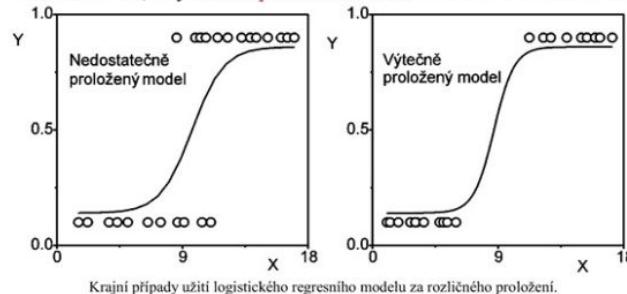


Logistická regrese LR

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.



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

2. Logistický regresní model

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

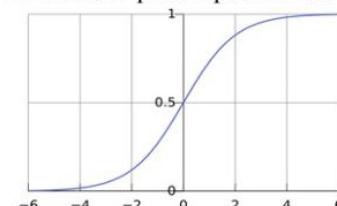
Dichotomická hodnota 0 - 1 závisle proměnné y predikuje odhad pravděpodobnosti, že se událost **stala (1)** či **nestala (0)**.

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** $L_{(1)}/L_{(0)}$, ve kterém pravděpodobnost $L_{(1)}$ je vyjádřena logistickou funkcí

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



Logitová transformace vede na sigmoidální vztah mezi závisle proměnnou y a nezávisle proměnnými 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 x obsahuje jednu či více spojitéch proměnných (*prediktory*) nebo diskrétních, kategorických (*faktory*).

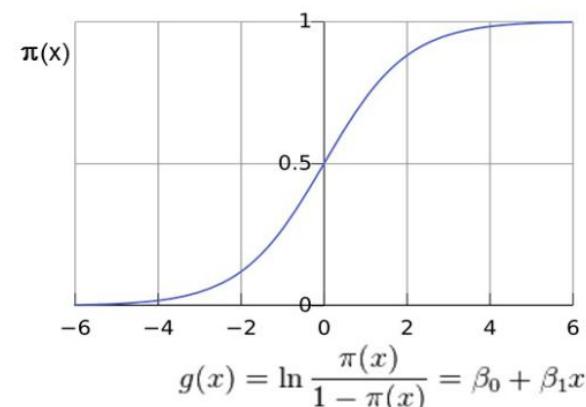
Ordinální logistická regrese: *ordinální závisle proměnná* nabývá tří a více možných stavů, např. silný nesouhlas, nesouhlas, souhlas, silný souhlas.

Vektor 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 x může obsahovat jak *prediktory*, tak i *faktory*.

Graf logistické funkce



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

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

Pravděpodobnostní poměr (zvaný „*poměr šancí*“) je vyjádřen

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

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

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

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

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

Dle klasifikačního postupu je

$$L_{(0)} = P(G = 1 | x) \quad \text{a}$$

$$L_{(1)} = P(G = 0 | x) = 1 - P(G = 1 | x).$$

a po úpravách bude

$$\ln\left(\frac{L(1)}{L(0)}\right) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p,$$

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

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 $\frac{L(1)}{L(0)} = \frac{0.75}{1-0.75} = \frac{3}{1}$.

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 y rozměru $n \times 1$ a matice X rozměru $n \times m$.

Pro i -tý objekt má y_i hodnotu buď 0, nebo 1 a x_i^T je i -tý řádek matice X .

Označme

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

a za předpokladu binomického rozdělení y lze zapsat logaritmus věrohodnostní funkce ve tvaru

$$\begin{aligned} \ln L(b) &= \sum_{i=1}^n (y_i \ln p(x_i, b) + (1 - y_i) \ln(1 - p(x_i, b))) \\ &= \sum_{i=1}^n (y_i b^T x_i - \ln(1 + \exp(b^T x_i))) \end{aligned}$$

kde $b^T = \{b_0, b_1\}$ a předpokládá se, že první sloupec matice X obsahuje pouze jedničky (absolutní člen).

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

$$\ln \frac{P(G = 1 | x)}{P(G = K | x)} = b_1, 0 + b_1^T x$$

$$\ln \frac{P(G = 2 | x)}{P(G = K | x)} = b_2, 0 + b_2^T x$$

...

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

Po zpětné transformaci vychází

$$P(G = j | x) = \frac{\exp(b_j, 0 + b_j^T x)}{1 + \sum_{l=1}^{K-1} \exp(b_l, 0 + b_l^T x)}$$

a

$$P(G = K | x) = \frac{1}{1 + \sum_{l=1}^{K-1} \exp(b_l, 0 + b_l^T x)}$$

Označíme pravděpodobnost $P(G = K | x) = p_k(x, b)$, aby se zvýraznilo, že jde o funkci regresních parametrů

(Pro $K = 2$ přechází model na logistický model pro binární proměnnou $y = G$).

Metoda odhadu parametrů:

Pro maximalizaci $\ln L(b)$ se využívá nulty prvních derivací

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

Jde o soustavu $m + 1$ nelineárních rovnic vzhledem k b .

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

$$H = \frac{d^2 L(b)}{db db^T} = -\sum_{i=1}^n x_i x_i^T p(x_i, b) (1 - p(x_i, b))$$

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

$$b_{(j+1)} = b_{(j)} - H_{(j)}^{-1} J_{(j)}$$

kde pro vektor pravděpodobnosti p rozměru $n \times 1$ s prvky $p(x_i, b_{(j)})$, a diagonální matici vah W rozměru $n \times m$ s prvky lze psát

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

Interpretace regresních koeficientů

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

$\ln(L_{(1)}/L_{(0)})$ je lineární funkcií 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

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

Obecně: Kladné znaménko koeficientu b_i zvyšuje pravděpodobnost $L_{(1)}$ 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 ∞ .

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 ∞ .

Po zlogaritmování poměru $L_{(1)}/L_{(0)}$ bude tato asymetrie odstraněna, neboť pro
pro $L_{(1)} = 0$ je $\ln(L_{(1)}/L_{(0)}) = \infty$, pro $L_{(1)} = 0.5$ je $\ln(L_{(1)}/L_{(0)}) = 0.0$,
pro $L_{(1)} = 1.0$ je $\ln(L_{(1)}/L_{(0)}) = -\infty$.

Test významnosti regresních koeficientů

Logistická regrese umožnuje **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 $W_{a,i} = (b/s(b_i))^2$ vyčísluje statistickou významnost pro odhad regresních koeficientů stejně jako ve vícenásobné regrese.

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 člena (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 biomedických aplikacích. Je mírou spojení binární proměnné, jako je faktor riziku 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 opaková 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 prediktivní 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 odlehle hodnoty. Rozptylové diagramy snadno odhalí **odlehle body**. Proměnné nemusí být **normálně** rozdeleny. **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.

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

Místo veličiny $L_{(1)}$ se používá tzv. **odchylka, deviance** $D = -2 \ln L_{(1)}$ čili

$D = -2LL$, 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 $D = -2LL = 0,0002$. Je-li 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 naprostě perfektní těsnosti proložení.

Rozdíl v odchylce je definován $G = D(\text{model bez proměnné}) - D(\text{model s proměnnou})$ 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 zaměstnání) 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 $\chi^2_{\text{exp}} = \sum_{i=1}^4 2E \left(\ln \frac{E}{V} \right)$, 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 regrese pro logitový model se vypočte dle

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

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**

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

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.

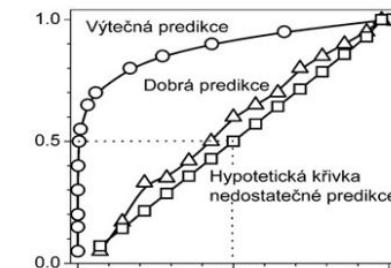
Na ose y je procento správně zařazených objektů “v události” nazvané **pozitivní podíl** (a v lékařském výzkumu nazývané **citlivost**).

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** představuje **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 zkušeností 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 výběru: **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 rizika 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 celulrita, 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 $W_{a,i}^2 = [b_i / s(b_i)]^2$.

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.

Rádek	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 vadou metabolismu Transient Neonatal Tyrosinemia (TNT) byl porovnán s kontrolním výběrem zdravých dětí Illinoiskými testy psycholinguální schopnosti ITPA. Testujte vztah kódů nemocnosti GP (1 nemocný a 0 zdravý) vůči 11 nezávislým proměnným, a to vypoštěním některých nezávisle proměnných které vrozené vady v metabolismu dětí lze posuzovat testem ITPA.

O Data: Výběr dat TNT obsahuje 54 děti 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 (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:

x1 (AUDITORY) značí sluchový přijímací záznam,

x2 (VISUAL) značí vizuální přijímací záznam,

x3 (VISMEM) značí vizuální paměť,

x4 (AUDASSOC) značí sluchové seskupení,

x5 (AUDMEMO) značí sluchová paměť,

x6 (VISASSOC) značí vizuální seskupení,

x7 (VISCLOS) značí vizuální uzavření,

x8 (VISEXPR) značí verbální projev,

x9 (GRAMCLOS) značí gramatické uzavření,

x10 (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

GP - Odhad parametrů (B29)						
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT						
Modelovaná pravděpodobnost, že GP = 0						
Efekt	Urovení Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 95, %
Abs,člen			1 -6,46735	3,145394	4,227685	-12,6322
SEX			2 0,81050	0,742746	1,190770	-0,6453
AUDITORY			3 0,11115	0,070430	2,490416	-0,0269
VISUAL			4 -0,01189	0,050823	0,054716	-0,1115
VISMEM			5 0,07274	0,051966	1,959094	-0,0291
AUDASSOC			6 -0,02514	0,045743	0,301975	-0,1148
AUDMEMO			7 0,02674	0,047616	0,315328	-0,0666
VISASSOC			8 -0,00693	0,056582	0,015007	-0,1178
VISCLOS			9 0,16040	0,089226	3,231441	-0,0145
VISEXPR			10 -0,01088	0,041853	0,067547	-0,0929
GRAMCLOS			11 -0,02802	0,061647	0,206569	-0,1488
MANUAL			12 -0,0979	0,061550	2,628476	-0,2204
Měřítko			1,00000	0,000000		1,0000
						1,00000

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 (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské, a výsledky psychologických testů ITPA:

x1 (AUDITORY) značí sluchový přijímací záznam,

x2 (VISUAL) značí vizuální přijímací záznam,

x3 (VISMEM) značí vizuální paměť,

x4 (AUDASSOC) značí sluchové seskupení,

x5 (AUDMEMO) značí sluchová paměť,

x6 (VISASSOC) značí vizuální seskupení,

x7 (VISCLOS) značí vizuální uzavření,

x8 (VISEXPR) značí verbální projev,

x9 (GRAMCLOS) značí gramatické uzavření,

x10 (MANUAL) značí manuální zručnost.

Průběh iteračního procesu zjemňování odhadů parametrů b

GP - Historie iterací (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0							
Efekt	Urovení 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	
Vérohodnost			-37,4299	-30,8522	-30,5934	-30,5885	-30,5885

Posouzení statistické významnosti odhadů parametrů dle jejich intervalu spolehlivosti

GP - Intervaly spolehlivost odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0				
Efekt	Urovení Efekt	Sloupec	Dolní LS	Horní LS
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	

Posouzení kvality těsnosti proložení experimentálních dat logistickým modelem

GP - Kvalita proložení: Hosmer-Lemeshow Test (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 6,5670, p hodn. = 0,5839								
Odezva	Skup1a	Skup2a	Skup3a	Skup4a	Skup5a	Skup6a	Skup7a	Skup8a
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

GP - Kvalita proložení: Hosmer-Lemeshow Test (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Hosmer Lemeshow = 6,5670, p hodn. = 0,5839				
Odezva	Skup1a	Skup10	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	

GP - Statistiky kvality proložení (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
Stat.	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-vérohodnost		-30,5885	

GP - Odhadnutá korelační matici odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0							
Efekt	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	0,135106	-0,071590	1,000000	-0,038633	
AUDMEMO	-0,180429	-0,143760	0,155892	-0,152214	0,077652	1,000000	
VISASSOC	-0,113032	-0,060806	0,144857	-0,096102	0,060292	0,057594	
VISCLOS	-0,348970	0,144857	0,006998	-0,079119	-0,152896	-0,146624	0,160934
VISEXPR	0,195515	-0,096102	-0,389055	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 (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské, a výsledky psychologických testů ITPA:

x1 (**AUDITORY**) značí sluchový přijímací záznam,

x2 (**VISUAL**) značí vizuální přijímací záznam,

x3 (**VISMEM**) značí vizuální paměť,

x4 (**AUDASSOC**) značí sluchové seskupení,

x5 (**AUDMEMO**) značí sluchová paměť,

x6 (**VISASSOC**) značí vizuální seskupení,

x7 (**VISCLOS**) značí vizuální uzavření,

x8 (**VISEXPR**) značí verbální projev,

x9 (**GRAMCLOS**) značí gramatické uzavření,

x10 (**MANUAL**) značí manuální zručnost.

GP - Odhadnutá korelační matici odhadů (B29)
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT
Modelovaná pravděpodobnost, že GP = 0

Efekt	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

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

GP - Test všech efektů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
Efekt	Stupně volnosti	Ln-vérohod.	P
Abs. člen	1	-37,4299	
SEX	1	-37,3927	0,074500
AUDITORY	1	-35,6439	3,497555
VISUAL	1	-35,5665	0,154770
VISMEM	1	-34,0424	3,048217
AUDASSOC	1	-33,8976	0,289663
AUDMEMO	1	-33,7563	0,282585
VISASSOC	1	-33,6946	0,123414
VISCLOS	1	-32,1339	3,121384
VISEXPR	1	-32,0116	0,244634
GRAMCLOS	1	-31,9782	0,066815
MANUAL	1	-30,5885	2,779401

GP - Vérohodn. test typ 1 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
Efekt	Stupně volnosti	Ln-vérohod.	Chi-kvadrát
Abs. člen	1	-37,4299	
SEX	1	-35,6439	3,497555
AUDITORY	1	-35,5665	0,154770
VISUAL	1	-34,0424	3,048217
VISMEM	1	-33,8976	0,289663
AUDASSOC	1	-33,7563	0,282585
AUDMEMO	1	-33,6946	0,123414
VISASSOC	1	-32,1339	3,121384
VISCLOS	1	-32,0116	0,244634
VISEXPR	1	-31,9782	0,066815

GP - Základní rezidua (B29)			
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0			
Případ	Čistá	Pearson.	Odhylka
1	-0.296291	-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	-0.96824
11	-0.140265	-0.48565	-0.61842
12	-0.370516	-0.77545	-0.96129
13	-0.763895	-1.90487	-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	-0.46208	-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	-1.39675
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	0.49964
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	0.345262	0.72621	0.92038
46	0.533262	1.06893	1.24523
47	0.20487	0.50767	0.67721
48	0.432245	0.87254	1.06402
49	0.599064	1.22236	1.35200

Statistická analýza reziduí

jj (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské,
a výsledky psychologických testů ITPA:
x1 (**AUDITORY**) značí sluchový přijímací záznam,
x2 (**VISUAL**) značí vizuální přijímací záznam,
x3 (**VISMEM**) značí vizuální paměť,
x4 (**AUDASSOC**) značí sluchové seskupení,
x5 (**AUDMEMO**) značí sluchová paměť,
x6 (**VISASSOC**) značí vizuální seskupení,
x7 (**VISCLOS**) značí vizuální uzavření,
x8 (**VISEXPR**) značí verbální projev,
x9 (**GRAMCLOS**) značí gramatické uzavření,
x10 (**MANUAL**) značí manuální zručnost.

GP - Kvalita proložení - ROC, (B29)						
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0						
Případ	Predikt. čisté	Správné předpisy	Správné neúspěchy	Neúspěch predik.	Jako úspěch	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	0.806968	8.00000	27.00000	0.00000	19.00000	0.296296
9	0.806983	9.00000	27.00000	0.00000	18.00000	0.333333
10	0.785063	10.00000	26.00000	0.00000	18.00000	0.333333
11	0.783995	9.00000	25.00000	0.00000	18.00000	0.333333
12	0.701889	9.00000	24.00000	0.00000	18.00000	0.333333
13	0.679869	10.00000	24.00000	0.00000	17.00000	0.370370
14	0.674683	11.00000	24.00000	0.00000	16.00000	0.407407
15	0.654718	12.00000	24.00000	0.00000	15.00000	0.444444
16	0.651785	13.00000	24.00000	0.00000	14.00000	0.481481
17	0.643784	13.00000	23.00000	0.00000	14.00000	0.481481
18	0.637644	14.00000	23.00000	0.00000	13.00000	0.518518
19	0.637514	14.00000	22.00000	0.00000	13.00000	0.518518
20	0.624031	14.00000	21.00000	0.00000	13.00000	0.518518
21	0.594812	14.00000	20.00000	0.00000	13.00000	0.518518
22	0.582125	15.00000	20.00000	0.00000	12.00000	0.555556
23	0.576374	15.00000	19.00000	0.00000	12.00000	0.555556
24	0.576698	16.00000	19.00000	0.00000	11.00000	0.592592
25	0.567755	17.00000	19.00000	0.00000	10.00000	0.629630
26	0.528293	17.00000	18.00000	0.00000	10.00000	0.629630
27	0.505041	17.00000	17.00000	0.00000	10.00000	0.629630
28	0.490592	18.00000	17.00000	0.00000	9.00000	0.666667
29	0.482818	19.00000	17.00000	0.00000	8.00000	0.703707
30	0.467506	20.00000	17.00000	0.00000	7.00000	0.740741
31	0.466718	21.00000	17.00000	0.00000	6.00000	0.777778
32	0.458280	21.00000	16.00000	0.00000	6.00000	0.777778
33	0.445445	22.00000	16.00000	0.00000	11.00000	0.814815
34	0.437712	22.00000	15.00000	0.00000	12.00000	0.814815
35	0.429473	23.00000	15.00000	0.00000	12.00000	0.851851
36	0.400938	24.00000	15.00000	0.00000	12.00000	0.888889
37	0.375516	24.00000	14.00000	0.00000	13.00000	0.925526
38	0.368156	24.00000	13.00000	0.00000	14.00000	0.962592
39	0.351570	24.00000	13.00000	0.00000	15.00000	0.999999
40	0.336089	25.00000	12.00000	0.00000	15.00000	0.999999
41	0.308683	25.00000	11.00000	0.00000	16.00000	0.999999
42	0.268686	25.00000	10.00000	0.00000	17.00000	0.999999
43	0.277246	25.00000	9.00000	0.00000	18.00000	0.999999
44	0.247269	26.00000	9.00000	0.00000	19.00000	1.00000
45	0.243558	26.00000	8.00000	0.00000	19.00000	1.00000
46	0.214659	26.00000	7.00000	0.00000	20.00000	1.00000
47	0.174053	26.00000	6.00000	0.00000	21.00000	1.00000
48	0.169694	26.00000	5.00000	0.00000	22.00000	1.00000
49	0.814865					

GP - Kvalita proložení - ROC, (B29)						
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0						
Případ	1-Specificita	1-Specificita				
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
16	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
17	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
18	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
19	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
20	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
21	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
22	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
23	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
24	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
26	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
27	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
28	0.000000	0.000000	0.00000			

GP - Odhad parametrů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0						
Efekt	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 (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské, a výsledky psychologických testů ITPA:
 x1 (AUDITORY) značí sluchový přijímací záznam,
 x2 (VISUAL) značí vizuální přijímací záznam,
 x3 (VISMEM) značí vizuální paměť,
 x4 (AUDASSOC) značí sluchové seskupení,
 x5 (AUDMEMO) značí sluchová paměť,
 x6 (VISASSOC) značí vizuální seskupení,
 x7 (VISCLOS) značí vizuální uzavření,
 x8 (VISEXPR) značí verbální projev,
 x9 (GRAMCLOS) značí gramatické uzavření,
 x10 (MANUAL) značí manuální zručnost.

GP - Vérohodn. test typ 1 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0				
Efekt	Stupně volnosti	Ln-vérohod.	Chi-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

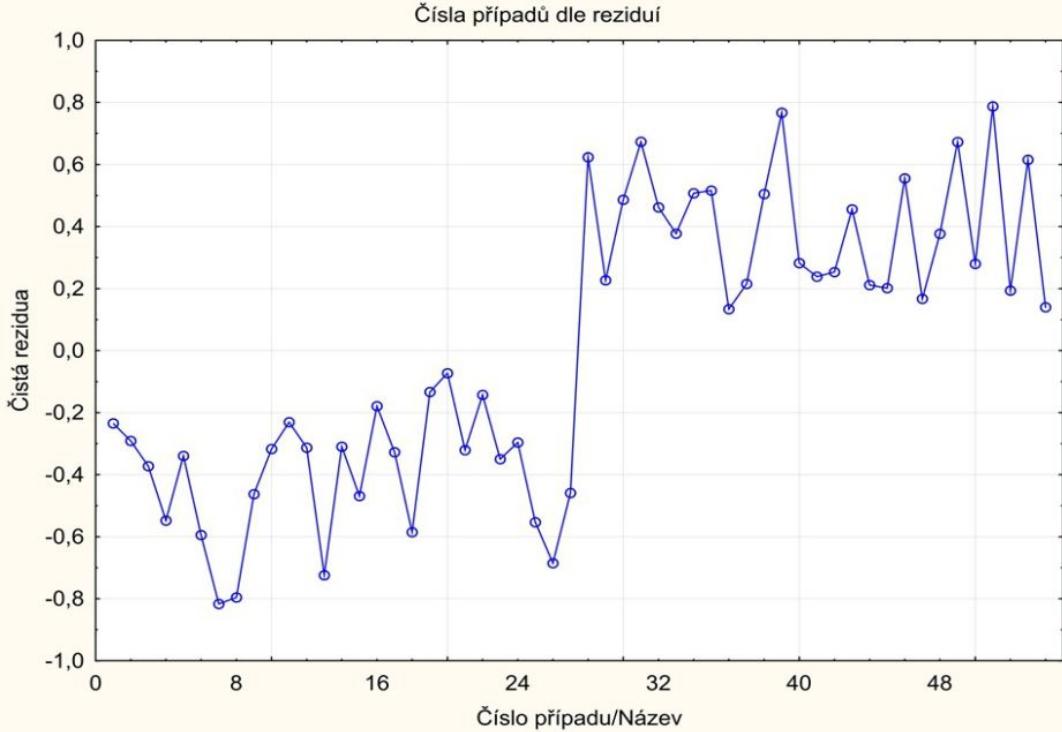
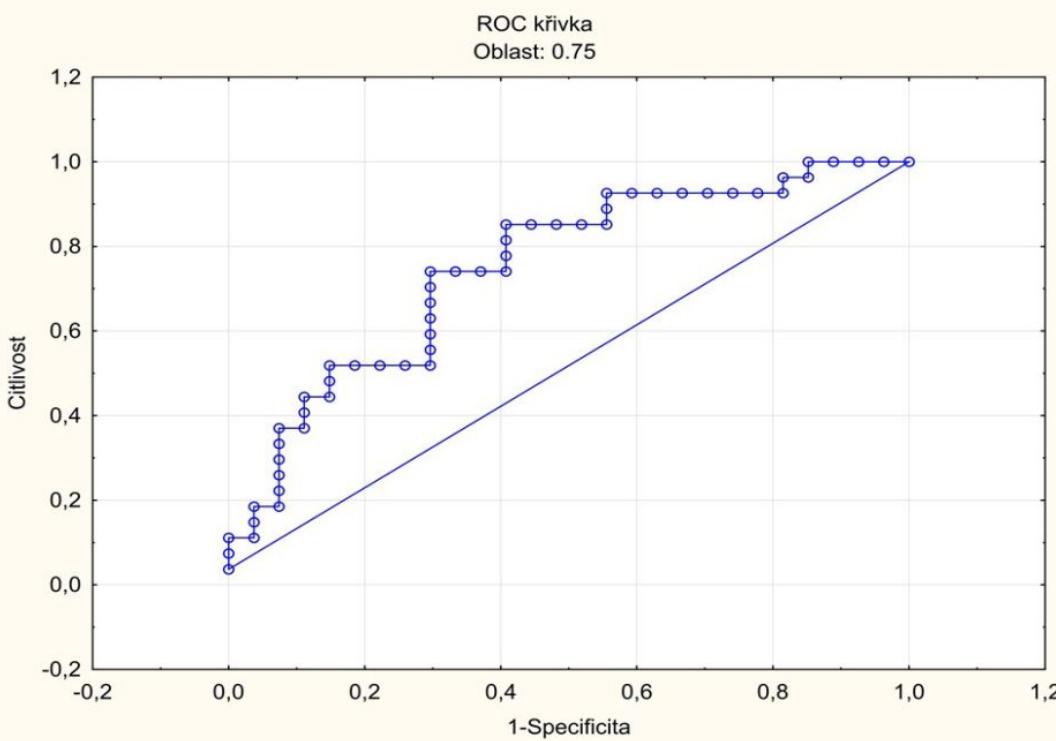
GP - Vérohodn. test typ 3 (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0				
Efekt	Stupně volnosti	Ln-vérohod.	Chi-kvadrát	p
AUDITORY	1	-32,6044	1,546600	0,213638
VISMEM	1	-32,3886	1,115168	0,290961
VISCLOS	1	-33,5297	3,397255	0,065305
MANUAL	1	-33,6473	3,632514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

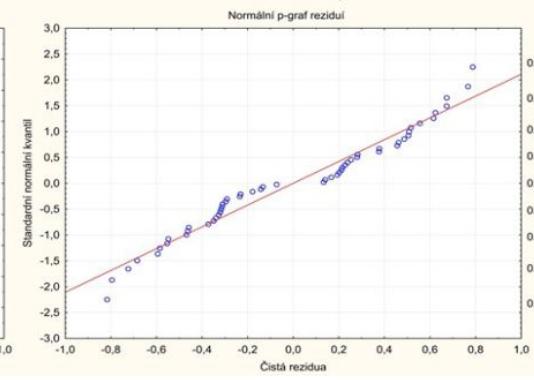
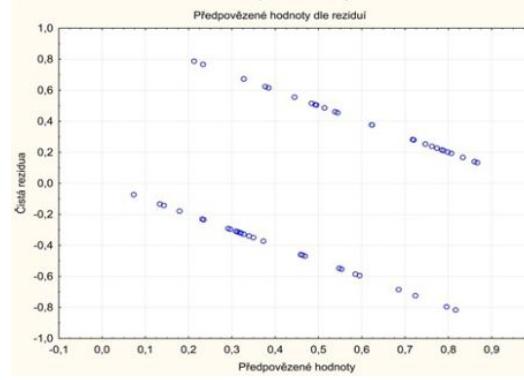
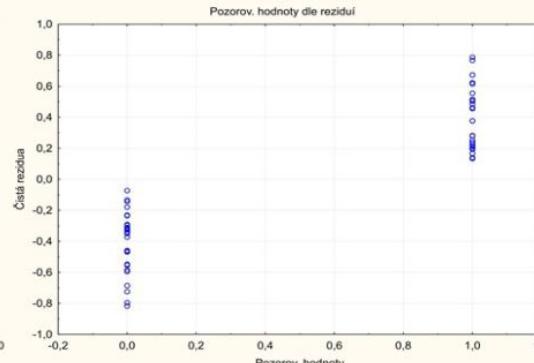
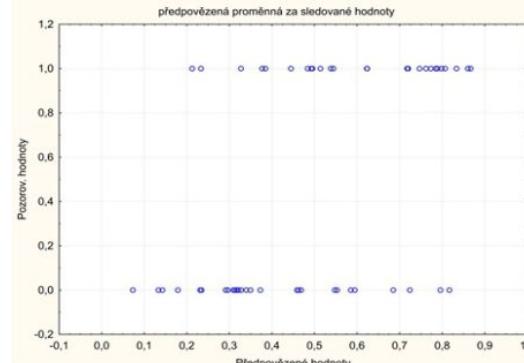
GP - Intervaly spolehlivost odhadů (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0				
Efekt	Úroveň Efekt	Sloupec	Dolní LS 80, %	Horní LS 80, %
Abs.člen			1 -7,55708	-1,50055
AUDITORY	2		-0,00342	0,12577
VISMEM	3		-0,01085	0,10483
VISCLOS	4		0,03753	0,24012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

GP - Historie iterací (B29) Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT Modelovaná pravděpodobnost, že GP = 0					
Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2
Abs.člen			1	0,0000	-3,9746
AUDITORY			2	0,0000	0,0484
VISMEM			3	0,0000	0,0425
VISCLOS			4	0,0000	0,1194
MANUAL			5	0,0000	-0,0836
Vérohodnost				-37,4299	-31,9418
				-31,8315	-31,8311

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

jj (SEX) značí pohlaví dítěte: 1 mužské, 0 ženské, a výsledky psychologických testů ITPA:
 x1 (AUDITORY) značí sluchový přijímací záznam,
 x2 (VISUAL) značí vizuální přijímací záznam,
 x3 (VISMEM) značí vizuální paměť,
 x4 (AUDASSOC) značí sluchové seskupení,
 x5 (AUDMEMO) značí sluchová paměť,
 x6 (VISASSOC) značí vizuální seskupení,
 x7 (VISCLOS) značí vizuální uzavření,
 x8 (VISEXPR) značí verbální projev,
 x9 (GRAMCLOS) značí gramatické uzavření,
 x10 (MANUAL) značí manuální zručnost.





Příklad 8.7/str. 308 (261)

Data PCS.start

CAPSULE - Odhad parametrů (87PCS.sta)						
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT						
Modelovaná pravděpodobnost, že CAPSULE = 0						
Efekt	Úroveň Efekt	Slopec	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	0.136396	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 - Věrohodn. test typ 3 (87PCS.sta)				
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT				
Modelovaná pravděpodobnost, že CAPSULE = 0				
Efekt	Stupně volnosti	Ln-věrohod.	Chi-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 - Historie iterací (87PCS sta)							
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT							
Modelovaná pravděpodobnost, že CAPSULE = 0							
Efekt	Úroveň	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4
Abs.člen	Efekt		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
Vérohodnost			-263.396	-196.529	-189.699	-189.325	-189.324

Efekt	CAPSULE - Věrohodn. test typ 1 (87PCS.sta)			
	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

Případ	CAPULE		Předpověď hodnoty (TSTS PC) a		Rozdělení BINOMICKÉ, Linkující funkce LOGIT		Modelovaná pravděpodobnost		z CAPULE = 0
	Odezva "1" = 0	Před. Hodnota	Před.	Lineár. Chyba	Smržed. Chyba	Dolní LS 95 %	Horní LS 95 %		
1	1.000000	0.768398	-1.19929	0.20915	-0.687373	0.833511			
1	1.000000	0.293666	-0.87765	0.474089	0.41020	0.512085			
1	1.000000	0.775030	1.23682	0.526399	0.550153	0.906680			
1	1.000000	0.460394	-0.15876	0.547913	0.225718	0.714049			
1	1.000000	0.894164	2.13421	0.396422	0.793511	0.948397			
1	0.000000	0.145174	-1.77296	0.522955	0.057436	0.321260			
1	1.000000	0.175247	-1.54809	0.603193	0.061162	0.409351			
1	1.000000	0.092677	-2.28138	0.715210	0.024526	0.293265			
1	1.000000	0.736801	0.12941	0.274594	0.62039	0.872442			
10.000000	1.000000	0.834106	1.61501	0.669565	0.576827	0.948839			
11.000000	0.000000	0.313334	-0.74845	0.637360	0.115071	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	0.06700			
14.000000	0.000000	0.082666	-2.40402	0.532013	0.030865	0.204030			
15.000000	1.000000	0.936222	2.68644	0.398105	0.870587	0.967926			
16.000000	1.000000	0.903767	2.23980	0.340511	0.82127	0.948201			
17.000000	1.000000	0.969159	3.44765	0.620369	0.930964	0.990655			
18.000000	1.000000	0.911819	2.33605	0.294658	0.653026	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	0.864540	0.007048	0.173740			
23.000000	0.000000	0.444019	-2.22487	0.255094	0.526327	0.568346			

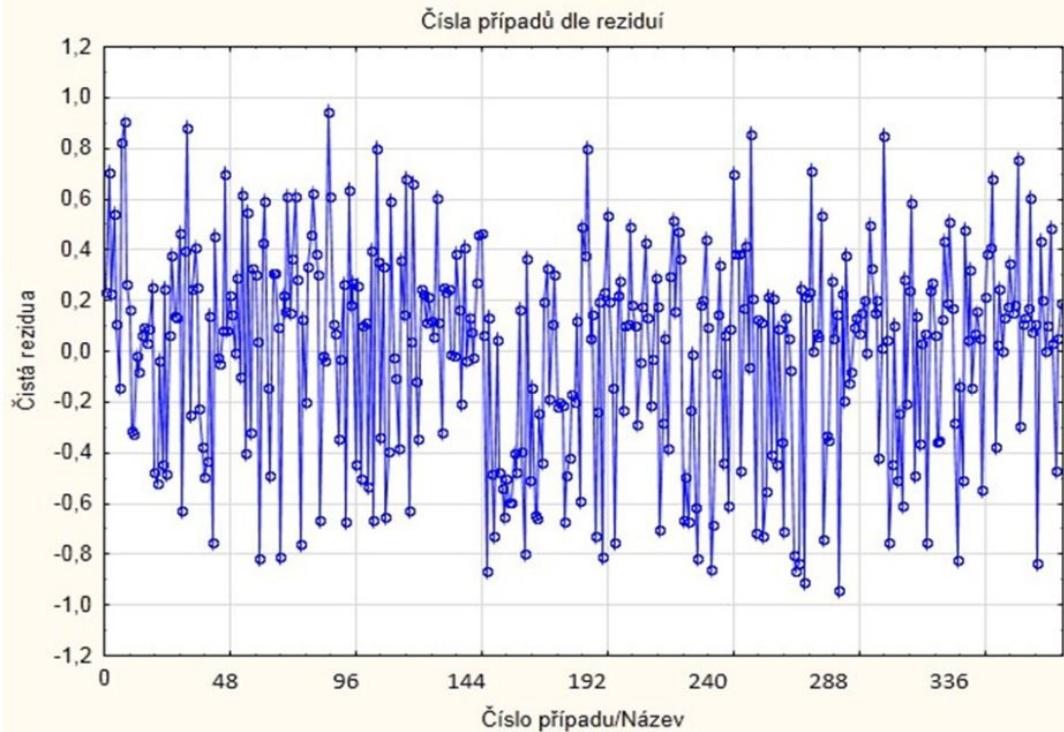
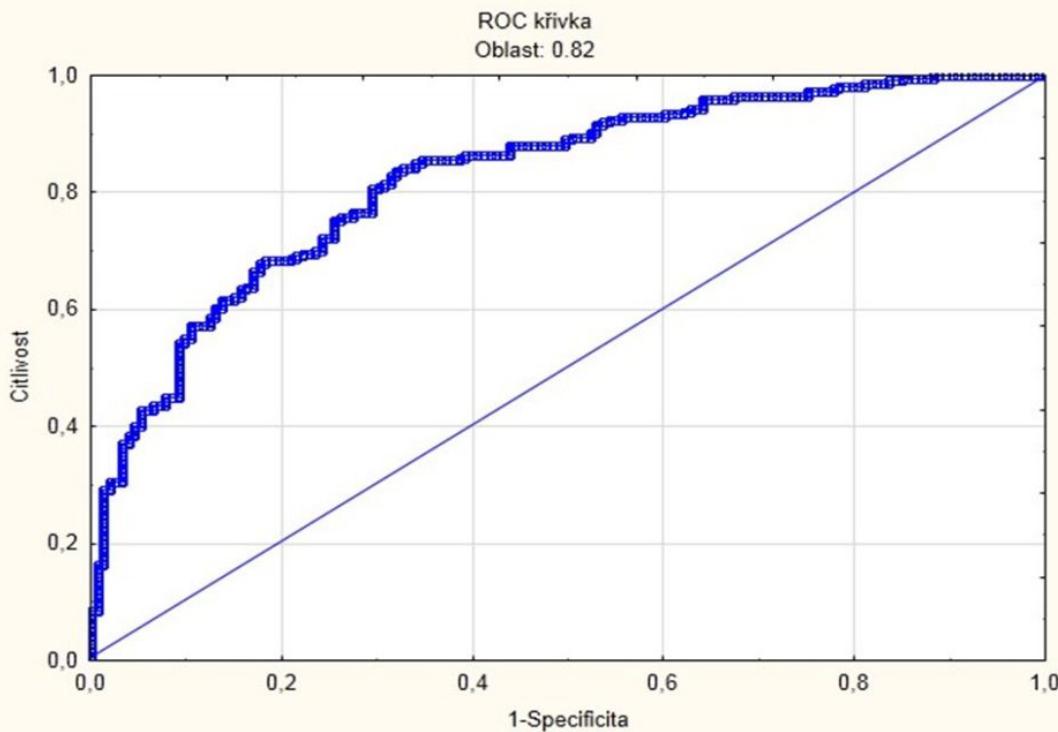
CAPSULE - Odhad parametrů (87PCS)								
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT								
Modelovaná pravděpodobnost, že CAPSULE = 0								
Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	Dolní LS 80, %	Horní LS 80, %	p
Abs. člen		1	7,279683	1,688841	18,58007	5,11535	9,444020	0,000016
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		1,000000	1,000000	

CAPSULE - Vérohodn. test typ 1 (87PCS)				
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT				
Modelovaná pravděpodobnost, že CAPSULE = 0				
Efekt	Stupně volnosti	Ln-vérohod.	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 - Vérohodn. test typ 3 (87PCS)				
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT				
Modelovaná pravděpodobnost, že CAPSULE = 0				
Efekt	Stupně volnosti	Ln-vérohod.	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 - Intervaly spolehlivost odhadů (87PCS)			
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT			
Modelovaná pravděpodobnost, že CAPSULE = 0			
Efekt	Úroveň	Sloupec	Dolní LS 80, % Horní LS 80, %
Abs. člen	Urovnění	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

CAPSULE - Historie iterací (87PCS)						
Rozdělení : BINOMICKE, Linkující funkce: LOGIT						
Modelovaná pravděpodobnost, že CAPSULE = 0						
Efekt	Úroveň	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3
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
Vérohodnost			-263,396	-196,529	-189,699	-189,325



Příklad 8.8/str. 311

Data Meexp.sta

A STATISTICA Cz - [Data: 88Meexp.sta (8 s kód 412F)]

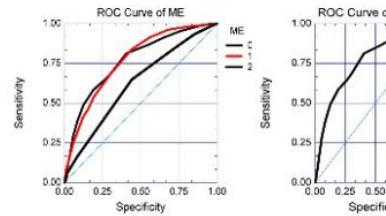
	1	2	3	4	5	6	7	8
ME	0	3	7	0	1	2	Var7	Var8
1.000000	0	3	7	0	1	2		
2.000000	0	2	11	0	1	3		
3.000000	0	3	8	1	1	3		
4.000000	1	3	11	0	1	3		
5.000000	2	4	7	0	1	3		
6.000000	0	3	7	0	1	3		
7.000000	2	4	6	0	1	2		
8.000000	0	4	6	0	1	3		
9.000000	0	2	6	0	1	3		
10.000000	1	4	6	0	1	3		
11.000000	0	4	8	0	1	2		
12.000000	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		

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

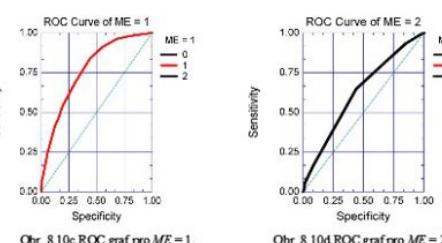
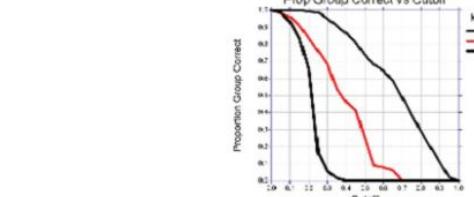
Efekt	Úroveň Efekt	Sloupec	Odhad	Standard chyba	Wald. Stat.	p
Abs.člen		1	-1.24385	1.053363	1.394382	0.237667
ME		2	-0.62219	0.258655	5.786336	0.016151
SYMPT		3	-0.07790	0.166829	0.218037	0.640540
PB		4	0.13412	0.072958	3.379245	0.066022
HIST		5	-0.18157	0.570330	0.101353	0.750212
DETC		6	-0.46355	0.252377	3.373580	0.066250
Měřítko			1.000000	0.000000		

Závěr: V logistickém regresním modelu jsou z původních 5 znaků nezávisle proměnné statisticky významné 3 znaky, a to pro závisle promennou $BSE = 0$ 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. 8.10c ROC graf pro ME = 1.
Obr. 8.10d ROC graf pro ME = 2.

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 - Historické iteraci (88Meexp.sta)					
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT					
Modelovaná pravděpodobnost, že $BSE = 0$					
Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2
Abs.člen		1	0.000000	-1.038	-1.204
ME		2	0.000000	-0.214	-0.447
SYMPT		3	0.000000	-0.036	-0.066
PB		4	0.000000	0.065	0.116
HIST		5	0.000000	-0.062	-0.132
DETC		6	0.000000	-0.258	-0.418
Věrohodnost			-285.577	-156.983	-148.625
				-147.981	-147.971
					-147.971

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT		3	0.000000	-0.036	-0.066	-0.077	-0.078	-0.078
PB		4	0.000000	0.065	0.116	0.133	0.134	0.134
HIST		5	0.000000	-0.062	-0.132	-0.175	-0.181	-0.182
DETC		6	0.000000	-0.258	-0.418	-0.461	-0.464	-0.464
Věrohodnost			-285.577	-156.983	-148.625	-147.981	-147.971	-147.971

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT		3	0.000000	-0.036	-0.066	-0.077	-0.078	-0.078
PB		4	0.000000	0.065	0.116	0.133	0.134	0.134
HIST		5	0.000000	-0.062	-0.132	-0.175	-0.181	-0.182
DETC		6	0.000000	-0.258	-0.418	-0.461	-0.464	-0.464
Věrohodnost			-285.577	-156.983	-148.625	-147.981	-147.971	-147.971

BSE - Statistiky kvality proložení (88Meexp.sta)					
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT					
Modelovaná pravděpodobnost, že $BSE = 0$					
Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2
Abs.člen		1	0.000000	-1.038	-1.204
ME		2	0.000000	-0.214	-0.447
SYMPT		3	0.000000	-0.036	-0.066
PB		4	0.000000	0.065	0.116
HIST		5	0.000000	-0.062	-0.132
DETC		6	0.000000	-0.258	-0.418
Věrohodnost			-285.577	-156.983	-148.625
				-147.981	-147.971
					-147.971

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT		3	0.000000	-0.036	-0.066	-0.077	-0.078	-0.078
PB		4	0.000000	0.065	0.116	0.133	0.134	0.134
HIST		5	0.000000	-0.062	-0.132	-0.175	-0.181	-0.182
DETC		6	0.000000	-0.258	-0.418	-0.461	-0.464	-0.464
Věrohodnost			-285.577	-156.983	-148.625	-147.981	-147.971	-147.971
				-147.981	-147.971			
					-147.971			

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT		3	0.000000	-0.036	-0.066	-0.077	-0.078	-0.078
PB		4	0.000000	0.065	0.116	0.133	0.134	0.134
HIST		5	0.000000	-0.062	-0.132	-0.175	-0.181	-0.182
DETC		6	0.000000	-0.258	-0.418	-0.461	-0.464	-0.464
Věrohodnost			-285.577	-156.983	-148.625	-147.981	-147.971	-147.971
				-147.981	-147.971			
					-147.971			

Efekt	Úroveň Efekt	Sloupec	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5
Abs.člen		1	0.000000	-1.038	-1.204	-1.241	-1.244	-1.244
ME		2	0.000000	-0.214	-0.447	-0.594	-0.622	-0.622
SYMPT								

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	6 - PTL	6 - PTL
5 - SMOKE	5 - SMOKE	6 - PTL	7 - HT	7 - HT	8 - UI	9 - FTV	10 - BWT
6 - PTL	6 - PTL	7 - HT	8 - UI	8 - UI	9 - FTV	10 - BWT	
7 - HT		9 - FTV		10 - BWT			
8 - UI							
9 - FTV							
10 - BWT							

Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit

Závislá proměnná: Kateg. nezáv. (faktory): Spojité nezáv. prom.: Čítací proměnná

1 2-4

Ukázat pouze odpovídající proměnné

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

Efekt	Úroveň	Sloupec	Odhad	Standard chyba	Wald Stat.	p
Abs. člen			1 -1 02012	1.13611	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 - Vérohodn. test typ 1 (86Lowbwt.sta)
Rozdělení: BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že **LOW = 0**

Efekt	Stupně volnosti	Ln-vérohod.	Chi-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 - Vérohodn. test typ 3 (86Lowbwt.sta)
Rozdělení: BINOMICKÉ, Linkující funkce: L
Modelovaná pravděpodobnost, že **LOW = 0**

Efekt	Stupně volnosti	Ln-vérohod.	Chi-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 2009)]

Vyberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - STA	11 - HRA	1 - STA	11 - HRA	11 - HRA	21 - Var2	1 - STA	11 - HRA	11 - HRA	21 - Var2
2 - AGE	12 - PRE	2 - AGE	12 - PRE	12 - PRE	22 - Var2	2 - AGE	12 - PRE	12 - PRE	22 - Var2
3 - SEX	13 - TYP	3 - SEX	13 - TYP	13 - TYP		3 - SEX	13 - TYP	13 - TYP	
4 - RACE	14 - FRA	4 - RACE	14 - FRA	14 - FRA		4 - RACE	14 - FRA	14 - FRA	
5 - SER	15 - PO2	5 - SER	15 - PO2	15 - PO2		5 - SER	15 - PO2	15 - PO2	
6 - CAN	16 - PH	6 - CAN	16 - PH	16 - PH		6 - CAN	16 - PH	16 - PH	
7 - CRN	17 - PCO	7 - CRN	17 - PCO	17 - PCO		7 - CRN	17 - PCO	17 - PCO	
8 - INF	18 - BIC	8 - INF	18 - BIC	18 - BIC		8 - INF	18 - BIC	18 - BIC	
9 - CPR	19 - CRE	9 - CPR	19 - CRE	19 - CRE		9 - CPR	19 - CRE	19 - CRE	
10 - SYS	20 - LOC	10 - SYS	20 - LOC	20 - LOC		10 - SYS	20 - LOC	20 - LOC	

Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit Roztažn. Přiblížit

Závislá proměnná: Kateg. nezáv. (faktory): Spojité nezáv. prom.: Čítací proměnná

1 2-20

Ukázat pouze odpovídající proměnné

Základní nastavení | Detaily | OK Storno Možnosti Otevři Data Vložit momenty Editor syntaxi

Proměnné
Závislá proměnná: STA
Kateg. nezáv. (faktory): Vložit Přiblížit OK Storno Možnosti

Kódy faktorů / zadání:
Mezikup. efekty: $AGE + SEX + RACE + SER + CAN + CRN + INF + CPR + SY'S + HRA + PRE + TYP +$

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

Efekt	Úroveň	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 055496	0.501193	1.22605	0.268176
RACE			4 0 00445	0.528255	0.00007	0.993131
SER			5 0 54534	0.575342	0.89844	0.343201
CAN			6 0 275774	0.980453	7.91140	0.004912
CRN			7 0 10191	0.762363	0.01787	0.893657
INF			8 0 05607	0.534731	0.01099	0.916495
CPR			9 0 97789	0.983501	0.98861	0.320082
SY'S			10 0 01154	0.007854	2.16075	0.141575
HRA			11 0 00364	0.009344	0.15188	0.696741
PRE			12 0 92926	0.628869	2.18361	0.139487
TYP			13 0 274427	0.995264	7.62084	0.005828
FRA			14 0 115208	0.999321	1.32909	0.248968
PO2			15 0 38667	0.851057	0.20643	0.649583
PH			16 0 241525	1.23111	3.84884	0.049780
PCO			17 0 317282	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 0 270693	0.751827	0.745530	0.000319
Měřítko			1.00000	0.000000		

GLZ Vlastní obecné schéma: 85ICU.sta

Základní nastavení | Detaily | OK Storno Možnosti

Proměnné
Závislá proměnná: STA
Cílová proměnná: Vložit Přiblížit OK Storno Možnosti

Kódy faktorů / zadání:
Mezikup. efekty: AGE + SEX + RACE + SER + CAN + CRN + INF + CPR + SY'S + HRA + PRE + TYP +

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

Efekt	Úroveň	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 055496	0.501193	1.22605	0.268176
RACE			4 0 00445	0.528255	0.00007	0.993131
SER			5 0 54534	0.575342	0.89844	0.343201
CAN			6 0 275774	0.980453	7.91140	0.004912
CRN			7 0 10191	0.762363	0.01787	0.893657
INF			8 0 05607	0.534731	0.01099	0.916495
CPR			9 0 97789	0.983501	0.98861	0.320082
SY'S			10 0 01154	0.007854	2.16075	0.141575
HRA			11 0 00364	0.009344	0.15188	0.696741
PRE			12 0 92926	0.628869	2.18361	0.139487
TYP			13 0 274427	0.995264	7.62084	0.005828
FRA			14 0 115208	0.999321	1.32909	0.248968
PO2			15 0 38667	0.851057	0.20643	0.649583
PH			16 0 241525	1.23111	3.84884	0.049780
PCO			17 0 317282	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 0 270693	0.751827	0.745530	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	Sloupec	Dolní LS	Horní LS
Abs. člen			1 2.5525	9.57235
AGE			2 0.08631	-0.01856
SEX			3 0.42736	1.53728
RACE			4 0.103081	0.103991
SER			5 0.58231	1.67299
CAN			6 0.47938	-0.83609
CRN			7 0.139229	1.59612
INF			8 0.099198	1.10412
CPR			9 0.205051	0.94974
SY'S			10 0.00385	0.02694
FRA			11 0.01467	0.02195
PO2			12 0.16184	0.30328
PH			13 0.69495	-0.79359
PCO			14 0.31071	0.80656
BIC			15 0.205471	1.28137
CRE			16 0.42818	-0.00232
LOC			17 0.45635	5.88929
Měřítko			18 0.100186	2.58953
			19 0.234094	1.87431
			20 0.47964	-1.23253

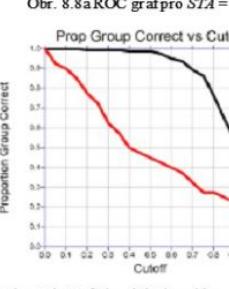
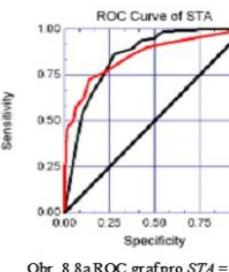
STA - Historie iteraci (85ICU.sta)								
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT								
Modelovaná pravděpodobnost, že STA = 0								
Úroveň	Efekt	Iterace 0	Iterace 1	Iterace 2	Iterace 3	Iterace 4	Iterace 5	Iterace 6
Abs.člen		1	0.000	2.362	3.9102	4.9616	5.3637	5.4131
SEX		2	-0.000	-0.0183	-0.0353	-0.0475	-0.0520	-0.0524
RACE		3	0.000	0.2156	0.3884	0.5059	0.5509	0.5549
AGE		4	0.000	0.0259	0.0286	0.0122	0.0056	0.0046
CAN		5	0.000	0.0183	0.3593	0.4891	0.5397	0.5453
CRN		6	0.000	-1.0249	-1.8844	-2.4880	-2.7283	-2.7574
INF		7	0.000	-0.0210	0.0266	0.0758	0.0985	0.1019
SYS		8	0.000	-0.0991	-0.0865	0.0040	0.0511	0.0560
FRA		9	0.000	-0.2621	-0.6009	-0.8640	-0.9659	-0.9777
BYS		10	0.000	0.0039	0.0073	0.0101	0.0119	0.0115
HRA		11	0.000	0.0018	0.0031	0.0036	0.0036	0.0036
NE		12	0.000	-0.4345	-0.7237	-0.8755	-0.9244	-0.9292
TP		13	0.000	-0.8794	-1.7325	-2.4110	-2.7051	-2.7443
IRA		14	0.000	-0.2734	-0.6687	-1.0072	-1.1373	-1.1519
POZ		15	0.000	-0.190	-0.3076	-0.3608	-0.3866	-0.3867
IN		16	0.000	-0.9865	-1.5848	-2.1411	-2.3843	-2.4149
CRC		17	0.000	0.9508	1.9106	2.7379	3.1220	3.1728
BIC		18	0.000	0.1664	0.4441	0.6829	0.7819	0.7937
CRE		19	0.000	-0.0213	-0.2434	-0.2372	-0.2336	-0.2333
ZOC		20	0.000	-1.3865	-1.9862	-2.4599	-2.6779	-2.7057
Vérohodnost			-138.629	-74.4996	-66.6317	-64.9551	-64.8017	-64.7998

STA - Vérohodn. test typ 1 (85ICU.sta)							
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT							
Modelovaná pravděpodobnost, že STA = 0							
Úroveň	Efekt	Stupně volnosti	Ln-vérohod.	Chi-kvadrát	p		
AGE		1	-100.080				
SEX		1	-96.153	7.85459	0.0005069		
RACE		1	-96.153	0.00093	0.975723		
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		
POZ		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		

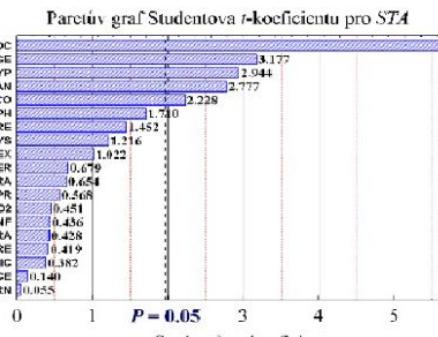
STA - Vérohodn. test typ 1 (85ICU.sta)							
Rozdělení : BINOMICKÉ, Linkující funkce: LOGIT							
Modelovaná pravděpodobnost, že STA = 0							
Úroveň	Efekt	Stupně volnosti	Ln-vérohod.	Chi-kvadrát	p		
Abs.člen		1	-100.080				
SEX		1	-96.153	7.85459	0.0005069		
RACE		1	-96.153	0.00093	0.975723		
SER		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		
POZ		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		



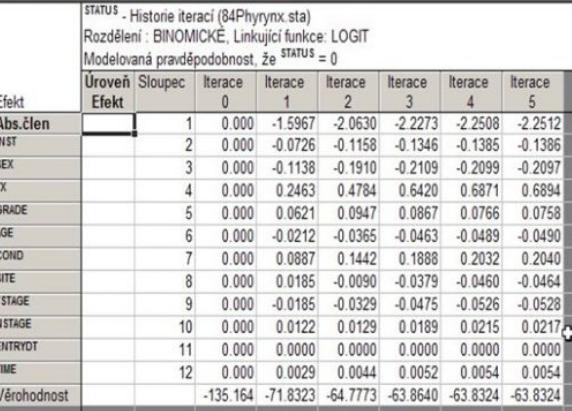
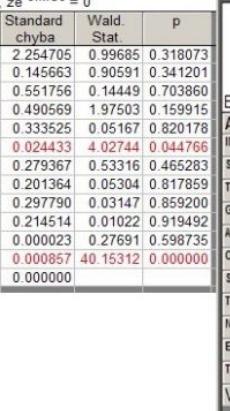
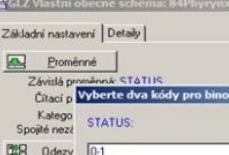
Závěr: V logistickém regresním modelu jsou z původních 18 znaků nezávisle proměnných statisticky významných 6 znaků (pro proměnnou **STA=0**), a to **AGE, CAN, TYP, PH, PCO a LOC**.



Obr. 8.8a ROC graf pro STA = 0 a 1.
Obr. 8.8b Graf obou krivék podél správně zařazených objektů proti P_c (cuto ff) určuje hodnotu prahového délečího bodu P_c .



Obr. 8.7 Pareto diagram vyšetření statistické významnosti parametrů logistického regresního modelu dat ICUna základě hodnot Studentova t-koefficientu.



STATUS - Vérohodn. test typ 1 (84Phrynx.sta) Rozdělení : BINOMICKE, Linkující funkce: L Modelovaná pravděpodobnost, že STATUS = 0				
Efekt	Stupně volnosti	Ln-vérohod.	Chi-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 - Vérohodn. test typ 3 (84Phrynx.sta) Rozdělení : BINOMICKE, Linkující funkce: LO Modelovaná pravděpodobnost, že STATUS = 0				
Efekt	Stupně volnosti	Ln-vérohod.	Chi-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 (84Phrynx.sta) Rozdělení : BINOMICKE, 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-vérohodnost		-63.8324	

Příklad 8.3/str. 287 (starší 243)

Data Phrynx.sta

STATISTICA Cz - Data: 83Koureni.sta (4s krát 93F)				
Soubor Upravit Zobrazit Vložit Formát Statistiky				
1	2	3	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

Výberte závislou proměnnou a spojité nezávislé proměnné (regresory):

1 - Objekt 2 - Pulz 3 - Kouri 4 - Vaha

1 - Objekt 2 - Pulz 3 - Kouri 4 - Vaha

1 - Objekt 2 - Pulz 3 - Kouri 4 - Vaha

1 - Objekt 2 - Pulz 3 - Kouri 4 - Vaha

OK Storno [Svažky] ...

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

Závislá proměnná: Kateg. nezáv. (faktory): Spojité nezávislé proměnné: Čítací proměnná

Ukázat pouze odpovídající proměnné

GZ Vlastní obecné schéma: 83Koureni.sta

Základní nastavení | Detaily |

Proměnné

Závislá proměnná: Pulz

Cílová proměnná: Pulz

Výberte dva kódy pro binomické odkazy

Kateg. nezáv.: Pulz: Vše Přiblížit OK Možnosti

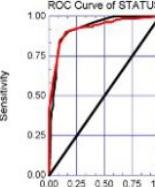
Spojité nezáv.: Odezvy: 0-1 Storno

Kódy faktoru: zadné Kouri + Vaha

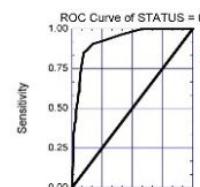
Kódů skupin: Kouri + Vaha

Editor syntaxe

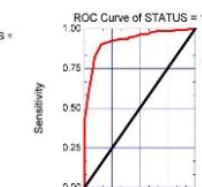
Závěr: V logistickém regresním modelu byly dokázány pouze dva statisticky významné znaky **AGE** a **TIME**, především ovlivňující přežití pacienta. Ostatní znaky jsou proto z lékařského hlediska prokázány za statisticky nevýznamné.



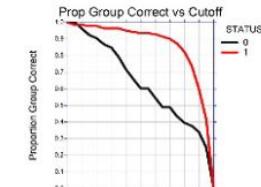
Obr. 8.6a ROC graf pro obě hodnoty STATUS = 0 a 1.



Obr. 8.6b ROC graf pro STATUS = 0.



Obr. 8.6c ROC graf pro STATUS = 1.



Obr. 8.6d Graf obou křivek podél 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 logistický 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 .

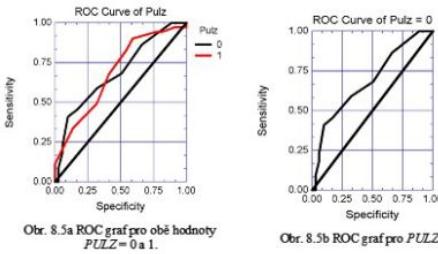
Pulz - Odhady parametrů (83Koureni.sta)					
Rozdělení : BINOMICKE, Linkující funkce: LOGIT					
Modelovaná pravděpodobnost, že Pulz = 0					
Efekt	Uroven	Sloupec	Odhad	Standard chyba	Wald. Stat.
Abs. člen		1	2.94291	1.885121	2.437119 0.118494
Kouri		2	-1.24740	0.551060	5.124017 0.023597
Vaha		3	-0.02323	0.012404	3.505869 0.061152
Měřítko			1.00000	0.000000	

Pulz - Intervaly spolehlivost odhadů (83Koureni.sta)					
Rozdělení : BINOMICKE, Linkující funkce: LOGIT					
Modelovaná pravděpodobnost, že Pulz = 0					
Efekt	Uroven	Sloupec	Iterace 0	Iterace 1	Iterace 2
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
Vérohodnost			-64.4627	-47.7906	-47.1722 -47.1611 -47.1611

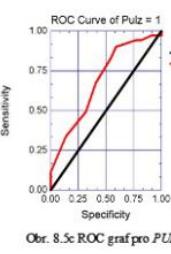
Pulz - Vérohodn. test typ 3 (83Koureni.sta)				
Rozdělení : BINOMICKE, Linkující funkce: LO				
Modelovaná pravděpodobnost, že Pulz = 0				
Efekt	Stupně volnosti	Ln-vérohod.	Chi-kvadrát	p
Kouri	1	-49.7638	5.205374	0.022517
Vaha	1	-49.0903	3.858456	0.049496

Pulz - Statistiky kvality proložení (83Koureni.sta)		
Rozdělení : BINOMICKE, Linkující funkce: LO		
Modelovaná pravděpodobnost, že Pulz = 0		
Stat.	SV	Stat.
Odhylka	90	94.3222
Deviance v měřítku	90	94.3222
Pearsonovo Chi2	90	90.0931
Scaled P. Chi2	90	90.0931
Log-vérohodnost		-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 podél 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* bliží 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 vycíslované pravděpodobnosti.

TREATMENT - Odhad parametrů (B35)						
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT						
Modelovaná pravděpodobnost, že TREATMENT = 0						
Efekt	Urovení 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	-0,001798	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	0,000000	0,000279	-0,000003	0,000028 0,986669
RELATION4	7	0,000120	0,000002	2,167761	-0,00004	0,000281 0,140931
Měřítko		1,000000	0,000000		1,000000	1,000000

x1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo zařazení placeba a léku D kódem 2,
x2 (PERIOD) značí sledovanou periodu 1 nebo 2,
x3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),
x4 (RELATION1) značí vztah EE - AUC (pg*hr/ml),
x5 (RELATION2) značí vztah EE - Cmax (pg/ml),
x6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),
x7 (RELATION4) značí vztah NET - Cmax (pg/ml).

TREATMENT - Intervaly spolehlivost odhadů (B35)			
Rozdělení: BINOMICKÉ, Linkující funkce: LOGIT			
Modelovaná pravděpodobnost, že TREATMENT = 0			
Efekt	Urovení 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

TREATMENT - Historie iterací (B35)						
Rozdělení: BINOMICKE, Linkující funkce: LOGIT						
Modelovaná pravděpodobnost, že TREATMENT = 0						
Efekt	Urovení 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
Vérohodnost				-30,4985	-25,4792	-25,2636 -25,2579 -25,2578

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 pilulkami orální antikoncepcie. 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 antikoncepcie ve dvou sledovaných časových periodách. Orální antikoncepcie se skládala ze dvou komponent, ethinyl estradiol EE a norethindron NET. Vyšetřete, zda závisle proměnná **TREATM**, tj. přítomnost léku D x4 ovlivnila míru absoruce substance do organismu u každé z obou komponent antikoncepcie pro vyjádření vztahu mezi závisle proměnnou léčením placebem nebo lékem D x4 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,

x1 (TREATM) značí pořadí zařazení léku **D** a placeba kódem 1 nebo placebo a léku **D** kódem 2,

x2 (PERIOD) značí sledovanou periodu 1 nebo 2,

x3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),

x4 (RELATION1) značí vztah EE - AUC (pg*hr/ml),

x5 (RELATION2) značí vztah EE - Cmax (pg/ml),

x6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),

x7 (RELATION4) značí vztah NET - Cmax (pg/ml).

x1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo zařazení placebo a léku D kódem 2,

x2 (PERIOD) značí sledovanou periodu 1 nebo 2,

x3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),

x4 (RELATION1) značí vztah EE - AUC (pg*hr/ml),

x5 (RELATION2) značí vztah EE - Cmax (pg/ml),

x6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),

x7 (RELATION4) značí vztah NET - Cmax (pg/ml).

x1 (TREATM) značí pořadí zařazení léku D a placeba kódem 1 nebo zařazení placebo a léku D kódem 2,

x2 (PERIOD) značí sledovanou periodu 1 nebo 2,

x3 (TREATMENT) léčení placebem (kód 0) nebo lékem D (kód 1),

x4 (RELATION1) značí vztah EE - AUC (pg*hr/ml),

x5 (RELATION2) značí vztah EE - Cmax (pg/ml),

x6 (RELATION3) značí vztah NET - AUC (pg*hr/ml),

x7 (RELATION4) značí vztah NET - Cmax (pg/ml).

TREATMENT - Vérohodn. test typ 1 (B35)

Rozdělení: BINOMICKE, Linkující funkce: LOGIT

Modelovaná pravděpodobnost, že TREATMENT = 0

Efekt	Stupně volnosti	Ln-vérohod.	Chi-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

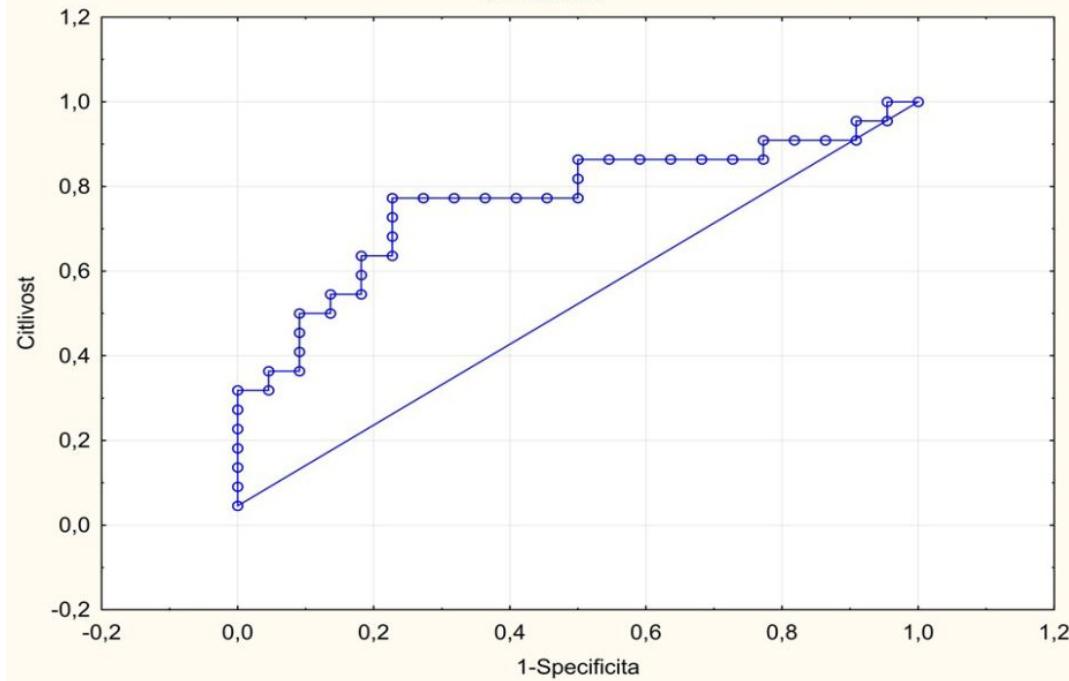
TREATMENT - Vérohodn. test typ 3 (B35)

Rozdělení: BINOMICKE, Linkující funkce: LOGIT

Modelovaná pravděpodobnost, že TREATMENT = 0

Efekt	Stupně volnosti	Ln-vérohod.	Chi-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



Čísla případů dle reziduí

