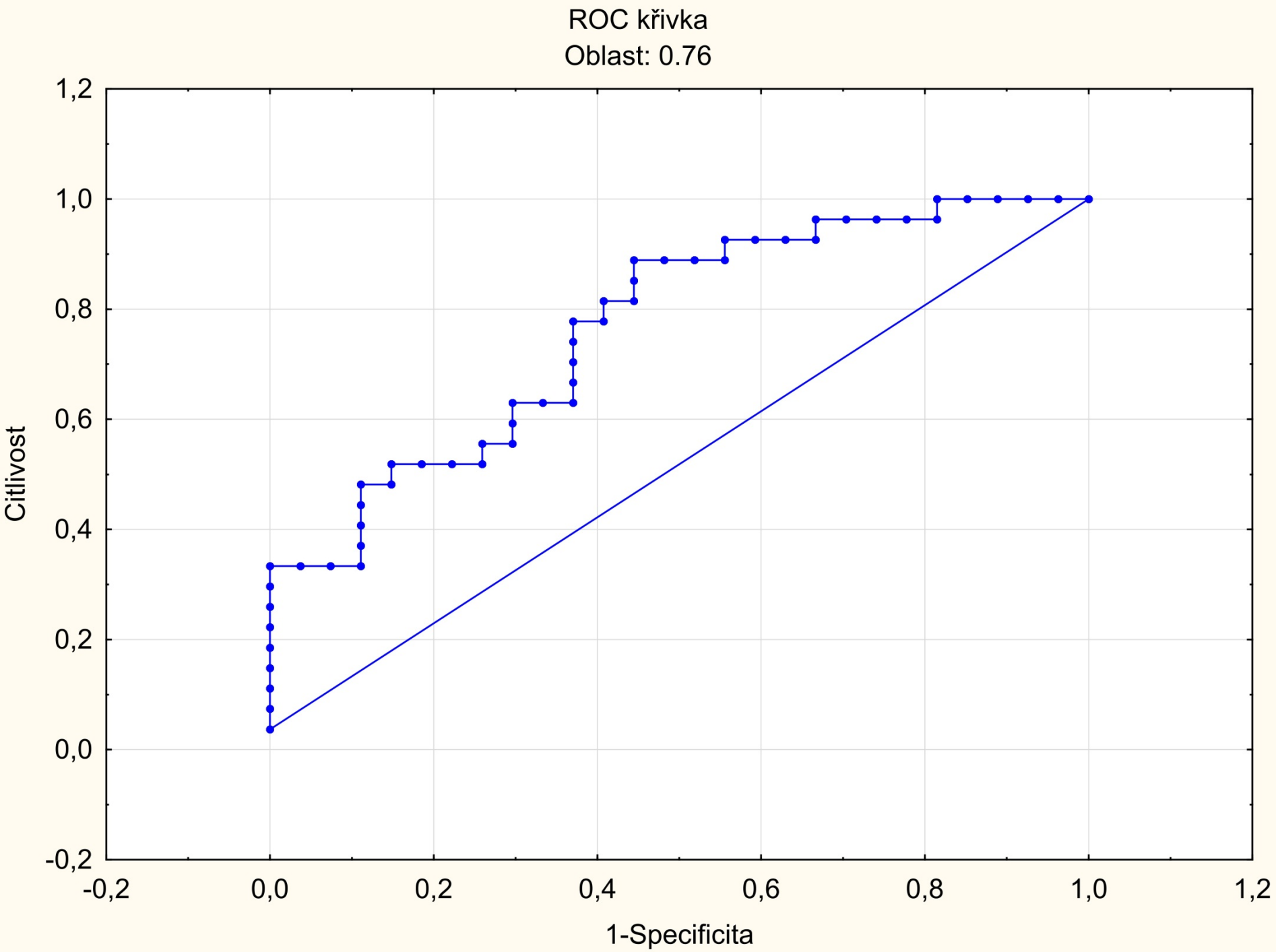
Případ	GP - Základní rezidua (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0		
	0 Čistá	0 Pearson.	0 Odchylka
1	-0,298291	-0,65199	-0,84171
2	-0,458280	-0,91977	-1,10725
3	-0,643764	-1,34430	-1,43678
4	-0,594812	-1,21161	-1,34418
5	-0,437712	-0,88230	-1,07307
6	-0,637514	-1,32617	-1,42462
7	-0,785096	-1,91134	-1,75360
8	-0,701889	-1,53443	-1,55582
9	-0,280895	-0,62499	-0,81209
10	-0,368156	-0,76333	-0,95824
11	-0,174053	-0,45905	-0,61842
12	-0,375516	-0,77545	-0,97039
13	-0,783895	-1,90457	-1,75042
14	-0,243556	-0,56743	-0,74716
15	-0,351570	-0,73633	-0,93081
16	-0,108147	-0,34823	-0,47844
17	-0,169694	-0,45208	-0,60985
18	-0,528293	-1,05828	-1,22589
19	-0,069853	-0,27404	-0,38056
20	-0,045644	-0,21869	-0,30567
21	-0,135797	-0,39640	-0,54027
22	-0,101524	-0,33615	-0,46272
23	-0,272745	-0,61240	-0,79809
24	-0,214859	-0,52312	-0,69555
25	-0,576374	-1,16643	-1,31065
26	-0,624031	-1,28833	-1,39875
27	-0,505041	-1,01013	-1,18599
28	0,509408	1,01900	1,19343
29	0,193057	0,48913	0,65499
30	0,570527	1,15258	1,30015
31	0,517184	1,03498	1,20675
32	0,348215	0,73092	0,92525
33	0,424302	0,85850	1,05088
34	0,554555	1,11577	1,27176
35	0,320011	0,68601	0,87827
36	0,054492	0,24007	0,33476
37	0,116026	0,36229	0,49664
38	0,532494	1,06724	1,23316
39	0,663911	1,40549	1,47674
40	0,131616	0,38931	0,53127
41	0,173548	0,45825	0,61744
42	0,362156	0,75351	0,94833
43	0,417875	0,84726	1,04026
44	0,192696	0,48856	0,65430
45	0,345282	0,72621	0,92038
46	0,533282	1,06893	1,23453
47	0,204917	0,50767	0,67721
48	0,432245	0,87254	1,06402
49	0,599064	1,22236	1,35200

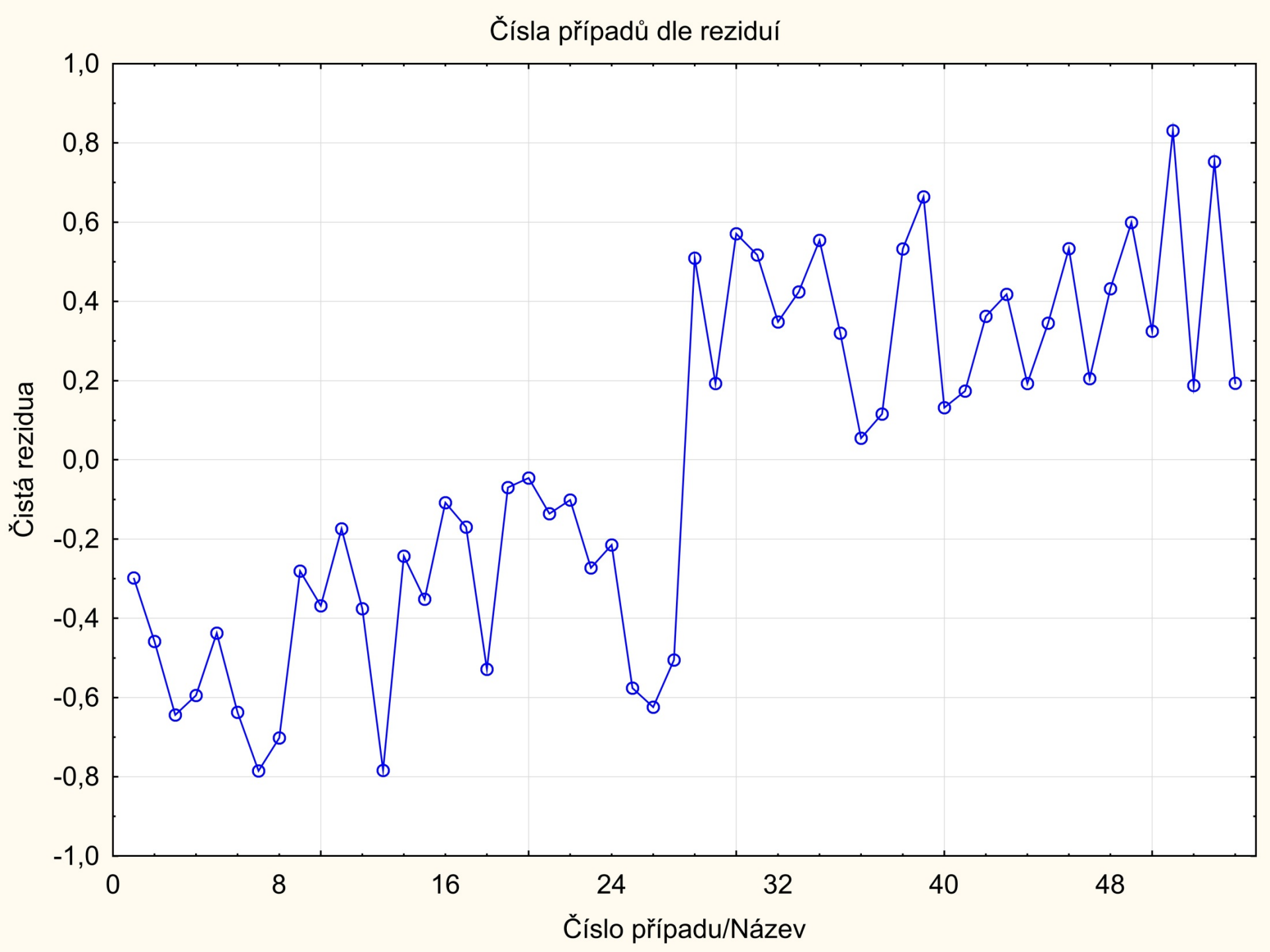
jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (**AUDITORY**) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (**VISMEM**) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

Případ	GP - Základní rezidua (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0		
	0 Čistá	0 Pearson.	0 Odchylka
50	0,325017	0,69392	0,88664
51	0,830980	2,21731	1,88560
52	0,187975	0,48113	0,64533
53	0,752731	1,74476	1,67169
54	0,193431	0,48971	0,65569

Případ	GP - Kvalita proložení - ROC; (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0					
	Předpovězená Pravděpodobnost	Správně Predik. úspěchy	Správně Predik. neúspěchy	Neúspěch predik. Jako úspěch	Úspěch predik. Jako neúspěch	Citlivost
1	0,945508	1,00000	27,00000	0,00000	26,00000	0,037037
2	0,883974	2,00000	27,00000	0,00000	25,00000	0,074074
3	0,868384	3,00000	27,00000	0,00000	24,00000	0,111111
4	0,826452	4,00000	27,00000	0,00000	23,00000	0,148148
5	0,812025	5,00000	27,00000	0,00000	22,00000	0,185185
6	0,807304	6,00000	27,00000	0,00000	21,00000	0,222222
7	0,806943	7,00000	27,00000	0,00000	20,00000	0,259259
8	0,806569	8,00000	27,00000	0,00000	19,00000	0,296296
9	0,795083	9,00000	27,00000	0,00000	18,00000	0,333333
10	0,785096	9,00000	26,00000	1,00000	18,00000	0,333333
11	0,783895	9,00000	25,00000	2,00000	18,00000	0,333333
12	0,701889	9,00000	24,00000	3,00000	18,00000	0,333333
13	0,679989	10,00000	24,00000	3,00000	17,00000	0,370370
14	0,674983	11,00000	24,00000	3,00000	16,00000	0,407407
15	0,654718	12,00000	24,00000	3,00000	15,00000	0,444444
16	0,651785	13,00000	24,00000	3,00000	14,00000	0,481481
17	0,643764	13,00000	23,00000	4,00000	14,00000	0,481481
18	0,637844	14,00000	23,00000	4,00000	13,00000	0,518519
19	0,637514	14,00000	22,00000	5,00000	13,00000	0,518519
20	0,624031	14,00000	21,00000	6,00000	13,00000	0,518519
21	0,594812	14,00000	20,00000	7,00000	13,00000	0,518519
22	0,582125	15,00000	20,00000	7,00000	12,00000	0,555556
23	0,576374	15,00000	19,00000	8,00000	12,00000	0,555556
24	0,575698	16,00000	19,00000	8,00000	11,00000	0,592593
25	0,567755	17,00000	19,00000	8,00000	10,00000	0,629630
26	0,528293	17,00000	18,00000	9,00000	10,00000	0,629630
27	0,505041	17,00000	17,00000	10,00000	10,00000	0,629630
28	0,490592	18,00000	17,00000	10,00000	9,00000	0,666667
29	0,482816	19,00000	17,00000	10,00000	8,00000	0,703704
30	0,467506	20,00000	17,00000	10,00000	7,00000	0,740741
31	0,466718	21,00000	17,00000	10,00000	6,00000	0,777778
32	0,458280	21,00000	16,00000	11,00000	6,00000	0,777778
33	0,445445	22,00000	16,00000	11,00000	5,00000	0,814815
34	0,437712	22,00000	15,00000	12,00000	5,00000	0,814815
35	0,429473	23,00000	15,00000	12,00000	4,00000	0,851852
36	0,400936	24,00000	15,00000	12,00000	3,00000	0,888889
37	0,375516	24,00000	14,00000	13,00000	3,00000	0,888889
38	0,368156	24,00000	13,00000	14,00000	3,00000	0,888889
39	0,351570	24,00000	12,00000	15,00000	3,00000	0,888889
40	0,336089	25,00000	12,00000	15,00000	2,00000	0,925926
41	0,298291	25,00000	11,00000	16,00000	2,00000	0,925926
42	0,280895	25,00000	10,00000	17,00000	2,00000	0,925926
43	0,272745	25,00000	9,00000	18,00000	2,00000	0,925926
44	0,247269	26,00000	9,00000	18,00000	1,00000	0,962963
45	0,243556	26,00000	8,00000	19,00000	1,00000	0,962963
46	0,214859	26,00000	7,00000	20,00000	1,00000	0,962963
47	0,174053	26,00000	6,00000	21,00000	1,00000	0,962963
48	0,169694	26,00000	5,00000	22,00000	1,00000	0,962963

Případ	GP - Kvalita proložení - ROC; (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0
	1-Specificita
1	0,000000
2	0,000000
3	0,000000
4	0,000000
5	0,000000
6	0,000000
7	0,000000
8	0,000000
9	0,000000
10	0,037037
11	0,074074
12	0,111111
13	0,111111
14	0,111111
15	0,111111
16	0,111111
17	0,148148
18	0,148148
19	0,185185
20	0,222222
21	0,259259
22	0,259259
23	0,296296
24	0,296296
25	0,296296
26	0,333333
27	0,370370
28	0,370370
29	0,370370
30	0,370370
31	0,370370
32	0,407407
33	0,407407
34	0,444444
35	0,444444
36	0,444444
37	0,481481
38	0,518519
39	0,555556
40	0,555556
41	0,592593
42	0,629630
43	0,666667
44	0,666667
45	0,703704
46	0,740741
47	0,777778
48	0,814815


ROC křivka
Oblast: 0.76
Citlivost
1-Specificita
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,2
0,0
0,2
0,4
0,6
0,8
1,0
1,2


Čísla případů dle reziduí
Čistá rezidua
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
0
8
16
24
32
40
48
Číslo případu/Název

Efekt	GP - Odhady parametrů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 80, %	Horní LS 80, %	p
Abs.člen		1	-4,52882	2,362968	3,673280	-7,55708	-1,50055	0,055291
AUDITORY		2	0,06118	0,050402	1,473230	-0,00342	0,12577	0,224837
VISMEM		3	0,04699	0,045130	1,084102	-0,01085	0,10483	0,297782
VISCLOS		4	0,13882	0,079041	3,084639	0,03753	0,24012	0,079035
MANUAL		5	-0,10200	0,055915	3,327670	-0,17366	-0,03034	0,068123
Měřítko			1,00000	0,000000		1,00000	1,00000	

jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (AUDITORY) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (VISMEM) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

Efekt	GP - Intervaly spolehlivost odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Úroveň Efekt	Sloupec	Dolní LS 80, %	Horní LS 80, %
Abs.člen		1	-7,55708	-1,50055
AUDITORY		2	-0,00342	0,12577
VISMEM		3	-0,01085	0,10483
VISCLOS		4	0,03753	0,24012
MANUAL		5	-0,17366	-0,03034

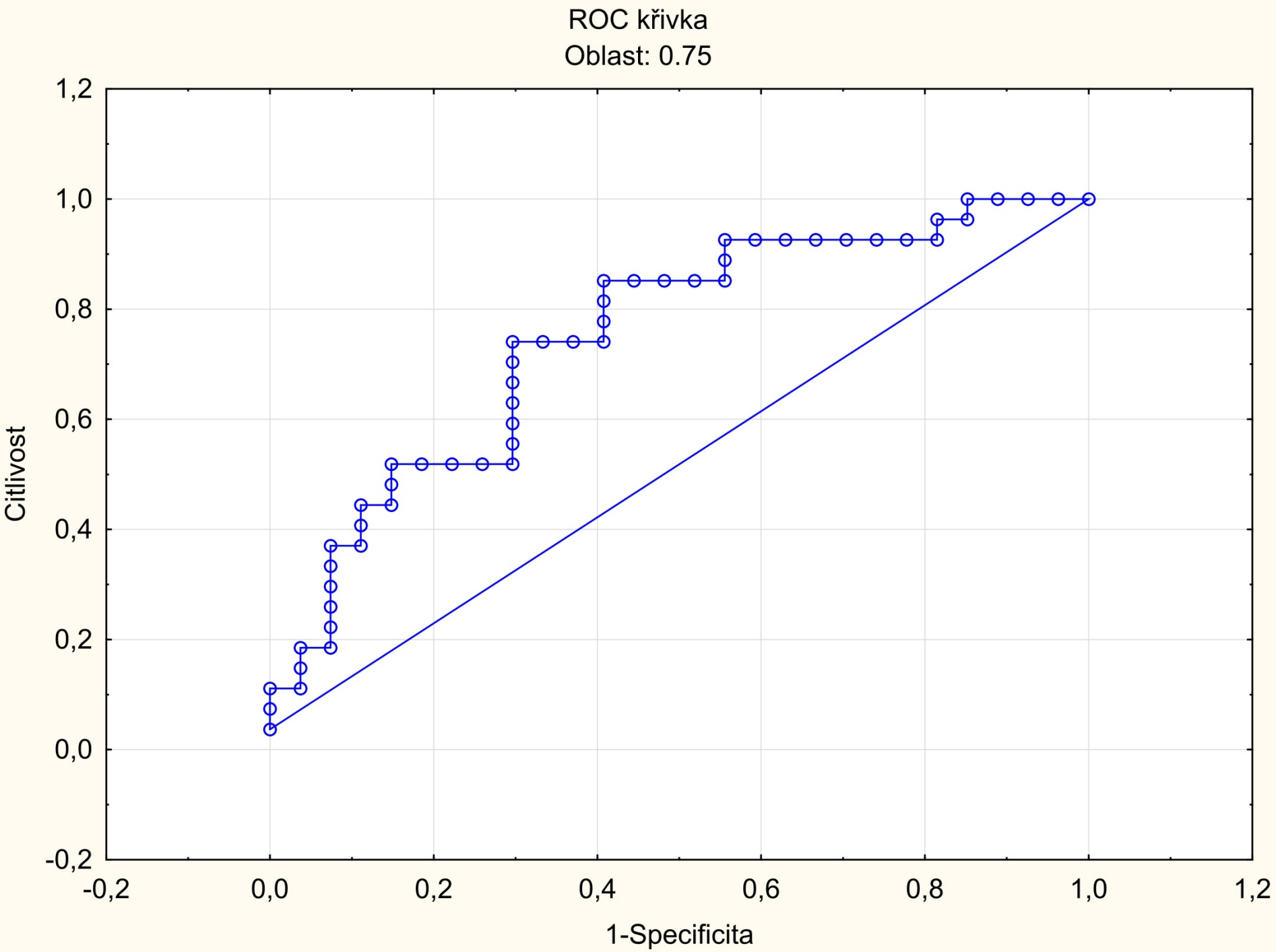
Efekt	GP - Historie iterací (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0					
	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3
Abs.člen		1	0,0000	-3,9746	-4,4990	-4,5287
AUDITORY		2	0,0000	0,0484	0,0603	0,0612
VISMEM		3	0,0000	0,0425	0,0468	0,0470
VISCLOS		4	0,0000	0,1194	0,1377	0,1388
MANUAL		5	0,0000	-0,0836	-0,1007	-0,1020
Věrohodnost			-37,4299	-31,9418	-31,8315	-31,8311

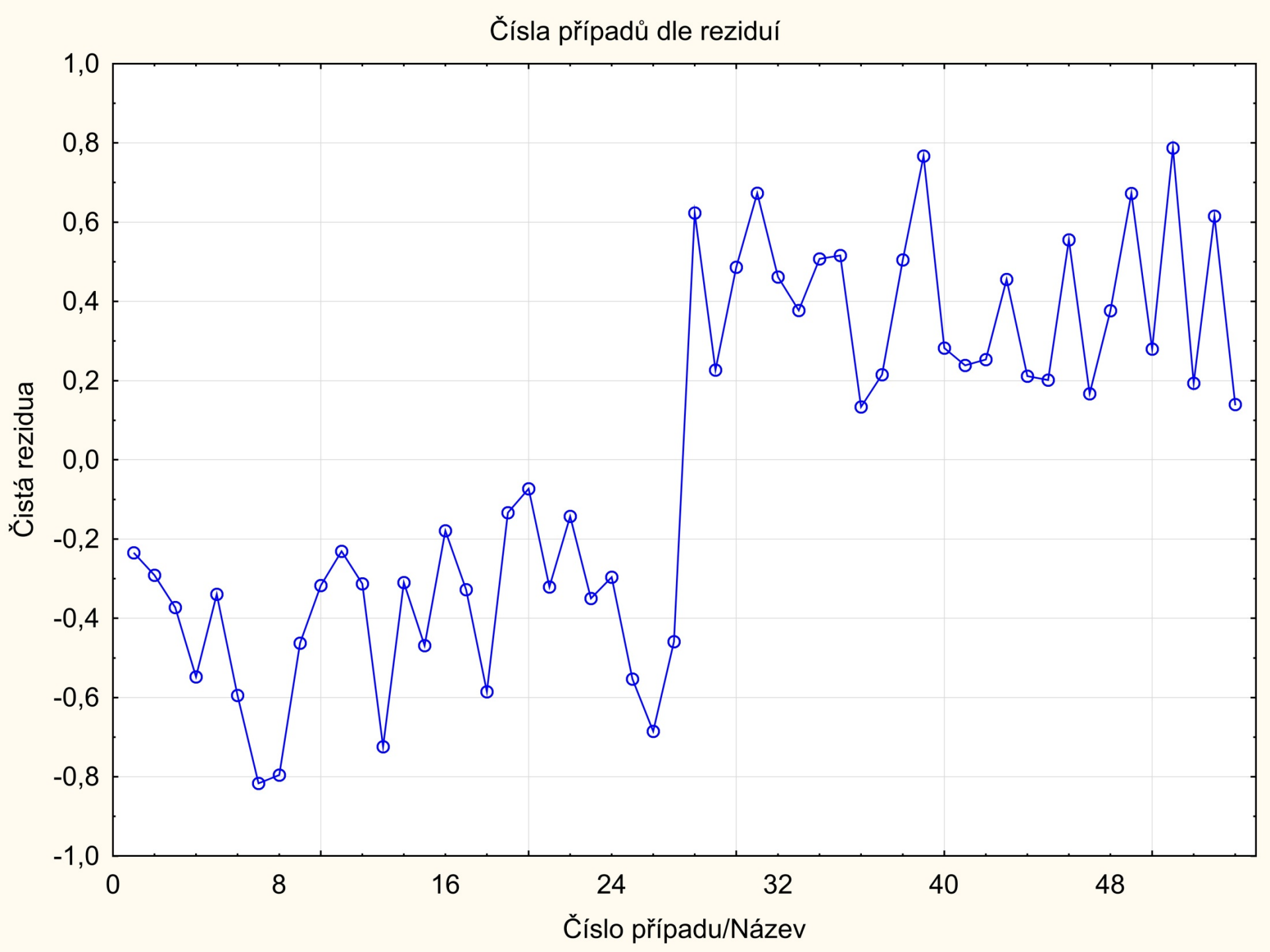
Efekt	GP - Věrohodn. test typ 1 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-37,4299		
AUDITORY	1	-35,9680	2,923851	0,087279
VISMEM	1	-34,8071	2,321881	0,127566
VISCLOS	1	-33,6473	2,319533	0,127758
MANUAL	1	-31,8311	3,632514	0,056661

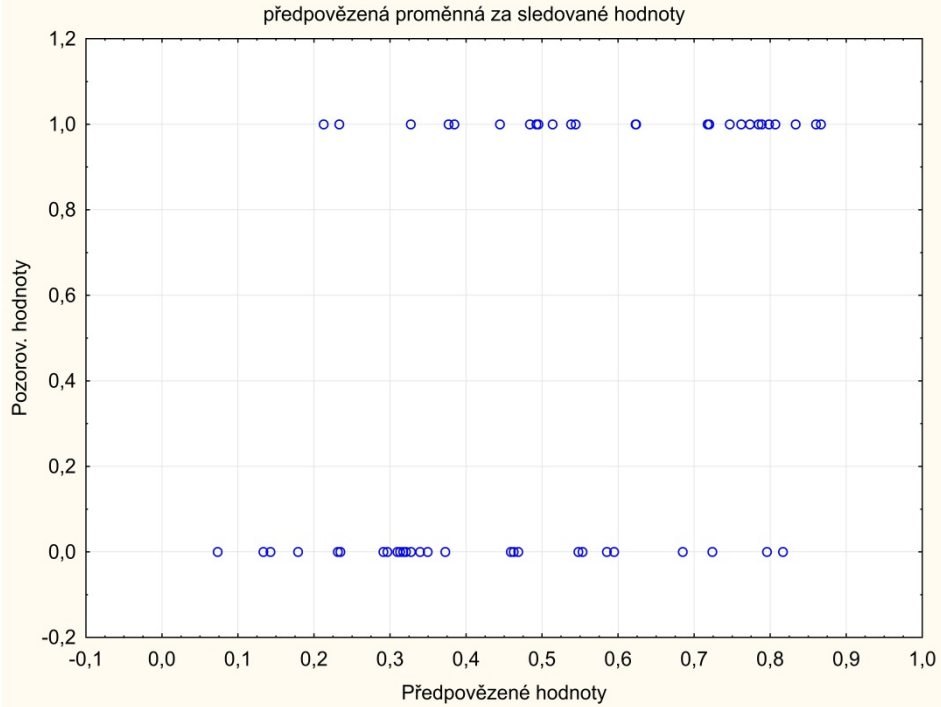
Efekt	GP - Věrohodn. test typ 3 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
AUDITORY	1	-32,6044	1,546600	0,213638
VISMEM	1	-32,3886	1,115168	0,290961
VISCLOS	1	-33,5297	3,397255	0,065305
MANUAL	1	-33,6473	3,632514	0,056661

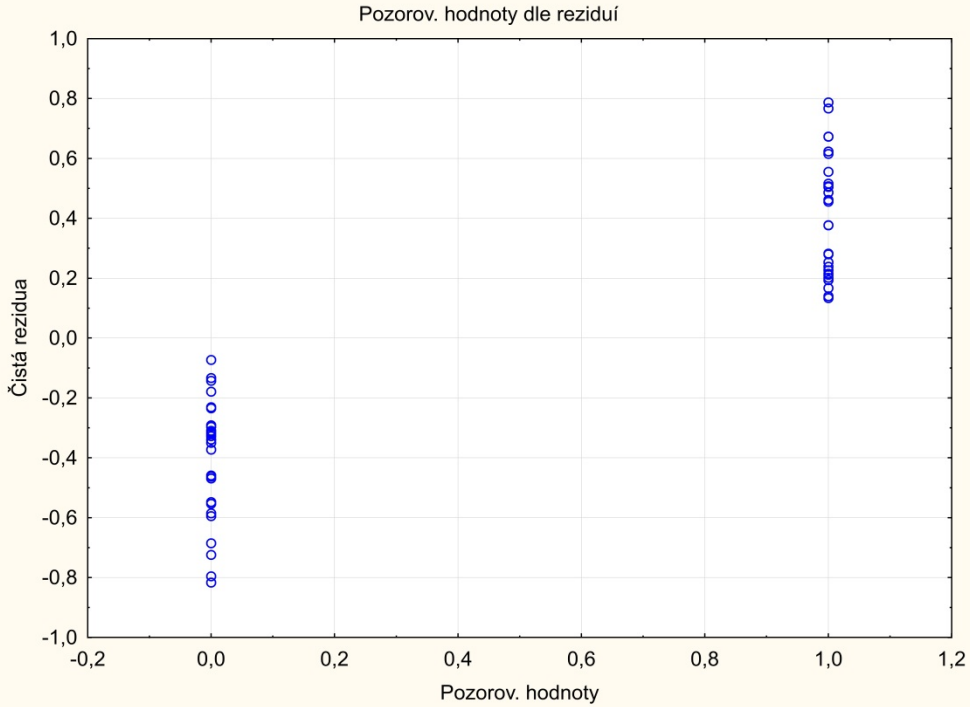
Efekt	GP - Test všech efektů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0		
	Stupně volnosti	Wald. Stat.	p
Abs. člen	1	3,673280	0,055291
AUDITORY	1	1,473230	0,224837
VISMEM	1	1,084102	0,297782
VISCLOS	1	3,084639	0,079035
MANUAL	1	3,327670	0,068123

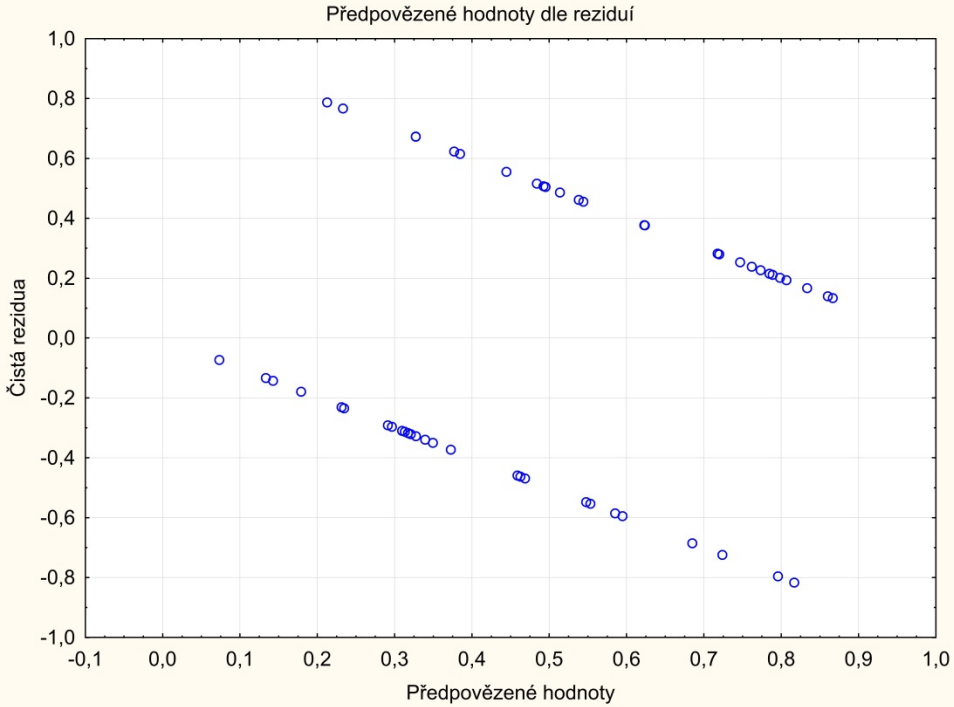
jj (SE*X*) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
*x*1 (AUDITORY) značí sluchový přijímací záznam,
*x*2 (VISUAL) značí vizuální přijímací záznam,
*x*3 (VISMEM) značí vizuální paměť,
*x*4 (AUDASSOC) značí sluchové seskupení,
*x*5 (AUDMEMO) značí sluchová paměť,
*x*6 (VISASSOC) značí vizuální seskupení,
*x*7 (**VISCLOS**) značí vizuální uzavření,
*x*8 (VISE*X*PR) značí verbální projev,
*x*9 (GRAMCLOS) značí gramatické uzavření,
*x*10 (**MANUAL**) značí manuální zručnost.

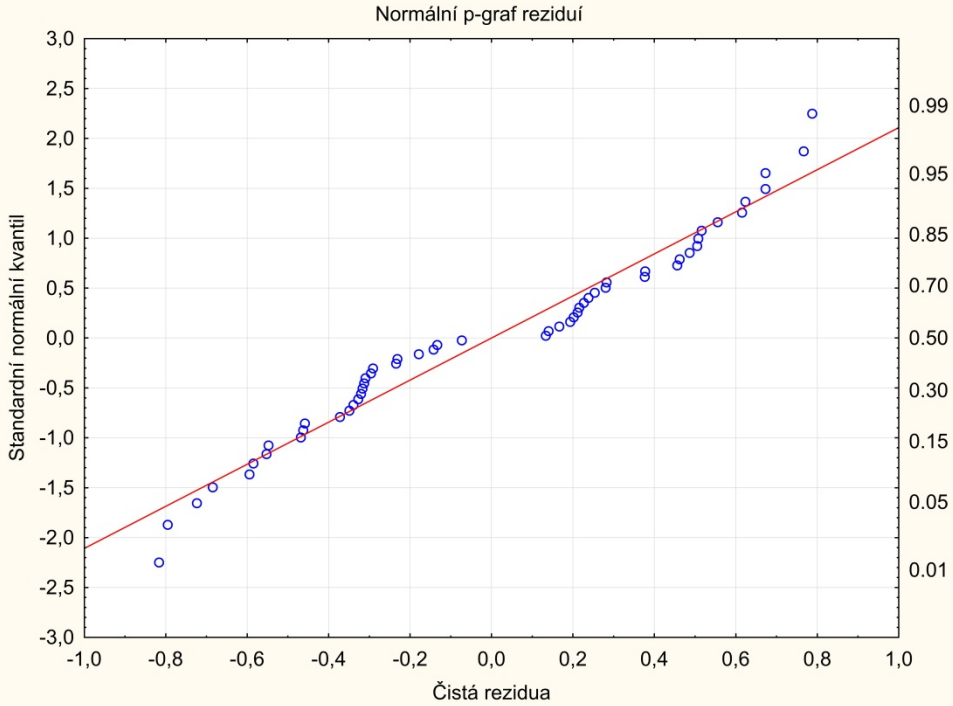

ROC křivka
Oblast: 0.75
Citlivost
1-Specificita
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,2
0,0
0,2
0,4
0,6
0,8
1,0
1,2


Čísla případů dle reziduí
Čistá rezidua
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
0
8
16
24
32
40
48
Číslo případu/Název


předpovězená proměnná za sledované hodnoty
Pozorov. hodnoty
Předpovězené hodnoty


Pozorov. hodnoty dle reziduí
Čistá rezidua
Pozorov. hodnoty


Předpovězené hodnoty dle reziduí
Čistá rezidua
Předpovězené hodnoty


Normální p-graf reziduí
Standardní normální kvantil
Čistá rezidua

Úloha: Predikce počtu samečků kraba usilujících o samičku logistickou regresí

Samice kraba obecného může být oplozena pouze v určitou, omezenou dobu, která nastává krátce po svlékání, dokud je její krunýř ještě měkký. Jakmile se samice kraba obecného zbaví starého krunýře, samec ihned vystříkne semeno do jejího těla, kde zůstává uchováno, dokud se v těle samice nevytvoří vajíčka. Teprve pak dochází k oplodnění. Oplozená vajíčka pak visí přichycena zvláštní lepkavou hmotou pod zadečkem samice. Po určitém čase se z vajíček vylíhnou průhledné, malým garnátům podobné larvy, které mají mezi očima dlouhý výrůstek. Larvy žijí po několik týdnů jako součást planktonu. Z těch, které se navzájem nepožerou, vyrostou larvy, které se již částečně podobají dospělým krabům. Ty klesnou na mořské dno, kde se krmí a rostou. Průměrná doba inkubace, než se z vajíček vylíhnou aktivně larvy trvá přibližně 12 až 18 týdnů.

Samičce kraba se dvoří řada samců, přibližujících se k samičce a jejímu hnízdu a které zde označíme **satelity**. Počet samičkou lákaných satelitů je řízen barvou samičky (Color), stavem hřbetu (Spine), šířkou těla v cm (Width) a její hmotností v kg (Weight) samičky. **Logitem** čili závisle proměnnou y (Satellite) je počet samečků kraba, dvořících se jedné samičce, který může být nulový ($y = 0$) nebo nenulový ($y = 1$). Kategorické (nespojité) nezávisle proměnné čili **faktory** jsou dva, Color o 4 hodnotách barvy a Spine o třech hodnotách kvality hřbetu. Spojité nezávisle proměnné čili **prediktory** jsou také dva, Width (šířka těla samičky) a Weight (hmotnost samičky).

Cíl úlohy: Cestou logitu (satellit) $y = 0$ či 1 budeme hledat minimální počet nezávisle proměnných (faktorů a prediktorů), které budou spolehlivě predikovat, zda jsou samci-sateliti vedle jednoho samce-partnera ve svém dvoření jedné samičce úspěšní.

STATISTICA Cz - [Data: Crabs (7s krát 173ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu

Arial 10 Proměnné Případy

Number of crab satellites by female's color, spine condition, width, and weight

	1 Y	2 COLOR	3 SPINE	4 WIDTH	5 SATELLTS	6 WEIGHT	7 CATWIDTH
1	1	medium	bothworn	28,3	8	3,05	28,75
2	0	darkmed	bothworn	22,5	0	1,55	22,75
3	1	lightmed	bothgood	26,0	9	2,30	25,75
4	0	darkmed	bothworn	24,8	0	2,10	24,75
5	1	darkmed	bothworn	26,0	4	2,60	25,75
6	0	medium	bothworn	23,8	0	2,10	23,75
7	0	lightmed	bothgood	26,5	0	2,35	26,75
8	0	darkmed	oneworn	24,7	0	1,90	24,75
9	0	medium	bothgood	23,7	0	1,95	
10	0	darkmed	bothworn	25,6	0	2,15	
11	0	darkmed	bothworn	24,3	0	2,15	
12	0	medium	bothworn	25,8	0	2,65	
13	1	medium	bothworn	28,2	11	3,05	
14	0	dark	oneworn	21,0	0	1,85	
15	1	medium	bothgood	26,0	14	2,30	
16	1	lightmed	bothgood	27,1	8	2,95	
17	1	medium	bothworn	25,2	1	2,00	
18	1	medium	bothworn	29,0	1	3,00	
19	0	dark	bothworn	24,7	0	2,20	
20	1	medium	bothworn	27,4	5	2,70	
21	1	medium	oneworn	23,2	4	1,95	
22	1	lightmed	oneworn	25,0	3	2,30	
23	1	medium	bothgood	22,5	1	1,60	
24	1	darkmed	bothworn	26,7	2	2,60	
25	1	dark	bothworn	25,8	3	2,00	
26	0	dark	bothworn	26,2	0	1,30	
27	1	medium	bothworn	28,7	3	3,15	
28	1	medium	bothgood	26,8	5	2,70	
29	0	dark	bothworn	27,5	0	2,60	
30	0	medium	bothworn	24,9	0	2,10	
31	1	lightmed	bothgood	29,3	4	3,20	
32	0	lightmed	bothworn	25,8	0	2,60	

Otevřít

Oblast hledání: Datasets

Poslední dokumenty, Plocha, Dokumenty, Tento počítač, Místa v síti

Aircraft	Boston2	Composit
Allergy	BostonHousing	Computers
AnalyzingSICCodes	Brake pad	Constrr
Ancova	Bulbs	Coupons
Animal Weights	Car Accidents	Cover
Ball bearings	Cars	Cover2
Barotro2	Cat Clinic	Crabs
Barotrop	Cat Clinic Measurements (2000)	Cracker promc
Barotropsnn	Cat Clinic Patient Info	Credit applicat
Baseball	Ccd	CreditSNN
BatchesForDeployment	Center	Crossover
Beverag2	Center2	Cusum
Beverage	Circuits	Cycling
Bid_prep	College	Damage
Bleach	Comfort	Demand

Název souboru: Crabs

Soubory typu: STATISTICA soubory (*.stw;*.sta;*.stg;*.str;*.svl

Otevřít

Storno

33. **LR:** Zdrojová matice *Crabs.sta* se načte z databáze *EXAMPLES* programu STATISTICA a obsahuje 173 řádků sledovaných případů satelitních krabů a 7 sloupců faktorů a prediktorů.

Cílem úlohy: odhalit počet faktorů, které predikují logit (ano či ne), zda přebyteční (satelitní) samci kraba jsou vedle jednoho partnerského samce u každé samice úspěšní ve svém dvoření se samičce.

STATISTICA Cz - [Data: Crabs (7s krát 173ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápověda

Obnovit... Ctrl+R
Analýza skupin
Základní statistiky/tabulky
Vícerozměrná regrese
ANOVA
Neparametrická statistika
Prokládání rozdělení
Pokročilé lineární/nelineární modely
Vícerozměrné průzkumné techniky
Průmyslová statistika & Six Sigma
Analýza síly testu
Neuronové sítě
Vytěžování dat
QC Data mining & Analýza hlavních příčin
Text & Document Mining, Web Crawling
Statistiky bloku dat
STATISTICA Visual Basic
Pravděpodobnostní kalkulátor

Obecné lineární modely
Zobecněné lineární/nelineární modely
Obecné regresní modely
Obecné modely s parciálními nejmenšími čtverci
NIPALS algoritmus (PCA/PLS)
Komponenty rozptylu
Analýza přežívání
Nelineární odhady
Jednoduchá nelineární regrese
Log-lineární analýza kontingenčních tabulek
Časové řady/Predikce
Modelování pomocí strukturálních rovnic

Number of crab satellites by female's

	1 Y	2 COLOR	3 SPINE	4 WIDTH	SAT		
1	1	medium	bothworn	28,3			
2	0	darkmed	bothworn	22,5			
3	1	lightmed	bothgood	26,0			
4	0	darkmed	bothworn	24,8			
5	1	darkmed	bothworn	26,0			
6	0	medium	bothworn	23,8			
7	0	lightmed	bothgood	26,5			
8	0	darkmed	oneworn	24,7			
9	0	medium	bothgood	23,7			
10	0	darkmed	bothworn	25,6			
11	0	darkmed	bothworn	24,3			
12	0	medium	bothworn	25,8			
13	1	medium	bothworn	28,2			
14	0	dark	oneworn	21,0			
15	1	medium	bothgood	26,0			
16	1	lightmed	bothgood	27,1			
17	1	medium	bothworn	25,2			
18	1	medium	bothworn	29,0	1	3,00	28,75
19	0	dark	bothworn	24,7	0	2,20	24,75
20	1	medium	bothworn	27,4	5	2,70	27,75
21	1	medium	oneworn	23,2	4	1,95	22,75
22	1	lightmed	oneworn	25,0	3	2,30	24,75
23	1	medium	bothgood	22,5	1	1,60	22,75
24	1	darkmed	bothworn	26,7	2	2,60	26,75
33	0	medium	oneworn	25,7	0	2,00	25,75
34	1	medium	bothgood	25,7	8	2,00	25,75
35	1	medium	bothgood	26,7	5	2,70	26,75
36	0	dark	bothworn	23,7	0	1,85	23,75
37	0	medium	bothworn	26,8	0	2,65	26,75
38	1	medium	bothworn	27,5	6	3,15	27,75
39	0	dark	bothworn	23,4	0	1,90	23,75
40	1	medium	bothworn	27,9	6	2,80	27,75
41	1	darkmed	bothworn	27,5	3	3,10	27,75
42	1	lightmed	bothgood	26,1	5	2,80	25,75
43	1	lightmed	bothgood	27,7	6	2,50	27,75
44	1	medium	bothgood	30,0	5	3,30	29,75
45	1	darkmed	bothgood	28,5	9	3,25	28,75
46	1	darkmed	bothworn	28,9	4	2,80	28,75

34. **LR:** V menu *Pokročilé lineární/nelineární modely* zvolíme *GLZ Zobecněné lineární/nelineární modely.*

Spustí modul Zobecněné lineární/nelineární modely

Start | STATISTICA Cz - [Dat... | STATISTICA Electronic M... | Microsoft PowerPoint - [... | 18:55

STATISTICA Cz - [Data: Crabs (7s krát 173ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu

Arial 10 Proměnné Případy

Number of crab satellites by female's color, spine condition, width, and weight

	1 Y	2 COLOR	3 SPINE	4 WIDTH	5 SATELLTS	6 WEIGHT	7 CATWIDTH
1	1	medium	bothworn	28,3	8	3,05	28,75
2	0	darkmed	bothworn	22,5	0	1,55	22,75
3	1	lightmed	bothgood	26,0	9	2,30	25,75
4	0	darkmed	bothworn	24,8	0	2,10	24,75
5	1	darkmed	bothworn	26,0	4	2,60	25,75
6	0	medium	bothworn	23,8	0	2,10	23,75
7	0	lightmed	bothgood	26,5	0	2,35	26,75
8	0	darkmed	oneworn	24,7	0	1,90	24,75
9	0	medium	bothgood	23,7	0	1,95	
10	0	darkmed	bothworn	25,6	0	2,15	
11	0	darkmed	bothworn	24,3	0	2,15	
12	0	medium	bothworn	25,8	0	2,65	
13	1	medium	bothworn	28,2	11	3,05	
14	0	dark	oneworn	21,0	0	1,85	
15	1	medium	bothgood	26,0	14	2,30	
16	1	lightmed	bothgood	27,1	8	2,95	
17	1	medium	bothworn	25,2	1	2,00	
18	1	medium	bothworn	29,0	1	3,00	
19	0	dark	bothworn	24,7	0	2,20	
20	1	medium	bothworn	27,4	5	2,70	
21	1	medium	oneworn	23,2	4	1,95	
22	1	lightmed	oneworn	25,0	3	2,30	
23	1	medium	bothgood	22,5	1	1,60	
24	1	darkmed	bothworn	26,7	2	2,60	
25	1	dark	bothworn	25,8	3	2,00	
26	0	dark	bothworn	26,2	0	1,30	
27	1	medium	bothworn	28,7	3	3,15	
28	1	medium	bothgood	26,8	5	2,70	
29	0	dark	bothworn	27,5	0	2,60	
30	0	medium	bothworn	24,9	0	2,10	
31	1	lightmed	bothgood	29,3	4	3,20	
32	0	lightmed	bothworn	25,8	0	2,60	

Zobecněné lineární/nelineární modely: Crabs

Základní nastavení | Detaily

Typ analýzy:
- Jednofaktorová ANOVA
- ANOVA hlavních efektů
- Vícefaktorová ANOVA
- Hierarchická ANOVA
- Jednorozměrná regrese
- Vícerozměrná regrese
- Faktorová regrese
- Polynomická regrese
- Regrese odezvové plochy
- Regrese odezv. plochy směsi
- Analýza kovariance
- Separovaný smíšený model
- Úplný smíšený model
- Vlastní obecná schémata

Metoda specifikace:
- Rychlé nastavení
- Průvodce analýzou
- Editor syntaxí analýz

Rozdělení:
- Normální
- Poissonovo
- Gama
- Binomické
- Multinomické
- Ordinální multinomické
- Inverzní normální

Linkující funkce:
- Logit
- Probit
- Ln-Ln

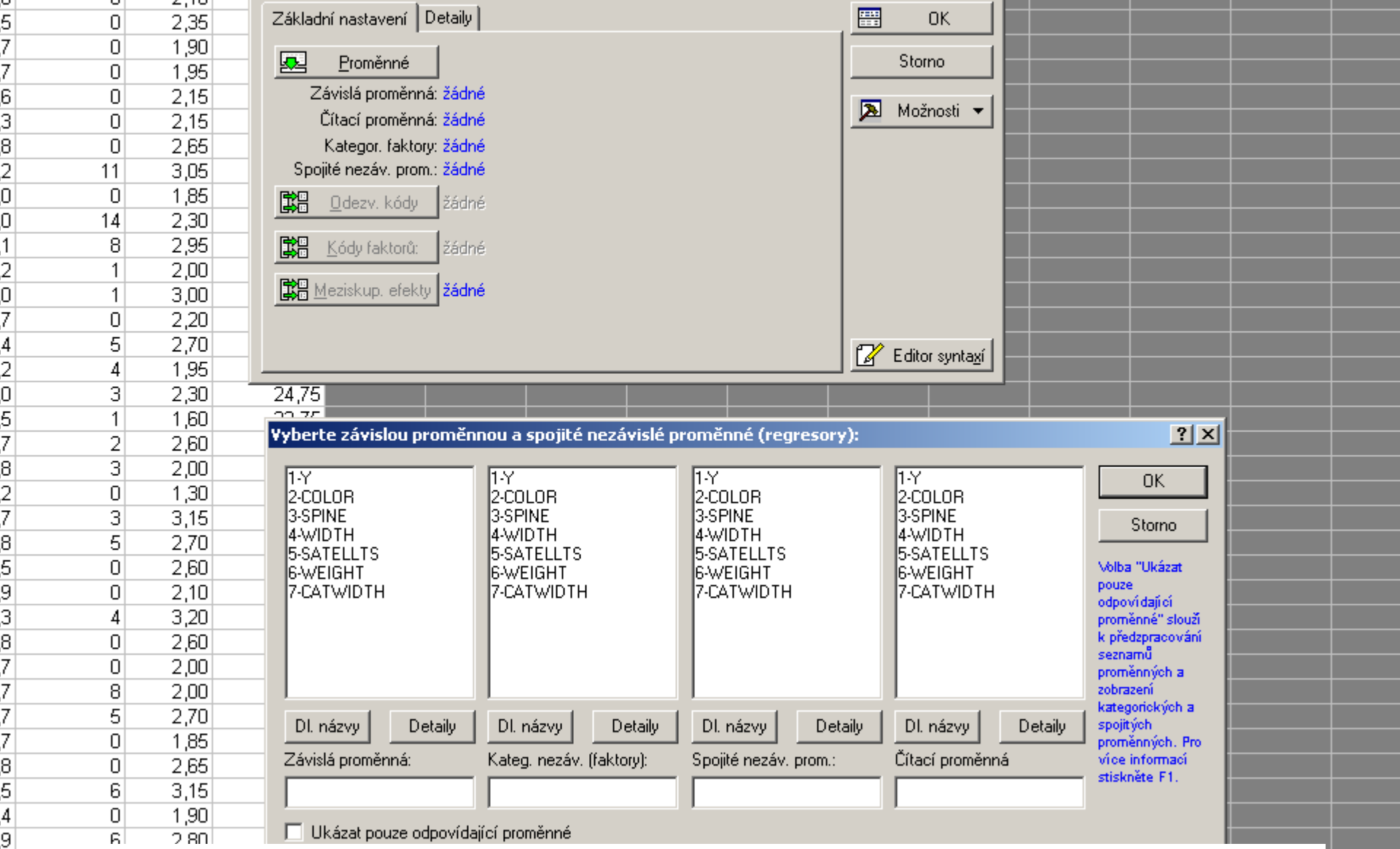
OK | Storno | Možnosti | Otevři Data | SELECT CASES | Vážené momenty | SV = W-1 N-1 | Mocnina | Param: 1.

35. **LR:** Otevřeme záložku *Detaily* a v *Typ analýzy* zvolíme *Vlastní obecná schémata*. V bloku *Metoda specifikace* zvolíme *Rychlé nastavení*. V bloku *Rozdělení* zvolíme *Binomické*. V bloku *Linkující funkce* zvolíme *Logit*. Potvrdíme *OK*.

Zobecněné lineární/...

Připravena | Ř1;S1 | 1 | Filtr - | Váhy: VYPN | ABC | 123 | ZÁZN

Start | STATISTICA Cz - [Dat... | STATISTICA Electronic M... | Microsoft PowerPoint - [... | 18:58



36. **LR:** V záložce *Základní nastavení* zvolíme *Proměnné* a otevře se okno *Vyberte závisle proměnnou a spojité nezávisle proměnné (regresory)*, které nyní uvážlivě naplníme.

Number of crab satellites by female's color, spine condition, width, and weight

	1 Y	2 COLOR	3 SPINE	4 WIDTH	5 SATELLTS	6 WEIGHT	7 CATWIDTH
1	1	medium	bothworn	28,3	8	3,05	28,75
2	0	darkmed	bothworn	22,5	0	1,55	22,75
3	1	lightmed	bothgood	26,0	9	2,30	25,75
4	0	darkmed	bothworn	24,8	0	2,10	24,75
5	1	darkmed	bothworn	26,0	4	2,60	
6	0	medium	bothworn	23,8	0	2,10	
7	0	lightmed	bothgood	26,5	0	2,35	
8	0	darkmed	oneworn	24,7	0	1,90	
9	0	medium	bothgood	23,7	0	1,95	
10	0	darkmed	bothworn	25,6	0	2,15	
11	0	darkmed	bothworn	24,3	0	2,15	
12	0	medium	bothworn	25,8	0	2,65	
13	1	medium	bothworn	28,2	11	3,05	
14	0	dark	oneworn	21,0	0	1,85	
15	1	medium	bothgood	26,0	14	2,30	
16	1	lightmed	bothgood	27,1	8	2,95	
17	1	medium	bothworn	25,2	1	2,00	
18	1	medium	bothworn	29,0	1	3,00	
19	0	dark	bothworn	24,7	0	2,20	
20	1	medium	bothworn	27,4	5	2,70	
21	1	medium	oneworn	23,2	4	1,95	
22	1	lightmed	oneworn	25,0	3	2,30	24,75
23	1	medium	bothgood	22,5	1	1,60	
24	1	darkmed	bothworn	26,7	2	2,60	
25	1	dark	bothworn	25,8	3	2,00	
26	0	dark	bothworn	26,2	0	1,30	
27	1	medium	bothworn	28,7	3	3,15	
28	1	medium	bothgood	26,8	5	2,70	
29	0	dark	bothworn	27,5	0	2,60	
30	0	medium	bothworn	24,9	0	2,10	
31	1	lightmed	bothgood	29,3	4	3,20	
32	0	lightmed	bothworn	25,8	0	2,60	
33	0	medium	oneworn	25,7	0	2,00	
34	1	medium	bothgood	25,7	8	2,00	
35	1	medium	bothgood	26,7	5	2,70	
36	0	dark	bothworn	23,7	0	1,85	
37	0	medium	bothworn	26,8	0	2,65	
38	1	medium	bothworn	27,5	6	3,15	
39	0	dark	bothworn	23,4	0	1,90	
40	1	medium	bothworn	27,9	6	2,80	

GLZ Vlastní obecné schéma: Crabs

Základní nastavení | Detaily

Proměnné
Závislá proměnná: Y
Čítací proměnná: žádné
Kategor. faktory: COLOR-SPINE
Spojité nezáv. prom.: WIDTH WEIGHT
Odezv. kódy: žádné
Kódy faktorů: žádné
Meziskup. efekty: WIDTH + WEIGHT + COLOR | SPINE

OK
Storno
Možnosti
Editor syntaxí

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

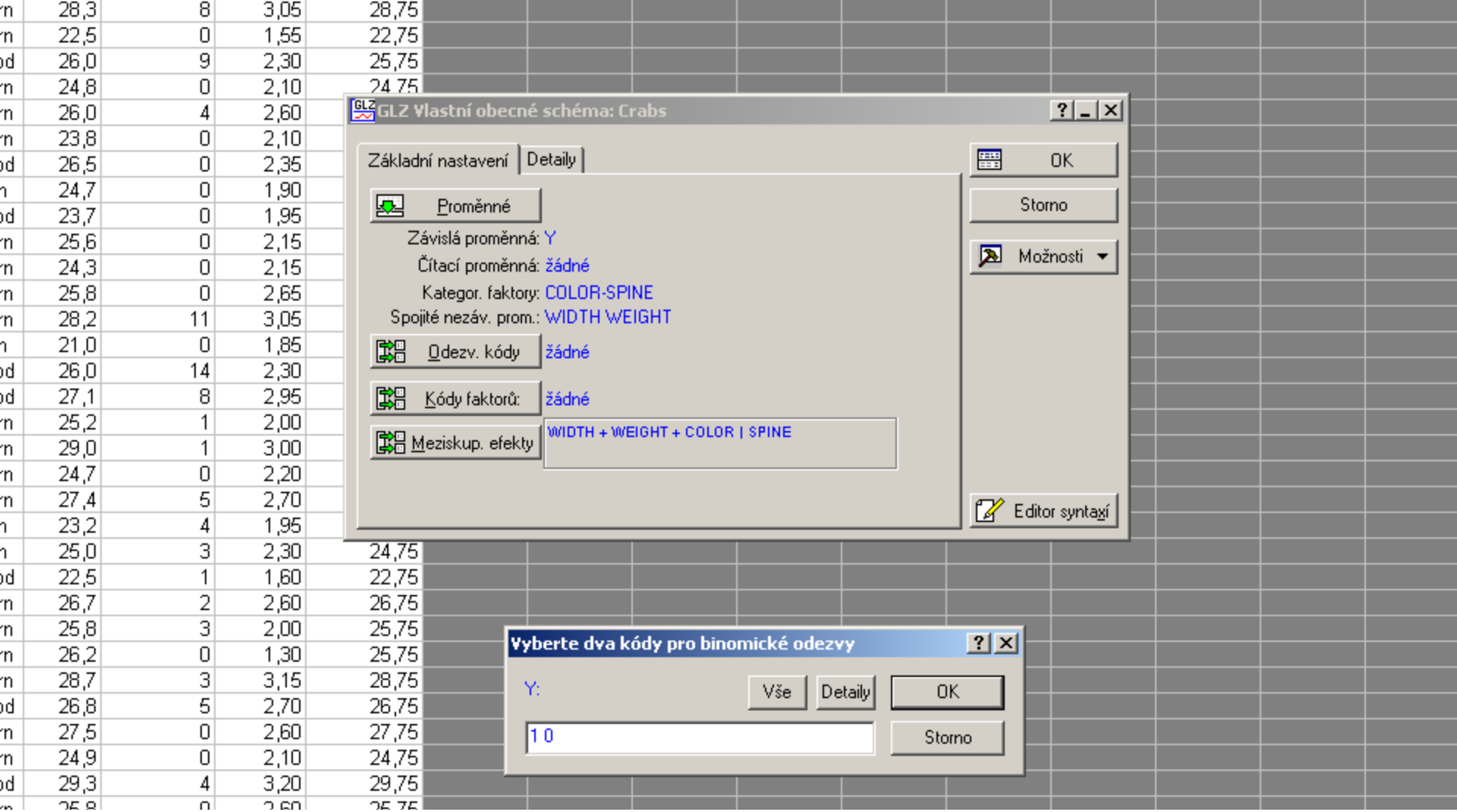
1-Y	1-Y	1-Y	1-Y
2-COLOR	2-COLOR	2-COLOR	2-COLOR
3-SPINE	3-SPINE	3-SPINE	3-SPINE
4-WIDTH	4-WIDTH	4-WIDTH	4-WIDTH
5-SATELLTS	5-SATELLTS	5-SATELLTS	5-SATELLTS
6-WEIGHT	6-WEIGHT	6-WEIGHT	6-WEIGHT
7-CATWIDTH	7-CATWIDTH	7-CATWIDTH	7-CATWIDTH
Dl. názvy / Detaily	Dl. názvy / Detaily	Dl. názvy / Detaily	Dl. názvy / Detaily
Závislá proměnná: 1	Kateg. nezáv. (faktory): 2-3	Spojité nezáv. prom.: 4 6	Čítací proměnná

OK
Storno

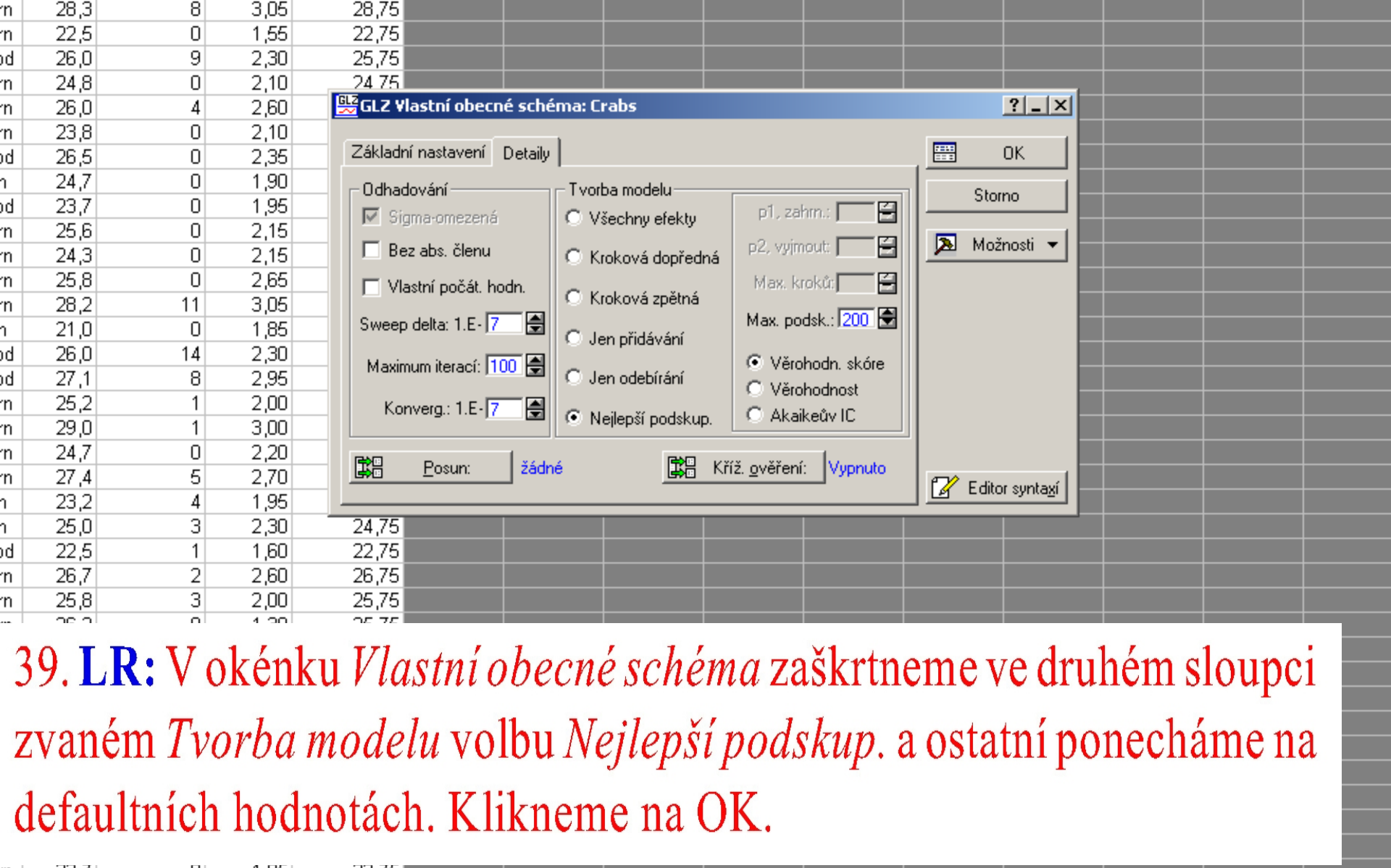
Volba "Ukázat pouze odpovídající proměnné" slouží k předzpracování seznamů proměnných a zobrazení kategorických a spojitých proměnných. Pro více informací stiskněte F1.

Ukázat pouze odpovídající proměnné

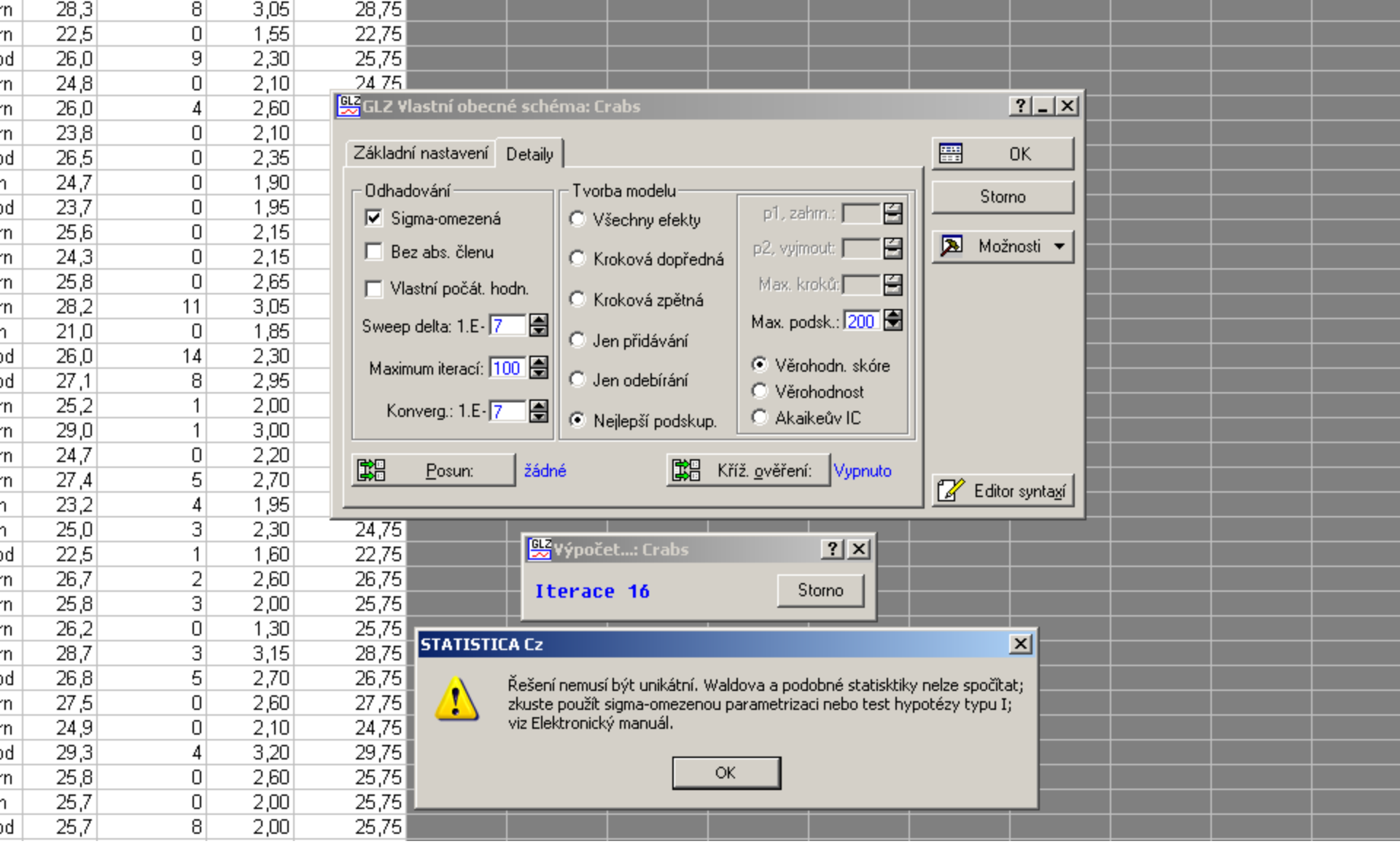
37. **LR:** Dle obrázku zvolíme za *Závisle proměnnou* znak 1-*y*, za *Faktory* zvolíme 2-Color a 3-Spine. Za *Spojité nezávisle proměnné* (prediktory) zvolíme 4-Width a 6-Weights. Klikneme na OK. Využijeme zde defaultní model, a proto budeme prokládat faktoriální návrh pro dva kategorické faktory Color a Spine a model hlavních efektů pro spojité prediktory Width a Weight.



38. **LR:** Otevřeme okno *Odezv. Kódy* kliknutím a v okénku nadepsaném *Vyberte dva kódy pro binomické odezvy* zadáme dvě možné hodnoty pro *Logit*, tj. 0 a 1 a uzavřeme OK.

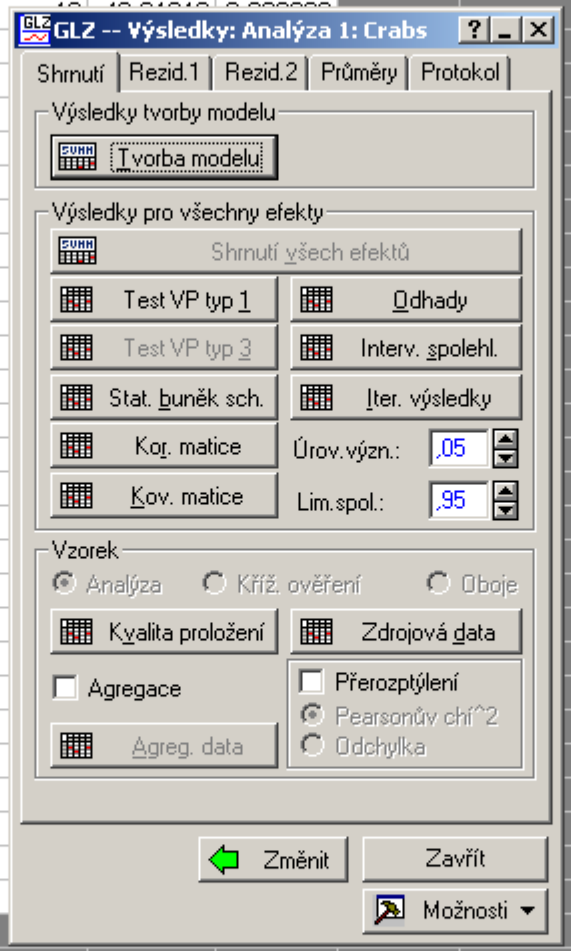


39. **LR:** V okénku *Vlastní obecné schéma* zaškrtneme ve druhém sloupci zvaném *Tvorba modelu* volbu *Nejlepší podskup.* a ostatní ponecháme na defaultních hodnotách. Klikneme na OK.



40. **LR:** Provede se 16 iterací a otevře se varovné upozornění o výpočtu. Potvrdíme OK.

Krok	Prom. 1	Prom. 2	Prom. 3	Prom. 4	Prom. 5	Stupně volnosti	věrohod. Skóre	p
1	WIDTH	WEIGHT	COLOR	SPINE	1*2	13	43,61142	0,000036
2	WIDTH	WEIGHT	COLOR	1*2		11	43,03263	0,000011
3	WEIGHT	COLOR	SPINE	1*2		12	42,81480	0,000024
4	WEIGHT	COLOR	1*2			10	42,42158	0,000006
5	WIDTH	COLOR	SPINE	1*2				
6	WIDTH	COLOR	1*2					
7	WIDTH	WEIGHT	SPINE	1*2				
8	WIDTH	SPINE	1*2					
9	WEIGHT	SPINE	1*2					
10	WIDTH	WEIGHT	1*2					
11	WIDTH	1*2						
12	WIDTH	WEIGHT	COLOR	SPINE				
13	WEIGHT	1*2						
14	WIDTH	COLOR	SPINE					
15	WIDTH	WEIGHT	COLOR					
16	WEIGHT	COLOR	SPINE					
17	WIDTH	COLOR						
18	WEIGHT	COLOR						
19	WIDTH	WEIGHT	SPINE					
20	WIDTH	WEIGHT						
21	WIDTH	SPINE						
22	WIDTH							
23	WEIGHT	SPINE						
24	COLOR	SPINE	1*2					
25	WEIGHT							
26	COLOR	1*2						
27	SPINE	1*2						
28	COLOR	SPINE						
29	1*2							
30	COLOR							
31	SPINE							



41. **LR:** Otevře se okno *Výsledky*, ve kterém klikneme na *Tvorbu modelu* v bloku *GLZ-Výsledky tvorby modelu.*

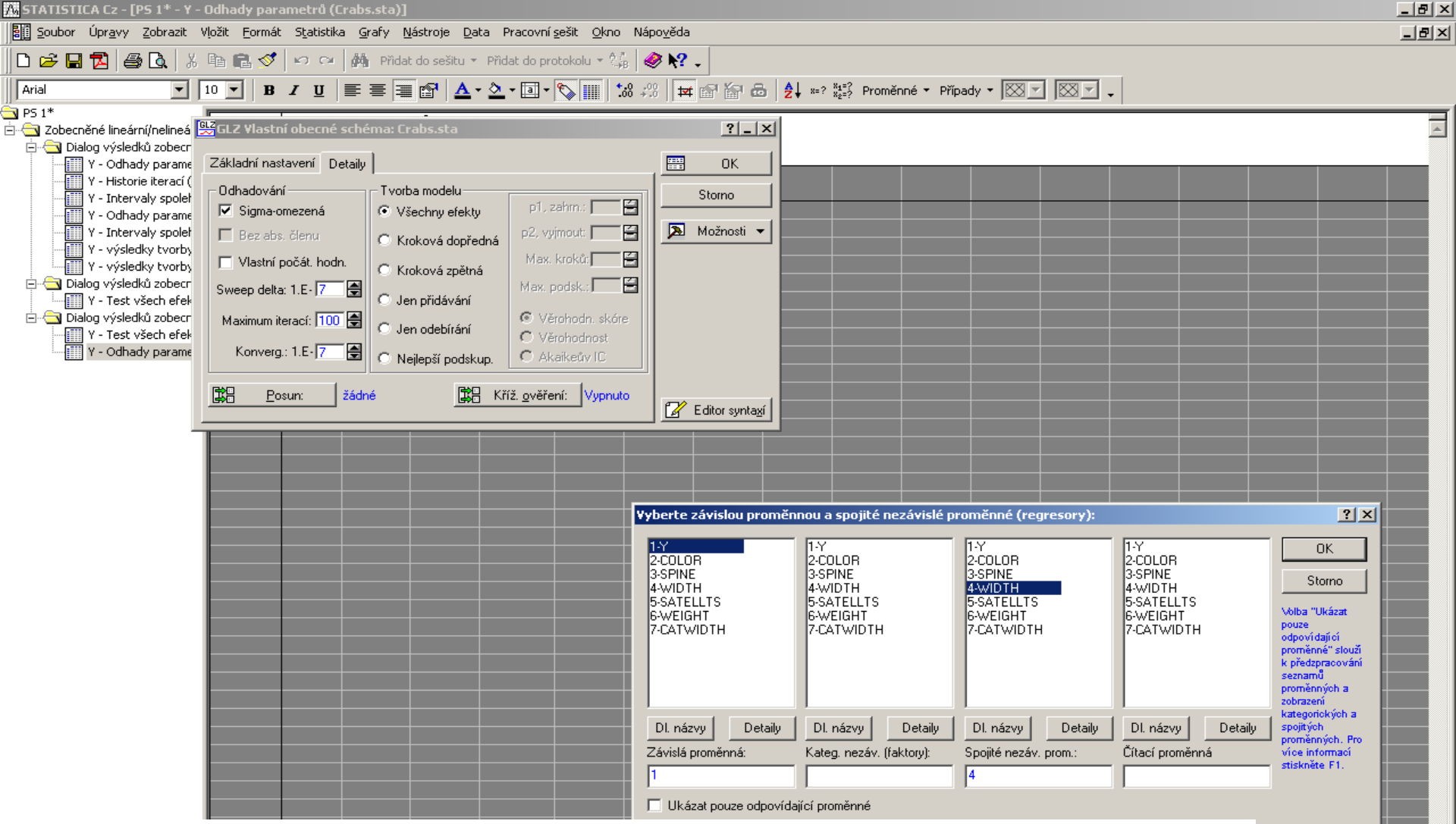
PS 1*
Zobecněné lineární/nelineárr
Dialog výsledků zobecně
Y - výsledky tvorby r
Y - výsledky tvorby r
Y - výsledky tvorby r

Y - výsledky tvorby modelu (Crabs)
Rozdělení : BINOMICKÉ
Linkující funkce: LOGIT

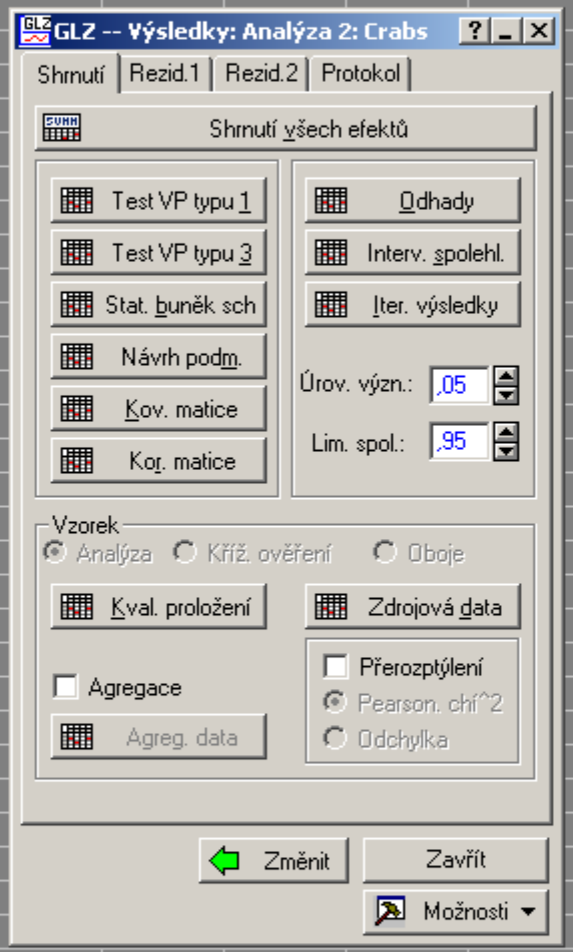
Krok	Prom. 1	Prom. 2	Prom. 3	Prom. 4	Prom. 5	Stupně volnosti	věrohod. Skóre	p
1	WIDTH	WEIGHT	COLOR	SPINE	1*2	13	43,61142	0,000036
2	WIDTH	WEIGHT	COLOR	1*2		11	43,03263	0,000011
3	WEIGHT	COLOR	SPINE	1*2		12	42,81480	0,000024
4	WEIGHT	COLOR	1*2			10	42,42158	0,000006
5	WIDTH	COLOR	SPINE	1*2		12	42,31810	0,000029
6	WIDTH	COLOR	1*2			10	41,47523	0,000009
7	WIDTH	WEIGHT	SPINE	1*2		10	40,71598	0,000013
8	WIDTH	SPINE	1*2			9	40,05314	0,000007
9	WEIGHT	SPINE	1*2			9	38,80766	0,000012
10	WIDTH	WEIGHT	1*2			8	37,59039	0,000009
11	WIDTH	1*2				7	36,69512	0,000005
12	WIDTH	WEIGHT	COLOR	SPINE		7	36,30684	0,000006
13	WEIGHT	1*2				7	35,90415	0,000008
14	WIDTH	COLOR	SPINE			6	35,34809	0,000004
15	WIDTH	WEIGHT	COLOR			5	35,00651	0,000002
16	WEIGHT	COLOR	SPINE			6	34,86763	0,000005
17	WIDTH	COLOR				4	34,33839	0,000001
18	WEIGHT	COLOR				4	32,95653	0,000001
19	WIDTH	WEIGHT	SPINE			4	28,91158	0,000008
20	WIDTH	WEIGHT				2	28,66396	0,000001
21	WIDTH	SPINE				3	27,99960	0,000004
22	WIDTH					1	27,87522	0,000000
23	WEIGHT	SPINE				3	26,56360	0,000007
24	COLOR	SPINE	1*2			11	26,38865	0,005679
25	WEIGHT					1	25,93596	0,000000
26	COLOR	1*2				9	25,44459	0,002517
27	SPINE	1*2				8	21,39145	0,006177
28	COLOR	SPINE				5	17,18215	0,004167
29	1*2					6	16,60959	0,010830
30	COLOR					3	14,07753	0,002802
31	SPINE					2	2,60183	0,272283

42. **LR:** Otevře se tabulka nejlepšího podsouboru výsledků s defaultně zvoleným věrohodnostním skóre jako kritériem věrohodnosti výsledků, seřazených zde v sestupném pořadí statistické významnosti celkového modelu. Všimněme si, že nejlepší model o jediném prediktoru je model se znakem Width. Nejlepší model o dvou znacích je model se znaky Width a dvoufaktorovou interakcí znaků Color a Spine. Protože je naším cílem nalézt nejjednodušší model, t.zn. model s jediným znakem Width v oblasti spojitých nezávisle proměnných. Navrátíme se zpět na *GLZ-Výsledky*.

GLZ -- Výsledky
Připravena
Start
ABC 123 ZÁZN
19:13



43. **LR:** V okně *Vlastní obecné schéma*, kam jsme se právě navrátili zadáme nyní ve *Tvorba modelu* návěští *Všechny efekty* a OK. Otevře se okno *Vyberte závisle proměnnou a spojité nezávisle proměnné (regresory)*. Za *Závisle proměnnou* zadáme 1-*y*. Za *Spojité nezávisle proměnné* zadáme 4-Width a OK.



44. **LR:** Otevře se okno *GLZ-Výsledky*. Otevřeme návěští *Odhady* v záložce *Shrnutí*.

metrů (Crabs)

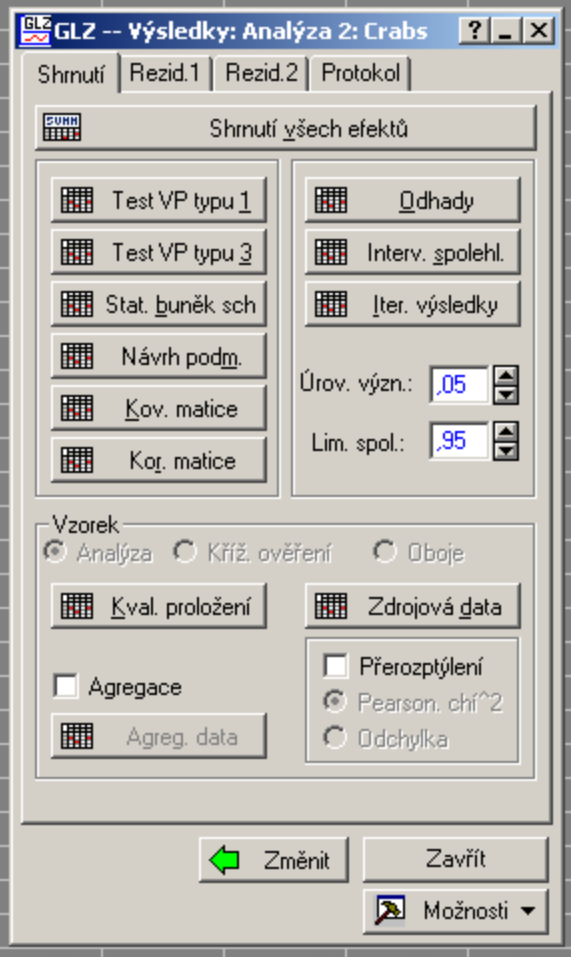
STATISTICA Cz - [PS 1* - Y - Odhady parametrů (Crabs)]

Y - Odhady parametrů (Crabs)
Rozdělení : BINOMICKÉ
Linkující funkce: LOGIT

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-12,3508	2,628731	22,07492	0,000003
WIDTH		2	0,4972	0,101736	23,88723	0,000001
Měřítko			1,0000	0,000000		

45. **LR:** Jak se dle předešlého dalo očekávat, odhad parametru znaku Width je statisticky významný. Znaménko tohoto odhadu ukazuje, že samička kraba je relativně široká, a proto má větší pravděpodobnost mít i více jak jednoho satelitního samečka. Proměnná *y* je zde indikátorem, zda existuje sameček navíc než jeden vlastní partner, kterého má každá samička vždy.

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-12,3508	2,628731	22,07492	0,000003
WIDTH		2	0,4972	0,101736	23,88723	0,000001
Měřítko			1,0000	0,000000		



46. **LR:** Otevřeme opět záložku na levém dolním okraji obrazovky zvanou *GLZ-Výsledky* a v nově otevřeném okně klikneme na *Kvalita proložení.*

- [PS 1* - Y - Statistiky kvality proložení (Crabs)]

Zobrazit Vložit Formát Statistika Grafy Nástroje Data Pracovní sešit Okno Nápověda

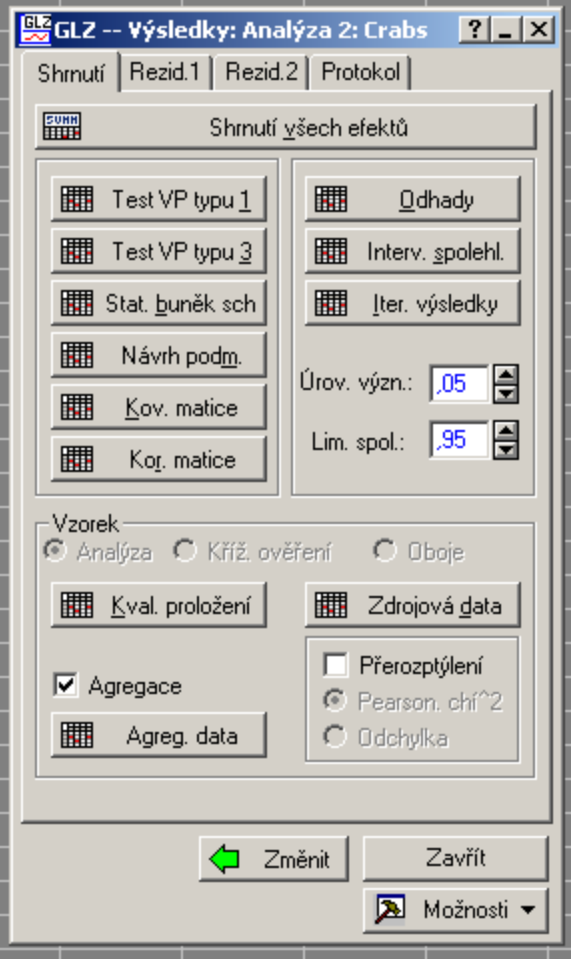
Přidat do sešitu Přidat do protokolu

Y - Statistiky kvality proložení (Crabs)
Rozdělení : BINOMICKÉ
Linkující funkce: LOGIT

Stat.	SV	Stat.	Stat/sv
deviance	171	194,4527	1,137150
Deviance v měřítku	171	194,4527	1,137150
Pearsonův chí^2	171	165,1434	0,965751
P. chí^2 v měř.	171	165,1434	0,965751
Log-věrohodnost		-97,2263	

47. **LR:** *Statistiky kvality proložení:* Objeví se rozličná kritéria těsnosti proložení jako jsou Deviance, Deviance v měřítku, Pearsonův Chi2, Pearsonův Chi2 v měřítku a log-věrohodnosti. Když tyto hodnoty podělíme počtem stupňů volnosti, obdržíme vždy číslo blízké 1. Neexistuje proto důkaz přerozptýlení a odhad znaku Width je zde vhodně škálován.

Stat.	SV	Stat.	Stat/sv
deviance	171	194,4527	1,137150
Deviance v měřítku	171	194,4527	1,137150
Pearsonův chí^2	171	165,1434	0,965751
P. chí^2 v měř.	171	165,1434	0,965751
Log-věrohodnost		-97,2263	



48. **LR:** Po návratu do tabulky *GLZ-Výsledky* zaškrtneme okénko *Agregace.*

STATISTICA Cz - [PS 1* - Y - Agregovaná data (Crabs.sta)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Pracovní sešit Okno Nápověda

Přidat do sešitu Přidat do protokolu

Arial 10 Proměnné Případy

PS 1*
- Zobecněné lineární/nelineárn
  - Dialog výsledků zobecněl
    - Y - Odhady parametr
    - Y - Historie iterací (Cr
    - Y - Intervaly spolehli
    - Y - Odhady parametr
    - Y - Intervaly spolehli
    - Y - výsledky tvorby r
    - Y - výsledky tvorby r
  - Dialog výsledků zobecněl
    - Y - Test všech efektů
  - Dialog výsledků zobecněl
    - Y - Test všech efektů
    - Y - Odhady parametr
    - Y - Statistiky kvality p
    - Y - Agregovaná data

Y - Agregovaná data (Crabs.sta)
Rozdělení : BINOMICKÉ
Linkující funkce: LOGIT

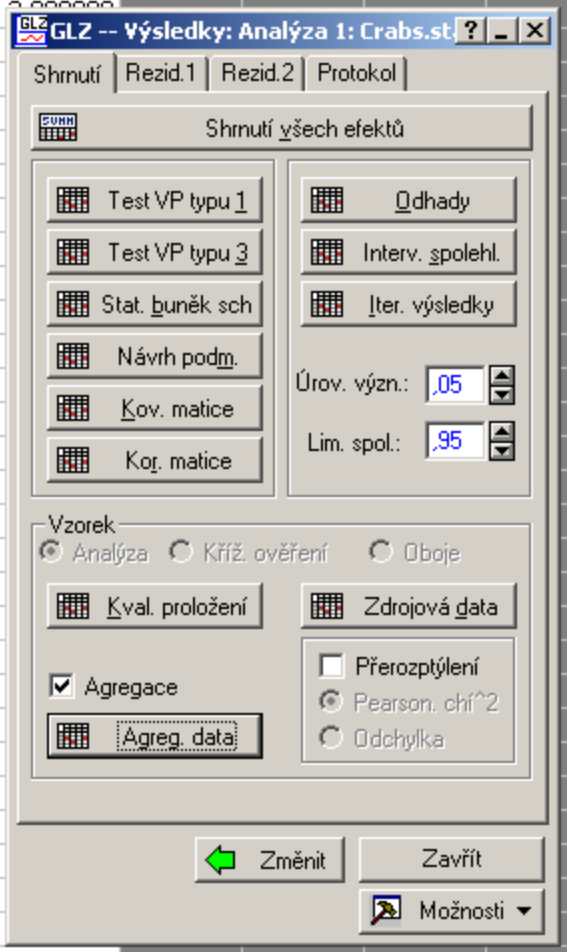
Ag. případ	Abs.člen	WIDTH 1	Kateg. 0	Kateg. 1	Kateg. Celkem
1	1,000000	28,30000	1,000000	2,000000	3,000000
2	1,000000	22,50000	1,000000	2,000000	3,000000
3	1,000000	26,00000	0,000000	6,000000	6,000000
4	1,000000	24,80000	1,000000	0,000000	1,000000
5	1,000000	23,80000	2,000000	1,000000	3,000000
6	1,000000	26,50000	2,000000	4,000000	6,000000
7	1,000000	24,70000	2,000000	3,000000	5,000000
8	1,000000	23,70000	3,000000	0,000000	3,000000
9	1,000000	25,60000	1,000000	1,000000	2,000000
10	1,000000	24,30000	2,000000	0,000000	2,000000
11	1,000000	25,80000	5,000000	2,000000	7,000000
12	1,000000	28,20000	0,000000	4,000000	4,000000
13	1,000000	21,00000	1,000000	0,000000	1,000000
14	1,000000	27,10000	1,000000	1,000000	2,000000
15	1,000000	25,20000	0,000000	2,000000	2,000000
16	1,000000	29,00000	0,000000	6,000000	6,000000
17	1,000000	27,40000	0,000000	3,000000	3,000000
18	1,000000	23,20000	0,000000	1,000000	1,000000
19	1,000000	25,00000	0,000000	6,000000	6,000000
20	1,000000	26,70000	1,000000	2,000000	3,000000
21	1,000000	26,20000	4,000000	4,000000	8,000000
22	1,000000	28,70000	1,000000	1,000000	2,000000
23	1,000000	26,80000	2,000000	1,000000	3,000000
24	1,000000	27,50000	2,000000	4,000000	6,000000
25	1,000000	24,90000	2,000000	1,000000	3,000000
26	1,000000	29,30000	0,000000	2,000000	2,000000
27	1,000000	25,70000	4,000000	2,000000	6,000000
28	1,000000	23,40000	1,000000	0,000000	1,000000
29	1,000000	27,90000	0,000000	2,000000	2,000000
30	1,000000	26,10000	0,000000	2,000000	2,000000
31	1,000000	27,70000	0,000000	2,000000	2,000000
32	1,000000	30,00000	0,000000	3,000000	3,000000
33	1,000000	28,50000	1,000000	3,000000	4,000000
34	1,000000	28,90000	0,000000	1,000000	1,000000
35	1,000000	30,30000	0,000000	1,000000	1,000000
36	1,000000	22,90000	2,000000	1,000000	3,000000
37	1,000000	27,20000	0,000000	2,000000	2,000000
38	1,000000	27,80000	1,000000	1,000000	2,000000
39	1,000000	25,50000	3,000000	0,000000	3,000000
40	1,000000	24,50000	3,000000	4,000000	7,000000
41	1,000000	27,00000	1,000000	4,000000	5,000000
42	1,000000	28,00000	1,000000	2,000000	3,000000

49. **LR:** Klikneme na *Agreg. data* a vidíme tabulku agregovaných dat.

Připravena ABC 123 ZÁZN

Start Doručená pošta - Micr... Total Commander 6.0 ... Microsoft PowerPoint -... STATISTICA Cz - [P... Anglicko-český - Lexicon 10:47

Ag. případ	Abs.člen	WIDTH 1	Kateg. 0	Kateg. 1	Kateg. Celkem
1	1,000000	28,30000	1,000000	2,000000	3,000000
2	1,000000	22,50000	1,000000	2,000000	3,000000
3	1,000000	26,00000	0,000000	6,000000	6,000000
4	1,000000	24,80000	1,000000	0,000000	1,000000
5	1,000000	23,80000	2,000000	1,000000	[illegible]
6	1,000000	26,50000	2,000000	4,000000	
7	1,000000	24,70000	2,000000	3,000000	
8	1,000000	23,70000	3,000000	0,000000	
9	1,000000	25,60000	1,000000	1,000000	
10	1,000000	24,30000	2,000000	0,000000	
11	1,000000	25,80000	5,000000	2,000000	
12	1,000000	28,20000	0,000000	4,000000	
13	1,000000	21,00000	1,000000	0,000000	
14	1,000000	27,10000	1,000000	1,000000	
15	1,000000	25,20000	0,000000	2,000000	
16	1,000000	29,00000	0,000000	6,000000	
17	1,000000	27,40000	0,000000	3,000000	
18	1,000000	23,20000	0,000000	1,000000	
19	1,000000	25,00000	0,000000	6,000000	
20	1,000000	26,70000	1,000000	2,000000	
21	1,000000	26,20000	4,000000	4,000000	
22	1,000000	28,70000	1,000000	1,000000	
23	1,000000	26,80000	2,000000	1,000000	
24	1,000000	27,50000	2,000000	4,000000	
25	1,000000	24,90000	2,000000	1,000000	
26	1,000000	29,30000	0,000000	2,000000	
27	1,000000	25,70000	4,000000	2,000000	
28	1,000000	23,40000	1,000000	0,000000	
29	1,000000	27,90000	0,000000	2,000000	
30	1,000000	26,10000	0,000000	2,000000	
31	1,000000	27,70000	0,000000	2,000000	
32	1,000000	30,00000	0,000000	3,000000	
33	1,000000	28,50000	1,000000	3,000000	4,000000
34	1,000000	28,90000	0,000000	1,000000	1,000000
35	1,000000	30,30000	0,000000	1,000000	1,000000



50. **LR:** Po návratu zpět do GLZ klikneme opět na *Kvalitu proložení* a uvidíme, že se data změnila.

TISTICA Cz - [PS 1* - Y - Statistiky kvality proložení (Crabs.sta)]

ubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Pracovní sešit Okno Nápověda

Přidat do sešitu Přidat do protokolu

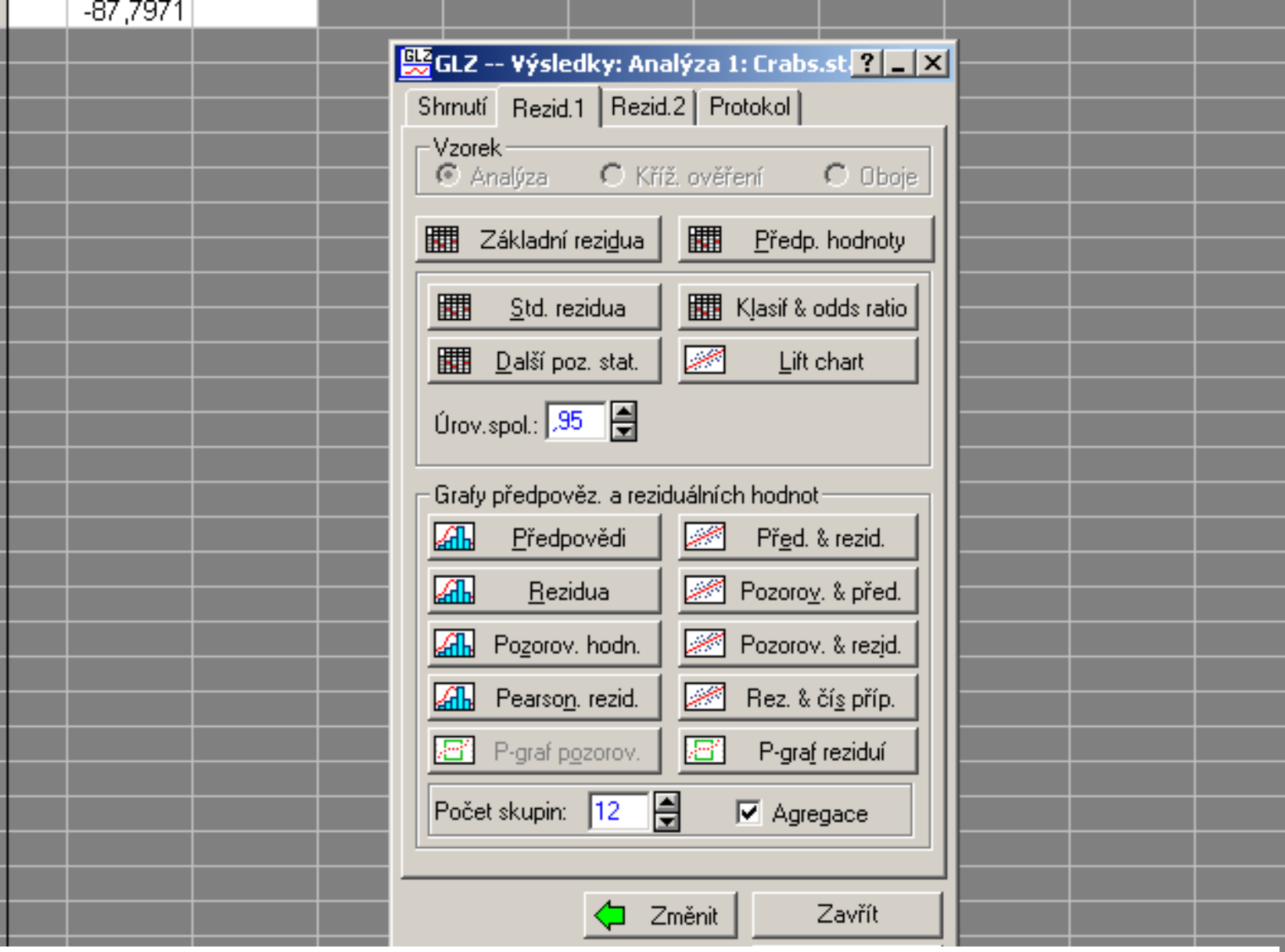
10

*
Zobecněné lineární/nelineárn
Dialog výsledků zobecněi
Y - Odhady parametr
Y - Historie iterací (Ci
Y - Intervaly spolehli
Y - Odhady parametr
Y - Intervaly spolehli
Y - výsledky tvorby r
Y - výsledky tvorby r
Dialog výsledků zobecněi
Y - Test všech efektů
Dialog výsledků zobecněi
Y - Test všech efektů

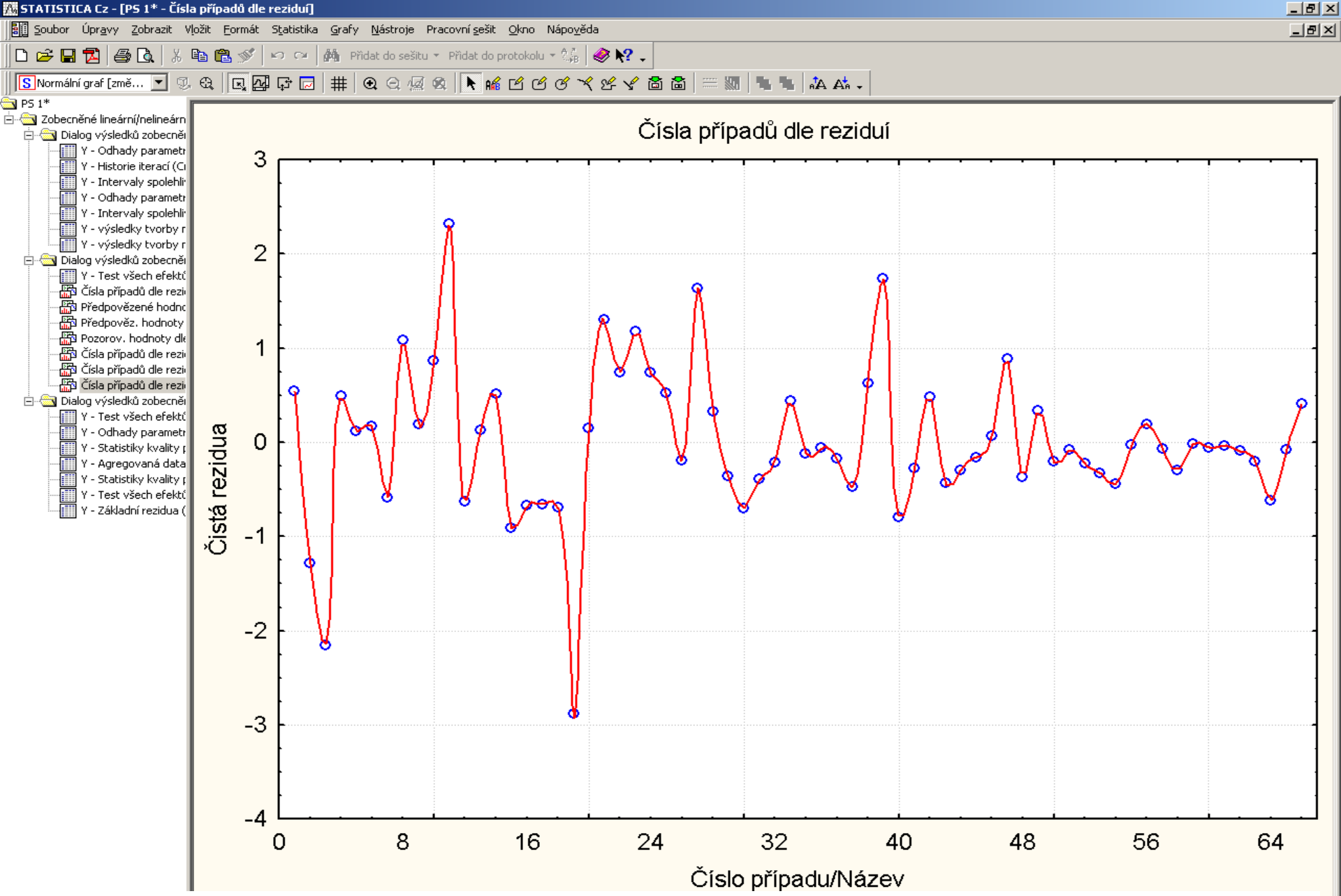
Y - Statistiky kvality proložení (Crabs.sta)
Rozdělení : BINOMICKÉ
Linkující funkce: LOGIT

Stat.	SV	Stat.	Stat/sv
deviance	64	69,7260	1,089469
Deviance v měřítku	64	69,7260	1,089469
Pearsonův chí^2	64	55,1779	0,862155
P. chí^2 v měř.	64	55,1779	0,862155
Log-věrohodnost		-97,2263	

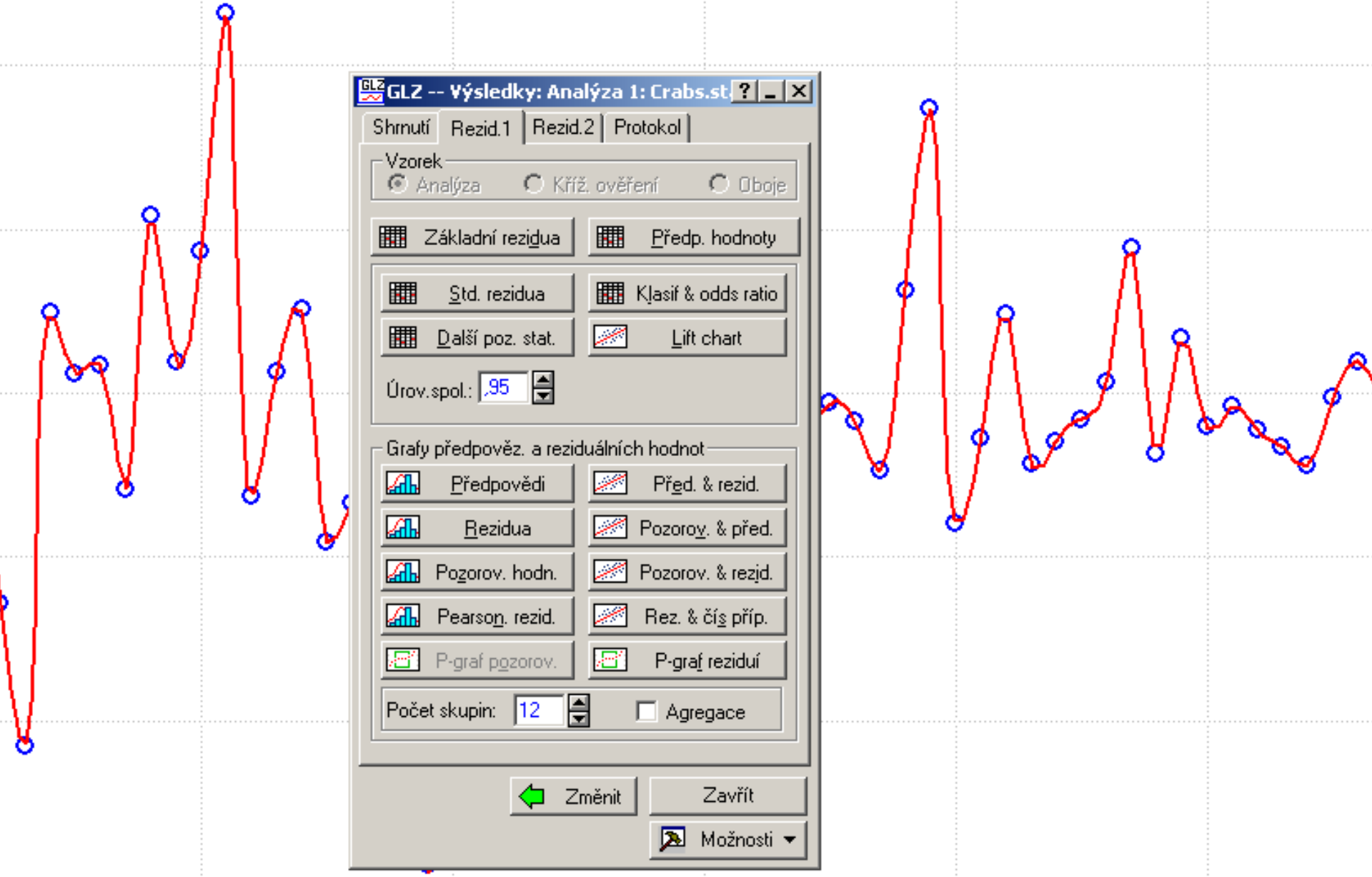
51. **LR:** Dostaneme poněkud jiné hodnoty statistiky těsnosti proložení. Také všechny hodnoty reziduí, predikovaných hodnot, atd. se změnily, můžeme si je prohlédnout v záložce *Rezid 1* a *Rezid 2*. Nyní máme pouze 66 řádků dat z původních 173.



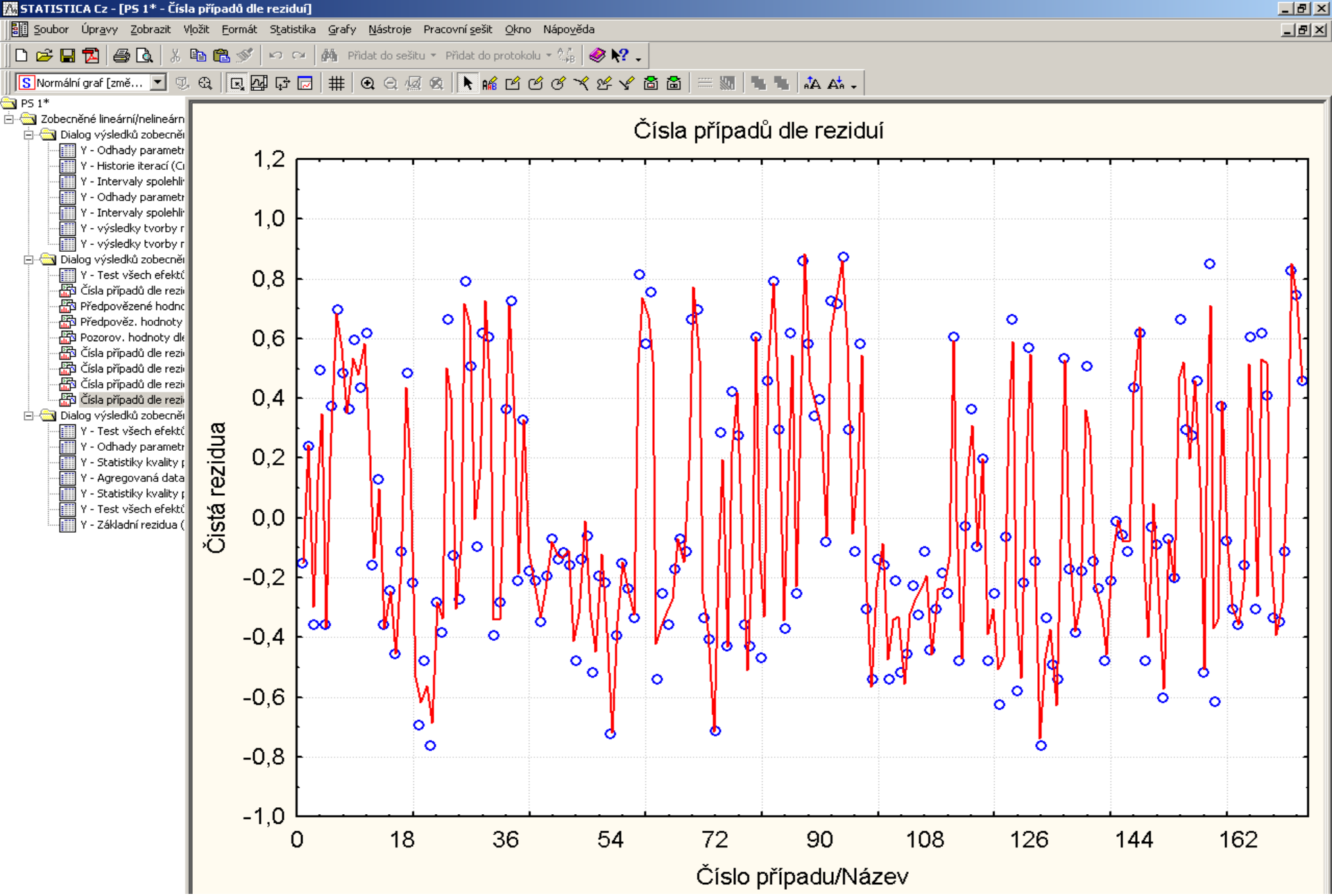
52. **LR:** Po návratu zpět do GLZ klikneme na záložku *Rezid 1* a zaškrtneme okénko *Agregace*. Pak si vybereme graf *Rezidua a číslo případu*.



53. **LR:** Otevře se graf reziduí. Všimněte si, že na x-ové ose máme pouze 66 případů.



54. **LR:** Když bychom nezaškrtli okénko *Agregace*, dostali bychom graf všech původních neagregovaných případů.



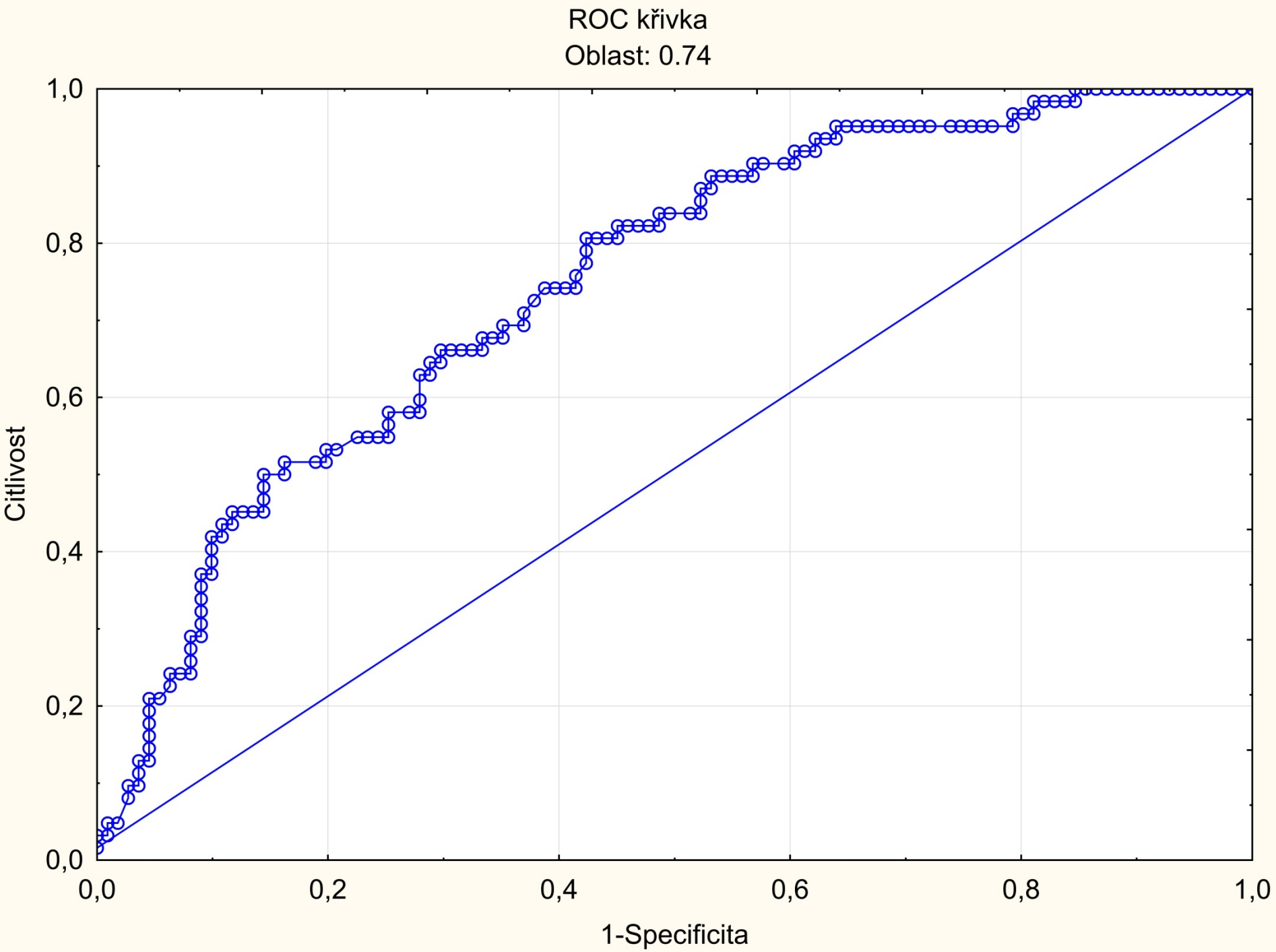
55. **LR:** Graf všech případů dle reziduí.

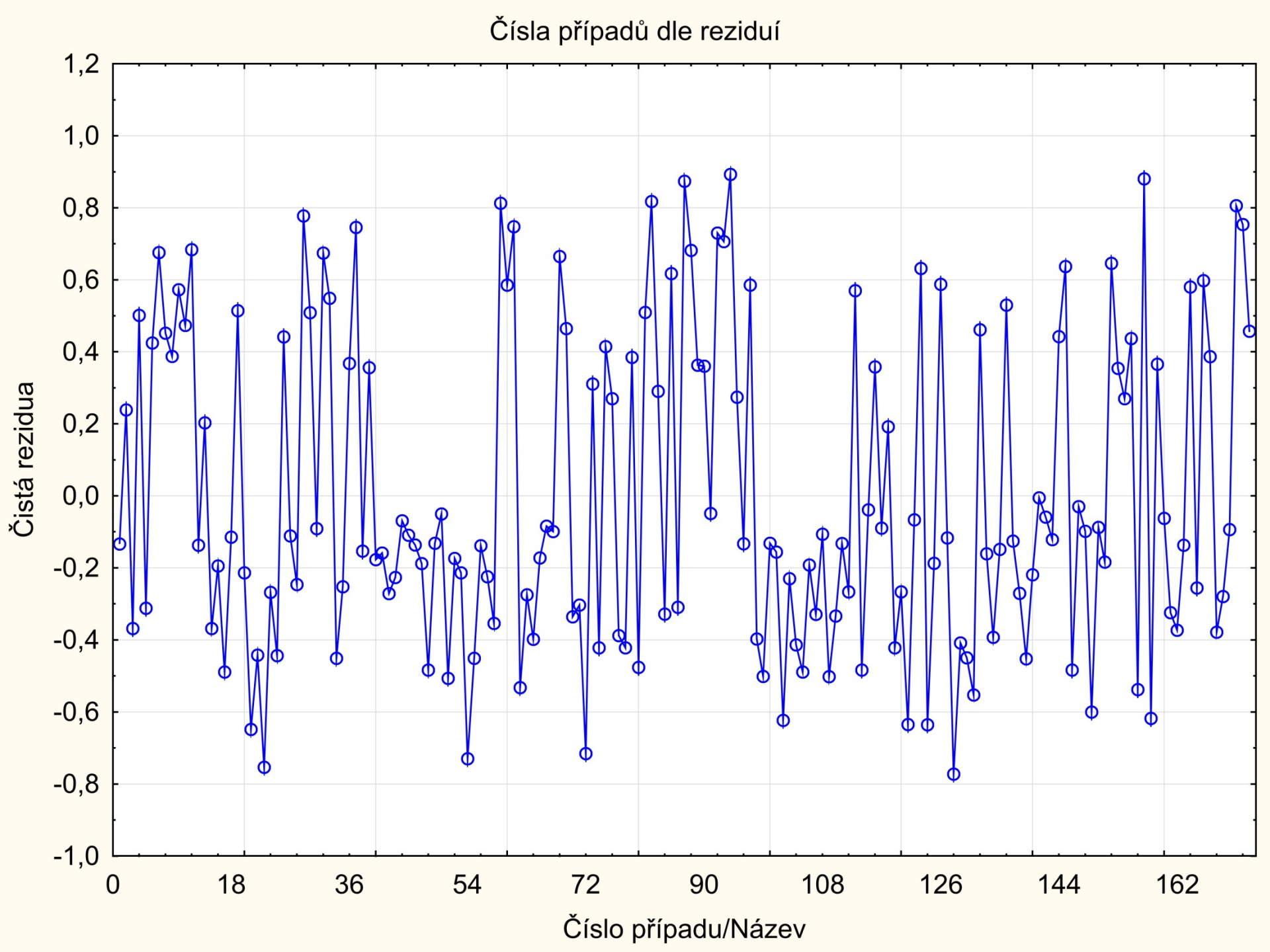
Efekt	Y - Odhady parametrů (Crabs) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že Y = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 80, %	Horní LS 80, %	p
Abs.člen		1	9,354726	3,528076	7,030502	4,83332	13,87614	0,008013
WIDTH		2	-0,306789	0,181948	2,843043	-0,53997	-0,07361	0,091770
WEIGHT		3	-0,833792	0,671648	1,541101	-1,69454	0,02696	0,214454
Měřítko			1,000000	0,000000		1,00000	1,00000	

Efekt	Y - Věrohodn. test typ 1 (Crabs) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že Y = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-112,879		
WIDTH	1	-97,226	31,30586	0,000000
WEIGHT	1	-96,446	1,56078	0,211552

Efekt	Y - Věrohodn. test typ 3 (Crabs) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že Y = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
WIDTH	1	-97,8686	2,845263	0,091644
WEIGHT	1	-97,2263	1,560777	0,211552

Efekt	Y - Test všech efektů (Crabs) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že Y = 0		
	Stupně volnosti	Wald. Stat.	p
Abs. člen	1	7,030502	0,008013
WIDTH	1	2,843043	0,091770
WEIGHT	1	1,541101	0,214454


ROC křivka
Oblast: 0.74
Citlivost
1-Specificita
0,0
0,2
0,4
0,6
0,8
1,0


Čísla případů dle reziduí
Čistá rezidua
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
0
18
36
54
72
90
108
126
144
162
Číslo případu/Název

Příklad 8.7/str. 308 (261)

Data PCS.sta

STATISTICA Cz - [Data: 87PCS.sta (11s krát 380ř)]

Soubor Upravit Zobrazit Vložit Formát Statistiky Data Mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS

Arial 10

	1 CAPSULE	2 AGE	3 RACE	4 DPROS	5 DCAPS	6 PSA	7 VOL	8 GLEASON
1.000000	0	65	1	2	1	1.4	0	6
2.000000	0	72	1	3	2	6.7	0	7
3.000000	0	70	1	1	2	4.9	0	6
4.000000	0	76	2	2	1	51.2	20	7
5.000000	0	69	1	1	1	12.3	55.9	6
6.000000	1	71	1	3	2	3.3	0	8
7.000000	0	68	2	4	2	31.9	0	7
8.000000	0	61	2	4	2	66.7	27.2	7
9.000000	0	69	1	1	1	3.9	24	7
10.000000	0	68	2	1	2	13	0	6
11.000000	1	68	2	4	2	4	0	7
12.000000	1	72	1	2	2	21.2	0	7
13.000000	1	72	1	4	2	22.7	0	9
14.000000	1	65	1	4	2	39	0	7
15.000000	0	75	1	1	1	7.5	0	5
16.000000	0	73	1	2	1	2.6	0	5
17.000000	0	75	2	1	1	2.5	0	5
18.000000	0	70	1	2	1	2.6	11.8	5
19.000000	0	54	1	1	2	2.8	0	6
20.000000	1	67	2	3	2	8.6	25.5	7
21.000000	1	58	1	2	1	3.1	0	7
22.000000	1	70	0	4	1	67.1	0	7
23.000000	1	74	1	3	1	12.7	27.5	7
24.000000	0	73	1	1	1	12.3	47.3	7
25.000000	1	77	1	1	1	61.1	58	7
26.000000	0	77	1	1	1	8.8	0	5
27.000000	0	67	1	2	1	2.8	25.6	7
28.000000	0	73	1	3	1	2.9	14.1	5
29.000000	0	64	1	3	1	5.6	34.1	5
30.000000	0	58	1	4	1	5.6	22.8	6
31.000000	1	54	1	3	1	8.4	18.3	6
32.000000	0	72	1	2	1	6.5	22.5	7
33.000000	0	77	1	4	1	11	0	8
34.000000	1	60	1	3	2	9.5	0	7
35.000000	0	65	1	2	1	11.1	17.7	6
36.000000	0	71	1	2	1	7.3	20	7
37.000000	0	54	1	2	1	1	0	6
38.000000	1	78	1	1	2	27.2	0	8
39.000000	1	63	1	2	1	35.1	18.7	7
40.000000	1	73	1	3	1	4.5	26.4	7
41.000000	1	66	1	3	1	7.9	20.8	7
42.000000	0	71	1	1	1	2	0	6
43.000000	1	71	1	2	1	7.5	0	6
44.000000	0	72	1	2	1	5.3	0	7

GLZ – Výsledky: Anal...

	4 DPROS	5 DCAPS	6 PSA	7 VOL	8 GLEASON	9 Var9	10 Var10	11 Var11
1	2	1	1.4	0	6			
1	3	2	6.7	0	7			
1	1	2	4.9	0	6			
2	2	1	51.2	20	7			
1	1	1	12.3	55.9	6			
1	3	2	3.3	0	8			
2	4	2						
2	4	2						
1	1	1						
2	1	2						
2	4	2						
1	2	2						
1	4	2						
1	4	2						
1	1	1						
1	2	1						
2	1	1						
1	2	1						
1	1	2						
2	3	2						
1	2	1						
0	4	1						
1	3	1						
1	1	1						
1	71	1	3	2	3.3	0	8	
0	68	2	4	2	31.9	0	7	
0	61	2	4	2	66.7	27.2	7	
0	69	1	1	1	3.9	24	7	
0	68	2	1	2	13	0	6	
1	68	2	4	2	4	0	7	
1	72	1	2	2	21.2	0	7	
1	72	1	4	2	22.7	0	9	
1	65	1	4	2	39	0	7	
0	75	1	1	1	7.5	0	5	
0	73	1	2	1	2.6	0	5	
0	75	2	1	1	2.5	0	5	
0	70	1	2	1	2.6	11.8	5	
0	54	1	1	2	2.8	0	6	

Zobecněné lineární/nelineární modely: 87PCS.sta

Základní nastavení | Detaily

Typ analýzy:
- Jednofaktorová ANOVA
- ANOVA hlavních efektů
- Vícefaktorová ANOVA
- Hierarchická ANOVA
- Jednorozměrná regrese
- Vícerozměrná regrese
- Faktorová regrese
- Polynomická regrese
- Regrese odezvové plochy
- Regrese odezv. plochy směsi

Metoda specifikace:
- Rychlé nastavení
- Průvodce analýzou
- Editor syntaxí analýz

Rozdělení:
- Normální
- Poissonovo
- Gama
- Binomické
- Multinomické
- Ordinální multinomické

OK | Storno | Možnosti | Otevři Data | SELECT CASES | Vážené momenty | SV = W-1 N-1

GLZ Vlastní obecné schéma:

Základní nastavení | Detaily

Proměnné

Závislá proměnná: žádné
Čítací proměnná: žádné
Kategor. faktory: žádné
Spojité nezáv. prom.: žádné

Odezv. kódy žádné
...aktorů: žádné
...efekty žádné

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - CAPSULE
2 - AGE
3 - RACE
4 - DPROS
5 - DCAPS
6 - PSA
7 - VOL
8 - GLEASON
9 - Var9
10 - Var10
11 - Var11

Roztáhn. | Přiblížit

Závislá proměnná: 1
Kateg. nezáv. (faktory):
Spojité nezáv. prom.: 2-8
Čítací proměnná

OK | Storno | [Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

Ukázat pouze odpovídající proměnné

CAPSULE - Odhady parametrů (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	7.279683	1.688841	18.58007	0.000016
AGE		2	0.012132	0.019029	0.40652	0.523741
RACE		3	0.624238	0.453722	1.89287	0.168878
DPROS		4	-0.556606	0.136395	16.65320	0.000045
DCAPS		5	-0.483745	0.463298	1.09022	0.296424
PSA		6	-0.027395	0.009863	7.71510	0.005476
VOL		7	0.011241	0.007750	2.10370	0.146943
GLEASON		8	-0.976322	0.167041	34.16171	0.000000
Měřítko			1.000000	0.000000		

CAPSULE - Intervaly spolehlivost odhadů (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Efekt	Úroveň Efekt	Sloupec	Dolní LS 95. %	Horní LS 95. %
Abs.člen		1	3.96962	10.58975
AGE		2	-0.02516	0.04943
RACE		3	-0.26504	1.51352
DPROS		4	-0.82394	-0.28928
DCAPS		5	-1.39179	0.42430
PSA		6	-0.04672	-0.00806
VOL		7	-0.00395	0.02643
GLEASON		8	-1.30372	-0.64893

CAPSULE - Historie iterací (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen		1	0.000	4.648	6.742	7.249	7.280
AGE		2	0.000	0.007	0.011	0.012	0.012
RACE		3	0.000	0.331	0.526	0.615	0.624
DPROS		4	0.000	-0.370	-0.510	-0.553	-0.557
DCAPS		5	0.000	-0.390	-0.441	-0.480	-0.484
PSA		6	0.000	-0.015	-0.023	-0.027	-0.027
VOL		7	0.000	0.008	0.010	0.011	0.011
GLEASON		8	0.000	-0.586	-0.894	-0.972	-0.976
Věrohodnost			-263.396	-196.529	-189.699	-189.325	-189.324

CAPSULE - Věrohodn. test typ 1 (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že CAPSULE =

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-256.144		
AGE	1	-256.012	0.26399	0.607392
RACE	1	-255.910	0.20436	0.651222
DPROS	1	-235.091	41.63944	0.000000
DCAPS	1	-228.590	13.00028	0.000311
PSA	1	-212.322	32.53650	0.000000
VOL	1	-210.527	3.58933	0.058152
GLEASON	1	-189.324	42.40788	0.000000

CAPSULE - Věrohodn. test typ 3 (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
AGE	1	-189.527	0.40647	0.523769
RACE	1	-190.301	1.95457	0.162095
DPROS	1	-198.101	17.55580	0.000028
DCAPS	1	-189.880	1.11215	0.291614
PSA	1	-193.919	9.19058	0.002433
VOL	1	-190.409	2.17088	0.140646
GLEASON	1	-210.527	42.40788	0.000000

CAPSULE - Statistiky kvality proložení (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Stat.	SV	Stat.	Stat/sv
Odchylka	372	378.647	1.017868
Deviance v měřítku	372	378.647	1.017868
Pearsonovo Chi2	372	364.215	0.979073
Scaled P. Chi2	372	364.215	0.979073
Log-věrohodnost		-189.324	

CAPSULE - Předpověz. hodnoty (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že CAPSULE = 0

Případ	Odezva '1' = 0	Před. Hodnota	Lineár. Před.	Směrod. Chyba	Dolní LS 95. %	Horní LS 95. %
\	1.000000	0.768398	1.19929	0.209915	0.687373	0.833511
\	1.000000	0.293666	-0.87765	0.474089	0.141020	0.512885
\	1.000000	0.775030	1.23693	0.528399	0.550153	0.906580
\	1.000000	0.460394	-0.15876	0.547913	0.225718	0.714049
\	1.000000	0.894184	2.13421	0.396422	0.795311	0.948397
\	0.000000	0.145174	-1.77296	0.522955	0.057436	0.321260
\	1.000000	0.175247	-1.54889	0.603193	0.061162	0.409351
\	1.000000	0.092677	-2.28138	0.715218	0.024526	0.293265
\	1.000000	0.736801	1.02941	0.274594	0.620392	0.827442
10.000000	1.000000	0.834106	1.61501	0.665955	0.576827	0.948839
11.000000	0.000000	0.313334	-0.78458	0.637360	0.115701	0.614110
12.000000	0.000000	0.327776	-0.71826	0.466610	0.163447	0.548912
13.000000	0.000000	0.021348	-3.82521	0.605465	0.006614	0.066700
14.000000	0.000000	0.082866	-2.40402	0.532013	0.030866	0.204030
15.000000	1.000000	0.936222	2.68644	0.398109	0.870587	0.969726
16.000000	1.000000	0.903767	2.23980	0.340511	0.828127	0.948201
17.000000	1.000000	0.969161	3.44765	0.620369	0.903064	0.990655
18.000000	1.000000	0.911819	2.33605	0.294658	0.853026	0.948513
19.000000	1.000000	0.750324	1.10034	0.566742	0.497387	0.901245
20.000000	0.000000	0.480143	-0.07947	0.605388	0.219942	0.751582
21.000000	0.000000	0.522851	0.09147	0.272743	0.391003	0.651590
22.000000	0.000000	0.037196	-3.25364	0.864450	0.007048	0.173740
23.000000	0.000000	0.444019	-0.22487	0.255094	0.326327	0.568346
24.000000	1.000000	0.752082	1.10975	0.356844	0.601171	0.859259

- Základní rezidua (87PCS.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOG
Modelovaná pravděpodobnost, že CAPSULE = 0

Případ	0 Čistá	0 Pearson.	0 Odchylka
\	0.231602	0.54901	0.72587
\	0.706334	1.55088	1.56545
\	0.224970	0.53877	0.71394
\	0.539606	1.08261	1.24553
\	0.105816	0.34400	0.47296
\	-0.145174	-0.41210	-0.56010
\	0.824753	2.16938	1.86631
\	0.907323	3.12893	2.18112
\	0.263199	0.59768	0.78158
10.000000	0.165894	0.44597	0.60232
11.000000	-0.313334	-0.67551	-0.86707
12.000000	-0.327776	-0.69828	-0.89125
13.000000	-0.021348	-0.14770	-0.20775
14.000000	-0.082866	-0.30059	-0.41594
15.000000	0.063778	0.26100	0.36305
16.000000	0.096233	0.32631	0.44985
17.000000	0.030839	0.17838	0.25030
18.000000	0.088181	0.31098	0.42968
19.000000	0.249676	0.57685	0.75796
20.000000	-0.480143	-0.96104	-1.14386
21.000000	-0.522851	-1.04680	-1.21649
22.000000	-0.037196	-0.19655	-0.27534

Efekt	CAPSULE - Odhady parametrů (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 80, %	Horní LS 80, %	p
Abs.člen		1	7,279683	1,688841	18,58007	5,11535	9,444020	0,000016
AGE		2	0,012132	0,019029	0,40652	-0,01225	0,036519	0,523741
RACE		3	0,624238	0,453722	1,89287	0,04277	1,205706	0,168878
DPROS		4	-0,556606	0,136395	16,65320	-0,73140	-0,381808	0,000045
DCAPS		5	-0,483745	0,463298	1,09022	-1,07749	0,109995	0,296424
PSA		6	-0,027395	0,009863	7,71510	-0,04003	-0,014755	0,005476
VOL		7	0,011241	0,007750	2,10370	0,00131	0,021174	0,146943
GLEASON		8	-0,976322	0,167041	34,16171	-1,19039	-0,762251	0,000000
Měřítko			1,000000	0,000000		1,00000	1,000000	

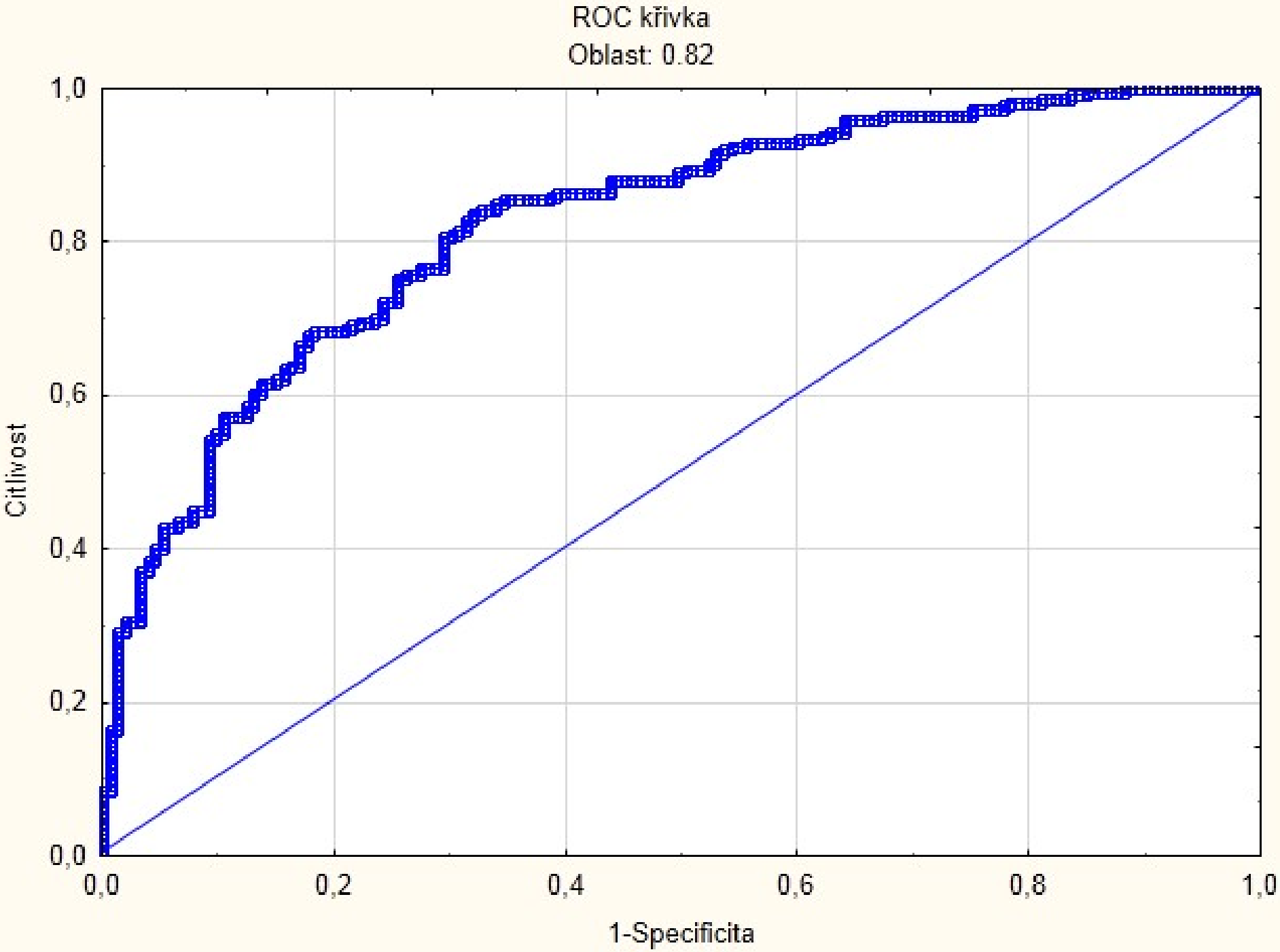
Efekt	CAPSULE - Intervaly spolehlivost odhadů (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0			
	Úroveň Efekt	Sloupec	Dolní LS 80, %	Horní LS 80, %
Abs.člen		1	5,11535	9,444020
AGE		2	-0,01225	0,036519
RACE		3	0,04277	1,205706
DPROS		4	-0,73140	-0,381808
DCAPS		5	-1,07749	0,109995
PSA		6	-0,04003	-0,014755
VOL		7	0,00131	0,021174
GLEASON		8	-1,19039	-0,762251

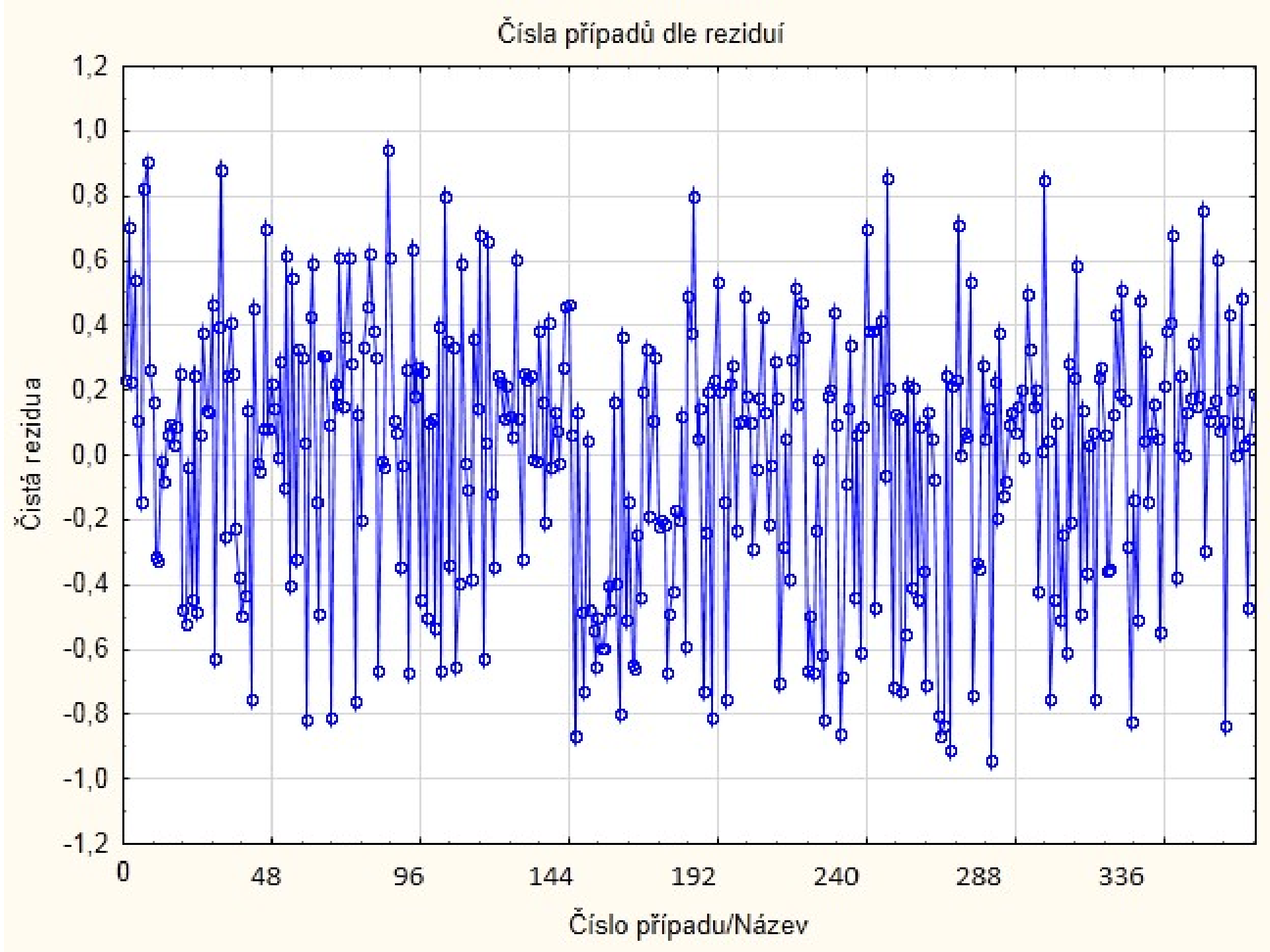
Efekt	CAPSULE - Historie iterací (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0						
	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen		1	0,000	4,648	6,742	7,249	7,280
AGE		2	0,000	0,007	0,011	0,012	0,012
RACE		3	0,000	0,331	0,526	0,615	0,624
DPROS		4	0,000	-0,370	-0,510	-0,553	-0,557
DCAPS		5	0,000	-0,390	-0,441	-0,480	-0,484
PSA		6	0,000	-0,015	-0,023	-0,027	-0,027
VOL		7	0,000	0,008	0,010	0,011	0,011
GLEASON		8	0,000	-0,586	-0,894	-0,972	-0,976
Věrohodnost			-263,396	-196,529	-189,699	-189,325	-189,324

Efekt	CAPSULE - Věrohodn. test typ 1 (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-256,144		
AGE	1	-256,012	0,26399	0,607392
RACE	1	-255,910	0,20436	0,651222
DPROS	1	-235,091	41,63944	0,000000
DCAPS	1	-228,590	13,00028	0,000311
PSA	1	-212,322	32,53650	0,000000
VOL	1	-210,527	3,58933	0,058152
GLEASON	1	-189,324	42,40788	0,000000

Efekt	CAPSULE - Věrohodn. test typ 3 (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
AGE	1	-189,527	0,40647	0,523769
RACE	1	-190,301	1,95457	0,162095
DPROS	1	-198,101	17,55580	0,000028
DCAPS	1	-189,880	1,11215	0,291614
PSA	1	-193,919	9,19058	0,002433
VOL	1	-190,409	2,17088	0,140646
GLEASON	1	-210,527	42,40788	0,000000

Efekt	CAPSULE - Test všech efektů (87PCS) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že CAPSULE = 0		
	Stupně volnosti	Wald. Stat.	p
Abs. člen	1	18,58007	0,000016
AGE	1	0,40652	0,523741
RACE	1	1,89287	0,168878
DPROS	1	16,65320	0,000045
DCAPS	1	1,09022	0,296424
PSA	1	7,71510	0,005476
VOL	1	2,10370	0,146943
GLEASON	1	34,16171	0,000000


ROC křivka
Oblast: 0.82
1,0
0,8
0,6
0,4
0,2
0,0
Citlivost
0,0
0,2
0,4
0,6
0,8
1,0
1-Specificita


Čísla případů dle reziduí
Čistá rezidua
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-1,2
0
48
96
144
192
240
288
336
Číslo případu/Název

Příklad 8.8/str. 311

Data Meexp.sta

STATISTICA Cz - [Data: 88Meexp.sta (8s krát 412ř)]

Soubor Upravit Zobrazit Vložit Formát Statistiky Data Mining

Přidat d

Arial 9

	1 ME	2 SYMPT	3 PB	4 HIST	5 BSE	6 DETC	7 Var7	8 Var8
1.000000	0	3	7	0	1	2		
2.000000	0	2	11	0	1	3		
3.000000	0	3	8	1	1	3		
4.000000	1	3	11	0	1	3		
5.000000	2	4	7	0	1	3		
6.000000	0	3	7	0	1	3		
7.000000	2	4	6	0	1	2		
8.000000	0	4	6	0	1	3		
9.000000	0	2	6	0	1	3		
10.000000	1	4	6	0	1	3		
11.000000	0	4	8	0	1	2		
12.000000	0	3	6	1	0	3		
13.000000	0	4	6	0	1	3		
14.000000	0	1	5	1	1	3		
15.000000	0	2	8	0	0	2		
16.000000	0	4	11	0	1	3		
17.000000	0	3	6	0	1	3		
18.000000	1	4	5	0	1	3		
19.000000	0	3	10	0	1	3		
20.000000	0	3	10	0	1	3		
21.000000	2	3	5	0	1	3		
22.000000	0	4	5	0	1	2		
23.000000	0	1	5	0	1	3		
24.000000	0	2	8	0	1	2		
25.000000	1	2	9	1	1	2		
26.000000	2	4	7	0	1	3		
27.000000	0	4	7	0	1	2		
28.000000	0	3	10	0	1	3		
29.000000	2	3	7	0	1	2		
30.000000	0	3	7	0	1	3		
31.000000	1	4	7	0	1	2		

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - ME	2 - SYMPT	3 - PB	4 - HIST	5 - BSE	6 - DETC	7 - Var7	8 - Var8

Roztáhn. Přiblížit

Závislá proměnná: 5

Kateg. nezáv. (faktory):

Spojité nezáv. prom.: 1-4 6

Čítací proměnná

OK

Storno

[Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

Ukázat pouze odpovídající proměnné

BSE - Odhady parametrů (88Meexp.sta)

Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT

Modelovaná pravděpodobnost, že BSE = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-1.24385	1.053363	1.394382	0.237667
ME		2	-0.62219	0.258655	5.786336	0.016151
SYMPT		3	-0.07790	0.166829	0.218037	0.640540
PB		4	0.13412	0.072958	3.379245	0.066022
HIST		5	-0.18157	0.570330	0.101353	0.750212
DETC		6	-0.46355	0.252377	3.373580	0.066250
Měřítko			1.00000	0.000000		

BSE - Intervaly spolehlivost odhadů (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

Efekt	Úroveň Efekt	Sloupec	Dolní LS 95. %	Horní LS 95. %
Abs.člen		1	-3.30840	0.820701
ME		2	-1.12914	-0.115235
SYMPT		3	-0.40488	0.249078
PB		4	-0.00888	0.277111
HIST		5	-1.29940	0.936257
DETC		6	-0.95820	0.031101

BSE - Věrohodn. test typ 1 (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-160.026		
ME	1	-153.302	13.44646	0.000245
SYMPT	1	-152.447	1.71186	0.190744
PB	1	-149.652	5.58832	0.018081
HIST	1	-149.611	0.08347	0.772646
DETC	1	-147.971	3.27898	0.070173

BSE - Historie iterací (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT		3	0.000	-0.036	-0.066	-0.077	-0.078	-0.078
PB		4	0.000	0.065	0.116	0.133	0.134	0.134
HIST		5	0.000	-0.062	-0.132	-0.175	-0.181	-0.182
DETC		6	0.000	-0.258	-0.418	-0.461	-0.464	-0.464
Věrohodnost			-285.577	-156.983	-148.625	-147.981	-147.971	-147.971

BSE - Věrohodn. test typ 3 (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
ME	1	-151.403	6.864224	0.008794
SYMPT	1	-148.079	0.216605	0.641639
PB	1	-149.661	3.379972	0.065993
HIST	1	-148.024	0.104914	0.746011
DETC	1	-149.611	3.278977	0.070173

BSE - Statistiky kvality proložení (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

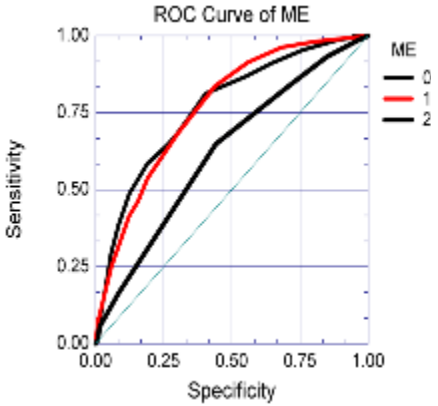
Stat.	SV	Stat.	Stat/sv
Odchylka	406	295.942	0.728922
Deviance v měřítku	406	295.942	0.728922
Pearsonovo Chi2	406	403.309	0.993372
Scaled P. Chi2	406	403.309	0.993372
Log-věrohodnost		-147.971	

BSE - Předpověz. hodnoty (88Meexp.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že BSE = 0

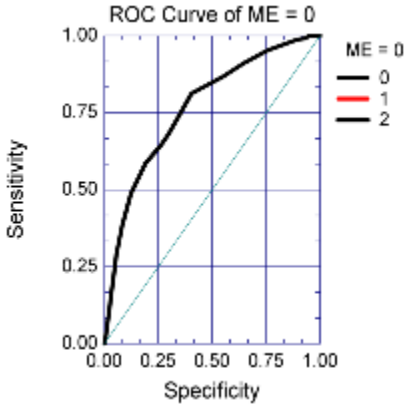
Případ	Odezva '1' = 0	Před. Hodnota	Lineár. Před.	Směrod. Chyba	Dolní LS 95. %	Horní LS 95. %
\	0.000000	0.187577	-1.46583	0.257538	0.122325	0.276664
\	0.000000	0.211649	-1.31502	0.328683	0.123551	0.338316
\	0.000000	0.121657	-1.97683	0.582480	0.042352	0.302548
\	0.000000	0.117626	-2.01510	0.365013	0.061196	0.214217
\	0.000000	0.037267	-3.25166	0.466306	0.015283	0.088047
\	0.000000	0.126819	-1.92938	0.227818	0.085029	0.184995
\	0.000000	0.051066	-2.92223	0.515474	0.019217	0.128766
\	0.000000	0.105138	-2.14140	0.314159	0.059685	0.178630

Závěr: V logistickém regresním modelu jsou z původních 5 znaků nezávisle proměnné statisticky významné 3 znaky, a to pro závisle proměnnou ***BSE* = 0** znaky ***ME, DETC, PB.***

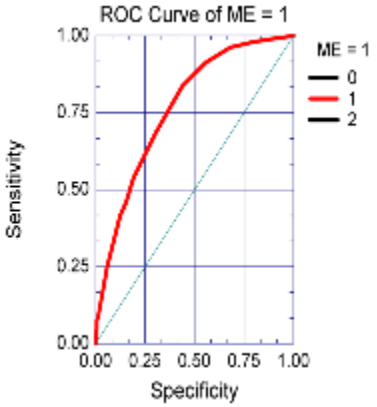
Grafy kvality logistické regresní analýzy: Graf prahové operační charakteristiky ROC vystihuje správnost diagnostického testu, zda modelem vypočtené Ano nebo Ne je správné.



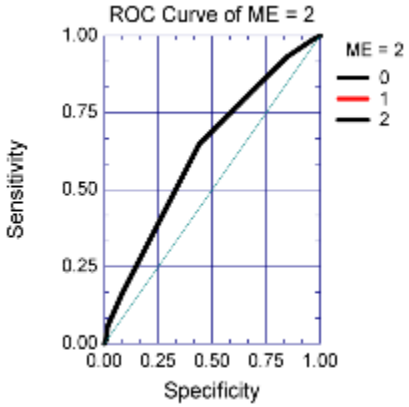
Obr. 8.10a ROC graf pro *ME* = 0, 1 a 2.



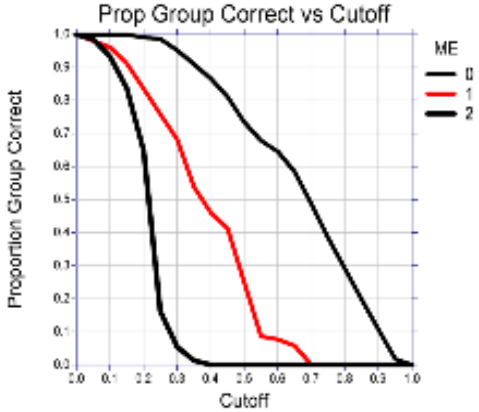
Obr. 8.10b ROC graf pro *ME* = 0.



Obr. 8.10c ROC graf pro *ME* = 1.

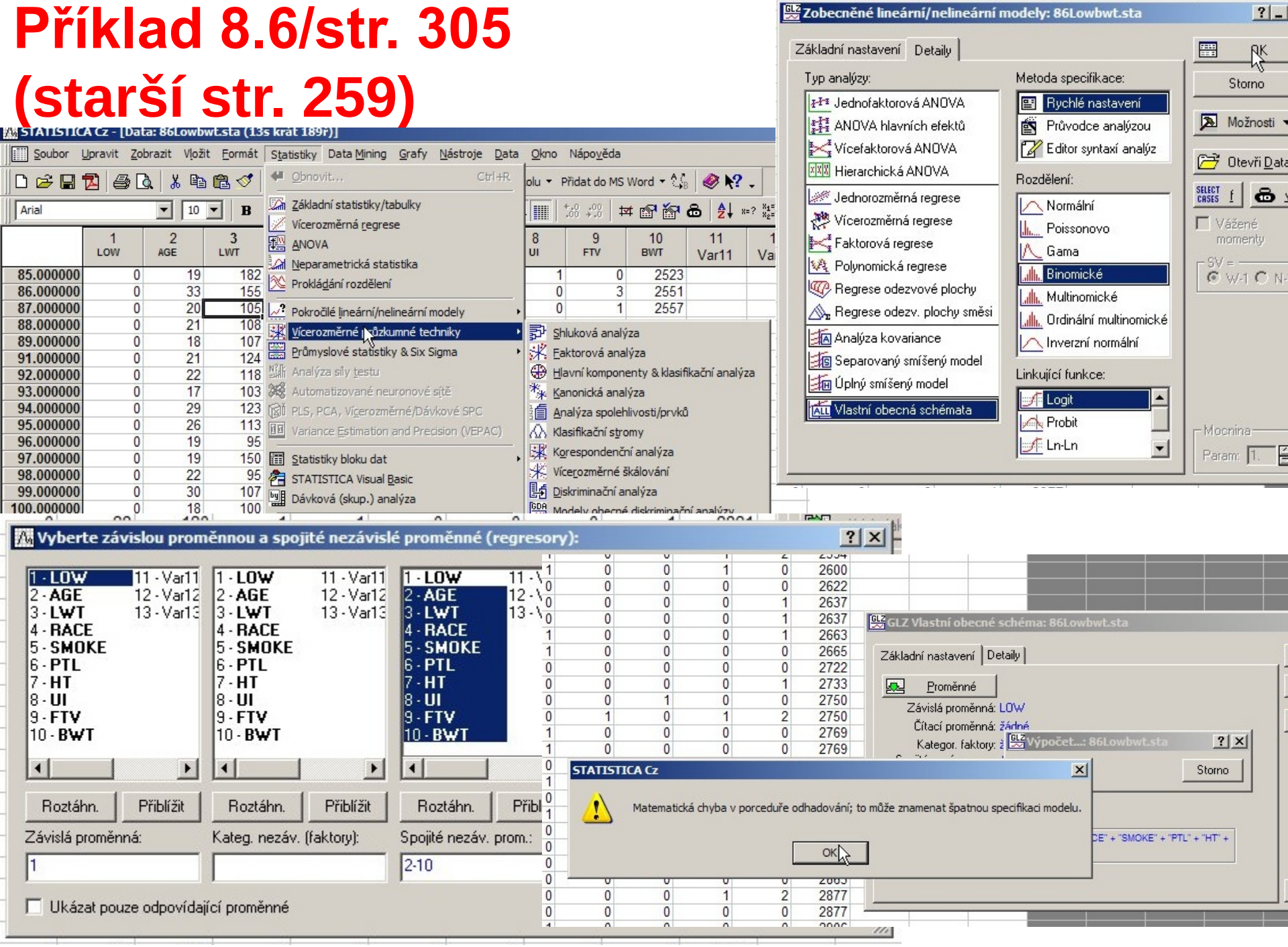


Obr. 8.10d ROC graf pro *ME* = 2.



Obr. 8.10e Graf obou křivek podílu správně zařazených objektů proti P_C (cutoff) určuje hodnotu prahového dělícího bodu P_C.

Příklad 8.6/str. 305 (starší str. 259)



Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - LOW 11 - Var11	1 - LOW 11 - Var11	1 - LOW 11 - Var11	1 - LOW 11 - Var11
2 - AGE 12 - Var12	2 - AGE 12 - Var12	2 - AGE 12 - Var12	2 - AGE 12 - Var12
3 - LWT 13 - Var13	3 - LWT 13 - Var13	3 - LWT 13 - Var13	3 - LWT 13 - Var13
4 - RACE	4 - RACE	4 - RACE	4 - RACE
5 - SMOKE	5 - SMOKE	5 - SMOKE	5 - SMOKE
6 - PTL	6 - PTL	6 - PTL	6 - PTL
7 - HT	7 - HT	7 - HT	7 - HT
8 - UI	8 - UI	8 - UI	8 - UI
9 - FTV	9 - FTV	9 - FTV	9 - FTV
10 - BWT	10 - BWT	10 - BWT	10 - BWT
Roztáhn. Přiblížit	Roztáhn. Přiblížit	Roztáhn. Přiblížit	Roztáhn. Přiblížit
Závislá proměnná:	Kateg. nezáv. (faktory):	Spojité nezáv. prom.:	Čítací proměnná
1		2-4	

OK

Storno

[Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

☐ Ukázat pouze odpovídající proměnné

LOW - Odhady parametrů (86Lowbwt.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že LOW = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-1.02012	1.133611	0.809792	0.368182
AGE		2	0.03411	0.032814	1.080404	0.298607
LWT		3	0.01155	0.006234	3.433950	0.063869
RACE		4	-0.23393	0.176638	1.753826	0.185396
Měřítko			1.00000	0.000000		

LOW - Historie iterací (86Lowbwt.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že LOW = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3
Abs.člen		1	0.000	-0.588	-0.981	-1.020
AGE		2	0.000	0.027	0.034	0.034
LWT		3	0.000	0.008	0.011	0.012
RACE		4	0.000	-0.206	-0.233	-0.234
Věrohodnost			-131.005	-112.991	-112.686	-112.685

LOW - Věrohodn. test typ 1 (86Lowbwt.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že LOW = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-117.336		
AGE	1	-115.956	2.760038	0.096646
LWT	1	-113.562	4.788570	0.028649
RACE	1	-112.685	1.754170	0.185353

LOW - Věrohodn. test typ 3 (86Lowbwt.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že LOW = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
AGE	1	-113.238	1.107180	0.292696
LWT	1	-114.604	3.839129	0.050070
RACE	1	-113.562	1.754170	0.185353

Závěr: V logistickém regresním modelu byly nalezeny pouze dva statisticky významné parametry ***AGE*** a ***BWT.***

Příklad 8.5/str. 298 (starší 253) Data ICU.sta

STATISTICA Cz - [Data: 85ICU.sta (22s krát 200ř)]

Soubor Upravit Zobrazit Vložit Formát Statistiky Data Mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS Word

Arial 9

	1 STA	2 AGE	3 SEX	4 RACE	5 SER	6 CAN	7 CRN	8 INF	9 CPR	10 SYS	11 HRA	12 PRE	13 TYP	14 FRA	15 PO2	16 PH	17 PCO	18 BIC	19 CRE	20 LOC
8.000000	0	27	1	1	0	0	0	1	0	142	88	0	1	0	0	0	0	0	0	0
12.000000	0	59	0	1	0	0	0	0	0	112	80	1	1	0	0	0	0	0	0	0
14.000000	0	77	0	1	1	0	0	0	0	100	70	0	0	0	0	0	0	0	0	0
28.000000	0	54	0	1	0	0	0	1	0	142	103	0	1	1	0	0	0	0	0	0
32.000000	0	87	1	1	1	0	0	1	0	110	154	1	1	0	0	0	0	0	0	0
38.000000	0	69	0	1	0	0	0	1	0	110	132	0	1	0	1	0	0	1	0	0
40.000000	0	63	0	1	1	0	0	0	0	104	66	0	0	0	0	0	0	0	0	0
41.000000	0	30	1	1	0	0	0	0	0	144	110	0	1	0	0	0	0	0	0	0
42.000000	0	35	0	2	0	0	0	0	0	108	60	0	1	0	0	0	0	0	0	0
50.000000	0	70	1	1	1	1	0	0	0	138	103	0	0	0	0	0	0	0	0	0
51.000000	0	55	1	1	1	0	0	1	0	188	86	1	0	0	0	0	0	0	0	0
53.000000	0	48	0	2	1	1	0	0	0	162	100	0	0	0	0	0	0	0	0	0
58.000000	0	66	1	1	1	0	0	0	0	160	80	1	0	0	0	0	0	0	0	0
61.000000	0	61	1	1	0	0	1	0	0	174	99	0	1	0	0	1	0	1	1	0
73.000000	0	66	0	1	0	0	0	0	0	206	90	0	1	0	0	0	0	0	1	0
75.000000	0	52	0	1	1	0	0	1	0	150	71	1	0	0	0	0	0	0	0	0
82.000000	0	55	0	1	1	0														
84.000000	0	59	0	1	0	0														
92.000000	0	63	0	1	0	0														
96.000000	0	72	0	1	1	0														
98.000000	0	60	0	1	0	0														
100.000000	0	78	0	1	1	0														
102.000000	0	16	1	1	0	0														
111.000000	0	62	0	1	1	0														
112.000000	0	61	0	1	0	0														
136.000000	0	35	0	1	0	0														
137.000000	0	74	1	1	1	0														
143.000000	0	68	0	1	1	0														
153.000000	0	69	1	1	1	0														
170.000000	0	51	0	1	0	0														
173.000000	0	55	0	1	1	0														
180.000000	0	64	1	3	1	0														
184.000000	0	88	1	1	1	0														
186.000000	0	23	1	1	1	0														
187.000000	0	73	1	1	1	1														
190.000000	0	53	0	3	1	0														
191.000000	0	74	0	1	1	0														
207.000000	0	68	0	1	1	0														
211.000000	0	66	1	1	0	0														
214.000000	0	60	0	1	1	1														
219.000000	0	64	0	1	1	0														

Zobecněné lineární/nelineární modely: 85ICU.sta

Základní nastavení | Detaily

Typ analýzy:
- Jednofaktorová ANOVA
- ANOVA hlavních efektů
- Vícefaktorová ANOVA
- Hierarchická ANOVA
- Jednorozměrná regrese
- Vícerozměrná regrese
- Faktorová regrese
- Polynomická regrese
- Regrese odezvové plochy
- Regrese odezv. plochy směsi
- Analýza kovariance
- Separovaný smíšený model
- Úplný smíšený model
- Vlastní obecná schémata

Metoda specifikace:
- Rychlé nastavení
- Průvodce analýzou
- Editor syntaxí analýz

Rozdělení:
- Normální
- Poissonovo
- Gama
- Binomické
- Multinomické
- Ordinální multinomické
- Inverzní normální

Linkující funkce:
- Logit
- Probit
- Ln-Ln

OK | Storno | Možnosti | Otevři Data | SELECT CASES | Vážené momenty | SV = W-1 N-1 | Mocnina | Param: 1.

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - STA	11 - HRA	21 - Var21	
2 - AGE	12 - PRE	22 - Var22	
3 - SEX	13 - TYP		
4 - RACE	14 - FRA		
5 - SER	15 - PO2		
6 - CAN	16 - PH		
7 - CRN	17 - PCO		
8 - INF	18 - BIC		
9 - CPR	19 - CRE		
10 - SYS	20 - LOC		

Roztáhn. | Přiblížit

Závislá proměnná: 1

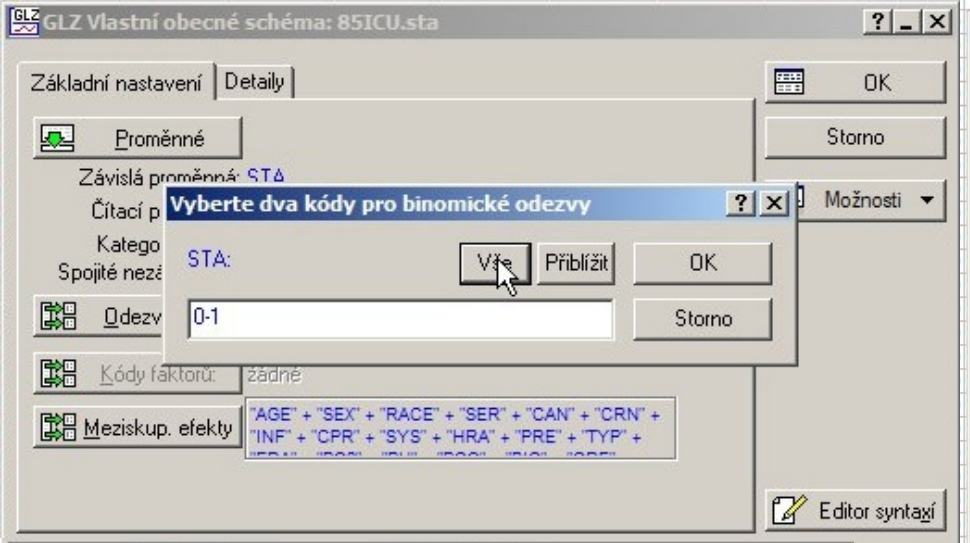
Kateg. nezáv. (faktory):

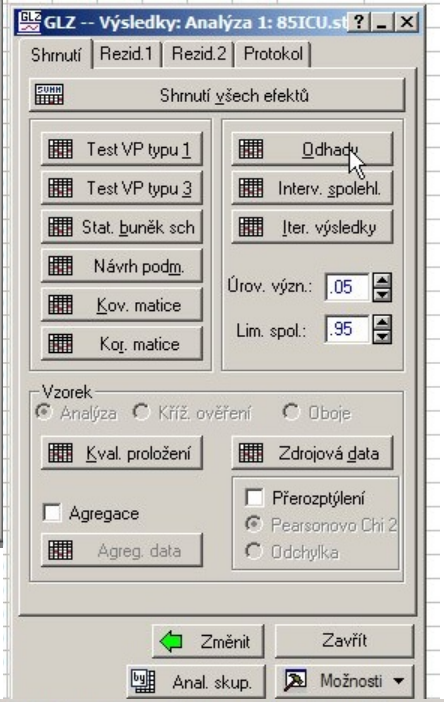
Spojité nezáv. prom.: 2-20

Čítací proměnná

OK | Storno | [Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.





STA - Odhady parametrů (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	5.41380	2.121749	6.51053	0.010724
AGE		2	-0.05243	0.017283	9.20314	0.002416
SEX		3	0.55496	0.501193	1.22605	0.268176
RACE		4	0.00455	0.528255	0.00007	0.993131
SER		5	0.54534	0.575342	0.89844	0.343201
CAN		6	-2.75774	0.980453	7.91140	0.004912
CRN		7	0.10191	0.762363	0.01787	0.893657
INF		8	0.05607	0.534731	0.01099	0.91649
CPR		9	-0.97789	0.983501	0.98861	0.320082
SYS		10	0.01154	0.007854	2.16075	0.141575
HRA		11	0.00364	0.009344	0.15188	0.696741
PRE		12	-0.92928	0.628869	2.18361	0.139487
TYP		13	-2.74427	0.995264	7.60284	0.005828
FRA		14	-1.15208	0.999321	1.32909	0.248968
PO2		15	-0.38667	0.851057	0.20643	0.649583
PH		16	-2.41525	1.231110	3.84884	0.049780
PCO		17	3.17282	1.385979	5.24056	0.022066
BIC		18	0.79384	0.916187	0.75074	0.386241
CRE		19	-0.23331	1.075337	0.04708	0.828233
LOC		20	-2.70609	0.751827	12.95530	0.000319
Měřítko			1.00000	0.000000		

STA - Intervaly spolehlivost odhadů (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Efekt	Úroveň Efekt	Sloupec	Dolní LS 95. %	Horní LS 95. %
Abs.člen		1	1.25525	9.57235
AGE		2	-0.08631	-0.01856
SEX		3	-0.42736	1.53728
RACE		4	-1.03081	1.03991
SER		5	-0.58231	1.67299
CAN		6	-4.67939	-0.83609
CRN		7	-1.39229	1.59612
INF		8	-0.99198	1.10412
CPR		9	-2.90551	0.94974
SYS		10	-0.00385	0.02694
HRA		11	-0.01467	0.02195
PRE		12	-2.16184	0.30328
TYP		13	-4.69495	-0.79359
FRA		14	-3.11071	0.80656
PO2		15	-2.05471	1.28137
PH		16	-4.82818	-0.00232
PCO		17	0.45635	5.88929
BIC		18	-1.00186	2.58953
CRE		19	-2.34094	1.87431
LOC		20	-4.17964	-1.23253

STA - Historie iterací (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5	Iterace 6
Abs.člen		1	0.000	2.3826	3.9102	4.9616	5.3637	5.4131	5.4138
AGE		2	0.000	-0.0183	-0.0353	-0.0475	-0.0520	-0.0524	-0.0524
SEX		3	0.000	0.2156	0.3884	0.5059	0.5500	0.5549	0.5550
RACE		4	0.000	0.0296	0.0286	0.0122	0.0052	0.0046	0.0045
SER		5	0.000	0.1837	0.3593	0.4891	0.5397	0.5453	0.5453
CAN		6	0.000	-1.0249	-1.8844	-2.4880	-2.7283	-2.7574	-2.7577
CRN		7	0.000	-0.0210	0.0268	0.0758	0.0985	0.1019	0.1019
INF		8	0.000	-0.0991	-0.0865	0.0040	0.0511	0.0560	0.0561
CPR		9	0.000	-0.2621	-0.6009	-0.8640	-0.9659	-0.9777	-0.9779
SYS		10	0.000	0.0039	0.0073	0.0101	0.0114	0.0115	0.0115
HRA		11	0.000	0.0018	0.0031	0.0036	0.0036	0.0036	0.0036
PRE		12	0.000	-0.4345	-0.7237	-0.8755	-0.9244	-0.9292	-0.9293
TYP		13	0.000	-0.8794	-1.7325	-2.4110	-2.7051	-2.7437	-2.7443
FRA		14	0.000	-0.2734	-0.6687	-1.0072	-1.1373	-1.1519	-1.1521
PO2		15	0.000	-0.1908	-0.3076	-0.3608	-0.3830	-0.3866	-0.3867
PH		16	0.000	-0.8685	-1.5848	-2.1411	-2.3843	-2.4149	-2.4152
PCO		17	0.000	0.9508	1.9106	2.7379	3.1220	3.1721	3.1728
BIC		18	0.000	0.1664	0.4441	0.6829	0.7819	0.7937	0.7938
CRE		19	0.000	-0.2123	-0.2434	-0.2372	-0.2336	-0.2333	-0.2333
LOC		20	0.000	-1.3865	-1.9862	-2.4599	-2.6779	-2.7057	-2.7061
Věrohodnost			-138.629	-74.4996	-66.6317	-64.9551	-64.8017	-64.7998	-64.7998

STA - Věrohodn. test typ 1 (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-100.080		
AGE	1	-96.153	7.85459	0.005069
SEX	1	-96.153	0.00093	0.975723
RACE	1	-96.127	0.05238	0.818966
SER	1	-92.303	7.64637	0.005689
CAN	1	-92.016	0.57388	0.448721
CRN	1	-90.761	2.51017	0.113114
INF	1	-89.572	2.37940	0.122945
CPR	1	-87.612	3.91940	0.047732
SYS	1	-84.644	5.93551	0.014839
HRA	1	-84.146	0.99670	0.318110
PRE	1	-83.755	0.78093	0.376858
TYP	1	-77.114	13.28350	0.000268
FRA	1	-76.905	0.41663	0.518625
PO2	1	-76.904	0.00162	0.967923
PH	1	-76.889	0.03070	0.860918
PCO	1	-76.195	1.38831	0.238691
BIC	1	-76.155	0.08041	0.776745
CRE	1	-76.005	0.29897	0.584529
LOC	1	-64.800	22.41094	0.000002

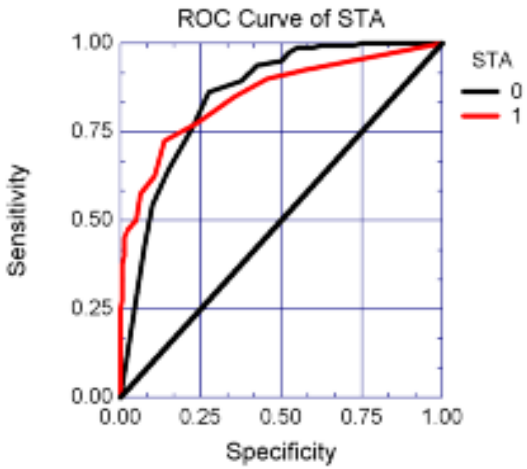
STA - Věrohodn. test typ 3 (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
AGE	1	-70.7259	11.85205	0.000576
SEX	1	-65.4318	1.26401	0.260893
RACE	1	-64.7999	0.00007	0.993129
SER	1	-65.2563	0.91284	0.339361
CAN	1	-68.8308	8.06195	0.004520
CRN	1	-64.8088	0.01797	0.893353
INF	1	-64.8053	0.01100	0.916458
CPR	1	-65.2743	0.94890	0.330000
SYS	1	-65.9547	2.30976	0.128564
HRA	1	-64.8767	0.15367	0.695057
PRE	1	-65.8565	2.11332	0.146022
TYP	1	-69.9842	10.36869	0.001282
FRA	1	-65.4029	1.20605	0.272115
PO2	1	-64.9022	0.20470	0.650954
PH	1	-66.9149	4.23020	0.039711
PCO	1	-68.4207	7.24175	0.007123
BIC	1	-65.1941	0.78855	0.374540
CRE	1	-64.8231	0.04652	0.829227
LOC	1	-76.0053	22.41094	0.000002

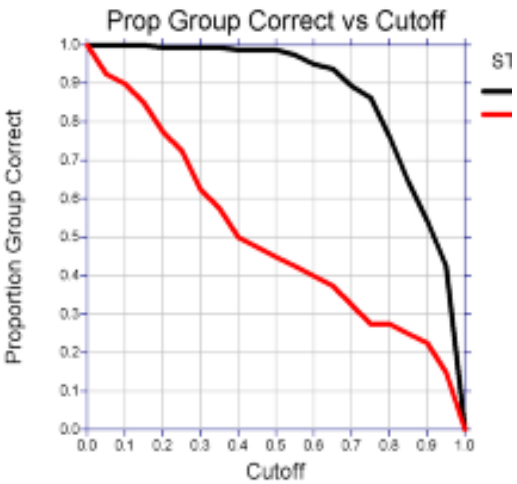
STA - Statistiky kvality proložení (85ICU.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STA = 0

Stat.	SV	Stat.	Stat/sv
Odchylka	180	129.5997	0.719998
Deviance v měřítku	180	129.5997	0.719998
Pearsonovo Chi2	180	214.4866	1.191592
Scaled P. Chi2	180	214.4866	1.191592
Log-věrohodnost		-64.7998	

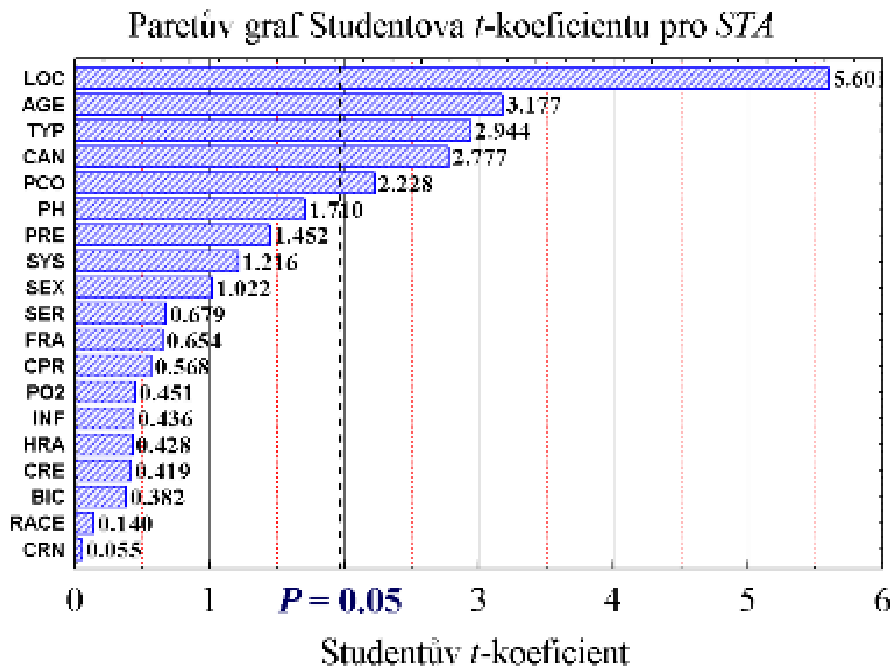
Závěr: V logistickém regresním modelu jsou z původních 18 znaků nezávisle proměnných statisticky významných 6 znaků (pro proměnnou ***STA***=0), a to ***AGE, CAN, TYP, PH, PCO*** **a** ***LOC.***



Obr. 8.8a ROC graf pro *STA* = 0 a 1.



Obr. 8.8b Graf obou křivek podílu správně zařazených objektů proti P_C (cutoff) určuje hodnotu prahového dělícího bodu P_C.



Obr. 8.7 Paretův diagram vyšetření statistické významnosti parametrů logistického regresního modelu dat *ICU* na základě hodnot Studentova *t*-koeficientu.

Příklad 8.4/str. 293 (249)

Data Phyrynx.sta

STATISTICA Cz - [Data: 84Phyrynx.sta (14s krát 195ř)]

Soubor Upravit Zobrazit Vložit Formát Statistiky Data Mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS Word

Arial 10

	1 INST	2 SEX	3 TX	4 GRADE	5 AGE	6 COND	7 SITE	8 TSTAGE	9 NSTAGE	10 ENTRYDT	11 STATUS	12 TIME
1.000000	2	2	1	1	51	1	2	3	1	2468	1	631
2.000000	2	1	2	1	65	1	4	2	3	2968	1	270
3.000000	2	1	1	2	64	2	1	3	3	3368	1	327
4.000000	2	1	1	1	73	1	1	4	0	5768	1	243
5.000000	5	1	2	2	64	1	1	4	3	9568	1	916
6.000000	4	1	2	1	61	1	2	3	0	10668	0	1823
7.000000	4	1	1	2	65	1	2	4	3	10768	1	637
8.000000	4	1	2	3	84	1	4	1	3	12068	1	235
9.000000	6	1	1	2	54	2	1	3	3	13368	1	255
10.000000	3	1	1	2	72	2	4	2	2	15468	1	184
11.000000	3	1	1	2	42	1	4	2	2	15468	1	1064
12.000000	2	1	1	2	61	1	1	4	3	18268	1	414
13.000000	3	1	2	1	71	1	2	3	1	18468	1	216
14.000000	4	1	2	2	83	3	4	3	1	19068	1	324
15.000000	2	1	1	3	43	1	2	4	3	20768	1	480
16.000000	5	1	2	2	52	1	4	4	3	21768	1	245
17.000000	4	2	1	3	68	1	4	2	3	22768	0	1565
18.000000	6	1	2	2	69	1	1	3	0	23368	1	560
19.000000	3	2	2	3	65	3	1	3	0	25968	1	376
20.000000	5	1	1	2	58	1	2					
21.000000	2	1	2	2	63	1	2					
22.000000	4	1	1	2	59	3	2					
23.000000	3	1	1	1	75	1	2					
24.000000	6	1	1	1	65	2	1					
25.000000	4	1	2	3	41	1	2					
26.000000	3	1	1	2	60	1	4					
27.000000	3	1	2	2	72	1	4					
28.000000	5	1	2	2	51	1	1					
29.000000	2	2	1	2	72	2	2					
30.000000	6	1	1	2	49	1	4					
31.000000	3	1	2	2	82	3	1					
32.000000	2	2	1	2	64	1	1					
33.000000	4	2	2	2	57	2	2					
34.000000	3	1	2	1	67	2	2					
35.000000	6	1	2	2	65	2	1					
36.000000	3	1	2	2	62	1	4					
37.000000	2	1	1	2	49	1	4					
38.000000	5	1	1	2	60	1	4					
39.000000	3	1	1	2	75	2	2					
40.000000	2	1	2	2	54	1	2					
41.000000	3	1	2	2	59	1	4					
42.000000	5	1	1	2	58	1	1					
43.000000	3	1	2	2	50	1	1					
44.000000	2	1	1	1	60	1	1					

Základní nastavení | Detaily

Typ analýzy:
- Jednofaktorová ANOVA
- ANOVA hlavních efektů
- Vícefaktorová ANOVA
- Hierarchická ANOVA
- Jednorozměrná regrese
- Vícerozměrná regrese
- Faktorová regrese
- Polynomická regrese
- Regrese odezvové plochy
- Regrese odezv. plochy směsi
- Analýza kovariance
- Separovaný smíšený model
- Úplný smíšený model
- Vlastní obecná schémata

Metoda specifikace:
- Rychlé nastavení
- Průvodce analýzou
- Editor syntaxí analýz

Rozdělení:
- Normální
- Poissonovo
- Gama
- Binomické
- Multinomické
- Ordinální multinomické
- Inverzní normální

Linkující funkce:
- Logit
- Probit
- Ln-Ln

OK | Storno | Možnosti | Otevři Data | SELECT CASES f | v

Vážené momenty

SV = W-1 N-1

Mocnina

Param: 1.

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

11 - STATUS
12 - TIME
13 - Var13
14 - Var14

1 - INST
2 - SEX
3 - TX
4 - GRADE
5 - AGE
6 - COND
7 - SITE
8 - TSTAGE
9 - NSTAGE
10 - ENT...
11 - ST
12 - TI
13 - Var
14 - Var

Roztáhn. | Přiblížit

OK | Storno | [Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

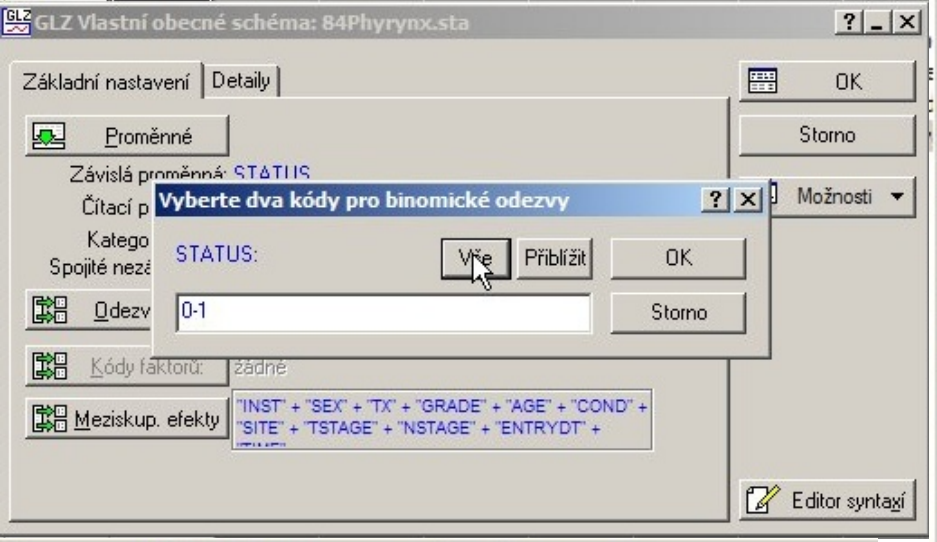
Závislá proměnná: 11

Kateg. nezáv. (faktory):

Spojité nezáv. prom.: 1-10 12

Čítací proměnná

Ukázat pouze odpovídající proměnné



STATUS - Intervaly spolehlivost odhadů (84Ph
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že STATUS = (

Efekt	Úroveň Efekt	Sloupec	Dolní LS 95. %	Horní LS 95. %
Abs.člen		1	-6.67030	2.167986
INST		2	-0.42413	0.146853
SEX		3	-1.29115	0.871691
TX		4	-0.27207	1.650922
GRADE		5	-0.57788	0.729513
AGE		6	-0.09692	-0.001146
COND		7	-0.34356	0.751535
SITE		8	-0.44104	0.348292
TSTAGE		9	-0.63648	0.530831
NSTAGE		10	-0.39876	0.442121
ENTRYDT		11	-0.00003	0.000058
TIME		12	0.00375	0.007113

STATUS - Odhady parametrů (84Phyrynx.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STATUS = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-2.25115	2.254705	0.99685	0.318073
INST		2	-0.13864	0.145663	0.90591	0.341201
SEX		3	-0.20973	0.551756	0.14449	0.703860
TX		4	0.68942	0.490569	1.97503	0.159915
GRADE		5	0.07582	0.333525	0.05167	0.820178
AGE		6	-0.04903	0.024433	4.02744	0.044766
COND		7	0.20399	0.279367	0.53316	0.465283
SITE		8	-0.04637	0.201364	0.05304	0.817859
TSTAGE		9	-0.05283	0.297790	0.03147	0.859200
NSTAGE		10	0.02168	0.214514	0.01022	0.919492
ENTRYDT		11	0.00001	0.000023	0.27691	0.598735
TIME		12	0.00543	0.000857	40.15312	0.000000
Měřítko			1.00000	0.000000		

STATUS - Historie iterací (84Phyrynx.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STATUS = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000	-1.5967	-2.0630	-2.2273	-2.2508	-2.2512
INST		2	0.000	-0.0726	-0.1158	-0.1346	-0.1385	-0.1386
SEX		3	0.000	-0.1138	-0.1910	-0.2109	-0.2099	-0.2097
TX		4	0.000	0.2463	0.4784	0.6420	0.6871	0.6894
GRADE		5	0.000	0.0621	0.0947	0.0867	0.0766	0.0758
AGE		6	0.000	-0.0212	-0.0365	-0.0463	-0.0489	-0.0490
COND		7	0.000	0.0887	0.1442	0.1888	0.2032	0.2040
SITE		8	0.000	0.0185	-0.0090	-0.0379	-0.0460	-0.0464
TSTAGE		9	0.000	-0.0185	-0.0329	-0.0475	-0.0526	-0.0528
NSTAGE		10	0.000	0.0122	0.0129	0.0189	0.0215	0.0217
ENTRYDT		11	0.000	0.0000	0.0000	0.0000	0.0000	0.0000
TIME		12	0.000	0.0029	0.0044	0.0052	0.0054	0.0054
Věrohodnost			-135.164	-71.8323	-64.7773	-63.8640	-63.8324	-63.8324

STATUS - Věrohodn. test typ 1 (84Phyrynx.sta
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že STATUS = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen		-114.082		
INST	1	-113.990	0.18345	0.668424
SEX	1	-113.831	0.31951	0.571905
TX	1	-113.827	0.00633	0.936582
GRADE	1	-113.636	0.38323	0.535877
AGE	1	-112.342	2.58697	0.107746
COND	1	-108.877	6.92954	0.008478
SITE	1	-108.700	0.35453	0.551557
TSTAGE	1	-107.421	2.55942	0.109639
NSTAGE	1	-106.315	2.21056	0.137069
ENTRYDT	1	-106.283	0.06403	0.800230
TIME	1	-63.832	84.90170	0.000000

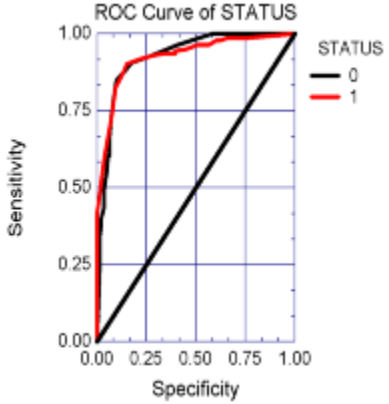
STATUS - Věrohodn. test typ 3 (84Phyrynx.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LO
Modelovaná pravděpodobnost, že STATUS = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
INST	1	-64.292	0.92014	0.337439
SEX	1	-63.906	0.14633	0.702067
TX	1	-64.862	2.05945	0.151264
GRADE	1	-63.858	0.05071	0.821834
AGE	1	-65.951	4.23749	0.039541
COND	1	-64.051	0.43825	0.507970
SITE	1	-63.859	0.05321	0.817576
TSTAGE	1	-63.848	0.03141	0.859330
NSTAGE	1	-63.837	0.01023	0.919429
ENTRYDT	1	-63.972	0.27836	0.597777
TIME	1	-106.283	84.90170	0.000000

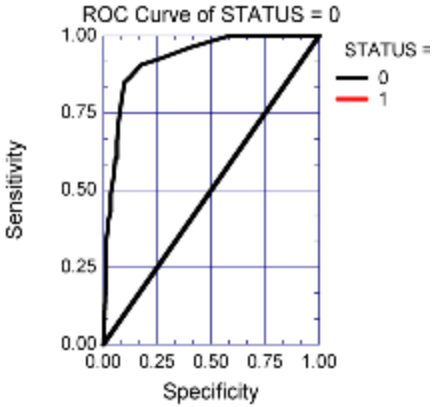
STATUS - Statistiky kvality proložení (84Phyrynx.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že STATUS = 0

Stat.	SV	Stat.	Stat/sv
Odchylka	183	127.6647	0.697621
Deviance v měřítku	183	127.6647	0.697621
Pearsonovo Chi2	183	222.2862	1.214678
Scaled P. Chi2	183	222.2862	1.214678
Log-věrohodnost		-63.8324	

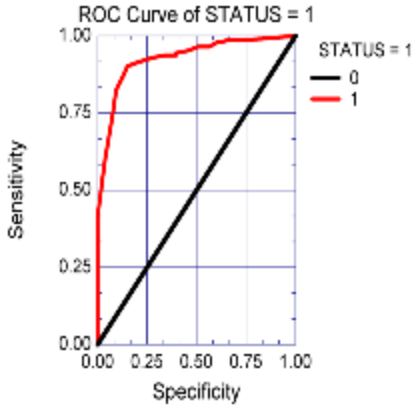
Závěr: V logistickém regresním modelu byly dokázány pouze dva statisticky významné znaky ***AGE*** a ***TIME*,** především ovlivňující přežití pacienta. Ostatní znaky jsou proto z lékařského hlediska prokázány za statisticky nevýznamné.



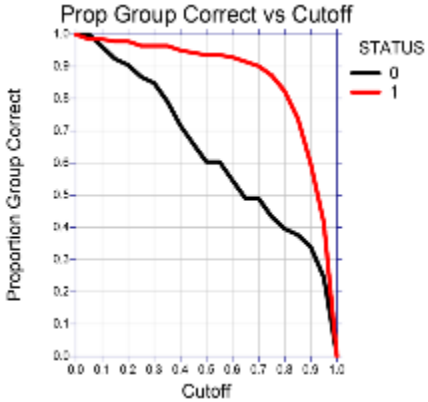
Obr. 8.6a ROC graf pro obě hodnoty *STATUS* = 0 a 1.



Obr. 8.6b ROC graf pro *STATUS* = 0.



Obr. 8.6c ROC graf pro *STATUS* = 1.



Obr. 8.6d Graf obou křivek podílu správně zařazených objektů proti P_C (cutoff) určuje hodnotu prahového dělícího bodu P_C.

Graf prahové operační charakteristiky ROC (obr. 8.6a, b, c) vystihuje správnost diagnostického testu, zda logistickým modelem vypočtené Ano nebo Ne je správné.

Graf podílu správně zařazených objektů proti P_C (obr. 8.6d) je velice užitečný graf k určení nejlepší hodnoty prahového dělícího bodu P_C.

Příklad 8.3/str. 287 (starší 243)

Data Phyrynx.sta

STATISTICA Cz - [Data: 83Koureni.sta (4s krát 93ř)]

Soubor Upravit Zobrazit Vložit Formát Statistiky

Arial 10

	1 Objekt	2 Pulz	3 Kouri	4 Vaha
1	1	1	1	140
2	2	1	1	145
3	3	1	0	160
4	4	1	0	190
5	5	1	1	155
6	6	1	1	165
7	7	0	1	150
8	8	1	1	190
9	9	1	1	195
10	10	1	1	138
11	11	0	0	160
12	12	1	1	155
13	13	0	0	153
14	14	1	1	145
15	15	1	1	170
16	16	1	1	175
17	17	1	0	175
18	18	1	0	170
19	19	1	0	180
20	20	1	1	135
21	21	1	1	170
22	22	1	1	157
23	23	1	1	130
24	24	1	0	185
25	25	0	1	140
26	26	1	1	120
27	27	1	0	130
28	28	0	1	138
29	29	0	0	121
30	30	1	1	125

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - Objekt	1 - Objekt	1 - Objekt	1 - Objekt
2 - Pulz	2 - Pulz	2 - Pulz	2 - Pulz
3 - Kouri	3 - Kouri	3 - Kouri	3 - Kouri
4 - Vaha	4 - Vaha	4 - Vaha	4 - Vaha
Roztáhn. Přiblížit	Roztáhn. Přiblížit	Roztáhn. Přiblížit	Roztáhn. Přiblížit
Závislá proměnná:	Kateg. nezáv. (faktory):	Spojité nezáv. prom.:	Čítací proměnná
2		3-4	

OK

Storno

[Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

☐ Ukázat pouze odpovídající proměnné

GLZ Vlastní obecné schéma: 83Koureni.sta

Základní nastavení | Detaily

OK

Storno

Proměnné

Závislá proměnná: Pulz

Čítací p

Kategorické

Spojité nezá

Možnosti

Odezv

Vyberte dva kódy pro binomické odezvy

Pulz: Vše Přiblížit OK

0-1

Storno

Kódy faktorů: žádné

Meziskup. efekty: Kouri + Vaha

Editor syntaxí

Pulz - Odhady parametrů (83Koureni.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že Pulz = 0

Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	2.94291	1.885121	2.437119	0.118494
Kouri		2	-1.24740	0.551060	5.124017	0.023597
Vaha		3	-0.02323	0.012404	3.505869	0.061152
Měřítko			1.00000	0.000000		

Pulz - Intervaly spolehlivost odhadů (83Koureni.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že Pulz = 0

Efekt	Úroveň Efekt	Sloupec	Dolní LS 95. %	Horní LS 95. %
Abs.člen		1	-0.75186	6.637684
Kouri		2	-2.32745	-0.167338
Vaha		3	-0.04754	0.001086

Pulz - Historie iterací (83Koureni.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že Pulz = 0

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen		1	0.0000	1.6431	2.7384	2.9383	2.9429
Kouri		2	0.0000	-0.8885	-1.2004	-1.2464	-1.2474
Vaha		3	0.0000	-0.0143	-0.0218	-0.0232	-0.0232
Věrohodnost			-64.4627	-47.7906	-47.1722	-47.1611	-47.1611

Pulz - Věrohodn. test typ 1 (83Koureni.sta)
Rozdělení : BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že Pulz = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-50.8785		
Kouri	1	-49.0903	3.576471	0.058604
Vaha	1	-47.1611	3.858456	0.049496

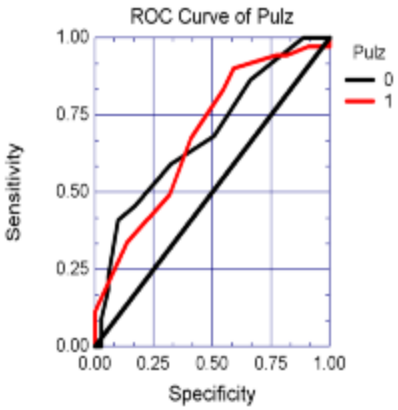
Pulz - Věrohodn. test typ 3 (83Koureni.sta)
Rozdělení : BINOMICKÉ, Linkující funkce:
Modelovaná pravděpodobnost, že Pulz = 0

Efekt	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Kouri	1	-49.7638	5.205374	0.022517
Vaha	1	-49.0903	3.858456	0.049496

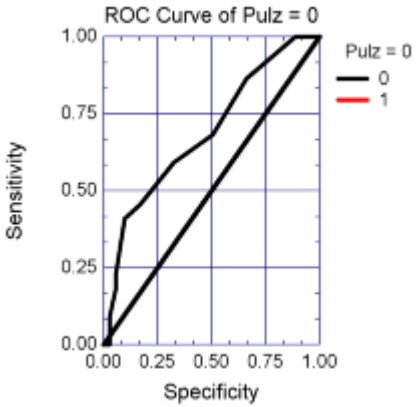
Pulz - Statistiky kvality proložení (83Koureni.st
Rozdělení : BINOMICKÉ, Linkující funkce: LOG
Modelovaná pravděpodobnost, že Pulz = 0

Stat.	SV	Stat.	Stat/sv
Odchylka	90	94.3222	1.048024
Deviance v měřítku	90	94.3222	1.048024
Pearsonovo Chi2	90	90.0931	1.001034
Scaled P. Chi2	90	90.0931	1.001034
Log-věrohodnost		-47.1611	

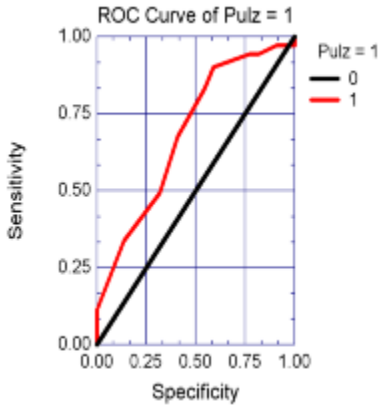
Závěr: Na obou grafech (obrázky 8.4a, b) lze identifikovat, že objekty 31 a 66 se týkají vysokého pulzu u nekuřáků, kteří mají hmotnost 58 a 68 kg, což je méně než průměr ostatních. Můžeme proto oba pacienty navrhnout k vyšetření, proč neodpovídají navrženému logistickému regresnímu modelu.



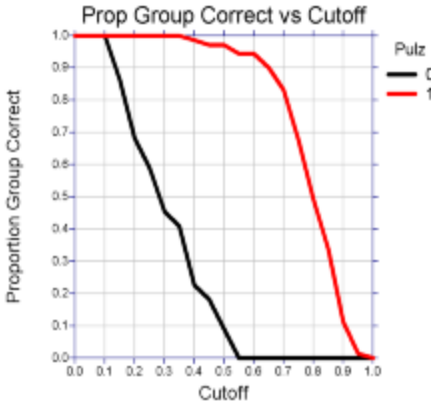
Obr. 8.5a ROC graf pro obě hodnoty *PULZ* = 0 a 1.



Obr. 8.5b ROC graf pro *PULZ* = 0.



Obr. 8.5c ROC graf pro *PULZ* = 1.



Obr. 8.5d Graf obou křivek podílu správně zařazených objektů proti P_C (cutoff) určuje hodnotu prahového dělícího bodu P_C.

Graf prahové operační charakteristiky ROC (obr. 8.5a, b, c)vystihuje správnost diagnostického testu, zda logistickým modelem vypočtené Ano nebo Ne je správné. Kvalita nalezeného logistického modelu se posuzuje také dle *plochy pod ROC křivkou AUC*. Čím více se *AUC* blíží jedné nebo 100%, tím je klasifikace objektů lepší.

Graf podílu správně zařazených objektů proti P_C (obr. 8.5d) je velice užitečný k určení nejlepší hodnoty prahového dělícího bodu P_C. Na *y*-nové ose se zobrazuje podíl (procento) správně zařazených objektů a na *x*-ové ose hodnoty pravděpodobnosti dělícího bodu P_C v jednotkách vyčíslované pravděpodobnosti.

Úloha B35 *Interakce léku s pilulkami pro orální antikoncepci* (str. 509)

Vyšetření vlivu dvou druhů terapie, a to vlivu dvou časových period a vlivu dvou pořadí zařazení nového léku D s placebem na interakci léků s pilukami orální antikoncepce. U 22 žen byla náhodně aplikována jedna ze dvou pilulek ve dvou překrývajících se obdobích. Pilulky obsahovaly jednak nový lék D a jednak placebo a byly podávány současně s pilulkami orální antikoncepce ve dvou sledovaných časových periodách. Orální antikoncepce se skládala ze dvou komponent, ethinyl estradiol EE a norethindron NET. Vyšetřete, zda závisle proměnná **TREATM**, tj. přítomnost léku D x4 ovlivnila míru absorpce substance do organismu u každé z obou komponent antikoncepce pro vyjádření vztahu mezi závisle proměnnou léčení placebem nebo lékem D *x*4 a sedmi znaky. Testujte statistickou významnost všech sedmi znaků a zjistěte, který znak podstatně ovlivňuje absorpci léku do organismu.

◯ *Data:* Výběr OCDRUG se týká 44 pozorování v řádcích 8 znaků ve sloupcích:

i (ID) identifikační číslo vyšetřované ženy,

***x*1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo placeba a léku D kódem 2,**

*x*2 (PERIOD) značí sledovanou periodu 1 nebo 2,

*x*3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),

*x*4 (RELATION1) značí vztah EE - AUC (pg*hr/ml),

*x*5 (RELATION2) značí vztah EE - Cmax (pg/ml),

*x*6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),

*x*7 (RELATION4) značí vztah NET - Cmax (pg/ml).

i	x1	x2	x3	x4	x5	x6	x7
1	2	1	0	2623	279	197525	21900
...	...	...	...	...	...	...	...
22	2	2	1	4333.3	362	253460	27200

Efekt	TREATMENT - Odhady parametrů (B35) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že TREATMENT = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 95, %	Horní LS 95, %	p
Abs.člen		1	0,505652	2,397553	0,044480	-4,19347	5,204770	0,832963
TREATM		2	0,289670	0,773007	0,140424	-1,22540	1,804736	0,707860
PERIOD		3	-0,281128	0,726066	0,149919	-1,70419	1,141935	0,698613
RELATION1		4	-0,001799	0,000958	3,524484	-0,00368	0,000079	0,060469
RELATION2		5	0,007603	0,007113	1,142493	-0,00634	0,021543	0,285126
RELATION3		6	-0,000000	0,000015	0,000279	-0,00003	0,000028	0,986669
RELATION4		7	0,000120	0,000082	2,167761	-0,00004	0,000281	0,140931
Měřítko			1,000000	0,000000		1,00000	1,000000	

*x*1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo zařazení placeba a léku D kódem 2,
*x*2 (PERIOD) značí sledovanou periodu 1 nebo 2,
*x*3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),
*x*4 (**RELATION1**) značí vztah EE - AUC (pg*hr/ml),
*x*5 (RELATION2) značí vztah EE - Cmax (pg/ml),
*x*6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),
*x*7 (**RELATION4**) značí vztah NET - Cmax (pg/ml).

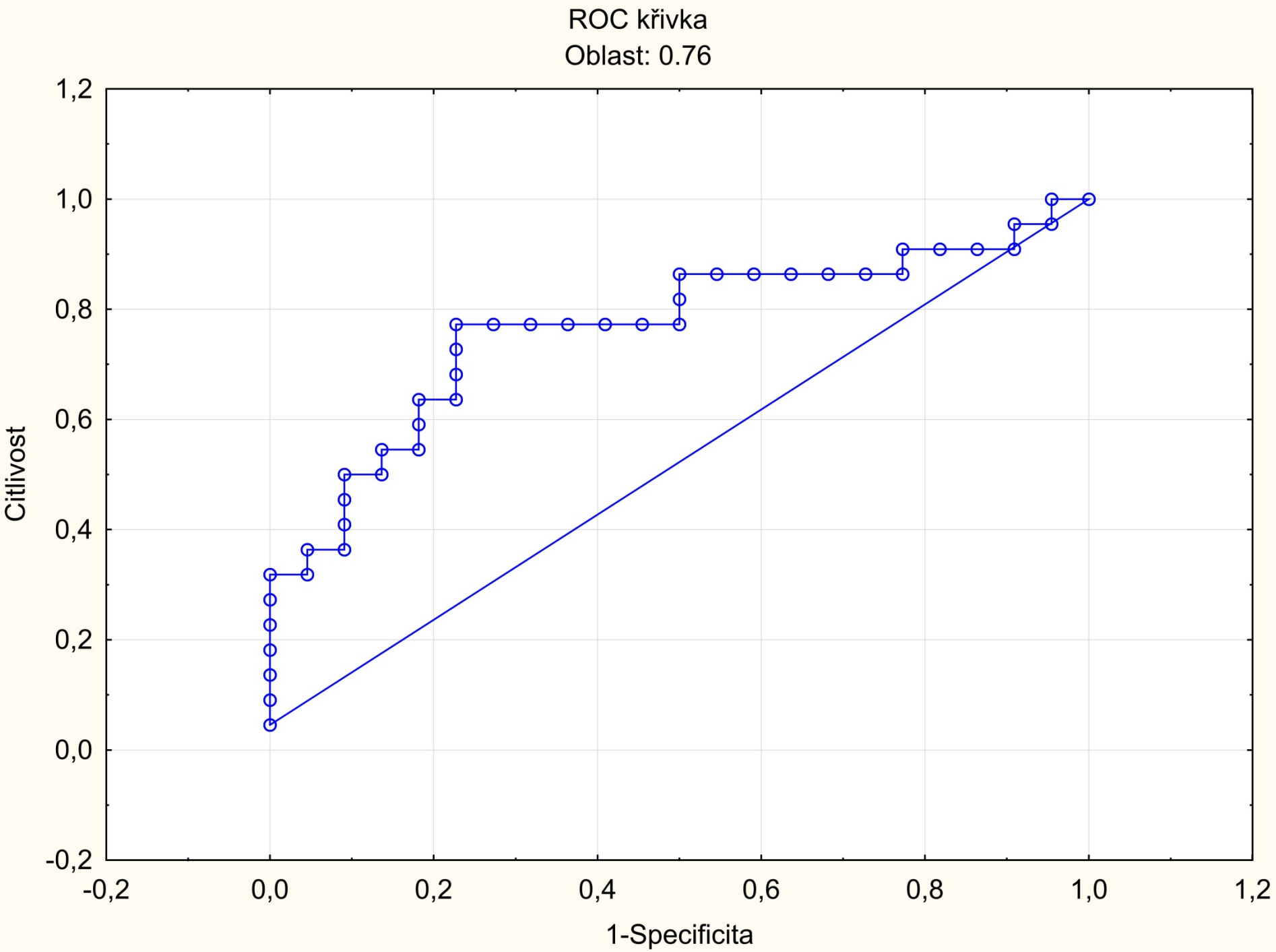
Efekt	TREATMENT - Intervaly spolehlivost odhadů (B35) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že TREATMENT = 0			
	Úroveň Efekt	Sloupec	Dolní LS 95, %	Horní LS 95, %
Abs.člen		1	-4,19347	5,204770
TREATM		2	-1,22540	1,804736
PERIOD		3	-1,70419	1,141935
RELATION1		4	-0,00368	0,000079
RELATION2		5	-0,00634	0,021543
RELATION3		6	-0,00003	0,000028
RELATION4		7	-0,00004	0,000281

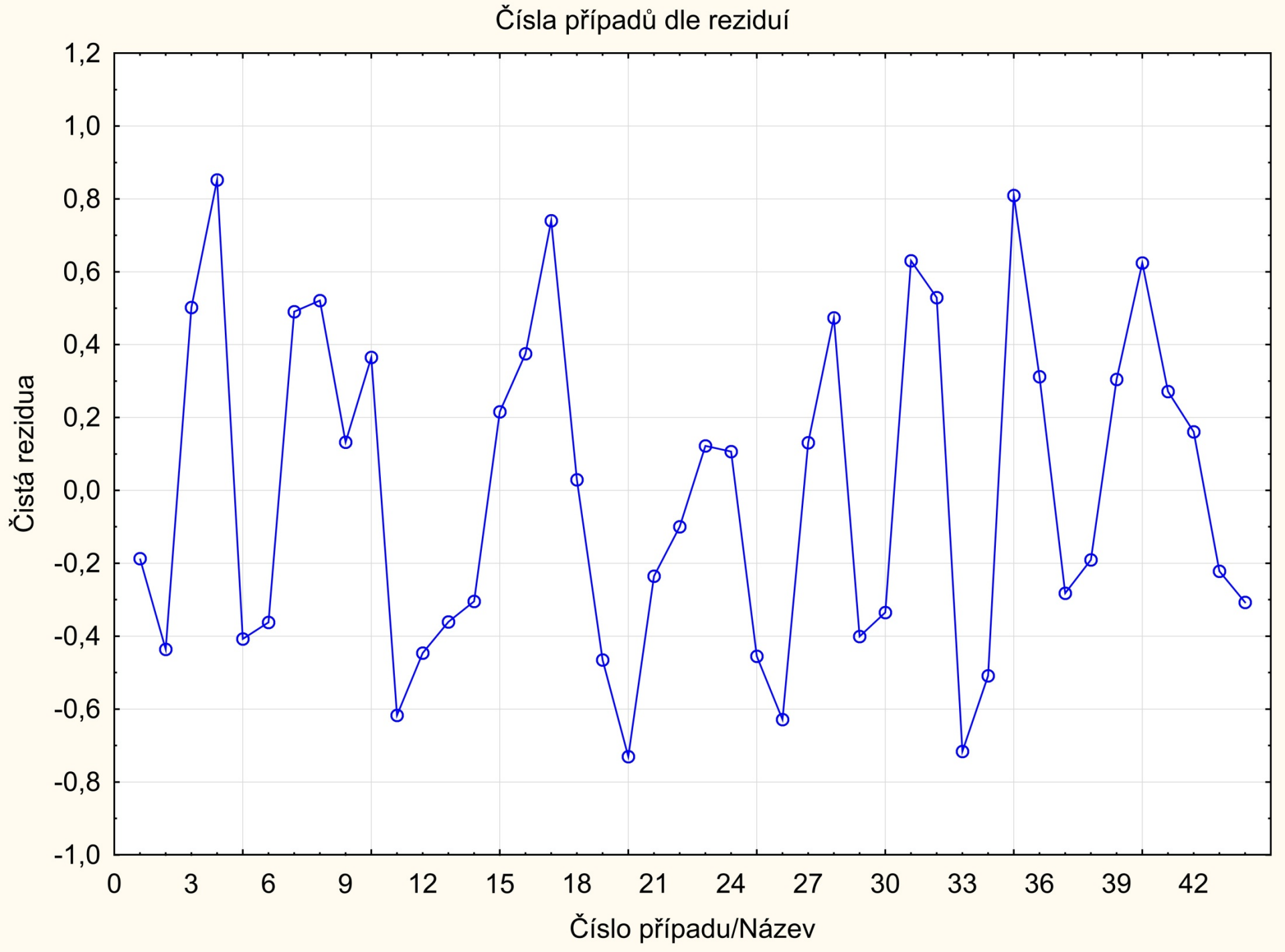
Efekt	TREATMENT - Historie iterací (B35) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že TREATMENT = 0						
	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen		1	0,0000	0,3967	0,5107	0,5063	0,5057
TREATM		2	0,0000	0,2875	0,2900	0,2896	0,2897
PERIOD		3	0,0000	-0,2652	-0,2852	-0,2813	-0,2811
RELATION1		4	0,0000	-0,0016	-0,0018	-0,0018	-0,0018
RELATION2		5	0,0000	0,0073	0,0077	0,0076	0,0076
RELATION3		6	0,0000	0,0000	0,0000	-0,0000	-0,0000
RELATION4		7	0,0000	0,0001	0,0001	0,0001	0,0001
Věrohodnost			-30,4985	-25,4792	-25,2636	-25,2579	-25,2578

*x*1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo zařazení placeba a léku D kódem 2,
*x*2 (PERIOD) značí sledovanou periodu 1 nebo 2,
*x*3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),
*x*4 (**RELATION1**) značí vztah EE - AUC (pg*hr/ml),
*x*5 (RELATION2) značí vztah EE - Cmax (pg/ml),
*x*6 (**RELATION3**) značí vztah NET - AUC (pg*hr/ml),
*x*7 (**RELATION4**) značí vztah NET - Cmax (pg/ml).

Efekt	TREATMENT - Věrohodn. test typ 1 (B35) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že TREATMENT = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-30,4985		
TREATM	1	-30,4985	-0,000000	1,000000
PERIOD	1	-30,4985	0,000000	1,000000
RELATION1	*1*	*-28,5558*	*3,885451*	*0,048706*
RELATION2	1	-27,8564	1,398777	0,236928
RELATION3	*1*	*-26,7754*	*2,161968*	*0,141463*
RELATION4	*1*	*-25,2578*	*3,035066*	*0,081483*

Efekt	TREATMENT - Věrohodn. test typ 3 (B35) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že TREATMENT = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
TREATM	1	-25,3285	0,141279	0,707013
PERIOD	1	-25,3330	0,150395	0,698158
RELATION1	*1*	*-27,2645*	*4,013358*	*0,045141*
RELATION2	1	-25,8499	1,184154	0,276512
RELATION3	1	-25,2580	0,000279	0,986666
RELATION4	*1*	*-26,7754*	*3,035066*	*0,081483*


ROC křivka
Oblast: 0.76
Citlivost
1-Specificita
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,2
0,0
0,2
0,4
0,6
0,8
1,0
1,2


Čísla případů dle reziduí
Čistá rezidua
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
Číslo případu/Název

Úloha B64 *Vliv polymorfizmů na inzulínovou sekreci u žen se syndromem polycystických ovárií a u kontrolních zdravých žen* (str. 537)

Varianty genu TCF7L2 jsou spojovány se sníženou sekrecí inzulínu a s rizikem diabetu 2. typu. Syndrom polycystických ovárií PCOS, který je častou endokrinopatií u žen fertilního věku, je také spojován mimo jiné i s poruchami sekrece inzulínu. Cílem úlohy je zjistit vliv polymorfizmů TCF7L2 genu na inzulínovou sekreci u žen s PCOS a kontrolních zdravých žen pod označením **Soubor** (P a K), kdy vliv rizikového haplotypu CTT (**riziko_hap** =1) genu TCF7L2 je závisle proměnnou a všechny ostatní znaky jsou nezávisle proměnnými.

◯ ***Data:*** bylo vyšetřeno 57 žen:

x1 body mass index (**BMI**),

x2 věk (**AGE**),

x3 rodinná anamnéza diabetu 2. typu (**RA_DM2**). Sledoval se efekt akutní odpovědi inzulínu a C-peptidu a glukagonová odpověď po intravenózním podání argininu nalačno (**AIRf, ACPf, AGRf**) a po glukózové stimulaci na 15 mmol/l (**AIR15, ACP15, AGR15**), dále glukózou stimulovaná sekrece inzulínu a C-peptidu (**slopeAIR, slopeCP**) a glukózou inhibovaná sekrece glukagonu (**slopeGR**). Byl studován vliv rizikového haplotypu x4 CTT (**riziko_hap** =1 jinak 0) genu TCF7L2 na tyto parametry.

Soubor	BMI	AGE	RA_DM2	riziko_hap	ACPf	AIRf
i	x1	x2	x3	x4	x5	x6
K	19.7	24	PK	0	0.1475	26.72688
.....	.....	.....	.....	.....	.....	.....
P	31.83	22	PD	1	1.3875	499.5594
K	42.8	34	PD	1	1.37	580.0988

AGRf	ACP15	AIR15	AGR15	slopeCP	slopeIR	slopeGR
x7	x8	x9	x10	x11	x12	x13
111.525	3.2325	1619.577	41.725	0.386591	199.6053	-8.74687
.....	.....	.....	.....	.....	.....	.....
60.8	5.215	3258.167	74.175	0.476057	343.1105	1.663558

Efekt	riziko_hap - Odhady parametrů (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0							
	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 80, %	Horní LS 80, %	p
Abs.člen		1	332,525	144,9953	5,259445	146,706	518,3436	0,021828
BMI		2	-0,211	0,1649	1,634981	-0,422	0,0005	0,201015
AGE		3	-0,058	0,0893	0,424228	-0,173	0,0563	0,514835
RA_DM2		4	-3,206	1,4131	5,147956	-5,017	-1,3953	0,023274
ACPf		5	-12,142	17,5038	0,481154	-34,573	10,2904	0,487900
AIRf		6	0,028	0,0314	0,770578	-0,013	0,0678	0,380038
AGRf		7	-1,284	0,4794	7,172157	-1,898	-0,6695	0,007404
"ACP15"		8	15,733	15,8276	0,988084	-4,551	36,0168	0,320211
"AIR15"		9	-0,014	0,0234	0,352919	-0,044	0,0161	0,552465
"AGR15"		10	1,407	0,5061	7,724608	0,758	2,0554	0,005447
slopeCP		11	-197,881	152,2972	1,688196	-393,057	-2,7040	0,193839
slopeIR		12	0,171	0,2186	0,615345	-0,109	0,4517	0,432782
slopeGR		13	-11,662	4,3005	7,354421	-17,174	-6,1512	0,006690
Měřítko			1,000	0,0000		1,000	1,0000	

Efekt	riziko_hap - Intervaly spolehlivost odhadů (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0			
	Úroveň Efekt	Sloupec	Dolní LS 80, %	Horní LS 80, %
Abs.člen		1	146,706	518,3436
BMI		2	-0,422	0,0005
AGE		3	-0,173	0,0563
RA_DM2		4	-5,017	-1,3953
ACPf		5	-34,573	10,2904
AIRf		6	-0,013	0,0678
AGRf		7	-1,898	-0,6695
"ACP15"		8	-4,551	36,0168
"AIR15"		9	-0,044	0,0161
"AGR15"		10	0,758	2,0554
slopeCP		11	-393,057	-2,7040
slopeIR		12	-0,109	0,4517
slopeGR		13	-17,174	-6,1512

Efekt	riziko_hap - Historie iterací (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0					
	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3
Abs.člen		1	0,5390	207,0132	271,316	308,219
BMI		2	0,0000	-0,0264	-0,095	-0,167
AGE		3	0,0000	-0,0511	-0,061	-0,059
RA_DM2		4	0,0000	-2,0096	-2,627	-2,976
ACPf		5	0,0000	0,7577	-3,806	-9,355
AIRf		6	0,0000	-0,0011	0,009	0,021
AGRf		7	0,0000	-0,4948	-0,794	-1,106
ACP15		8	0,0000	2,6587	7,504	12,890
AIR15		9	0,0000	-0,0006	-0,005	-0,011
AGR15		10	0,0000	0,5369	0,871	1,212
slopeCP		11	0,0000	-42,7334	-100,281	-163,727
slopeIR		12	0,0000	0,0202	0,073	0,136
slopeGR		13	0,0000	-4,3707	-7,124	-10,019
Věrohodnost			-37,5123	-24,8756	-21,531	-20,382

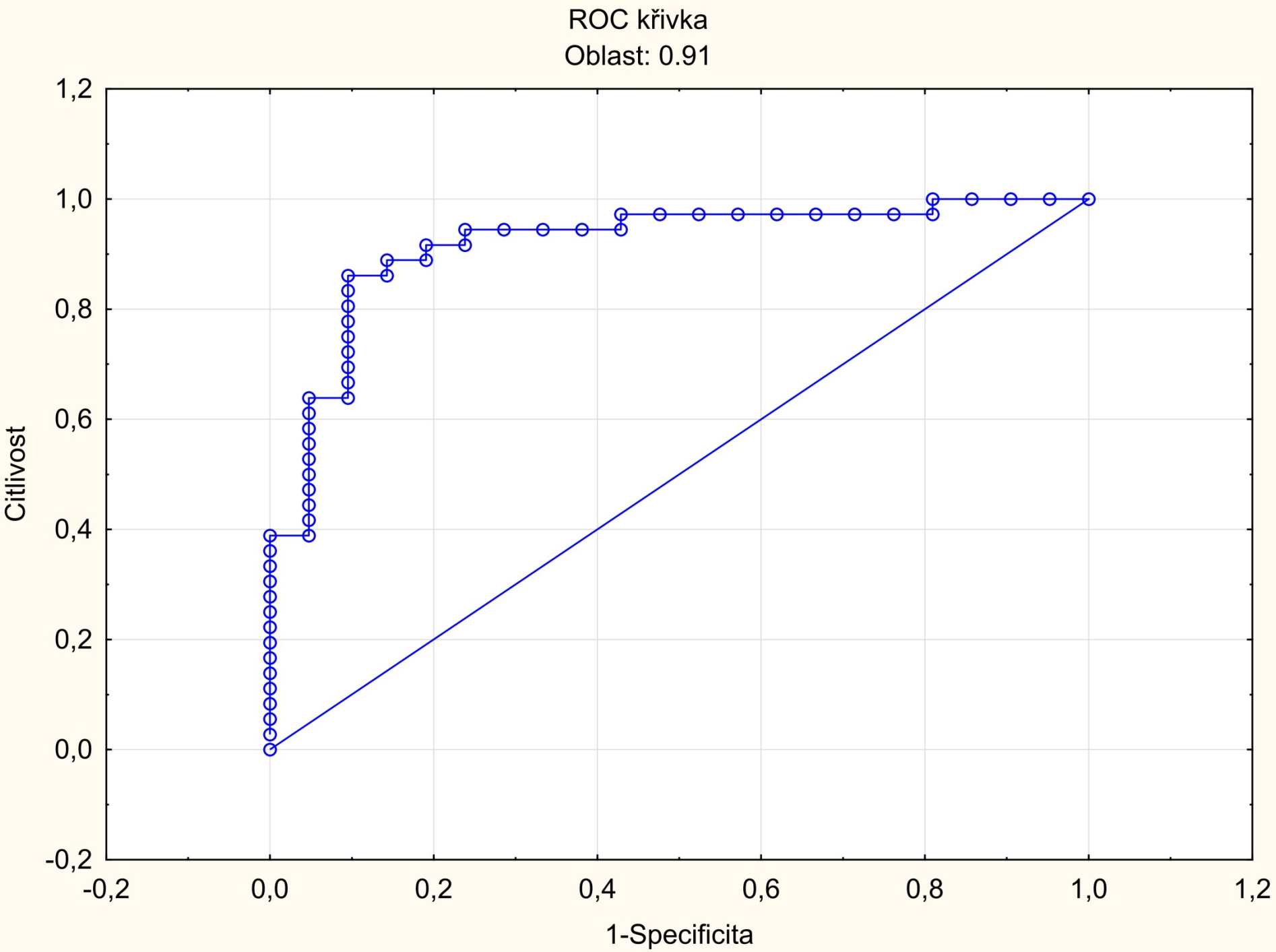
Efekt	riziko_hap - Historie iterací (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0			
	Iterace 4	Iterace 5	Iterace 6	Iterace 7
Abs.člen	328,212	332,404	332,525	332,525
BMI	-0,204	-0,211	-0,211	-0,211
AGE	-0,058	-0,058	-0,058	-0,058
RA_DM2	-3,165	-3,205	-3,206	-3,206
ACPf	-11,773	-12,133	-12,142	-12,142
AIRf	0,027	0,028	0,028	0,028
AGRf	-1,258	-1,283	-1,284	-1,284
ACP15	15,336	15,723	15,733	15,733
AIR15	-0,013	-0,014	-0,014	-0,014
AGR15	1,378	1,406	1,407	1,407
slopeCP	-193,013	-197,762	-197,881	-197,881
slopeIR	0,166	0,171	0,171	0,171
slopeGR	-11,421	-11,657	-11,662	-11,662
Věrohodnost	-20,251	-20,249	-20,249	-20,249

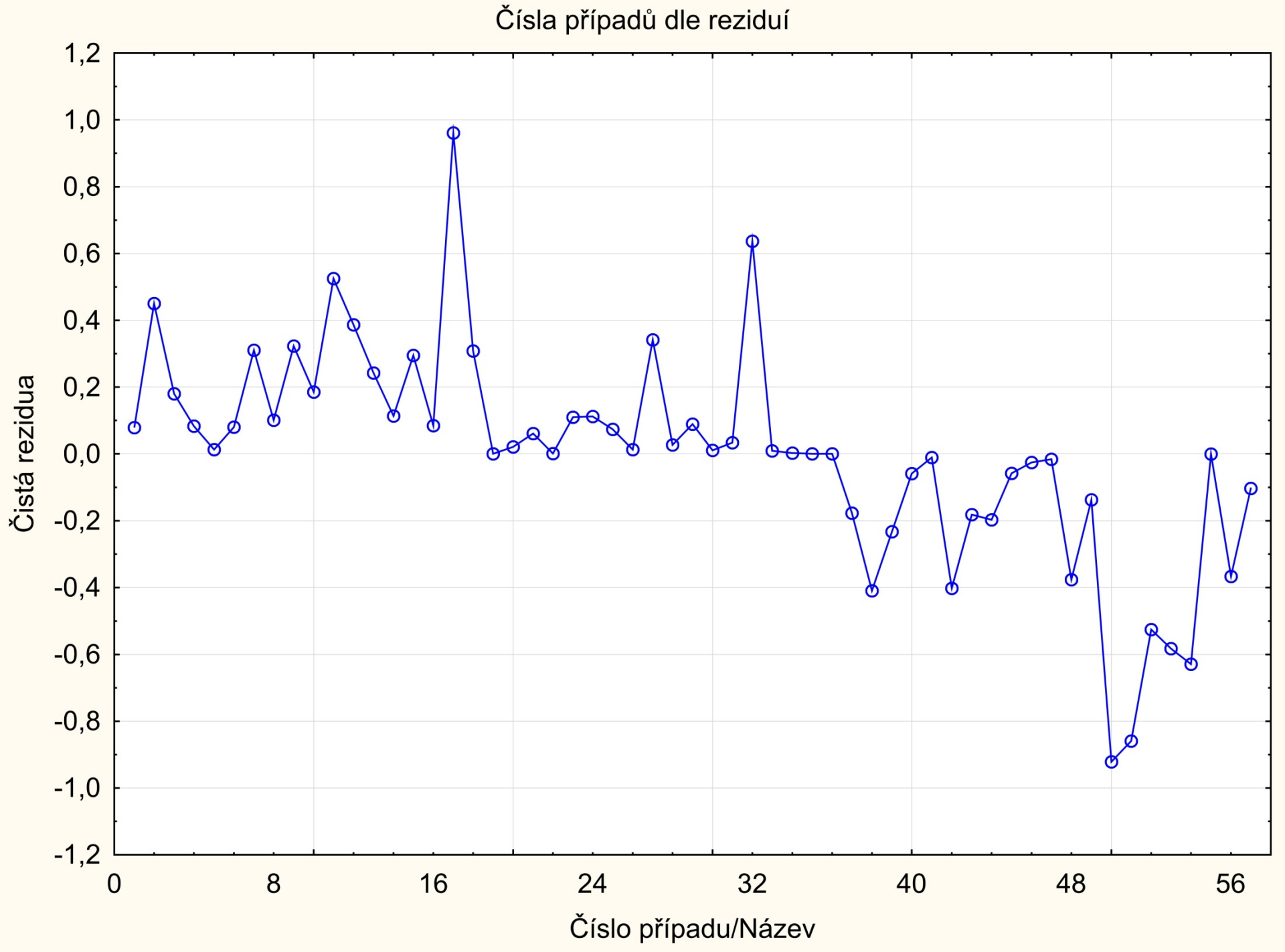
Efekt	riziko_hap - Věrohodn. test typ 1 (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
Abs. člen	1	-37,5123		
BMI	1	-37,3706	0,28329	0,594552
AGE	1	-37,3417	0,05794	0,809778
RA_DM2	1	-36,1233	2,43679	0,118519
ACPf	1	-33,3167	5,61308	0,017827
AIRf	1	-32,3986	1,83629	0,175386
AGRf	1	-30,9602	2,87673	0,089868
ACP15	1	-29,2881	3,34425	0,067440
AIR15	1	-28,9248	0,72649	0,394024
AGR15	1	-28,8923	0,06499	0,798770
slopeCP	1	-27,9217	1,94133	0,163524
slopeIR	1	-27,0493	1,74479	0,186533
slopeGR	1	-20,2488	13,60101	0,000226

Efekt	riziko_hap - Věrohodn. test typ 3 (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že riziko_hap = 0			
	Stupně volnosti	Ln-věrohod.	Chí-kvadrát	p
BMI	1	-21,1584	1,81920	0,177409
AGE	1	-20,4647	0,43189	0,511063
RA_DM2	1	-23,4227	6,34776	0,011753
ACPf	1	-20,4678	0,43803	0,508075
AIRf	1	-20,6031	0,70871	0,399874
AGRf	1	-26,7189	12,94030	0,000322
ACP15	1	-20,6878	0,87814	0,348712
AIR15	1	-20,4060	0,31448	0,574945
AGR15	1	-27,4807	14,46394	0,000143
slopeCP	1	-21,0223	1,54697	0,213584
slopeIR	1	-20,5224	0,54735	0,459401
slopeGR	1	-27,0493	13,60101	0,000226

Odezva	riziko_hap - Kvalita proložení: Hosmer-Lemeshow Test (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 15,1794, p hodn. = 0,0557492						
	Skupi1a	Skupi2a	Skupi3a	Skupi4a	Skupi5a	Skupi6a	Skupi7a
0: Pozorov.	4,00	5,00	4,00	4,00	2,00	0,00	1,00
Očekáv.	4,91	4,47	3,66	2,81	1,97	1,48	0,73
1: Pozorov.	1,00	0,00	1,00	1,00	3,00	5,00	4,00
Očekáv.	0,09	0,53	1,34	2,19	3,03	3,52	4,27
Vš. skup.	5,00	5,00	5,00	5,00	5,00	5,00	5,00

Odezva	riziko_hap - Kvalita proložení: Hosmer-Lemeshow Test (B64) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 15,1794, p hodn. = 0,0557492			
	Skupi8a	Skupi9a	Skupi10	Row Tot.
0: Pozorov.	0,00	1,00	0,0	21,0
Očekáv.	0,47	0,37	0,1	
1: Pozorov.	5,00	4,00	12,0	36,0
Očekáv.	4,53	4,63	11,9	
Vš. skup.	5,00	5,00	12,0	57,0


ROC křivka
Oblast: 0.91
Citlivost
1-Specificita
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,2
0,0
0,2
0,4
0,6
0,8
1,0
1,2


Čísla případů dle reziduí
Čistá rezidua
1,2
1,0
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
-1,0
-1,2
0
8
16
24
32
40
48
56
Číslo případu/Název