



Testování předpokladů pro metodu chain-ladder

Seminář z aktuárských věd

4. 11. 2016

Petra Španihelová



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Obsah

- ☐ Datová struktura
- ☐ Posouzení dat
- ☐ Předpoklady metody chain-ladder dle T. Macka
- ☐ Běžná lineární regrese
- ☐ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převezení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Ukázka na datech
- ☐ Shrnutí



Datová struktura



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Kumulativní trojúhelník

- ▶ Sloupce reprezentují zpoždění v nahlášení škody

$$j = 0, \dots, n-1 \quad \text{vývojové roky}$$

- ▶ Řádky reprezentují rok vzniku škody

$$i = 1, \dots, n \quad \text{vznikové roky}$$

- ▶ Prvky trojúhelníku $X_{i,j}$ reprezentují celkový úhrn výší škod vzniklých v roce i a nahlášených do roku $i+j$

- ▶ Vývoj škod je ukončen po n letech

$$X_{i,n-1} = X_{i,\infty}$$

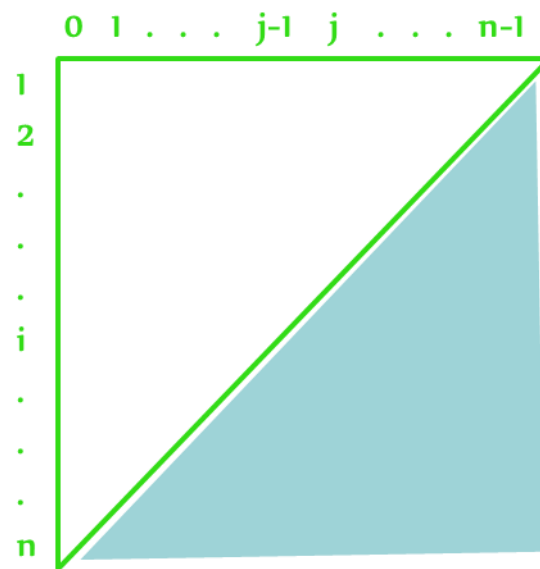
- ▶ Typy trojúhelníku

- ▶ Výplatní – výší škody rozumíme plnění

- ▶ Incurred – výší škody rozumíme plnění a RBNS rezervu

- ▶ Cílem metod rezervování je doplnit trojúhelník na čtverec

Tvorba IBNR rezervy



Obsah

- ✓ Datová struktura
- ☐ Posouzení dat
- ☐ Předpoklady metody chain-ladder dle T. Macka
- ☐ Běžná lineární regrese
- ☐ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převedení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Posouzení dat



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Posouzení dat

- ▶ Malé trojúhelníky
 - ▶ $n = 2, \dots, 5 \Rightarrow$ nadále budeme uvažovat $n > 5$
- ▶ Děravé trojúhelníky
 - ▶ Některé prvky $X_{i,j}$ jsou nulové
- ▶ Ukončený vývoj trojúhelníku
- ▶ Diagonální efekty
 - ▶ Inlace
 - ▶ Legislativní změny
- ▶ Odlehlá pozorování
 - ▶ Pro výpočet IBNR rezervy je možné očišťovat data o odlehlá pozorování – jedná se o škodu, která je odlehlá buď ve smyslu výše škody, nebo jejím umístěním v trojúhelníku, a která svým dopadem poškozuje vypovídající hodnotu rezervy

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ☐ Předpoklady metody chain-ladder dle T. Macka
- ☐ Běžná lineární regrese
- ☐ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převezení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Předpoklady metody chain-ladder dle T. Macka



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Předpoklady metody chain-ladder

Dle *Mack (1993)* považujeme $X_{i,j}$ za náhodné veličiny, kde známe jejich realizace pro $i+j \leq n$, a předpokládáme 3 vlastnosti:

1. Nezávislost roků vzniku

$\{X_{i,0}, \dots, X_{i,n-1}\}$ a $\{X_{j,0}, \dots, X_{j,n-1}\}$ jsou vzájemně nezávislé $i \neq j$

2. Úměrnost sloupců kumulativního trojúhelníku

Existují $b_0, \dots, b_{n-2} > 0$ takové, že pro $i = 1, \dots, n$ a $j = 0, \dots, n-1$ platí

$$E[X_{i,j+1} | X_{i,0}, \dots, X_{i,j}] = X_{i,j} b_j$$

Za platnosti 1. a 2. předpokladu lze odvodit nestranný odhad parametrů b_j pro $j = 0, \dots, n-2$ vztahem

$$\hat{b}_j = \frac{\sum_{i=1}^{n-j-1} X_{i,j+1}}{\sum_{i=1}^{n-j-1} X_{i,j}}$$

Předpoklady metody chain-ladder

3. Předpoklad o tvaru rozptylu

Existují $\sigma_0, \dots, \sigma_{n-1} > 0$ takové, že
pro $i = 1, \dots, n$ a $j = 0, \dots, n-1$ platí

$$\text{Var}[X_{i,j+1} | X_{i,0}, \dots, X_{i,j}] = X_{i,j} \sigma_j^2$$

Za platnosti 1., 2. a 3. předpokladu je nestranný odhad parametrů σ_j^2 pro $j = 0, \dots, n-2$ dán vztahem

$$\hat{\sigma}_j^2 = \frac{1}{n-j-2} \sum_{i=1}^{n-j-1} X_{i,j} \left(\frac{X_{i,j+1}}{X_{i,j}} - \hat{b}_j \right)^2$$
$$\hat{\sigma}_{n-1}^2 = 0$$

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ☐ Běžná lineární regrese
- ☐ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převezení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Běžná lineární regrese



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Běžná lineární regrese

Dle Anděl (2002) je běžný lineární regresní model popsán následovně:

- ▶ Mějme náhodné veličiny $Y^s_1, \dots, Y^s_n$ a matici daných čísel $X = (x_{i,j})$ typu $n \times k$, kde $k < n$.
- ▶ Nechť pro náhodný vektor $Y^s = (Y^s_1, \dots, Y^s_n)'$ platí $Y^s = Xb^s + \varepsilon$, kde
 - ▶ $b^s = (b^s_0, \dots, b^s_{k-1})'$ je vektor neznámých parametrů
 - ▶ $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ je náhodný vektor splňující $E\varepsilon = 0$ a $\text{Var}\varepsilon = \sigma^2 I$
 - ▶ $\sigma > 0$ je neznámý parametr
- ▶ Požaduje se, aby matice X měla lineárně nezávislé sloupce
 - ▶ $X'X$ je regulární matice typu $k \times k$
- ▶ Běžný lineární regresní model značíme

$$Y^s \sim (Xb^s, \sigma^2 I)$$

- ▶ Platí $E\mathbf{Y}^s = X\mathbf{b}^s$ a $\text{Var}\mathbf{Y}^s = \sigma^2 I$
- ▶ Hledáme-li odhady $\hat{\mathbf{b}}^s = (\hat{b}_{0^s}^s, \dots, \hat{b}_{k-1}^s)'$ parametrů $b_{0^s}^s, \dots, b_{k-1}^s$, pak metodou nejmenších čtverců získáme nestranný odhad

$$\hat{\mathbf{b}}^s = (X'X)^{-1}X'\mathbf{Y}^s$$
- ▶ Vektor $\hat{\mathbf{Y}}^s = X\hat{\mathbf{b}}^s$ značí nejlepší aproximaci vektoru $\mathbf{Y}^s$
- ▶ Vektor $(\mathbf{Y}^s - \hat{\mathbf{Y}}^s)$ nazýváme vektorem reziduí a reziduální součet čtverců značíme

$$S_{\varepsilon}^s = (\mathbf{Y}^s - \hat{\mathbf{Y}}^s)'(\mathbf{Y}^s - \hat{\mathbf{Y}}^s)$$

- ▶ Nestranným odhadem parametru σ^2 je náhodná veličina

$$\hat{\sigma}^{s2} = \frac{1}{n-k} S_{\varepsilon}^s$$

- ▶ Za předpokladu, že $\varepsilon \sim N(0, \sigma^2 I)$, platí

$$\mathbf{Y}^s \sim N(X\mathbf{b}^s, \sigma^2 I)$$

$$\hat{\mathbf{b}}^s \sim N(\mathbf{b}^s, \sigma^2 (X'X)^{-1})$$

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ✓ Běžná lineární regrese
- ☐ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převedení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Zobecněná lineární regrese (Aitkenův model)



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Zobecněná lineární regrese (Aitkenův model)

- ▶ Zobecněný (Aitkenův) lineární regresní model budeme značit

$$\mathbf{Y} \sim (Xb, \sigma^2 W^{-1})$$

- ▶ W je pozitivně definitní matice konstant
- ▶ Platí $E\mathbf{Y} = Xb$ a $\text{Var}\mathbf{Y} = \sigma^2 W^{-1}$
- ▶ W pozitivně definitní $\Rightarrow$ existuje regulární C : $C'C = W$
- ▶ Označíme $X^* = CX$ a $Y^* = CY$. Využijeme-li $I = CW^{-1}C'$, pak Y^* vyhovuje běžnému lineárnímu modelu

$$\mathbf{Y}^* \sim (CXb, \sigma^2 CW^{-1}C') = (X^*b, \sigma^2 I)$$

- ▶ Reziduální součet čtverců lze zapsat ve tvaru

$$S_\epsilon = \|\mathbf{Y}^* - \hat{\mathbf{Y}}^*\|^2 = \|C\mathbf{Y} - C\hat{\mathbf{Y}}\|^2 = (\mathbf{Y} - \hat{\mathbf{Y}})'W(\mathbf{Y} - \hat{\mathbf{Y}})$$

- ▶ Pro parametry $b_0, \dots, b_{k-1}$ získáme zobecněnou metodou nejmenších čtverců nestranný odhad $\hat{b} = (\hat{b}_0, \dots, \hat{b}_{k-1})'$

$$\hat{b} = (X'WX)^{-1}X'WY \quad \text{Aitkenův odhad}$$

- ▶ Odhad $\hat{Y}$ střední hodnoty EY je

$$\hat{Y} = X\hat{b} = X(X'WX)^{-1}X'WY$$

- ▶ Nestranným odhadem parametru σ^2 je náhodná veličina

$$\hat{\sigma}^2 = \frac{1}{n-k} S_\varepsilon$$

- ▶ Pro některé typy matice lze výpočet statistik zjednodušit

- ▶ Uvažujme diagonální matici

$$W = (w_{i,j}) = 0 \text{ pro } i \neq j$$

Speciální případ: diagonální matice W

- ▶ Reziduální součet čtverců lze zjednodušit na

$$S_{\varepsilon} = \sum_{i=1}^n w_{i,i} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n w_{i,i} \left(Y_i - \sum_{j=0}^{k-1} x_{i,j} \hat{b}_j \right)^2$$

- ▶ Odhad metodou vážených nejmenších čtverců je pro $j = 0, \dots, k-1$

$$\hat{b}_j = \frac{\sum_{i=1}^n w_{i,i} Y_i x_{i,j}}{\sum_{i=1}^n w_{i,i} x_{i,j}^2}$$

- ▶ Pro odhad rozptylu pak platí

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n w_{i,i} (Y_i - \hat{Y}_i)^2 = \frac{1}{n-k} \sum_{i=1}^n w_{i,i} \left(Y_i - \sum_{j=0}^{k-1} x_{i,j} \hat{b}_j \right)^2$$

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ✓ Běžná lineární regrese
- ✓ Zobecněná lineární regrese (Aitkenův model)
- ☐ Převedení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Převedení metody chain-ladder na úlohu lineární regrese

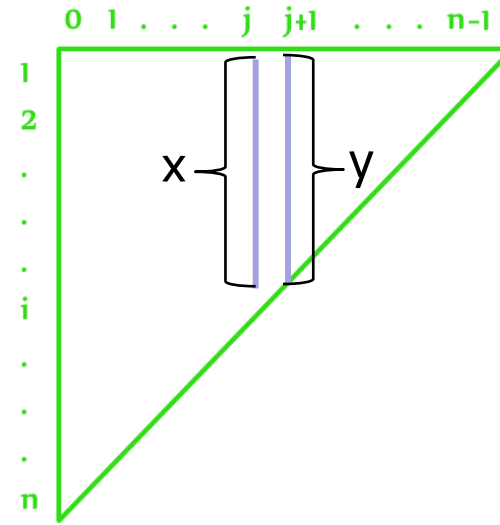


Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Převedení metody chain-ladder na úlohu lineární regrese

- ▶ Dle *Barnett & Zehnwirth (2000)* definujeme
 - ▶ $x = (x(1), \dots, x(n-j-1))'$
 - ▶ $y = (y(1), \dots, y(n-j-1))'$
- ▶ Pak pro $j = 1, \dots, n-3$ pevně značí
 - ▶ $x(i)$, $i = 1, \dots, n-j-1$, kumulativní úhrny škod ve vývojovém roce j a roce vzniku i
 - ▶ $y(i)$ odpovídají kumulativní úhrny škod ve vývojovém roce $j+1$ a roce vzniku i
- ▶ V obvyklém značení je tedy
 - ▶ $x = (X_{1,j}, \dots, X_{n-j-1,j})'$
 - ▶ $y = (X_{1,j+1}, \dots, X_{n-j-1,j+1})'$



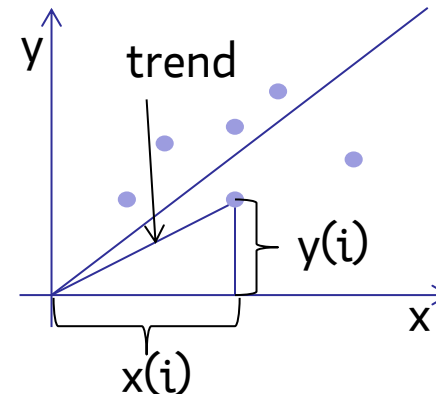
Graf závislosti kumulativních škod na předchozím vývojovém roce

- Vykreslíme graf závislosti y na x
- Poměr $y(i) / x(i)$ udává sklon přímky procházející počátkem a bodem $[x(i), y(i)]$, tj. každý poměr $y(i) / x(i)$ udává trend
- K hledání průměrného trendu pro všechna $i = 1, \dots, n - j - 1$ použijeme lineární regresi

$$y(i) = b x(i) + \varepsilon(i),$$

$$\text{kde } \text{Var}[\varepsilon(i)] = \sigma^2 x(i)^\delta$$

- Parametr b reprezentuje sklon „nejlepší“ přímky procházející počátkem a body $[x(i), y(i)]$, pro všechna $i = 1, \dots, n - j - 1$
- Rozptyl $y(i)$ okolo přímky procházející počátkem závisí na $x(i)$ prostřednictvím funkce $x(i)^\delta$, kde δ je vážící parametr
- Pro $j = n - 2$ odhadneme sklon b jako $\hat{b} = y(i) / x(i)$



Zpět k Aitkenově modelu

- ▶ $Y \sim (Xb, \sigma^2 W^{-1})$
 - ▶ X je matice typu $n \times k$, $k < n$
 - ▶ b je vektor o k prvcích
 - ▶ W je matice typu $n \times n$
- ▶ Konkrétně pro pevné j :
 - ▶ X je sloupec vývojového trojúhelníku, tedy typu $(n - j - 1) \times 1$, kde $1 < n - j - 1$
 - ▶ b je jeden prvek (b_j)
 - ▶ W je diagonální matice, kde

$$W = (w_{i,k}) = \begin{cases} x(i)^{-\delta}, & i = k \\ 0, & i \neq k \end{cases}$$

Odhady parametru b

- Do Aitkenova odhadu

$$\hat{b}_j = \frac{\sum_{i=1}^n w_{i,i} Y_i x_{i,j}}{\sum_{i=1}^n w_{i,i} x_{i,j}^2}$$

dosadíme za váhy $w_{i,i}$

$$\hat{b} = \frac{\sum_{i=1}^{n-j-1} x(i)^{1-\delta} y(i)}{\sum_{i=1}^{n-j-1} x(i)^{2-\delta}}$$

- Pro $\delta = 0$ dostaneme odhad pomocí běžné lineární regrese

$$\hat{b}_j = \frac{\sum_{i=1}^{n-j-1} X_{i,j} X_{i,j+1}}{\sum_{i=1}^{n-j-1} X_{i,j}^2}$$

- Pro $\delta = 1$ dostaneme odhad metodou chain-ladder

$$\hat{b}_j = \frac{\sum_{i=1}^{n-j-1} X_{i,j+1}}{\sum_{i=1}^{n-j-1} X_{i,j}}$$

Odhady parametru b a σ

- Pro $\delta = 2$ dostaneme aritmetický průměr

$$\hat{b}_j = \frac{1}{n-j-1} \sum_{i=1}^{n-j-1} \frac{X_{i,j+1}}{X_{i,j}}$$

- Do odhadu rozptylu z Aitkenova modelu

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n w_{i,i} \left(Y_i - \sum_{j=0}^{k-1} x_{i,j} \hat{b}_j \right)^2$$

dosadíme $\delta = 1$ do váhy $w_{i,i} = 1/X_{i,j}$ pro pevné j

$$\hat{\sigma}_j^2 = \frac{1}{n-j-2} \sum_{i=1}^{n-j-1} X_{i,j} \left(\frac{X_{i,j+1}}{X_{i,j}} - \hat{b}_j \right)^2$$

Tvar matice C z Aitkenova modelu

- ▶ Najdeme matici C, aby platilo $C'C = W$
 - ▶ Pro pevné j hledáme odmocninovou matici C_j
- ▶ Pro $\delta = 1$ máme:

$$X = \begin{pmatrix} X_{1,j} \\ \vdots \\ X_{n-j-1,j} \end{pmatrix}, Y = \begin{pmatrix} X_{1,j+1} \\ \vdots \\ X_{n-j-1,j+1} \end{pmatrix}, C = \begin{pmatrix} \frac{1}{\sqrt{X_{1,j}}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\sqrt{X_{n-j-1,j}}} \end{pmatrix}$$

- ▶ Vektory CX a CY jsou pak

$$CX = \left(\frac{X_{1,j}}{\sqrt{X_{1,j}}}, \dots, \frac{X_{n-j-1,j}}{\sqrt{X_{n-j-1,j}}} \right)^T, CY = \left(\frac{X_{1,j+1}}{\sqrt{X_{1,j}}}, \dots, \frac{X_{n-j-1,j+1}}{\sqrt{X_{n-j-1,j}}} \right)^T$$

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ✓ Běžná lineární regrese
- ✓ Zobecněná lineární regrese (Aitkenův model)
- ✓ Převedení metody chain-ladder na úlohu lineární regrese
- ☐ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Testování předpokladů



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Testování předpokladů

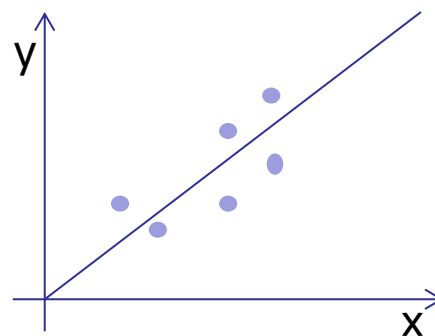
- Pro všechna $j = 0, \dots, n-1$ si upravíme vývojový trojúhelník tak, aby pro jednotlivé sloupce trojúhelníku platilo $X^* = CX$

i/j	0	...	$j+1$	...	$n-1$
1	$\frac{X_{1,0}}{\sqrt{X_{1,0}}}$	...	$\frac{X_{1,j+1}}{\sqrt{X_{1,j+1}}}$	...	$\frac{X_{1,n-1}}{\sqrt{X_{1,n-1}}}$
$\vdots$	$\vdots$	...	$\vdots$	...	
i	$\frac{X_{i,0}}{\sqrt{X_{i,0}}}$	...	$\frac{X_{i,j+1}}{\sqrt{X_{i,j+1}}}$	...	
$\vdots$	$\vdots$	...			
n	$\frac{X_{n,0}}{\sqrt{X_{n,0}}}$				

- Pro následující kroky budeme uvažovat výše popsany trojúhelník

Úměrnost sloupců

- ▶ Vykreslíme graf závislosti y na x a sledujeme, zda regresní přímka prochází počátkem.
- ▶ V lineární regresi bez absolutního členu, kde odhad parametru b odpovídá chain-ladder odhadu, budeme testovat významnost koeficientu na dané hladině. Požadujeme, aby koeficient byl významný.
- ▶ Uvažujeme obvyklé hladiny
 $\alpha = 0,05$, nebo $\alpha = 0,01$
- ▶ Požadujeme, aby b byl významný,
tj. $p\text{-value}_b < \alpha$



Úměrnost sloupců

- ▶ V lineární regresi s absolutním členem testujeme na dané hladině významnost absolutního členu. Jeho významnost ve více obdobích může svědčit o vychýlení výsledného odhadu rezerv, viz *Murphy (1993)*.

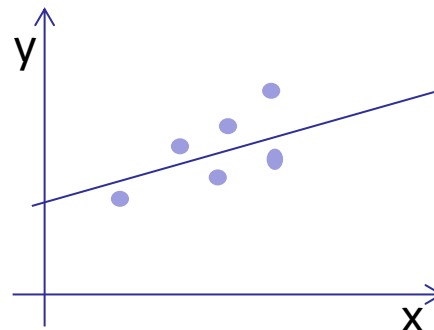
- ▶ Regresní model má pro pevné j a $i = 1, \dots, n-j-1$, tvar

$$y(i) = a + b x(i) + \varepsilon(i)$$

- ▶ V Aitkenově modelu je pak nevážená matice X tvaru

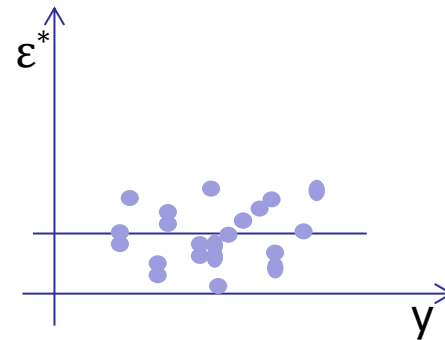
$$X = \begin{pmatrix} 1 & X_{1,j} \\ \vdots & \vdots \\ 1 & X_{n-j-1,j} \end{pmatrix}$$

- ▶ Požadujeme, aby a byl nevýznamný, tj. $p\text{-value}_a > \alpha$



Tvar rozptylu

- ▶ Vykreslíme graf, kde na vodorovné ose vyznačíme predikované hodnoty a na svislé ose odpovídající hodnoty standardizovaných reziduí.
- ▶ Případný trend je testován pomocí lineární regrese. Není-li v grafu pozorovatelný trend, můžeme předpoklad o tvaru rozptylu považovat za splněný.
- ▶ Požadujeme, aby trend byl nevýznamný, tj. $p\text{-value} > \alpha$

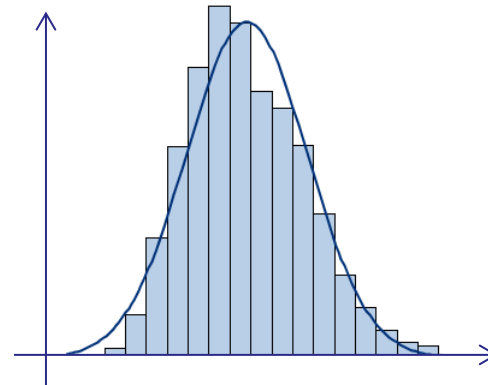


Tvar rozptylu

- ▶ Standardizovaná rezidua by měla přibližně vykazovat normalitu, tj.

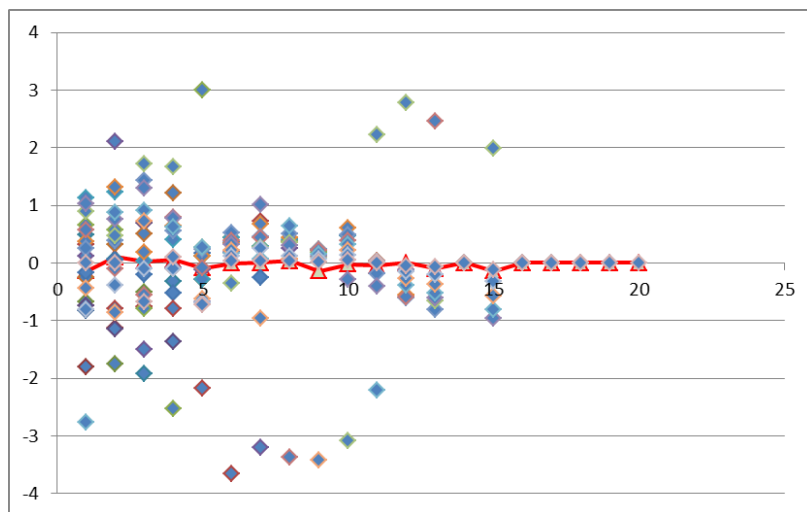
$$\frac{\varepsilon(i)}{\sigma \sqrt{x(i)}} \sim N(0,1)$$

- ▶ Lze testovat několika způsoby, např.:
 - ▶ Histogram: Porovnáme pozorované četnosti s očekávanými četnostmi při standardním normálním rozdělení
 - ▶ Shapiro-Wilkův test: Při nulové hypotéze má výběr standardní normální rozdělení



Celková vhodnost

- ▶ O celkové vhodnosti metody chain-ladder se přesvědčíme analýzou standardizovaných reziduí. Rezidua vykreslíme v grafech dle období vzniku, vývoje a výplatního období
- ▶ Vyznačíme a čarou spojíme průměry reziduí v jednotlivých obdobích
- ▶ Sledujeme, zda se v průměrech neprojevuje trend
- ▶ Je-li model vhodně zvolen, pak standardizovaná rezidua i jejich průměry by měly kmitat náhodně okolo horizontální osy



Problémy

Dle *Barnett & Zehnwirth (2000)*

- ▶ Graf závislosti vážených standardizovaných reziduí na predikovaných hodnotách vykazuje trend
 - ▶ Např. pro klesající trend graf indikuje, že roky vzniku, které mají vysoké kumulativní hodnoty, jsou nadhodnoceny, naopak roky vzniku s nízkými kumulativními hodnotami jsou podhodnoceny
 - ▶ Zjišťujeme významnost trendu
- ▶ Absolutní člen lineární regrese je významný
 - ▶ Pro kolik vývojových let a které?
 - ▶ Jak moc je významný?
 - ▶ Možnost jiné „poměrové“ metody (Cape Cod)

Problémy

- ▶ Rezidua nevykazují normalitu
- ▶ Rezidua a jejich průměry vykazují trend podle roků vzniku, vývoje nebo kalendářních roků
- ▶ Oba výše popsané problémy mohou vést k velmi špatným odhadům
- ▶ V těchto případech se doporučuje přejít na jinou metodu, která analyzuje přírůstky jednotlivých vývojových let, nejlépe pokud navíc zohledňuje diagonální trend

Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ✓ Běžná lineární regrese
- ✓ Zobecněná lineární regrese (Aitkenův model)
- ✓ Převod metody chain-ladder na úlohu lineární regrese
- ✓ Testování předpokladů
- ☐ Numerický příklad
- ☐ Shrnutí



Ukázka na datech



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Incurred trojúhelník úrazového pojištění dospělých osob

- Data níže jsou zamaskována
- Provedeme převážení dat – každou hodnotu vydělíme její odmocninou

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	3	5	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
5	8	16	20	22	22	23	23	23	23	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
6	9	17	24	28	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
7	11	22	31	36	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
8	11	25	36	41	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
9	15	31	46	56	58	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59
10	23	47	75	93	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98
11	30	59	97	114	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125
12	38	72	109	127	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131	131
13	48	89	138	158	165	165	166	166	166	167	167	167	167	167	167	167	167	167	167	167	167	167	167	167
14	50	95	125	148	153	154	154	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157
15	52	108	144	167	178	178	179	179	179	179	179	179	179	179	179	179	179	179	179	179	179	179	179	179
16	51	103	143	167	172	173	173	173	173	173	173	173	173	173	173	173	173	173	173	173	173	173	173	173
17	53	110	151	171	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
18	59	116	177	197	207	207	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
19	59	125	179	201	207	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208	208
20	47	95	129	146	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160	160
21	42	84	112	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125
22	37	72	97	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125
23	28	56	97	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125
24	27	56	97	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125

Úměrnost sloupců

- Pro jednotlivé vývojové roky $j = 0, \dots, n-1$ budeme postupně testovat koeficienty lineární regrese bez absolutního členu a s absolutním členem
- Zvolíme si hladinu významnosti $\alpha = 0,01$
- Lineární regrese bez absolutního členu:

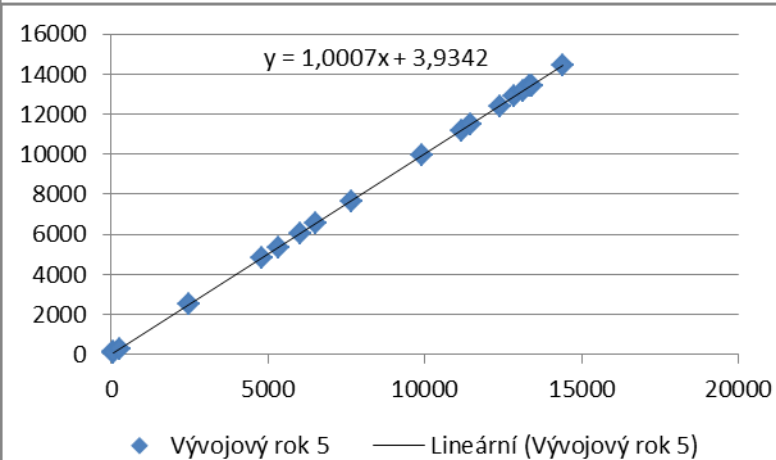
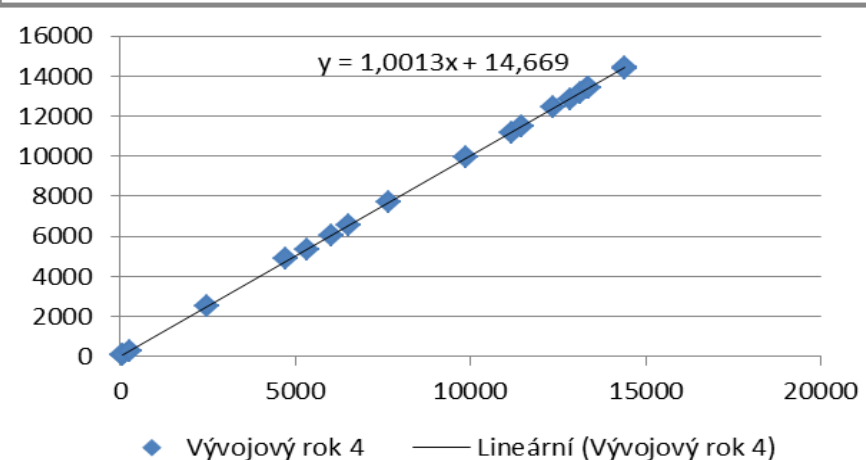
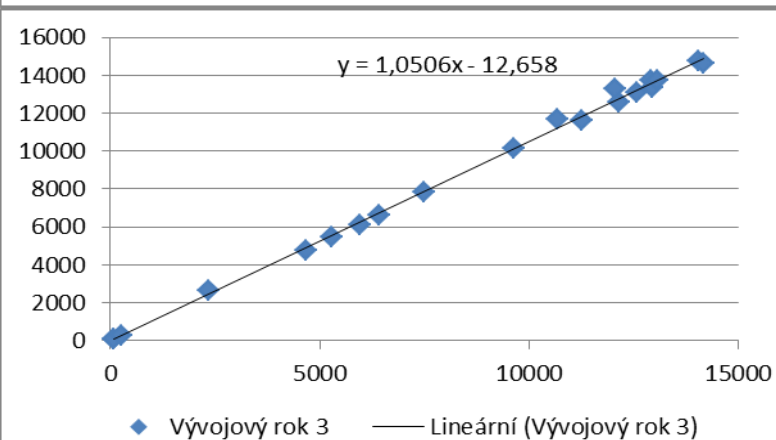
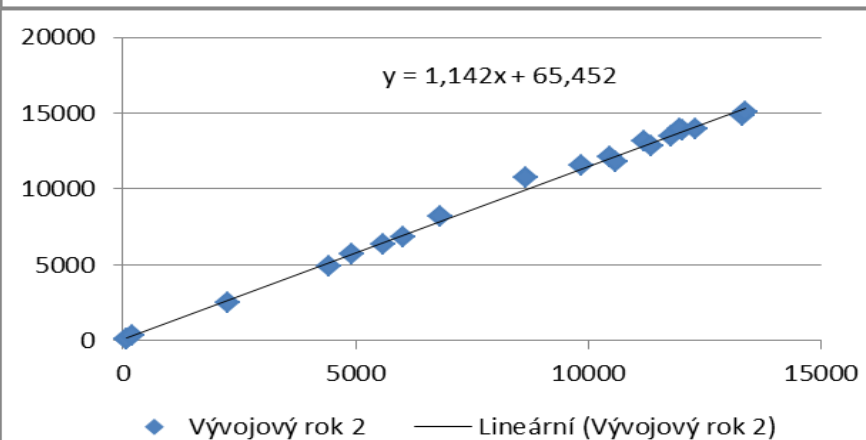
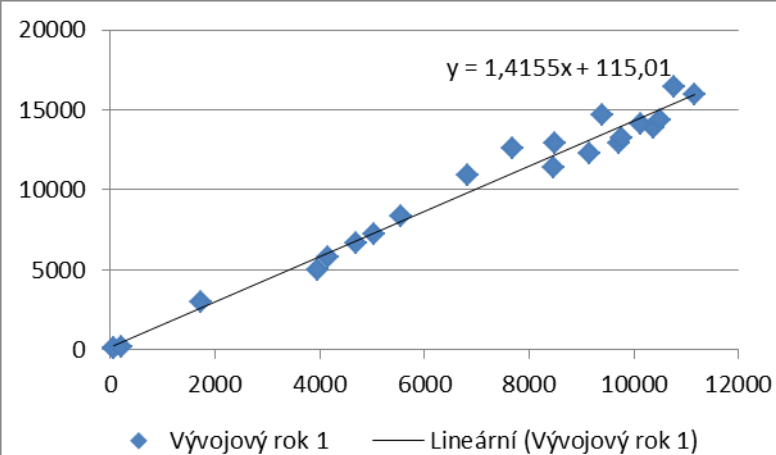
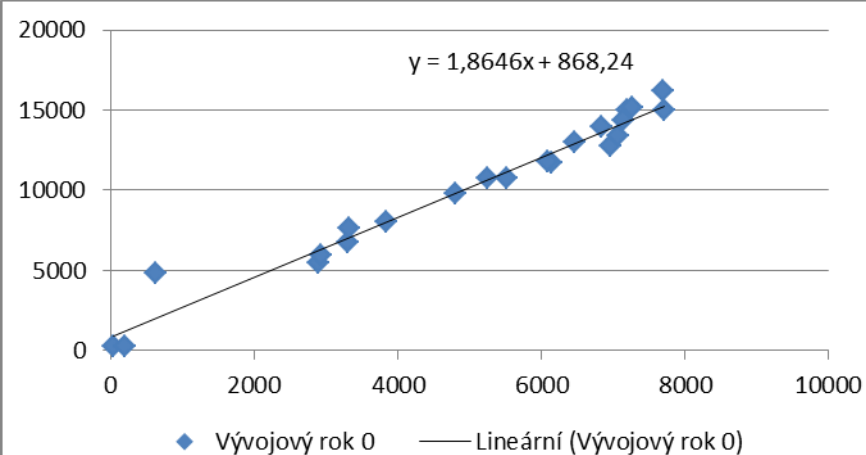
Chceme, aby $p\text{-value} < \alpha$

j	0	1	2	3	4	5	6	7	8	9
$\hat{b}$	2,01	1,43	1,15	1,05	1,00	1,00	1,00	1,00	1,00	1,00
$\hat{\sigma}$	0,04	0,02	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00
p-value	2,91E-24	5,96E-26	8,23E-33	2,37E-33	1,29E-45	4,45E-48	4,08E-36	6,70E-42	3,02E-30	6,81E-58

- Lineární regrese s absolutním členem:

Chceme, aby pro absolutní člen platilo $p\text{-value} > \alpha$

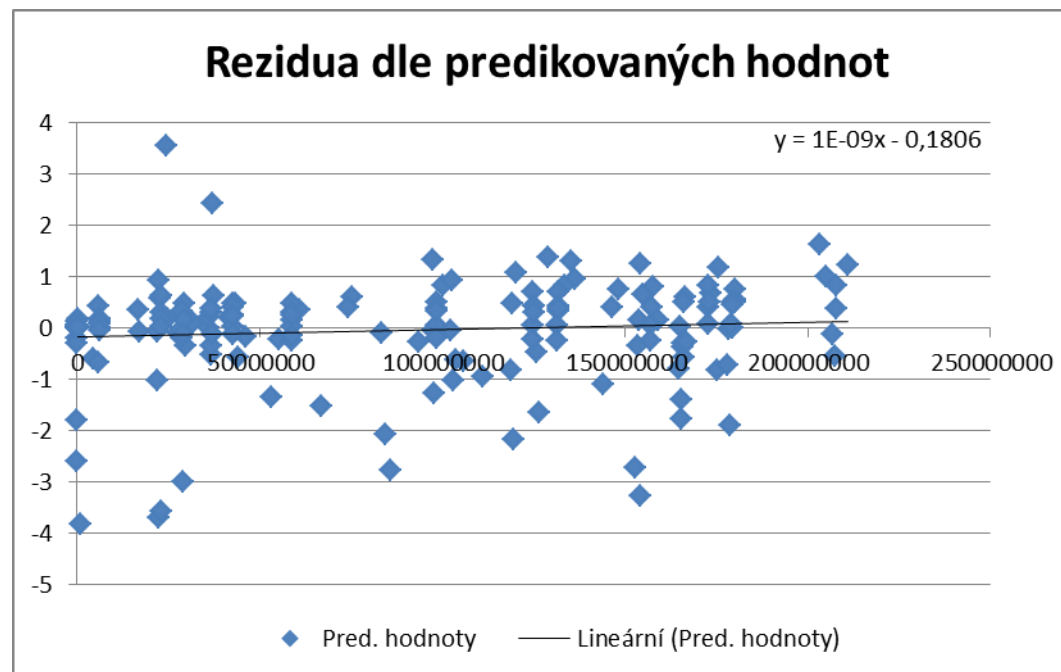
j	0	1	2	3	4	5	6	7	8	9	10
$\hat{b}$	1,86	1,42	1,14	1,05	1,00	1,00	1,00	1,00	1,00	1,00	1,00
$\hat{\sigma}_b$	0,08	0,05	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
p-value_b	2,76E-16	2,42E-18	2,68E-25	3,25E-26	2,43E-38	8,22E-41	6,86E-30	1,93E-35	5,77E-25	1,11E-50	1,91E-173
$\hat{a}$	868,24	115,01	65,45	-12,66	14,67	3,93	23,02	1,45	31,19	-0,15	0,00
$\hat{\sigma}_a$	425,32	351,09	133,16	101,18	16,02	8,18	29,66	8,66	35,68	0,18	0,00
p-value_a	5,5%	74,7%	62,9%	90,2%	37,3%	63,7%	45,0%	86,9%	39,8%	42,2%	10,1%



Tvar rozptylu

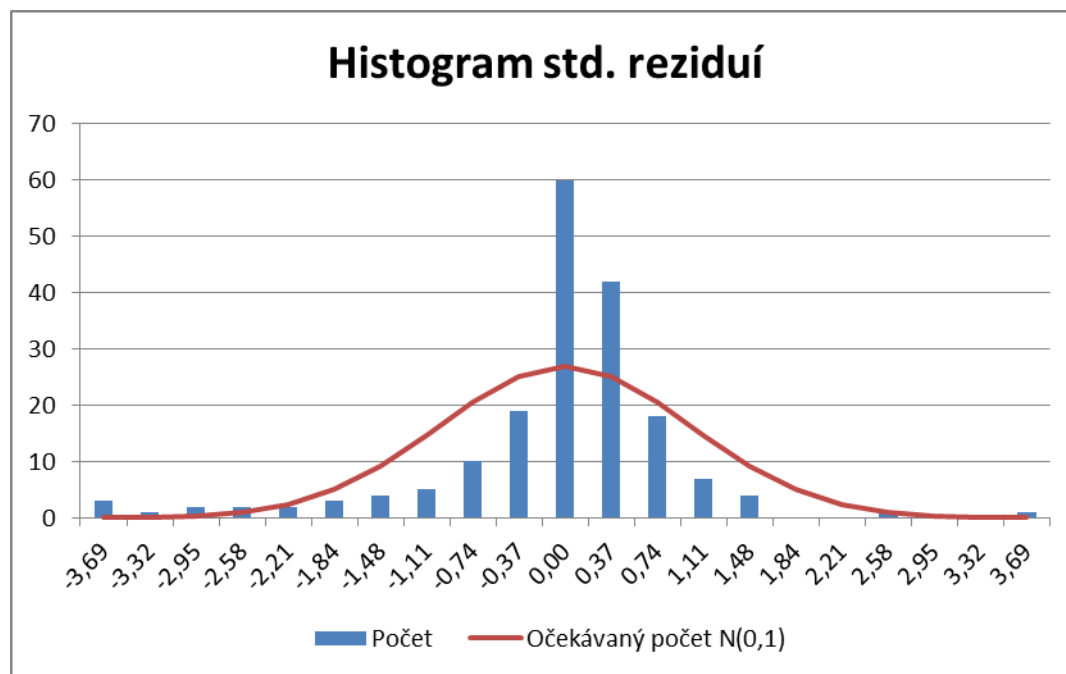
- Vykreslíme graf, kde na vodorovné ose vyznačíme predikované hodnoty a na svislé ose odpovídající hodnoty standardizovaných reziduí
- Požadujeme, aby trend byl nevýznamný, tj. chceme, aby $p\text{-value}_b > \alpha$

$\hat{b}$	1,48E-09
$\hat{\sigma}_b$	1,07E-09
p-value_b	16,7%
$\hat{a}$	-1,81E-01
$\hat{\sigma}_a$	1,13E-01



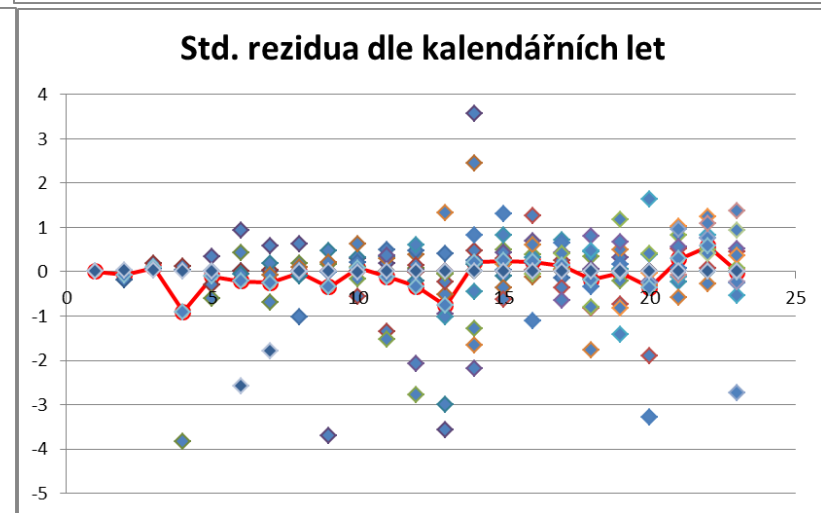
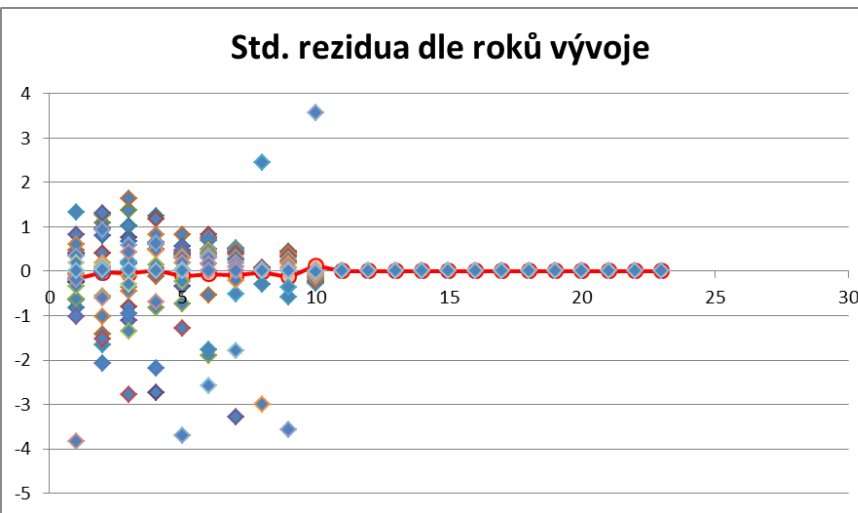
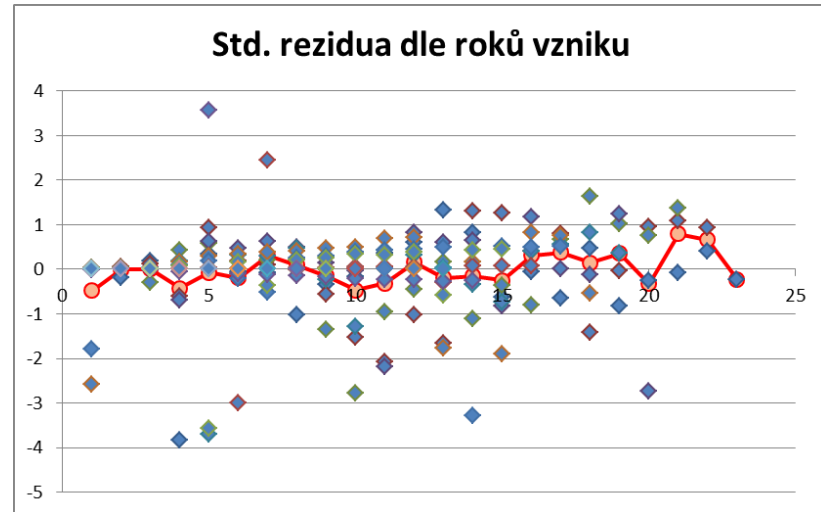
Tvar rozptylu – test normality

- ▶ Standardizovaná rezidua by měla přibližně vykazovat normalitu
- ▶ Četnosti výskytů rozděleny do 21 intervalů



Celková vhodnost

- ▶ Sledujeme, zda rezidua a jejich průměry náhodně kmitají kolem horizontální osy
- ▶ Graf dle roků vzniku
- ▶ Graf dle roků vývoje
- ▶ Graf dle kalendářních let

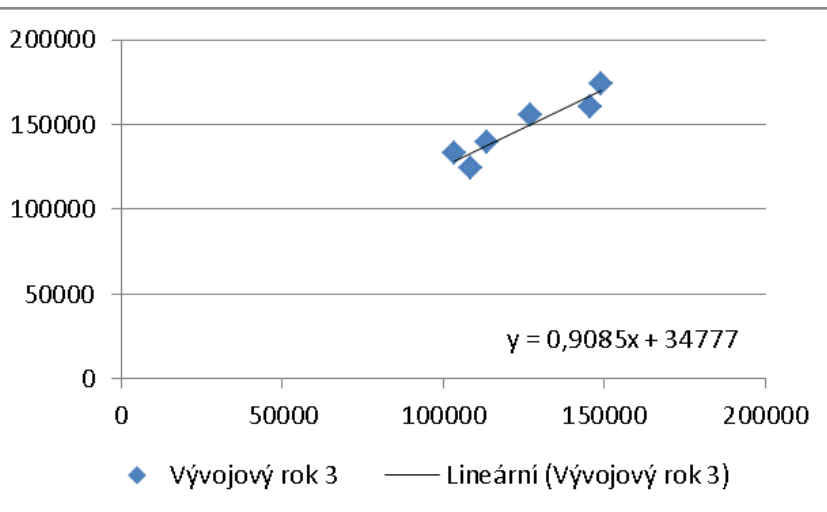
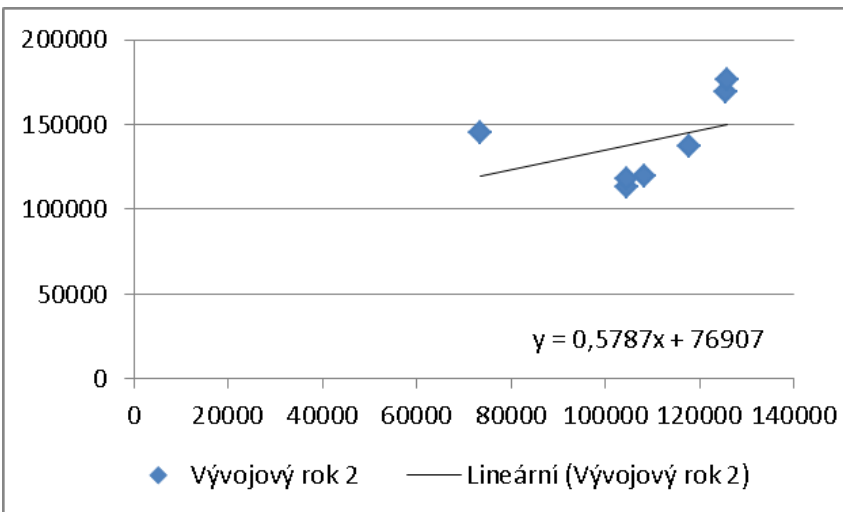
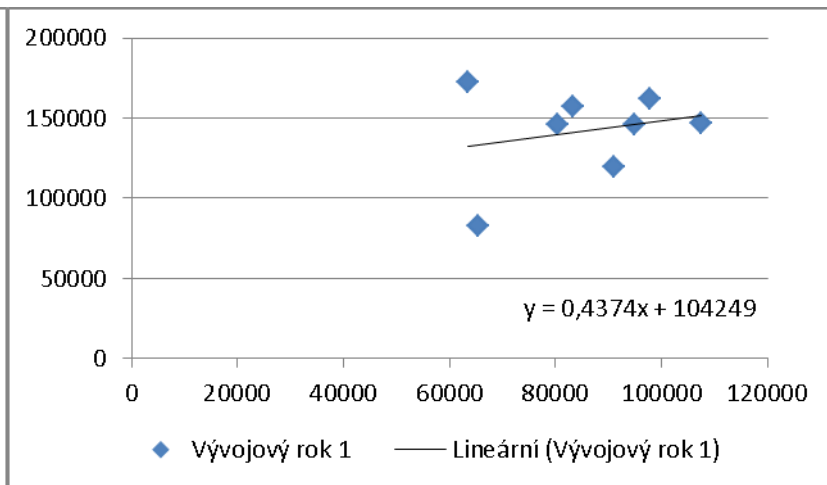
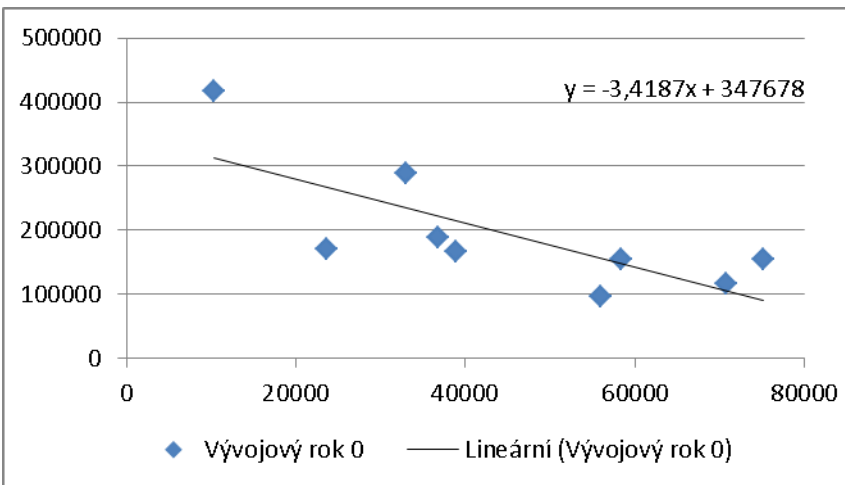


Mackova data

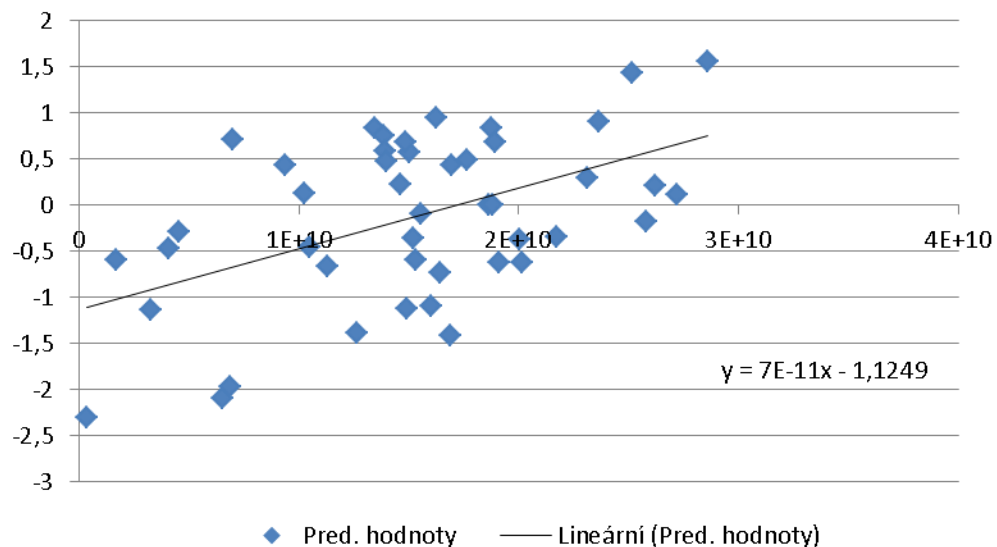
Mack	0	1	2	3	4	5	6	7	8	9
1981	5 012	8 269	10 907	11 805	13 539	16 181	18 009	18 608	18 662	18 834
1982	106	4 285	5 396	10 666	13 782	15 599	15 496	16 169	16 704	
1983	3 410	8 992	13 873	16 141	19 735	22 214	22 863	23 466		
1984	5 655	11 555	15 766	21 266	23 425	26 083	27 067			
1985	1 092	9 565	15 836	22 169	25 955	26 180				
1986	1 513	6 445	11 702	12 935	15 852					
1987	557	4 020	10 946	12 314						
1988	1 351	6 947	13 112							
1989	3 133	5 395								
1990	2 063									

j	0	1	2	3
$\hat{b}$	3,00	1,62	1,27	1,18
$\hat{\sigma}$	1,13	0,14	0,09	0,03
p-value	2,91E-02	6,53E-06	8,13E-06	1,09E-07

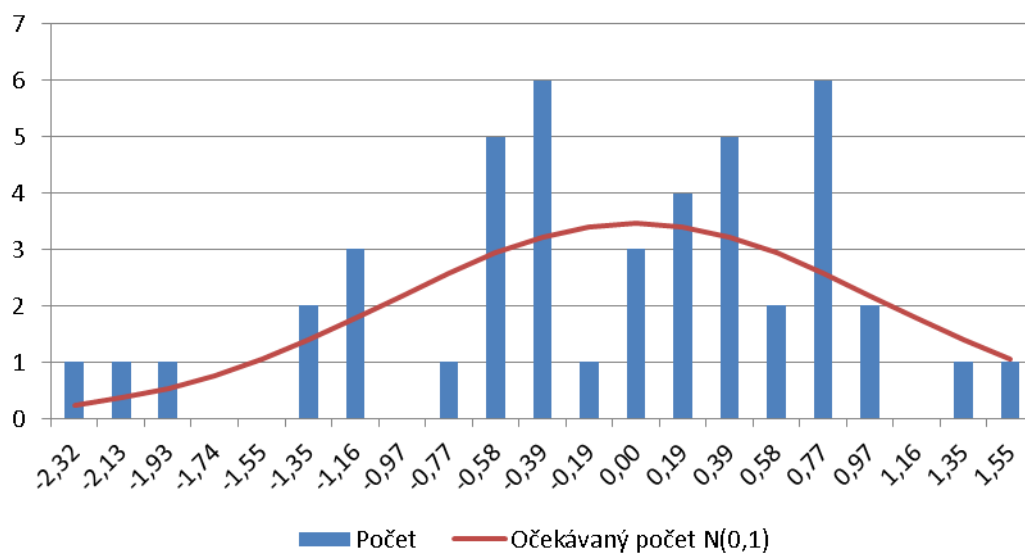
j	0	1	2	3
$\hat{b}$	-3,42	0,44	0,58	0,91
$\hat{\sigma}_b$	1,14	0,73	0,57	0,16
p-value_b	1,98E-02	5,71E-01	3,60E-01	4,77E-03
$\hat{a}$	347677,65	104249,13	76906,70	34776,66
$\hat{\sigma}_a$	56033,52	63190,46	63100,94	20161,56
p-value_a	0,0%	15,0%	27,7%	16,0%



Rezidua dle predikovaných hodnot



Histogram std. reziduí



Obsah

- ✓ Datová struktura
- ✓ Posouzení dat
- ✓ Předpoklady metody chain-ladder dle T. Macka
- ✓ Běžná lineární regrese
- ✓ Zobecněná lineární regrese (Aitkenův model)
- ✓ Převedení metody chain-ladder na úlohu lineární regrese
- ✓ Testování předpokladů
- ✓ Numerický příklad
- ☐ Shrnutí



Otázky?



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je



Shrnutí



Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je

Shrnutí

- ▶ Definice trojúhelníku a připomenutí předpokladů metody chain-ladder
- ▶ Běžná lineární regrese
- ▶ Zobecněná lineární regrese
 - ▶ Speciální diagonální matice rozptylu
- ▶ Přejít z metody chain-ladder na lineární regresi
- ▶ Testování předpokladů
 - ▶ Přímka procházející počátkem
 - ▶ Test normality pro standardizovaná rezidua
 - ▶ Graf standardizovaných reziduí v závislosti na roce vzniku, vývoje i účtování nemá vykazovat trend
 - ▶ Graf závislosti standardizovaných reziduí na predikovaných hodnotách nemá vykazovat trend

Literatura

- Anděl, J.: Základy matematické statistiky, Preprint, 2002
- Barnett, G., Zehnwirth, B.: Best Estimates for Reserves, Proceedings of the CAS, Vol. LXXXVII, No. 167, 2000
- Mack, T.: Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates, Astin Bulletin, Vol. 23, No. 2, 1993
- Murphy, D. M.: Unbiased Loss Development Factors, FCAS, MAAA, 1993
- Zvára, K.: Regrese, Matfyzpress, 2008



Krásný zbytek týdne!

pspanihelova@koop.cz


Kooperativa
VIENNA INSURANCE GROUP

Pro život jaký je