

Lineárny statistické modely I

1 Úvod

Prednášky z predmetu *Lineárny statistické modely I* nadväzujú na predmety *Pravdepodobnosť a statistika I, II* a predpokladajú sa znalosti získané v týchto predmetoch. Odporúčaná literatúra k štúdiu je

Anděl, J., *Matematická statistika*, SNTL, Praha, 1985.

Rao, C. R., *Lineárny metódy statistické indukce a jejich aplikace*, ACADEMIA, Praha, 1978.

Zvára, K., *Regresní analýza*, ACADEMIA, Praha, 1989.

Niektoré poznatky si zopakujeme (považujeme ich za "vzorce").

Majme náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)'$, ktorý má distribučnú funkciu $F_{\mathbf{X}}(x_1, \dots, x_n)$. Nech $T : \mathcal{R}^n \rightarrow \mathcal{R}^1$ je merateľné zobrazenie, potom stredná hodnota

$$\begin{aligned} \mathcal{E}(T) &= \mathcal{E}(T(X_1, \dots, X_n)) = \int_{\mathcal{R}^n} T(x_1, \dots, x_n) dF_{\mathbf{X}}(x_1, \dots, x_n) = \\ &= \begin{cases} \sum_{(x_1, \dots, x_n)} T(x_1, \dots, x_n) P_{(x_1, \dots, x_n)}, & \text{v prípade diskrétného náhodného vektora,} \\ \int_{\mathcal{R}^n} T(x_1, \dots, x_n) f_{\mathbf{X}}(x_1, \dots, x_n) dx_1 \dots dx_n, & \text{v prípade spojitého náhodného vektora.} \end{cases} \end{aligned}$$

Stredná hodnota

$$\mathcal{E}(X_i) = \int_{-\infty}^{\infty} x_i dF_i(x_i),$$

disperzia (rozptyl)

$$\mathcal{D}(X_i) = \mathcal{E}[(X_i - \mathcal{E}(X_i))^2] = \int_{-\infty}^{\infty} (x_i - \mathcal{E}(X_i))^2 dF_i(x_i),$$

kovariancia

$$\text{cov}(X_i, X_j) = \mathcal{E}[(X_i - \mathcal{E}(X_i))(X_j - \mathcal{E}(X_j))] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mathcal{E}(X_i))(x_j - \mathcal{E}(X_j)) dF_{i,j}(x_i, x_j)$$

a korelácia

$$\text{cor}(X_i, X_j) = \rho(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{\mathcal{D}(X_i)\mathcal{D}(X_j)}}$$

ak $0 < \mathcal{D}(X_i) < \infty$, $0 < \mathcal{D}(X_j) < \infty$.

Len na pripomenutie

$$\begin{aligned} F_i(x_i) &= \lim_{\substack{x_1 \rightarrow \infty \\ \vdots \\ x_{i-1} \rightarrow \infty \\ x_{i+1} \rightarrow \infty \\ \vdots \\ x_n \rightarrow \infty}} F_{\mathbf{X}}(x_1, \dots, x_n), \quad f_i(x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n. \end{aligned}$$

Matica náhodných veličín (náhodná matica) $\mathbf{Z}_{n,m}$ je taká matica, ktorej prvky $\{\mathbf{Z}\}_{i,j} = Z_{i,j}$ sú náhodné veličiny. Jej stredná hodnota $\mathcal{E}(\mathbf{Z})$ je matica, ktorej (i,j) -ty prvok je $\mathcal{E}(Z_{i,j})$ (predpokladáme, že všetky stredné hodnoty existujú a sú konečné).

Veta 1.1: Nech $\mathbf{Z}_1, \mathbf{Z}_2$ sú náhodné matice, nech existujú matice $\mathcal{E}(\mathbf{Z}_1), \mathcal{E}(\mathbf{Z}_2)$ a $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2, \mathbf{C}_1, \mathbf{C}_2$ sú (nenáhodné) reálne matice vhodných rozmerov. Platí

$$\mathcal{E}(\mathbf{A} + \mathbf{B}_1\mathbf{Z}_1\mathbf{C}_1 + \mathbf{B}_2\mathbf{Z}_2\mathbf{C}_2) = \mathbf{A} + \mathbf{B}_1\mathcal{E}(\mathbf{Z}_1)\mathbf{C}_1 + \mathbf{B}_2\mathcal{E}(\mathbf{Z}_2)\mathbf{C}_2, \quad [\mathcal{E}(\mathbf{Z}_1)]' = \mathcal{E}(\mathbf{Z}_1').$$

Dôkaz: Spravte ako cvičenie. Využite linearitu integrálu, teda platnosť vzťahov

$$\mathcal{E}(a + bX) = a + \mathcal{E}(X), \quad \mathcal{E}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i \mathcal{E}(X_i)$$

pre ľubovoľné reálne a, b, c_1, \dots, c_n a náhodné premenné X, X_1, \dots, X_n (ktoré majú konečné stredné hodnoty) Q.E.D.

Ak $\mathbf{X}_{n,1}$ je náhodný vektor $(X_1, \dots, X_n)'$, potom jeho stredná hodnota $\mathcal{E}(\mathbf{X}) = (\mathcal{E}(X_1), \dots, \mathcal{E}(X_n))'$ (ak všetky stredné hodnoty existujú a sú konečné). Ak všetky X_1, \dots, X_n majú konečné disperzie, tak kovariančná matica vektora \mathbf{X} je

$$\text{cov}(\mathbf{X}) = \begin{pmatrix} \mathcal{D}(X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_n) \\ \text{cov}(X_2, X_1) & \mathcal{D}(X_2) & \dots & \text{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \text{cov}(X_n, X_2) & \dots & \mathcal{D}(X_n) \end{pmatrix}.$$

Z Vety 1.1 ľahko dostávame

Lema 1.2:

$$\mathcal{E}(\mathbf{a}_{p,1} + \mathbf{B}_{p,n}\mathbf{X}) = \mathbf{a} + \mathbf{B}\mathcal{E}(\mathbf{X}).$$

Veta 1.3: (Vlastnosti kovariančnej matice $\text{cov}(\mathbf{X})$.) Nech \mathbf{X} je náhodný vektor s konečnými druhými momentmi. Platí:

1. $\text{cov}(\mathbf{X}) = \mathcal{E}[(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))'] = \mathcal{E}(\mathbf{X}\mathbf{X}') - \mathcal{E}(\mathbf{X})[\mathcal{E}(\mathbf{X})]'$,
2. $(\text{cov}(\mathbf{X}))' = \text{cov}(\mathbf{X})$,
3. $\text{cov}(\mathbf{a}_{m,1} + \mathbf{B}_{m,n}\mathbf{X}) = \mathbf{B}\text{cov}(\mathbf{X})\mathbf{B}'$, ak \mathbf{a} , \mathbf{B} sú nenáhodný vektor resp. nenáhodná matica,
4. $\text{cov}(\mathbf{X})$ je pozitívne semidefinitná matica.

Dôkaz:

1. Spravte ako cvičenie.
2. Vyplýva z vlastnosti kovariancie $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$.
3. $\text{cov}(\mathbf{a} + \mathbf{B}\mathbf{X}) = \mathcal{E}\{(\mathbf{a} + \mathbf{B}\mathbf{X} - \mathcal{E}(\mathbf{a} + \mathbf{B}\mathbf{X}))(\mathbf{a} + \mathbf{B}\mathbf{X} - \mathcal{E}(\mathbf{a} + \mathbf{B}\mathbf{X}))'\} = \mathcal{E}(\mathbf{B}(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{B}') = \mathbf{B}\text{cov}(\mathbf{X})\mathbf{B}'$.
4. Podľa predchádzajúceho bodu pre ľubovoľné $\mathbf{c}_{n,1}$ platí pre disperziu náhodnej veličiny $Y = \mathbf{c}'\mathbf{X}$ nerovnosť $0 \leq \mathcal{D}(Y) = \mathcal{D}(\mathbf{c}'\mathbf{X}) = \mathbf{c}'\text{cov}(\mathbf{X})\mathbf{c}$. Využívame aj poznatok, že disperzia každej náhodnej veličiny je nezáporné číslo (ak disperzia existuje). Q.E.D.

Niektoré aplikácie predchádzajúcich tvrdení:

Majme náhodné premenné X_1, X_2 a poznáme ich kovariančnú maticu $cov(\mathbf{X})$. Potom disperzie a kovariancie medzi náhodnými premennými $Y_1 = c_1X_1 + c_2X_2$, $Y_2 = d_1X_1 + d_2X_2$, $Y_3 = e_1X_1 + e_2X_2$ ľahko dostaneme z kovariančnej matice náhodného vektora

$$\mathbf{Y} = \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \\ e_1 & e_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \mathbf{B}\mathbf{X}, \quad \text{keďže} \quad cov(\mathbf{Y}) = \mathbf{B}cov(\mathbf{X})\mathbf{B}'.$$

Napríklad ak $Y_1 = X_1 - X_2$, $Y_2 = X_1 + X_2$, tak

$$\begin{aligned} cov(\mathbf{Y}) &= cov\left(\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}\right) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{D}(X_1) & cov(X_1, X_2) \\ cov(X_2, X_1) & \mathcal{D}(X_2) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} \mathcal{D}(X_1) - cov(X_2, X_1) & cov(X_1, X_2) - \mathcal{D}(X_2) \\ \mathcal{D}(X_1) + cov(X_2, X_1) & cov(X_1, X_2) + \mathcal{D}(X_2) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} \mathcal{D}(X_1) - 2cov(X_1, X_2) + \mathcal{D}(X_2) & \mathcal{D}(X_1) - \mathcal{D}(X_2) \\ \mathcal{D}(X_1) - \mathcal{D}(X_2) & \mathcal{D}(X_1) + 2cov(X_1, X_2) + \mathcal{D}(X_2) \end{pmatrix}. \end{aligned}$$

Majme dva náhodné vektory $\mathbf{X} = (X_1, \dots, X_n)'$, $\mathbf{Y} = (Y_1, \dots, Y_m)'$. Nech existujú konečné kovariancie $cov(X_i, Y_j)$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$. Matica

$$cov(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} cov(X_1, Y_1) & cov(X_1, Y_2) & \dots & cov(X_1, Y_m) \\ cov(X_2, Y_1) & cov(X_2, Y_2) & \dots & cov(X_2, Y_m) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, Y_1) & cov(X_n, Y_2) & \dots & cov(X_n, Y_m) \end{pmatrix}_{n,m}$$

je kovariančná matica vektorov \mathbf{X} , \mathbf{Y} .

Veta 1.4: (Vlastnosti kovariančnej matice $cov(\mathbf{X}, \mathbf{Y})$.) Platí

1. $cov(\mathbf{X}, \mathbf{Y}) = \mathcal{E}[(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))'] = \mathcal{E}(\mathbf{X}\mathbf{Y}') - \mathcal{E}(\mathbf{X})[\mathcal{E}(\mathbf{Y})]'$,
2. $cov(\mathbf{X}, \mathbf{X}) = cov(\mathbf{X})$,
3. $cov(\mathbf{a}_{k,1} + \mathbf{B}_{k,n}\mathbf{X}, \mathbf{b}_{l,1} + \mathbf{C}_{l,m}\mathbf{Y}) = \mathbf{B}cov(\mathbf{X}, \mathbf{Y})\mathbf{C}'$, ak \mathbf{a} , \mathbf{b} , \mathbf{B} , \mathbf{C} sú nenáhodné vektory resp. nenáhodné matice,
4. ak $m = n$, tak $cov(\mathbf{X} + \mathbf{Y}) = cov(\mathbf{X}) + cov(\mathbf{Y}) + cov(\mathbf{X}, \mathbf{Y}) + cov(\mathbf{Y}, \mathbf{X})$,
5. $cov(\sum_{i=1}^t \mathbf{X}_i) = \sum_{i=1}^t \sum_{j=1}^t cov(\mathbf{X}_i, \mathbf{X}_j)$, ak $\mathbf{X}_i \in \mathcal{R}^n$, $i = 1, 2, \dots, t$,
6. $[cov(\mathbf{X}, \mathbf{Y})]' = cov(\mathbf{Y}, \mathbf{X})$.
7. $\mathcal{E}Tr \mathbf{A}_{k,l} \mathbf{Z}_{l,k} = Tr \mathbf{A} \mathcal{E}(\mathbf{Z})$, ak \mathbf{A} je nenáhodná a \mathbf{Z} náhodná matica.

Dôkaz:

1. Spravte ako cvičenie.
2. Vyplýva z definície.
3. $cov(\mathbf{a} + \mathbf{B}\mathbf{X}, \mathbf{b} + \mathbf{C}\mathbf{Y}) = \mathcal{E}\{(\mathbf{a} + \mathbf{B}\mathbf{X} - \mathcal{E}(\mathbf{a} + \mathbf{B}\mathbf{X}))(\mathbf{b} + \mathbf{C}\mathbf{Y} - \mathcal{E}(\mathbf{b} + \mathbf{C}\mathbf{Y}))'\} =$
 $= \mathcal{E}\{[\mathbf{B}(\mathbf{X} - \mathcal{E}(\mathbf{X}))][\mathbf{C}(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))]'\} = \mathcal{E}\{\mathbf{B}(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))'\mathbf{C}'\} =$
 $= \mathbf{B}\mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))'\}\mathbf{C}' = \mathbf{B}cov(\mathbf{X}, \mathbf{Y})\mathbf{C}'.$

$$\begin{aligned}
4. \text{ cov}(\mathbf{X} + \mathbf{Y}) &= \mathcal{E}\{(\mathbf{X} + \mathbf{Y} - \mathcal{E}(\mathbf{X} + \mathbf{Y}))(\mathbf{X} + \mathbf{Y} - \mathcal{E}(\mathbf{X} + \mathbf{Y}))'\} = \\
&= \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}) + \mathbf{Y} - \mathcal{E}(\mathbf{Y}))(\mathbf{X} - \mathcal{E}(\mathbf{X}) + \mathbf{Y} - \mathcal{E}(\mathbf{Y}))'\} = \\
&= \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))' + (\mathbf{Y} - \mathcal{E}(\mathbf{Y}))(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))' + (\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{Y} - \mathcal{E}(\mathbf{Y}))' + (\mathbf{Y} - \mathcal{E}(\mathbf{Y}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\} = \\
&= \text{cov}(\mathbf{X}) + \text{cov}(\mathbf{Y}, \mathbf{X}) + \text{cov}(\mathbf{X}, \mathbf{Y}) + \text{cov}(\mathbf{Y}).
\end{aligned}$$

5. Dokážte ako cvičenie.

$$6. [\text{cov}(\mathbf{X}, \mathbf{Y})]' = \{\mathcal{E}(\mathbf{X}\mathbf{Y}') - \mathcal{E}(\mathbf{X})[\mathcal{E}(\mathbf{Y})]'\}' = \mathcal{E}(\mathbf{Y}\mathbf{X}') - \mathcal{E}(\mathbf{Y})[\mathcal{E}(\mathbf{X})]' = \text{cov}(\mathbf{Y}, \mathbf{X}), \text{ lebo } [\mathcal{E}(\mathbf{X}\mathbf{Y}')]' = \mathcal{E}(\mathbf{Y}\mathbf{X}').$$

$$7. \mathcal{E}\text{Tr}\mathbf{AZ} = \mathcal{E}\{\sum_{i=1}^k \sum_{j=1}^l \{\mathbf{A}\}_{i,j} \{\mathbf{Z}\}_{j,i}\} = \sum_{i=1}^k \sum_{j=1}^l \{\mathbf{A}\}_{i,j} \mathcal{E}\{\mathbf{Z}\}_{j,i} = \text{Tr}\mathbf{A}\mathcal{E}(\mathbf{Z}). \quad \text{Q.E.D.}$$

Nech existujú všetky korelačné koeficienty

$$\text{cor}(X_i, Y_j) = \varrho(X_i, Y_j) = \varrho_{i,j} = \frac{\text{cov}(X_i, Y_j)}{\sqrt{\mathcal{D}(X_i)\mathcal{D}(Y_j)}}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Matica

$$\text{cor}(\mathbf{X}, \mathbf{Y}) = (\varrho_{i,j})_{\substack{i=1,2,\dots,n \\ j=1,2,\dots,m}} = \begin{pmatrix} \text{cor}(X_1, Y_1) & \text{cor}(X_1, Y_2) & \dots & \text{cor}(X_1, Y_m) \\ \text{cor}(X_2, Y_1) & \text{cor}(X_2, Y_2) & \dots & \text{cor}(X_2, Y_m) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cor}(X_n, Y_1) & \text{cor}(X_n, Y_2) & \dots & \text{cor}(X_n, Y_m) \end{pmatrix}_{n,m}$$

sa volá korelačná matica vektorov \mathbf{X}, \mathbf{Y} . Špeciálne ak $\mathbf{X} = \mathbf{Y}$, píšeme

$$\text{cor}(\mathbf{X}) = \begin{pmatrix} 1 & \text{cor}(X_1, X_2) & \dots & \text{cor}(X_1, X_n) \\ \text{cor}(X_2, X_1) & 1 & \dots & \text{cor}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cor}(X_n, X_1) & \text{cor}(X_n, X_2) & \dots & 1 \end{pmatrix}$$

namiesto $\text{cor}(\mathbf{X}, \mathbf{X})$. Niekedy sa kovariančná matica píše

$$\text{cov}(\mathbf{X}) = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_{2,2} & \dots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma_{n,n} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \dots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma_n^2 \end{pmatrix},$$

teda $\sigma_{i,i} = \sigma_i^2 = \mathcal{D}(X_i)$. Pri označení

$$\mathbf{D} = \mathbf{D}_\mathbf{X} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{pmatrix} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$$

platí

$$\text{cor}(\mathbf{X}) = \mathbf{D}^{-1} \text{cov}(\mathbf{X}) \mathbf{D}^{-1} \quad \text{a} \quad \text{cov}(\mathbf{X}) = \mathbf{D} \text{cor}(\mathbf{X}) \mathbf{D},$$

lebo

$$\begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \dots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \dots & \sigma_n^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{pmatrix} =$$

$$\begin{aligned}
&= \begin{pmatrix} \sigma_1 & \frac{\sigma_{1,2}}{\sigma_1} & \dots & \frac{\sigma_{1,n}}{\sigma_1} \\ \frac{\sigma_{2,1}}{\sigma_2} & \sigma_2 & \dots & \frac{\sigma_{2,n}}{\sigma_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{n,1}}{\sigma_n} & \frac{\sigma_{n,2}}{\sigma_n} & \dots & \sigma_n \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{pmatrix} = \\
&= \begin{pmatrix} 1 & \frac{\sigma_{1,2}}{\sigma_1\sigma_2} & \dots & \frac{\sigma_{1,n}}{\sigma_1\sigma_n} \\ \frac{\sigma_{2,1}}{\sigma_2\sigma_1} & 1 & \dots & \frac{\sigma_{2,n}}{\sigma_2\sigma_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{n,1}}{\sigma_n\sigma_1} & \frac{\sigma_{n,2}}{\sigma_n\sigma_2} & \dots & 1 \end{pmatrix} = \text{cor}(\mathbf{X}).
\end{aligned}$$

Analogicky

$$\text{cor}(\mathbf{X}, \mathbf{Y}) = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n} \end{pmatrix} \text{cov}(\mathbf{X}, \mathbf{Y}) \begin{pmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_m} \end{pmatrix},$$

čiže

$$\text{cor}(\mathbf{X}, \mathbf{Y}) = \mathbf{D}_{\mathbf{X}}^{-1} \text{cov}(\mathbf{X}, \mathbf{Y}) \mathbf{D}_{\mathbf{Y}}^{-1} \quad \text{a} \quad \text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{D}_{\mathbf{X}} \text{cor}(\mathbf{X}, \mathbf{Y}) \mathbf{D}_{\mathbf{Y}}.$$

Veta 1.5: Ak $\mathbf{A}_{n,n}$ je reálna matica (nemusí byť ani symetrická) a existuje matica $\text{cov}(\mathbf{X})$, tak

$$\mathcal{E}(\mathbf{X}'\mathbf{A}\mathbf{X}) = [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X}) + \text{Tr}\mathbf{A}\text{cov}(\mathbf{X}).$$

Dôkaz: Pomocou Vety 1.4 dostávame

$$\begin{aligned}
\mathcal{E}(\mathbf{X}'\mathbf{A}\mathbf{X}) &= \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}) + \mathcal{E}(\mathbf{X}))'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X}) + \mathcal{E}(\mathbf{X}))\} = \\
&= \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X})) + [\mathcal{E}(\mathbf{X})]'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X})) + (\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{A}\mathcal{E}(\mathbf{X}) + [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X})\} = \\
&= [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X}) + \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X}))\} + \mathcal{E}\{[\mathcal{E}(\mathbf{X})]'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X}))\} + \mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{A}\mathcal{E}(\mathbf{X})\} = \\
&= [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X}) + \mathcal{E}\text{Tr}(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X})) + [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}[\mathbf{X} - \mathcal{E}(\mathbf{X})] + [\mathcal{E}(\mathbf{X} - \mathcal{E}(\mathbf{X}))]'\mathbf{A}\mathcal{E}(\mathbf{X}) = \\
&= [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X}) + \mathcal{E}\text{Tr}\mathbf{A}(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))' = [\mathcal{E}(\mathbf{X})]'\mathbf{A}\mathcal{E}(\mathbf{X}) + \text{Tr}\mathbf{A}\mathcal{E}\{(\mathbf{X} - \mathcal{E}(\mathbf{X}))(\mathbf{X} - \mathcal{E}(\mathbf{X}))'\} = \\
&= \mathcal{E}(\mathbf{X}) + \text{Tr}\mathbf{A}\text{cov}(\mathbf{X}). \quad \text{Q.E.D.}
\end{aligned}$$

Náhodný výber rozsahu n je n -tica X_1, X_2, \dots, X_n náhodných veličín, ktoré sú nezávislé a rovnako rozdelené. Ak sú rozdelené ako náhodná veličina X , teda ak $X_i \sim X$, $i = 1, 2, \dots, n$, tak povieme, že X_1, X_2, \dots, X_n je náhodný výber z rozdelenia, aké má náhodná veličina X , napr. z $N(\mu, \sigma^2)$.

Príklad 1.6: Nech $\mathbf{X} = (X_1, \dots, X_n)'$, $n \geq 2$, je náhodný výber z rozdelenia s konečným rozptylom (disperziou) σ^2 . Označme

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (\text{výberový rozptyl}) \quad \text{a} \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (\text{výberový priemer}).$$

Ukážte, že $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} (\sum_{i=1}^n X_i^2 - n\bar{X}^2)$ a určte $\mathcal{E}(S^2)$.

Riešenie:

$$\begin{aligned}
 S^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2) = \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i \frac{1}{n} \sum_{j=1}^n X_j + \bar{X}^2) = \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - 2n \frac{1}{n} \sum_{i=1}^n X_i \left(\frac{1}{n} \sum_{j=1}^n X_j \right) + n\bar{X}^2 \right) = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) = \\
 &= \frac{1}{n-1} \left(\mathbf{X}'\mathbf{X} - n \frac{1}{n^2} \left(\sum_{i=1}^n X_i \right) \left(\sum_{j=1}^n X_j \right) \right) = \frac{1}{n-1} \left(\mathbf{X}'\mathbf{X} - \frac{1}{n} \mathbf{X}'\mathbf{1}\mathbf{1}'\mathbf{X} \right) = \\
 &= \mathbf{X}' \left\{ \frac{1}{n-1} \mathbf{I} - \frac{1}{n(n-1)} \mathbf{E} \right\} \mathbf{X} = \mathbf{X}'\mathbf{A}\mathbf{X},
 \end{aligned}$$

pričom vektor (príslušného rozmeru), ktorého zložky sú samé jedničky budeme značiť $\mathbf{1}$ a štvorcovú maticu (príslušného rozmeru), ktorá má všetky prvky rovné jednej budeme značiť \mathbf{E} . Platí teda $\mathbf{E} = \mathbf{1}\mathbf{1}'$.

Keď na náhodný výber X_1, X_2, \dots, X_n pozeráme ako na náhodný vektor $\mathbf{X} = (X_1, X_2, \dots, X_n)'$, tak jeho stredná hodnota a kovariančná matica sú

$$\mathcal{E}(\mathbf{X}) = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix} = \mu\mathbf{1}, \quad \text{cov}(\mathbf{X}) = \begin{pmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma^2 \end{pmatrix} = \sigma^2\mathbf{I},$$

($\mathcal{E}(X_i) = \mu$, $\mathcal{D}(X_i) = \sigma^2$). Podľa Vety 1.5 je

$$\begin{aligned}
 \mathcal{E}(S^2) &= [\mathcal{E}(\mathbf{X})]' \mathbf{A} \mathcal{E}(\mathbf{X}) + \text{Tr} \mathbf{A} \text{cov}(\mathbf{X}) = \mu^2 \mathbf{1}' \left\{ \frac{1}{n-1} \mathbf{I} - \frac{1}{n(n-1)} \mathbf{E} \right\} \mathbf{1} + \text{Tr} \left[\left\{ \frac{1}{n-1} \mathbf{I} - \frac{1}{n(n-1)} \mathbf{E} \right\} \sigma^2 \mathbf{I} \right] = \\
 &= \mu^2 \left\{ \frac{1}{n-1} \mathbf{1}'\mathbf{1} - \frac{1}{n(n-1)} \mathbf{1}'\mathbf{1}\mathbf{1}'\mathbf{1} \right\} + \text{Tr} \left\{ \sigma^2 \left[\frac{1}{n-1} \mathbf{I} \right] \right\} - \text{Tr} \sigma^2 \left[\frac{1}{n(n-1)} \mathbf{E} \right] = \sigma^2 \frac{1}{n-1} n - \sigma^2 \frac{1}{n-1} = \sigma^2.
 \end{aligned}$$

2 Mnohorozmerné normálne rozdelenie

Ak má náhodná veličina X hustotu

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in (-\infty, \infty),$$

kde $\mu \in \mathcal{R}$, $\sigma^2 > 0$, tak X má *regulárne normálne rozdelenie* s parametrami μ a σ^2 . Píšeme $X \sim N(\mu, \sigma^2)$.

Veta 2.1: Nech $X \sim N(0, 1)$ a $\mu \in (-\infty, \infty)$, $\alpha \neq 0$. Platí $U = \mu + \alpha X \sim N(\mu, \alpha^2)$.

Dôkaz: X má hustotu $f_X = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

(i) ak $\alpha > 0$, tak distribučná funkcia náhodnej veličiny $U = \mu + \alpha X$ je

$$F_U(x) = F_{\mu+\alpha X}(x) = P\{\mu + \alpha X < x\} = P\left\{X < \frac{x-\mu}{\alpha}\right\} = \int_{-\infty}^{\frac{x-\mu}{\alpha}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt =$$

(substitúcia $t = \frac{u-\mu}{\alpha}$, $dt = \frac{du}{\alpha}$)

$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\alpha} e^{-\frac{(u-\mu)^2}{2\alpha^2}} du,$$

teda hustota U je v tomto prípade $\frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{(u-\mu)^2}{2\alpha^2}}$.

(i) ak $\alpha < 0$, tak distribučná funkcia náhodnej veličiny $U = \mu + \alpha X$ je

$$F_U(x) = F_{\mu+\alpha X}(x) = P\{\mu + \alpha X < x\} = P\left\{X > \frac{x-\mu}{\alpha}\right\} = \int_{\frac{x-\mu}{\alpha}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt =$$

(substitúcia $t = \frac{u-\mu}{\alpha}$, $dt = \frac{du}{\alpha}$)

$$= \int_x^{-\infty} \frac{1}{\sqrt{2\pi\alpha}} e^{-\frac{(u-\mu)^2}{2\alpha^2}} du = \int_{-\infty}^x \frac{1}{\sqrt{2\pi(-\alpha)}} e^{-\frac{(u-\mu)^2}{2\alpha^2}} du,$$

teda hustota U je v tomto prípade $\frac{1}{\sqrt{2\pi(-\alpha)}} e^{-\frac{(u-\mu)^2}{2\alpha^2}}$.

V oboch prípadoch je hustota U rovná $f_U(x) = \frac{1}{\sqrt{2\pi|\alpha|}} e^{-\frac{(u-\mu)^2}{2\alpha^2}} = \frac{1}{\sqrt{2\pi}\sqrt{\alpha^2}} e^{-\frac{(u-\mu)^2}{2\alpha^2}}$, a teda $U \sim N(\mu, \alpha^2)$. Q.E.D.

Charakteristická funkcia náhodnej veličiny $X \sim N(0, 1)$ je

$$\begin{aligned} \psi_X(t) = \mathcal{E}(e^{itX}) &= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \underbrace{(\cos tx)e^{-\frac{x^2}{2}}}_{\text{párna (sudá) funkcia}} dx + \int_{-\infty}^{\infty} \underbrace{(\sin tx)e^{-\frac{x^2}{2}}}_{\text{nepárna (lichá) funkcia}} dx \right\} = \\ &= \frac{1}{\sqrt{2\pi}} 2 \int_0^{\infty} (\cos tx)e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} 2 \frac{\sqrt{\pi}}{2\frac{1}{\sqrt{2}}} e^{-\frac{t^2}{4\frac{1}{2}}} = e^{-\frac{t^2}{2}}. \end{aligned}$$

Tu sme využili výsledok z analýzy

$$\int_0^{\infty} (\cos bx)e^{-a^2x^2} dx = \frac{\sqrt{\pi}}{2a} e^{-\frac{b^2}{4a^2}}, \quad \text{ak } a > 0.$$

Teda $X \sim N(0, 1)$ má charakteristickú funkciu $\psi_X(t) = e^{-\frac{t^2}{2}}$.

Ak $U \sim N(\mu, \sigma^2)$, $\sigma > 0$, tak $U = \mu + \sigma X$ (kde $X \sim N(0, 1)$) a preto

$$\psi_U(t) = \mathcal{E}(e^{it(\mu+\sigma X)}) = \mathcal{E}(e^{it\mu} e^{it\sigma X}) = e^{it\mu} \mathcal{E}(e^{i(t\sigma)X}) = e^{it\mu} \psi_X(t\sigma) = e^{it\mu - \frac{t^2\sigma^2}{2}}.$$

(Poznamenávame len, že rovnakú charakteristickú funkciu (teda aj rovnaké rozdelenie pravdepodobnosti) má aj náhodná veličina $U = \mu - \sigma X$).

Nech $\sigma^2 > 0$. Náhodná veličina X má regulárne normálne $N(\mu, \sigma^2)$ rozdelenie, keď má hustotu $f(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ a charakteristickú funkciu $e^{i\mu t - \frac{t^2\sigma^2}{2}}$.

Nech $\sigma = 0$. Ak má náhodná veličina X charakteristickú funkciu $e^{i\mu t}$, tak to je diskrétna náhodná veličina, ktorá nadobúda (jedinú) hodnotu μ s pravdepodobnosťou 1 (čiže $P\{X = \mu\} = 1$). Povieme, že v tomto prípade má X singulárne normálne $N(\mu, 0)$ rozdelenie.

Dospeli sme k (všeobecnej) definícii normálneho rozdelenia

Definícia 2.2: Nech $\mu \in (-\infty, \infty)$, $\sigma^2 \geq 0$. Povieme, že $X \sim N(\mu, \sigma^2)$ (teda, že X má jednorozmerné normálne rozdelenie s parametrami $\mu \in (-\infty, \infty)$, $\sigma^2 \geq 0$), ak má charakteristickú funkciu $\psi_X(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}}$, $t \in (-\infty, \infty)$.

Dokážte ako príklad, že ak $X \sim N(\mu, \sigma^2)$, tak $\mathcal{E}(X) = \mu$ a $\mathcal{D}(X) = \sigma^2$.

Definícia 2.3: Nech $\boldsymbol{\mu} \in \mathcal{R}^n$ a $\boldsymbol{\Sigma} = (\sigma_{i,j})_{\substack{i=1,2,\dots,n \\ j=1,2,\dots,n}}$ je pozitívne semidefinitná symetrická matica. Povieme, že náhodný vektor \mathbf{X} má n -rozmerné normálne rozdelenie, ak pre každý vektor $\mathbf{c} \in \mathcal{R}^n$ je $\mathbf{c}'\mathbf{X} \sim N(\mathbf{c}'\boldsymbol{\mu}, \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c})$. Píšeme $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Veta 2.4: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Potom $\mathcal{E}(\mathbf{X}) = \boldsymbol{\mu}$ a $\text{cov}\mathbf{X} = \boldsymbol{\Sigma}$.

Dôkaz: Označme $\mathbf{e}_j \in \mathcal{R}^n$ j -ty jednotkový vektor (čiže vektor, ktorý má na j -tom mieste 1, inde všade 0). Potom

$$X_j = \mathbf{e}_j'\mathbf{X} \sim N(\mathbf{e}_j'\boldsymbol{\mu} = \mu_j, \mathbf{e}_j'\boldsymbol{\Sigma}\mathbf{e}_j = \sigma_{j,j}),$$

teda

$$\mathcal{E}(X_j) = \mu_j, \quad \mathcal{D}(X_j) = \sigma_{j,j}.$$

Platí, že $\mathcal{E}(\mathbf{X}) = \boldsymbol{\mu}$ a $\{\boldsymbol{\Sigma}\}_{i,i} = \mathcal{D}(X_i)$, $i = 1, 2, \dots, n$.

Uvažujme teraz vektor $\mathbf{c}_{j,k} \in \mathcal{R}^n$, $\mathbf{c}_{j,k} = \mathbf{e}_j + \mathbf{e}_k$, $j \neq k$. Platí

$$\mathbf{c}'_{j,k}\mathbf{X} = X_j + X_k \sim N(\mathbf{c}'_{j,k}\boldsymbol{\mu} = \mu_j + \mu_k, \mathbf{c}'_{j,k}\boldsymbol{\Sigma}\mathbf{c}_{j,k} = \sigma_{j,j} + \sigma_{j,k} + \sigma_{k,j} + \sigma_{k,k}),$$

teda

$$\mathcal{D}(\mathbf{c}'_{j,k}\mathbf{X}) = \sigma_{j,j} + \sigma_{k,k} + 2\sigma_{j,k} \quad (1)$$

Podľa Vety 1.3, bod 3 ($\mathbf{a} = 0$, $\mathbf{B} = \mathbf{c}'_{j,k}$) a Vety 1.4, bod 5 je

$$\mathcal{D}(\mathbf{c}'_{j,k}\mathbf{X}) = \mathcal{D}(X_j + X_k) = \mathcal{D}(X_j) + \mathcal{D}(X_k) + \text{cov}(X_j, X_k) + \text{cov}(X_k, X_j) = \mathcal{D}(X_j) + \mathcal{D}(X_k) + 2\text{cov}(X_j, X_k). \quad (2)$$

Pretože $\mathcal{D}(X_j) = \sigma_{j,j}$, $\mathcal{D}(X_k) = \sigma_{k,k}$, dostávame z (1) a (2) $\sigma_{j,k} = \text{cov}(X_j, X_k)$ a teda $\boldsymbol{\Sigma} = \text{cov}\mathbf{X}$. Q.E.D.

Veta 2.5: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Potom charakteristická funkcia $\psi_{\mathbf{X}}(\mathbf{t}) = e^{it'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$, $\mathbf{t} \in \mathcal{R}^n$.

Dôkaz: Pre dané $\mathbf{t} \in \mathcal{R}^n$ je $\mathbf{t}'\mathbf{X} \sim N_1(\mathbf{t}'\boldsymbol{\mu}, \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$ a

$$\psi_{\mathbf{X}}(\mathbf{t}) = \mathcal{E}\left(e^{it'\mathbf{X}}\right) = \mathcal{E}\left(e^{i1(\mathbf{t}'\mathbf{X})}\right) = \psi_{\mathbf{t}'\mathbf{X}}(1) = e^{i1(\mathbf{t}'\boldsymbol{\mu}) - \frac{1^2\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2}} = e^{it'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}. \quad \text{Q.E.D.}$$

Veta 2.6: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{a} \in \mathcal{R}^m$, $\mathbf{B}_{m,n}$ reálna (pevná) matica. Potom $\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X} \sim N_m(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$.

Dôkaz: Pre dané $\mathbf{t} \in \mathcal{R}^m$ je

$$\begin{aligned} \psi_{\mathbf{Y}}(\mathbf{t}) &= \mathcal{E}\left(e^{it'\mathbf{Y}}\right) = \mathcal{E}\left(e^{it'(\mathbf{a} + \mathbf{B}\mathbf{X})}\right) = e^{it'\mathbf{a}}\mathcal{E}\left(e^{i(\mathbf{t}'\mathbf{B})\mathbf{X}}\right) = e^{it'\mathbf{a}}\psi_{\mathbf{X}}(\mathbf{B}'\mathbf{t}) = e^{it'\mathbf{a}}e^{it'\mathbf{B}\boldsymbol{\mu} - \frac{1}{2}(\mathbf{t}'\mathbf{B})\boldsymbol{\Sigma}(\mathbf{B}'\mathbf{t})} = \\ &= e^{it'(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}) - \frac{1}{2}\mathbf{t}'(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')\mathbf{t}}, \end{aligned}$$

čo je charakteristická funkcia m -rozmernej normálne rozdelenej náhodnej veličiny so strednou hodnotou $\mathbf{a} + \mathbf{B}\boldsymbol{\mu}$ a kovariančnou maticou $\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}'$, čiže $\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X} \sim N_m(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$. Q.E.D.

Označme

$$\mathbf{X}_{n,1} = \begin{pmatrix} X_1 \\ \vdots \\ X_k \\ X_{k+1} \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad 1 \leq k \leq n, \quad \text{čiže } \mathbf{X}_1 = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \in \mathcal{R}^k, \quad \mathbf{X}_2 = \begin{pmatrix} X_{k+1} \\ \vdots \\ X_n \end{pmatrix} \in \mathcal{R}^{n-k}. \quad (3)$$

Veta 2.7: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Potom $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, kde

$$\boldsymbol{\mu}_1 = \begin{pmatrix} \mathcal{E}(X_1) \\ \vdots \\ \mathcal{E}(X_k) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}, \quad \boldsymbol{\Sigma}_{11} = \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \dots & \sigma_{1,k} \\ \sigma_{2,1} & \sigma_{2,2} & \dots & \sigma_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k,1} & \sigma_{k,2} & \dots & \sigma_{k,k} \end{pmatrix}.$$

(Poznamenávame len, že $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$, teda $\boldsymbol{\Sigma}_{12}$ je $k \times (n-k)$ matica, $\boldsymbol{\Sigma}_{21}$ je $(n-k) \times k$ matica a $\boldsymbol{\Sigma}_{22}$ je $(n-k) \times (n-k)$ matica.)

Dôkaz: $\mathbf{X}_1 = (\mathbf{I}_{k,k} \mathbf{0}_{k,n-k})\mathbf{X} \implies \mathbf{X}_1 \sim N_k\left(\begin{pmatrix} \mathbf{I} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \boldsymbol{\mu}, \begin{pmatrix} \mathbf{I} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} \mathbf{I} \\ \mathbf{0} \end{pmatrix}\right)$, čiže $\mathbf{X}_1 \sim N_k(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$. Q.E.D.

Poznámka: Môžeme za \mathbf{X}_1 vybrať ľubovoľnú k -ticu náhodných premenných z X_1, \dots, X_n , marginálne rozdelenie náhodného vektora \mathbf{X}_1 je vždy normálne s "prirodzenými parametrami".

Veta 2.8: Nech $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (podľa (3)), potom platí

$$\mathbf{X}_1 \text{ a } \mathbf{X}_2 \text{ sú nezávislé} \iff \text{cov}(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{0}_{k,n-k} \text{ (sú nekorelované)}.$$

Dôkaz: Ak sú \mathbf{X}_1 a \mathbf{X}_2 sú nezávislé $\implies X_i$ a X_j sú nezávislé pre všetky $i \in \{1, \dots, k\}$, $j \in \{k+1, \dots, n\}$, teda $\text{cov}(X_i, X_j) = 0$ pre všetky $i \in \{1, \dots, k\}$, $j \in \{k+1, \dots, n\}$, čiže $\text{cov}(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{0}$.

Naopak, ak $\text{cov}(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{0} \implies \boldsymbol{\Sigma}_{12} = \mathbf{0}$, teda pre charakteristickú funkciu náhodného vektora \mathbf{X} platí pre každé $\mathbf{t} = (\mathbf{t}'_1, \mathbf{t}'_2)'$, pričom $\mathbf{t}'_1 \in \mathcal{R}^k$, $\mathbf{t}'_2 \in \mathcal{R}^{n-k}$,

$$\begin{aligned} \psi_{\mathbf{X}}(\mathbf{t}) &= e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}} = e^{i(\mathbf{t}'_1, \mathbf{t}'_2) \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} - \frac{1}{2}(\mathbf{t}'_1, \mathbf{t}'_2) \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{pmatrix}} = \\ &= e^{i\mathbf{t}'_1\boldsymbol{\mu}_1 - \frac{1}{2}\mathbf{t}'_1\boldsymbol{\Sigma}_{11}\mathbf{t}_1} e^{i\mathbf{t}'_2\boldsymbol{\mu}_2 - \frac{1}{2}\mathbf{t}'_2\boldsymbol{\Sigma}_{22}\mathbf{t}_2} = \psi_{\mathbf{X}_1}(\mathbf{t}_1)\psi_{\mathbf{X}_2}(\mathbf{t}_2) \end{aligned}$$

(je súčinom charakteristických funkcií subvektorov \mathbf{X}_1 a \mathbf{X}_2), teda (pozri Rényi, A., *Teória pravdepodobnosti*, ACADEMIA, Praha, 1972, str. 300) \mathbf{X}_1 a \mathbf{X}_2 sú nezávislé. Q.E.D.

Skôr ako sa dostaneme k faktorizácii kovariančnej matice, zopakujme si niekoľko poznatkov z algebry.

Nech \mathbf{A} je symetrická $m \times m$ (štvorcová) reálna matica.

$$\det(\mathbf{A} - \lambda\mathbf{I}) \stackrel{\text{ozn.}}{=} |\mathbf{A} - \lambda\mathbf{I}| = 0$$

je charakteristická rovnica matice \mathbf{A} . Je to rovnica m -tého stupňa. Jej korene sú $\lambda_1, \lambda_2, \dots, \lambda_m$. Voláme ich vlastné (alebo charakteristické) čísla matice \mathbf{A} . Ku každému vlastnému číslu existuje nenulový vlastný (alebo charakteristický) vektor \mathbf{P}_i , že platí $\mathbf{A}\mathbf{P}_i = \lambda_i\mathbf{P}_i$.

Platí:

I. Ak $h(\mathbf{A}) = r$ (hodnosť matice \mathbf{A}), tak nula je $(m-r)$ -násobným koreňom rovnice $|\mathbf{A} - \lambda\mathbf{I}| = 0$.

II. Všetky vlastné čísla sú reálne a aj všetky vlastné vektory sú reálne. Môžeme vlastné čísla teda prečíslovať tak, aby $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$.

III. $\mathbf{P}_i, \mathbf{P}_j$ prislúchajúce $\lambda_i \neq \lambda_j$ sú navzájom ortogonálne. Bez újmy na všeobecnosti môžeme zvoliť $\mathbf{P}'_i\mathbf{P}_j = 1$ (ortonormálne).

IV. Ak \mathbf{A} je pozitívne definitná, tak $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0$. Ak \mathbf{A} je pozitívne semidefinitná a $h(\mathbf{A}) = r$, tak $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_m = 0$.

V. Existuje ortogonálna matica $\mathbf{P}_{m,m}$ ($\mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}$), že $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{\Lambda}$. Matica $\mathbf{P} = (\mathbf{P}_1 \dot{\vdots} \mathbf{P}_2 \dot{\vdots} \dots \dot{\vdots} \mathbf{P}_m)$ a

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{pmatrix}. \text{ Teda } \mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{\Lambda} \quad \text{a} \quad \mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'.$$

Rovniciam

$$\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \lambda_1\mathbf{P}_1\mathbf{P}_1' + \dots + \lambda_m\mathbf{P}_m\mathbf{P}_m'$$

$$\mathbf{I} = \mathbf{P}\mathbf{P}' = \mathbf{P}_1\mathbf{P}_1' + \dots + \mathbf{P}_m\mathbf{P}_m'$$

hovoríme spektrálny rozklad matice \mathbf{A} (bližšie pozri napr. v Rao, C., R., *Lineárny metódy statistické indukce a jejích aplikace*, ACADEMIA, Praha, 1978).

Nech má náhodný vektor $\mathbf{X}_{n,1}$ kovariančnú maticu $\text{cov}(\mathbf{X}) = \mathbf{\Sigma}$. Ak $h(\mathbf{\Sigma}) = r \geq 1$, tak z predchádzajúceho $\mathbf{\Sigma} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$, kde

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \lambda_r & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

čiže

$$\begin{aligned} \mathbf{\Sigma} = (\mathbf{P}_1 \dot{\vdots} \mathbf{P}_2 \dot{\vdots} \dots \dot{\vdots} \mathbf{P}_m) & \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sqrt{\lambda_r} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \sqrt{\lambda_r} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1' \\ \mathbf{P}_2' \\ \vdots \\ \mathbf{P}_m' \end{pmatrix} = \\ & = (\sqrt{\lambda_1}\mathbf{P}_1 \dot{\vdots} \sqrt{\lambda_2}\mathbf{P}_2 \dot{\vdots} \dots \dot{\vdots} \sqrt{\lambda_r}\mathbf{P}_r) \begin{pmatrix} \sqrt{\lambda_1}\mathbf{P}_1' \\ \sqrt{\lambda_2}\mathbf{P}_2' \\ \vdots \\ \sqrt{\lambda_r}\mathbf{P}_r' \end{pmatrix} = \mathbf{B}_{n,r}\mathbf{B}'_{r,n} \end{aligned}$$

a platí $h(\mathbf{B}_{n,r}) = r$. Rozklad kovariančnej matice $\mathbf{\Sigma}_{n,n} = \mathbf{B}_{n,r}\mathbf{B}'_{r,n}$, pričom $h(\mathbf{B}_{n,r}) = r$, sa volá faktorizácia kovariančnej matice.

Veta 2.9: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{\Sigma})$, $h(\mathbf{\Sigma}) = r \geq 1$, $\mathbf{\Sigma} = \mathbf{B}_{n,r}\mathbf{B}'$, $h(\mathbf{B}) = r$. Nech $\mathbf{U} = (U_1, \dots, U_r)' \sim N_r(\mathbf{0}, \mathbf{I}_{r,r})$. Potom \mathbf{X} a $\boldsymbol{\mu} + \mathbf{B}\mathbf{U}$ majú rovnaké rozdelenie pravdepodobnosti (z pravdepodobnostného hľadiska sú ekvivalentné, nerozoznáme ich).

Dôkaz: $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}')$, $\mathbf{U} \sim N_r(\mathbf{0}, \mathbf{I})$. Podľa Vety 2.6 $\boldsymbol{\mu} + \mathbf{B}\mathbf{U} \sim N_n(\boldsymbol{\mu} + \mathbf{B}\mathbf{0} = \boldsymbol{\mu}, \mathbf{B}\text{cov}(\mathbf{U})\mathbf{B}' = \boldsymbol{\Sigma})$. Q.E.D.

Poznámka: Ak $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $h(\boldsymbol{\Sigma}) = r < n$, tak povieme, že \mathbf{X} má singulárne normálne rozdelenie (nemá napr. hustotu na celom \mathcal{R}^n). Popíšeme ho pomocou vektora $\mathbf{U} = (U_1, \dots, U_r)'$, kde $U_i \sim N(0, 1)$ sú nezávislé (pomocou Vety 2.9).

Skôr ako si odvodíme hustotu mnohorozmerného normálneho rozdelenia, zopakujme si vetu o hustote transformovaného náhodného vektora.

Veta 2.10: (O hustote transformovaného náhodného vektora) Nech náhodný vektor $\mathbf{X} = (X_1, \dots, X_n)'$ má hustotu $p(\mathbf{x})$ vzhľadom k Lebesgueovej miere v \mathcal{R}^n . Nech $\mathbf{t} : \mathcal{R}^n \rightarrow \mathcal{R}^n$ je regulárne a prosté zobrazenie na otvorenej množine $G \subset \mathcal{R}^n$, pre ktorú $\int_G p(\mathbf{x})d\mathbf{x} = 1$, t.j.

1. (prosté) $\mathbf{x}_1 \in G, \mathbf{x}_2 \in G, \mathbf{x}_1 \neq \mathbf{x}_2 \implies \mathbf{t}(\mathbf{x}_1) \neq \mathbf{t}(\mathbf{x}_2)$,

2. (regulárne) G je otvorená podmnožina \mathcal{R}^n ,

3. (regulárne) pre každé $\mathbf{x} \in G$ existuje spojitá $\frac{\partial t_i(\mathbf{x})}{\partial x_j}, i, j \in \{1, 2, \dots, n\}$,

4. (regulárne) pre každé $\mathbf{x} \in G$ je Jakobián $D_{\mathbf{t}}(\mathbf{x}) \stackrel{\text{ozn.}}{=} \det\left(\frac{\partial \mathbf{t}}{\partial \mathbf{x}'}\right) = \det\left(\begin{array}{ccc} \frac{\partial t_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial t_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial t_2(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial t_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial t_n(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial t_n(\mathbf{x})}{\partial x_n} \end{array}\right) \neq 0$.

Označme $\boldsymbol{\tau} : \mathbf{t}(G) \rightarrow G$ inverzné zobrazenie k $\boldsymbol{\tau}$ (teda $(\boldsymbol{\tau}(\mathbf{t}(\mathbf{x}))) = \mathbf{x}$ pre všetky $\mathbf{x} \in G$). Náhodný vektor $\mathbf{Y} = \mathbf{t}(\mathbf{X})$ má hustotu vzhľadom k Lebesgueovej miere rovnú

$$q(\mathbf{y}) = \begin{cases} p[\boldsymbol{\tau}(\mathbf{y})]|D_{\boldsymbol{\tau}}(\mathbf{y})|, & \text{pre } \mathbf{y} \in \mathbf{t}(G), \\ 0 & \text{pre } \mathbf{y} \notin \mathbf{t}(G). \end{cases}$$

Dôkaz: nájdeme v Jarník, V., *Integrální počet I. II.* Praha, NČSAV, 1955-1956, alebo Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 47.

Veta 2.11: Nech $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $h(\boldsymbol{\Sigma}) = n$. Potom existuje hustota vektora \mathbf{Y} a rovná sa

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \boldsymbol{\Sigma}}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}, \quad \mathbf{y} \in \mathcal{R}^n.$$

Dôkaz: Kovariančnú maticu faktorizujeme a dostávame $\boldsymbol{\Sigma} = \mathbf{B}_{n,n}\mathbf{B}'$, $h(\mathbf{B}) = n$. Podľa Vety 2.9 \mathbf{Y} sa dá písať ako $\boldsymbol{\mu} + \mathbf{B}\mathbf{U}$, $\mathbf{U} = (U_1, \dots, U_n)' \sim N_n(\mathbf{0}, \mathbf{I}_{n,n})$, a podľa Vety 2.8 sú $U_j \sim N_1(0, 1)$, $j = 1, 2, \dots, n$ nezávislé. Majú hustotu $\varphi_{U_j}(u_j) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u_j^2}{2}}$, $j = 1, 2, \dots, n$. Hustota náhodného vektora \mathbf{U} je

$$\psi_{\mathbf{U}}(\mathbf{u}) = \varphi_{U_1}(u_1) \dots \varphi_{U_n}(u_n) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\mathbf{u}'\mathbf{u}}, \quad \mathbf{u} \in \mathcal{R}^n.$$

Náhodný vektor $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{B}\mathbf{U}$, teda \mathbf{Y} je transformovaný vektor \mathbf{U} . Zobrazenie $\mathbf{t} : \mathcal{R}^n \rightarrow \mathcal{R}^n$ dané predpisom $\mathbf{t}(\mathbf{u}) = \boldsymbol{\mu} + \mathbf{B}\mathbf{u}$ je prosté a regulárne na celom \mathcal{R}^n a $\mathbf{Y} = \mathbf{t}(\mathbf{U})$.

Naozaj, ak $\mathbf{u}_1 \neq \mathbf{u}_2 \implies \boldsymbol{\mu} + \mathbf{B}\mathbf{u}_1 \neq \boldsymbol{\mu} + \mathbf{B}\mathbf{u}_2$ (lebo \mathbf{B} je regulárna matica).

$$\frac{\partial t_j(\mathbf{u})}{\partial u_i} = \frac{\partial(\mu_j + \{\mathbf{B}\}_{j1}u_1 + \dots + \{\mathbf{B}\}_{jn}u_n)}{\partial u_i} = \{\mathbf{B}\}_{ji}, \quad i, j \in \{1, 2, \dots, n\}$$

existuje a je spojité pre všetky $\mathbf{u} \in \mathcal{R}^n$.

$$\text{Jakobián } D_{\mathbf{t}}(\mathbf{u}) = \det \begin{pmatrix} \frac{\partial t_1(\mathbf{u})}{\partial u_1} & \cdots & \frac{\partial t_1(\mathbf{u})}{\partial u_n} \\ \frac{\partial t_2(\mathbf{u})}{\partial u_1} & \cdots & \frac{\partial t_2(\mathbf{u})}{\partial u_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial t_n(\mathbf{u})}{\partial u_1} & \cdots & \frac{\partial t_n(\mathbf{u})}{\partial u_n} \end{pmatrix} = \det \mathbf{B} \neq 0 \quad \text{pre všetky } \mathbf{u} \in \mathcal{R}^n.$$

Inverzné zobrazenie k \mathbf{t} je $\boldsymbol{\tau} : \mathcal{R}^n \rightarrow \mathcal{R}^n$ dané predpisom $\boldsymbol{\tau}(\mathbf{y}) = \mathbf{B}^{-1}(\mathbf{y} - \boldsymbol{\mu})$ (lebo $\boldsymbol{\tau}(\mathbf{t}(\mathbf{u})) = \mathbf{B}^{-1}(\mathbf{t}(\mathbf{u}) - \boldsymbol{\mu}) = \mathbf{B}^{-1}([\boldsymbol{\mu} + \mathbf{B}\mathbf{u}] - \boldsymbol{\mu}) = \mathbf{u}$ pre všetky $\mathbf{u} \in \mathcal{R}^n$).

Náhodný vektor $\mathbf{Y} = \mathbf{t}(\mathbf{U})$ má podľa Vety 2.10 hustotu rovnú

$$f_{\mathbf{Y}}(\mathbf{y}) = \psi(\boldsymbol{\tau}(\mathbf{y})) |D_{\boldsymbol{\tau}}(\mathbf{y})| = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}[\boldsymbol{\tau}(\mathbf{y})]'[\boldsymbol{\tau}(\mathbf{y})]} |D_{\boldsymbol{\tau}}(\mathbf{y})| = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'(\mathbf{B}^{-1})'\mathbf{B}^{-1}(\mathbf{y} - \boldsymbol{\mu})} |\det \mathbf{B}^{-1}| =$$

$$\frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|\det \boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})},$$

lebo $\mathbf{B}\mathbf{B}' = \boldsymbol{\Sigma} \implies \boldsymbol{\Sigma}^{-1} = (\mathbf{B}')^{-1}\mathbf{B}^{-1}$, ale $(\mathbf{B}')^{-1} = (\mathbf{B}^{-1})'$ (čo vyplýva z rovnosti $\mathbf{B}'(\mathbf{B}^{-1})' = (\mathbf{B}^{-1}\mathbf{B})' = \mathbf{I}$), teda

$$\boldsymbol{\Sigma}^{-1} = (\mathbf{B}^{-1})'\mathbf{B}^{-1}$$

a $\det \mathbf{B} = \det \mathbf{B}'$, teda $(0 <) \det \boldsymbol{\Sigma} = \det \mathbf{B}\mathbf{B}' = (\det \mathbf{B})^2$, čiže $|\det \mathbf{B}| = \sqrt{|\det \boldsymbol{\Sigma}|}$, z čoho dostávame $|\det \mathbf{B}^{-1}| = \frac{1}{|\det \mathbf{B}|} = \frac{1}{\sqrt{|\det \boldsymbol{\Sigma}|}}$. Q.E.D.

Dôsledok 2.12: V prípade $n = 2$ dostávame hustotu dvojrozmerného normálneho rozdelenia

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1 - \rho^2}\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)}} e^{-\frac{1}{2(1-\rho^2)} \left[\frac{(x_1 - \mu_1)^2}{\mathcal{D}(X_1)} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)}} + \frac{(x_2 - \mu_2)^2}{\mathcal{D}(X_2)} \right]},$$

lebo $\boldsymbol{\Sigma} = \begin{pmatrix} \mathcal{D}(X_1) & \text{cov}(X_1, X_2) \\ \text{cov}(X_2, X_1) & \mathcal{D}(X_2) \end{pmatrix}$ je pozitívne definitná práve vtedy ak $\mathcal{D}(X_1) > 0$, $\mathcal{D}(X_2) > 0$ a

$\rho^2 = \frac{\text{cov}^2(X_1, X_2)}{\mathcal{D}(X_1)\mathcal{D}(X_2)} \neq 1$ a v takomto prípade (podľa vzorca $\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$) je

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\mathcal{D}(X_1)\mathcal{D}(X_2) - \text{cov}^2(X_1, X_2)} \begin{pmatrix} \mathcal{D}(X_2) & -\text{cov}(X_1, X_2) \\ -\text{cov}(X_1, X_2) & \mathcal{D}(X_1) \end{pmatrix} =$$

$$= \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\mathcal{D}(X_1)} & -\frac{\rho}{\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)}} \\ -\frac{\rho}{\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)}} & \frac{1}{\mathcal{D}(X_2)} \end{pmatrix}.$$

Poznámka 2.13: Ak má náhodný vektor $\mathbf{X}_{n,1} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}$, kde $\mathbf{X}_1 \in \mathcal{R}^k$, $\mathbf{X}_2 \in \mathcal{R}^{n-k}$, hustotu $f(\mathbf{x}_1, \mathbf{x}_2)$, tak podmienené rozdelenie $\mathbf{X}_1/\mathbf{X}_2 = \mathbf{x}_2$ (v tomto absolútne spojitom prípade) má hustotu

$$\varphi(\mathbf{x}_1/\mathbf{x}_2) = \frac{f(\mathbf{x}_1, \mathbf{x}_2)}{\int_{\mathcal{R}^k} f(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1}.$$

Lema 2.14: Nech $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ je pozitívne definitná, tak

- (i) Σ_{22} je pozitívne definitná matica,
- (ii) $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ je pozitívne definitná matica.

Dôkaz:

(i) Σ je pozitívne definitná, teda $\forall \mathbf{y} \in \mathcal{R}^n \quad \mathbf{y}'\Sigma\mathbf{y} \geq 0$ a $\mathbf{y}'\Sigma\mathbf{y} = 0 \iff \mathbf{y} = \mathbf{0}$. Z toho vyplýva, že $\forall \mathbf{x} \in \mathcal{R}^{n-k} \quad (\mathbf{0}', \mathbf{x}')\Sigma \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix} = \mathbf{x}'\Sigma_{22}\mathbf{x} \geq 0$ a $(\mathbf{0}', \mathbf{x}')\Sigma \begin{pmatrix} \mathbf{0} \\ \mathbf{x} \end{pmatrix} = \mathbf{x}'\Sigma_{22}\mathbf{x} = 0 \iff \mathbf{x} = \mathbf{0}$.

(ii) $\mathbf{L} = \begin{pmatrix} \mathbf{I} & -\Sigma_{12}\Sigma_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ je regulárna (lebo $\det \mathbf{L} = 1$), teda

$$\begin{aligned} \mathbf{L}\Sigma\mathbf{L}' &= \begin{pmatrix} \mathbf{I} & -\Sigma_{12}\Sigma_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\Sigma_{22}^{-1}\Sigma_{21} & \mathbf{I} \end{pmatrix} = \\ &= \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & \mathbf{0} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\Sigma_{22}^{-1}\Sigma_{21} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} \end{pmatrix} \end{aligned}$$

je pozitívne definitná, teda

$$0 < \det \Sigma = \det(\mathbf{L}\Sigma\mathbf{L}') = \det(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) \det(\Sigma_{22}) \quad (4)$$

a preto $\det(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}) > 0$ a $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = (\mathbf{I} - \Sigma_{12}\Sigma_{22}^{-1}) \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ -\Sigma_{22}^{-1}\Sigma_{21} \end{pmatrix}$ je pozitívne definitná matica. Q.E.D.

Lema 2.15: Platí

$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{I} & \Sigma_{12}\Sigma_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \quad \text{a} \quad (\mathbf{L}')^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \Sigma_{22}^{-1}\Sigma_{21} & \mathbf{I} \end{pmatrix} = (\mathbf{L}^{-1})'.$$

Dôkaz: je jednoduchý, spravte ho ako cvičenie.

Lema 2.16: Nech $\mathbf{x}_1, \boldsymbol{\mu}_1 \in \mathcal{R}^k$, $\mathbf{x}_2, \boldsymbol{\mu}_2 \in \mathcal{R}^{n-k}$. Platí

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) &= (\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)])'(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)]) + \\ &\quad + (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2). \end{aligned}$$

Dôkaz: Počítajme

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) &= (\mathbf{x}'_1 - \boldsymbol{\mu}'_1, \mathbf{x}'_2 - \boldsymbol{\mu}'_2) \underbrace{\mathbf{L}'(\mathbf{L}')^{-1}\Sigma^{-1}(\mathbf{L})^{-1}\mathbf{L}}_{(\mathbf{L}\Sigma\mathbf{L}')^{-1}} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix} = \\ &= (\mathbf{x}'_1 - \boldsymbol{\mu}'_1, \mathbf{x}'_2 - \boldsymbol{\mu}'_2) \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\Sigma_{22}^{-1}\Sigma_{21} & \mathbf{I} \end{pmatrix} \begin{pmatrix} (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\Sigma_{12}\Sigma_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix} = \\ &= (\mathbf{x}'_1 - \boldsymbol{\mu}'_1 - (\mathbf{x}'_2 - \boldsymbol{\mu}'_2)\Sigma_{22}^{-1}\Sigma_{21}, \mathbf{x}'_2 - \boldsymbol{\mu}'_2) \begin{pmatrix} (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{pmatrix} = \end{aligned}$$

$$= (\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)])'(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1}(\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)]) + \\ + (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2). \quad \text{Q.E.D.}$$

Využívajúc (4) a Lemu 2.16 ľahko dostaneme

Dôsledok 2.17: Nech $\mathbf{X}_{n,1} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right)$ (označenie podľa (3) a Vety 2.7), potom hustota \mathbf{X} je

$$f_{\mathbf{X}}(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det \boldsymbol{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})} = \\ f_{\mathbf{X}}(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{(2\pi)^{\frac{k}{2}} \sqrt{\det(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})}} e^{-\frac{1}{2}(\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)])'(\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})^{-1}(\mathbf{x}_1 - [\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)])} \times \\ \frac{1}{(2\pi)^{\frac{n-k}{2}} \sqrt{\det \boldsymbol{\Sigma}_{22}}} e^{-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)} = \psi(\mathbf{x}_1, \mathbf{x}_2) f_2(\mathbf{x}_2).$$

Podmienené rozdelenie $\mathbf{X}_1/\mathbf{X}_2 = \mathbf{x}_2$ má podľa Poznámky 2.13 hustotu

$$\varphi(\mathbf{x}_1/\mathbf{x}_2) = \frac{f(\mathbf{x}_1, \mathbf{x}_2)}{\int_{\mathcal{R}^k} f(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1} = \frac{\psi(\mathbf{x}_1, \mathbf{x}_2) f_2(\mathbf{x}_2)}{\int_{\mathcal{R}^k} \psi(\mathbf{x}_1, \mathbf{x}_2) f_2(\mathbf{x}_2) d\mathbf{x}_1} = \frac{\psi(\mathbf{x}_1, \mathbf{x}_2)}{\int_{\mathcal{R}^k} \psi(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1}.$$

Vidíme, že pri pevnom \mathbf{x}_2 je $\psi(\mathbf{x}_1, \mathbf{x}_2)$ hustotou náhodného vektora $\boldsymbol{\xi} \sim N_k(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$ a teda $\int_{\mathcal{R}^k} \psi(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1 = 1$, čiže hustota $\varphi(\mathbf{x}_1/\mathbf{x}_2)$ podmieneného rozdelenia $\mathbf{X}_1/\mathbf{X}_2 = \mathbf{x}_2$ je $\psi(\mathbf{x}_1, \mathbf{x}_2)$. Dokázali sme, že

$$\mathbf{X}_1/\mathbf{X}_2 = \mathbf{x}_2 \sim N_k(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}). \quad (5)$$

Dôležitý je špeciálny prípad $n = 2$. Ak $\mathbf{X}_{2,1}$ má regulárne dvojrozmerné normálne rozdelenie, teda $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \mathcal{D}(X_1) & \varrho\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)} \\ \varrho\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)} & \mathcal{D}(X_2) \end{pmatrix} \right)$, tak

$$X_1/X_2 = x_2 \sim N_1 \left(\mu_1 + \varrho\sqrt{\frac{\mathcal{D}(X_1)}{\mathcal{D}(X_2)}}(x_2 - \mu_2), \mathcal{D}(X_1)(1 - \varrho^2) \right). \quad (6)$$

3 Rozdelenie kvadratických foriem

Definícia 3.1: Náhodná veličina X má Gama rozdelenie s parametrami a, b , $a > 0$, $b > 0$ (označme $X \sim \Gamma(a, b)$), ak má hustotu

$$h(x) = \begin{cases} \frac{a^b}{\Gamma(b)} e^{-ax} x^{b-1}, & \text{ak } x > 0, \\ 0, & \text{ak } x \leq 0. \end{cases}$$

Ponamenávame len, že pre funkciu gama platí $\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx$, $a > 0$, $\Gamma(n) = (n-1)!$ pre n prirodzené číslo, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $a\Gamma(a) = \Gamma(a+1)$ pre každé kladné číslo a . S funkciou gama je úzko spätá funkcia beta (označujeme $B(a, b)$, $a > 0, b > 0$), $B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$. Medzi funkciami gama beta platí vzťah $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Definícia 3.2: Náhodná veličina Y má χ_n^2 rozdelenie (chíkvadrát rozdelenie s n stupňami voľnosti), ak má $\Gamma\left(\frac{1}{2}, \frac{n}{2}\right)$ rozdelenie. Teda Y má hustotu

$$f(y) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{y}{2}} y^{\frac{n}{2}-1}, & \text{ak } y > 0, \\ 0, & \text{ak } y \leq 0. \end{cases}$$

Veta 3.3: Nech X_1, X_2, \dots, X_n sú nezávislé $N(0, 1)$ rozdelené náhodné veličiny. Náhodná veličina

$$Y = X_1^2 + X_2^2 + \dots + X_n^2$$

má χ_n^2 rozdelenie.

Dôkaz: (indukciou) Pre $n = 1$, nech $X_1 \sim N(0, 1)$ a $x > 0$, tak distribučná funkcia

$$F_{X_1^2}(x) = P\{X_1^2 < x\} = P\{-\sqrt{x} \leq X_1 < \sqrt{x}\} - P\{\sqrt{x} = X_1\} = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

teda hustota

$$\begin{aligned} f_{X_1^2}(x) &= \frac{d}{dx} F_{X_1^2}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{x})^2}{2}} (\sqrt{x})' - \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{x})^2}{2}} (-\sqrt{x})' = \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}} \frac{1}{2\sqrt{x}} - \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}} \left(-\frac{1}{2\sqrt{x}}\right) = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{x}{2}} x^{-\frac{1}{2}} = \frac{1}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)} e^{-\frac{x}{2}} x^{-\frac{1}{2}} \end{aligned}$$

a $f_{X_1^2}(x) = 0$ pre $x \leq 0$.

Nech teraz $X_1^2 + X_2^2 + \dots + X_n^2$ má pre $x > 0$ hustotu $\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}$. Potom

$$\begin{aligned} f_{X_1^2+X_2^2+\dots+X_{n+1}^2}(x) &= \int_0^\infty f_{X_1^2+X_2^2+\dots+X_n^2}(x-u) f_{X_{n+1}^2}(u) du = \\ &= \int_0^x \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x-u}{2}} (x-u)^{\frac{n}{2}-1} \frac{1}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)} e^{-\frac{u}{2}} u^{-\frac{1}{2}} du = \\ &= \frac{e^{-\frac{x}{2}}}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} = \int_0^x (x-u)^{\frac{n}{2}-1} u^{-\frac{1}{2}} du = \quad (\text{substitúcia } \frac{u}{x} = w, \quad du = xdw) \\ &= \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1} x^{-\frac{1}{2}} x}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_0^1 (1-w)^{\frac{n}{2}-1} w^{\frac{1}{2}-1} dw = \frac{e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1}}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} B\left(\frac{n}{2}, \frac{1}{2}\right) = \\ &= \frac{e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1}}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} = \frac{1}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n+1}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n+1}{2}-1}. \quad \text{Q.E.D.} \end{aligned}$$

Poznámka 3.4: Veta 3.3 je alternatívnou definíciou χ_n^2 rozdelenia.

Veta 3.5: Nech $Y \sim \chi_r^2$, $Z \sim \chi_s^2$ sú nezávislé náhodné veličiny. Potom $Y + Z \sim \chi_{r+s}^2$.

Dôkaz: $Y = X_1^2 + X_2^2 + \dots + X_r^2$, $Z = X_{r+1}^2 + X_{r+2}^2 + \dots + X_{r+s}^2$, pričom všetky X_1, X_2, \dots, X_{r+s} sú nezávislé a $N(0, 1)$ rozdelené. Preto $Y + Z = X_1^2 + \dots + X_{r+s}^2 \sim \chi_{r+s}^2$. Q.E.D.

Teraz si niečo zopakujeme z teórie zovšeobecnených inverzií matíc. Najprv si dokážeme vetu, ktorú budeme často používať.

Veta 3.6: Pre každú maticu $\mathbf{D}_{k,l}$ platí $\mathcal{M}(\mathbf{D}) = \mathcal{M}(\mathbf{D}\mathbf{D}')$, kde $\mathcal{M}(\mathbf{D}) = \{\mathbf{D}\mathbf{u} : \mathbf{u} \in \mathcal{R}^l\}$ je vektorový priestor generovaný stĺpcami matice \mathbf{D} (podpriestor priestora \mathcal{R}^k).

Dôkaz: Označme $[\mathcal{M}(\mathbf{D})]^\perp$ ortogonálny doplnok priestora $\mathcal{M}(\mathbf{D})$ v (celom) priestore \mathcal{R}^k . Platí $\mathcal{M}(\mathbf{D}) = \mathcal{M}(\mathbf{D}\mathbf{D}') \iff [\mathcal{M}(\mathbf{D})]^\perp = [\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp$. Budeme dokazovať rovnosť priestorov $[\mathcal{M}(\mathbf{D})]^\perp$ a $[\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp$.

Ak $\mathbf{z} \in [\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp \implies \mathbf{z}'\mathbf{D}\mathbf{D}' = \mathbf{0} \implies \mathbf{z}'\mathbf{D}\mathbf{D}'\mathbf{z} = 0 \implies (\mathbf{D}'\mathbf{z})(\mathbf{D}'\mathbf{z})' = 0 \implies \mathbf{D}'\mathbf{z} = \mathbf{0} \implies \mathbf{z}'\mathbf{D} = \mathbf{0} \implies \mathbf{z} \in [\mathcal{M}(\mathbf{D})]^\perp$, teda $[\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp \subset [\mathcal{M}(\mathbf{D})]^\perp$.

Ak $\mathbf{z} \in [\mathcal{M}(\mathbf{D})]^\perp \implies \mathbf{z}'\mathbf{D} = \mathbf{0} \implies \mathbf{z}'\mathbf{D}\mathbf{D}' = \mathbf{0} \implies \mathbf{z} \in [\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp$, teda $[\mathcal{M}(\mathbf{D})]^\perp \subset [\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp$

Dostávame, že $[\mathcal{M}(\mathbf{D})]^\perp = [\mathcal{M}(\mathbf{D}\mathbf{D}')]^\perp$. Q.E.D.

Definícia 3.7: Majme maticu $\mathbf{A}_{m,n}$. Maticu \mathbf{A}^- typu rozmerov $n \times m$ nazývame g-inverziou (zovšeobecnenou inverziou, pseudoinverziou) matice \mathbf{A} , ak platí $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$.

Veta 3.8: Pseudoinverzná matica k matici $\mathbf{A}_{m,n}$ vždy existuje. Nemusí byť jediná.

Dôkaz: (i) Ak $\mathbf{A} = \mathbf{0}$, tak ľubovoľná matica typu $n \times m$ je \mathbf{A}^- .

(ii) Ak $h(\mathbf{A}) = r \geq 1$, $r \leq \min\{m, n\}$, tak \mathbf{A} má r lineárne nezávislých stĺpcov. Vezmime tieto nezávislé stĺpce (v ľubovoľnom poradí) a dostaneme maticu $\mathbf{B}_{m,r}$. Každý stĺpec matice \mathbf{A} , teda $\{\mathbf{A}\}_{\cdot i}$, $i = 1, 2, \dots, n$ dostaneme ako lineárnu kombináciu stĺpcov matice \mathbf{B} , teda $\{\mathbf{A}\}_{\cdot i} = \mathbf{B}\mathbf{c}_i$, $i = 1, 2, \dots, n$. Maticu $(\mathbf{c}_1 : \mathbf{c}_2 : \dots : \mathbf{c}_n)_{r,n}$ označme \mathbf{C} . Teda $\mathbf{A} = \mathbf{B}\mathbf{C}$, pričom $h(\mathbf{B}) = r$. Pretože $h(\mathbf{A}) = r \leq \min\{h(\mathbf{B}), h(\mathbf{C})\} \leq h(\mathbf{C})$ a naopak $h(\mathbf{C}_{r,n}) \leq \min\{r, n\}$, dostávame $h(\mathbf{C}) = r$. Práve opísaný rozklad matice \mathbf{A}

$$\mathbf{A}_{m,n} = \mathbf{B}_{m,r}\mathbf{C}_{r,n}, \quad h(\mathbf{A}) = h(\mathbf{B}) = h(\mathbf{C}) = r \quad (7)$$

budeme často používať. Matica $\mathbf{B}'\mathbf{B}$ je rozmerov $r \times r$, pričom podľa Vety 3.6 je $\mathcal{M}(\mathbf{B}'\mathbf{B}) = \mathcal{M}(\mathbf{B}')$, teda aj $h(\mathbf{B}'\mathbf{B}) = h(\mathbf{B}') = h(\mathbf{B}) = r$. Matica $\mathbf{B}'\mathbf{B}$ je regulárna. Existuje inverzná matica $(\mathbf{B}'\mathbf{B})^{-1}$. Úplne analogicky dostaneme, že existuje matica $(\mathbf{C}\mathbf{C}')^{-1}$. Preto existuje aj matica $\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'$ (typu $n \times m$) a platí

$$\mathbf{A}\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{A} = \mathbf{B}\mathbf{C}\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{B}\mathbf{C} = \mathbf{B}\mathbf{C} = \mathbf{A},$$

čiže $\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-1}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}' = \mathbf{A}^-$. Q.E.D.

Veta 3.9: Ak $h(\mathbf{A}_{n,r}) = r \geq 1$, tak $\mathbf{A}^-\mathbf{A} = \mathbf{I}_r$ pre ľubovoľnú \mathbf{A}^- .

Dôkaz: Pre ľubovoľnú \mathbf{A}^- je $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$, teda pre každé $\mathbf{x} \in \mathcal{R}^r$ platí $\mathbf{A}(\mathbf{A}^-\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$. Pretože $h(\mathbf{A}) = r$, sú stĺpce matice \mathbf{A} lineárne nezávislé a z toho dostávame

$$\forall \mathbf{x} \in \mathcal{R}^r \quad \mathbf{A}[(\mathbf{A}^-\mathbf{A} - \mathbf{I})\mathbf{x}] = \mathbf{0} \implies \forall \mathbf{x} \in \mathcal{R}^r \quad (\mathbf{A}^-\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0},$$

čiže $\mathbf{A}^-\mathbf{A} - \mathbf{I} = \mathbf{0}$, alebo $\mathbf{A}^-\mathbf{A} = \mathbf{I}$. Q.E.D.

Veta 3.10: Ak $h(\mathbf{A}_{n,r}) = r \geq 1$, tak $\mathbf{A}'(\mathbf{A}')^- = \mathbf{I}_r$ pre ľubovoľnú $(\mathbf{A}')^-$.

Dôkaz: Pre ľubovoľnú $(\mathbf{A}')^-$ je $\mathbf{A}'(\mathbf{A}')^-\mathbf{A}' = \mathbf{A}'$, teda pre každé $\mathbf{x} \in \mathcal{R}^r$ platí $\mathbf{x}'(\mathbf{A}'(\mathbf{A}')^- - \mathbf{I})\mathbf{A}' = \mathbf{0}$. Pretože $h(\mathbf{A}') = r$, sú riadky matice \mathbf{A}' lineárne nezávislé a z toho dostávame

$$\forall \mathbf{x} \in \mathcal{R}^r \quad \mathbf{x}'[\mathbf{A}'(\mathbf{A}')^- - \mathbf{I}]\mathbf{A}' = \mathbf{0} \implies \forall \mathbf{x} \in \mathcal{R}^r \quad \mathbf{x}'(\mathbf{A}'(\mathbf{A}')^- - \mathbf{I}) = \mathbf{0},$$

čiže $\mathbf{A}'(\mathbf{A}')^- - \mathbf{I} = \mathbf{0}$, alebo $\mathbf{A}'(\mathbf{A}')^- = \mathbf{I}$. Q.E.D.

Veta 3.11: Ak $h(\mathbf{A}_{n,r}) = r \geq 1$, tak $\mathbf{A}'(\mathbf{A}\mathbf{A}')^-\mathbf{A} = \mathbf{I}_r$ pre ľubovoľnú $(\mathbf{A}\mathbf{A}')^-$.

Dôkaz: Pre ľubovoľnú $(\mathbf{A}\mathbf{A}')^-$ platí

$$\mathbf{A}\mathbf{A}'(\mathbf{A}\mathbf{A}')^-\mathbf{A}\mathbf{A}' = \mathbf{A}\mathbf{A}',$$

teda

$$\mathbf{A}^-\mathbf{A}\mathbf{A}'(\mathbf{A}\mathbf{A}')^-\mathbf{A}\mathbf{A}'(\mathbf{A}')^- = \mathbf{A}^-\mathbf{A}\mathbf{A}'(\mathbf{A}')^-.$$

Využijeme Vetu 3.9 a Vetu 3.10 a dostávame $\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-}\mathbf{A} = \mathbf{I}_r$. Q.E.D.

Veta 3.12: Ak $\mathbf{A}_{k,k}$ je idempotentná (teda $\mathbf{A}\mathbf{A} = \mathbf{A}$), tak $h(\mathbf{A}) = Tr\mathbf{A}$.

Dôkaz: (i) Ak $h(\mathbf{A}) = 0$, tak $\mathbf{A} = \mathbf{0}$ a tvrdenie je zřejmé.

(ii) Ak $h(\mathbf{A}) = r \geq 1$, tak podľa rozkladu (7) $\mathbf{A}_{k,k} = \mathbf{B}_{k,r}\mathbf{C}_{r,k}$, pričom $h(\mathbf{B}) = h(\mathbf{C}) = r$. Pretože \mathbf{A} je idempotentná, postupne platí

$$\mathbf{A}\mathbf{A} = \mathbf{A}$$

$$\mathbf{B}\mathbf{C}\mathbf{B}\mathbf{C} = \mathbf{B}\mathbf{C}$$

$$\mathbf{B}^{-}\mathbf{B}\mathbf{C}\mathbf{B}\mathbf{C}\mathbf{C}^{-} = \mathbf{B}^{-}\mathbf{B}\mathbf{C}\mathbf{C}^{-}$$

a podľa Vety 3.9 a Vety 3.10

$$\mathbf{C}\mathbf{B} = \mathbf{I}_r.$$

Dostávame $Tr\mathbf{A} = Tr\mathbf{B}\mathbf{C} = Tr\mathbf{C}\mathbf{B} = Tr\mathbf{I}_r = r = h(\mathbf{A})$. Q.E.D.

Teraz sa vráťme k rozdeleniam kvadratických foriem.

Veta 3.13: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $h(\boldsymbol{\Sigma}) = r \geq 1$ a nech $\boldsymbol{\Sigma}^{-}$ je ľubovoľná g-inverzia matice $\boldsymbol{\Sigma}$. Platí

$$(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_r^2.$$

Dôkaz: Faktorizujeme kovariančnú maticu a dostaneme $\boldsymbol{\Sigma} = \mathbf{B}_{n,r}\mathbf{B}'$, pričom $h(\mathbf{B}) = r$. Ak $\mathbf{U} \sim N_r(\mathbf{0}, \mathbf{I})$ tak podľa Vety 2.9 $\mathbf{X} = \boldsymbol{\mu} + \mathbf{B}\mathbf{U} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, teda

$$(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-} (\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{B}\mathbf{U})' \boldsymbol{\Sigma}^{-} \mathbf{B}\mathbf{U} = \mathbf{U}' \mathbf{B}' (\mathbf{B}\mathbf{B}')^{-} \mathbf{B}\mathbf{U} = \mathbf{U}' \mathbf{U} \sim \chi_r^2$$

podľa Vety 3.3 a Vety 3.11. Q.E.D.

Veta 3.14: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a $\mathbf{A}_{n,n}$ je symerická pozitívne semidefinitná matica, $\mathbf{A}\boldsymbol{\Sigma} \neq \mathbf{0}$, $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Sigma}$ (idempotentná). Potom

$$Y = (\mathbf{X} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_{Tr\mathbf{A}\boldsymbol{\Sigma}}^2.$$

Dôkaz: Pretože $\mathbf{A}\boldsymbol{\Sigma} \neq \mathbf{0}$, je $h(\mathbf{A}) = r \geq 1$ a existuje matica $\mathbf{B}_{n,r}$ s hodnotou $h(\mathbf{B}) = r$, že $\mathbf{A} = \mathbf{B}\mathbf{B}'$ (faktorizácia matice \mathbf{A}). Teda

$$Y = (\mathbf{X} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{X} - \boldsymbol{\mu})' \mathbf{B}\mathbf{B}' (\mathbf{X} - \boldsymbol{\mu}) = (\mathbf{B}'(\mathbf{X} - \boldsymbol{\mu}))' (\mathbf{B}'(\mathbf{X} - \boldsymbol{\mu})) = \boldsymbol{\xi}' \boldsymbol{\xi},$$

kde $\boldsymbol{\xi} = \mathbf{B}'\mathbf{X} - \mathbf{B}'\boldsymbol{\mu} \sim N_r(\mathbf{0}, \mathbf{B}'\boldsymbol{\Sigma}\mathbf{B})$. Pretože $\mathbf{A}\boldsymbol{\Sigma}$ je idempotentná, pomocou Vety 3.9 dostávame

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Sigma}$$

$$\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}$$

$$\mathbf{B}^{-}\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B} = \mathbf{B}^{-}\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}$$

$$\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B} = \mathbf{B}'\boldsymbol{\Sigma}\mathbf{B},$$

čiže matica $\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}$ je idempotentná a teda $(\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B})^{-} = \mathbf{I}_r$ (jedna jej g-inverzia). Je zřejmé, že $h(\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B}) = Tr\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B} = Tr\mathbf{B}\mathbf{B}'\boldsymbol{\Sigma} = Tr\mathbf{A}\boldsymbol{\Sigma} \geq 1$ (lebo $\mathbf{A}\boldsymbol{\Sigma} \neq \mathbf{0}$). Aplikujeme Vetu 3.13 na náhodný vektor $\boldsymbol{\xi} = \mathbf{B}'\mathbf{X} - \mathbf{B}'\boldsymbol{\mu}$ a dostávame

$$(\mathbf{B}'\mathbf{X} - \mathbf{B}'\boldsymbol{\mu} - \mathbf{0})' (\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B})^{-} (\mathbf{B}'\mathbf{X} - \mathbf{B}'\boldsymbol{\mu} - \mathbf{0}) \sim \chi_{h(\mathbf{B}'\boldsymbol{\Sigma}\mathbf{B})}^2$$

$$(\mathbf{X} - \boldsymbol{\mu})' \mathbf{B} \mathbf{I}_r \mathbf{B}' (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_{Tr \mathbf{A} \boldsymbol{\Sigma}}^2$$

$$(\mathbf{X} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_{Tr \mathbf{A} \boldsymbol{\Sigma}}^2. \quad \text{Q.E.D.}$$

Veta 3.15: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a $\mathbf{A}_{n,n}$, $\mathbf{B}_{m,n}$ reálne matice taká, že $\mathbf{B} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{0}$, \mathbf{A} je symetrická a pozitívne semidefinitná. Potom náhodný vektor $\mathbf{Y} = \mathbf{B} \mathbf{X} + \mathbf{b}$ a náhodná veličina $V = (\mathbf{X} - \mathbf{a})' \mathbf{A} (\mathbf{X} - \mathbf{a})$ sú nezávislé pre ľubovoľné vektory $\mathbf{b} \in \mathcal{R}^m$ a $\mathbf{a} \in \mathcal{R}^n$.

Dôkaz: Ak $\mathbf{A} = \mathbf{0}$, tak V je nezávislá s (ľubovoľným) \mathbf{Y} .

Ak $\mathbf{A} \neq \mathbf{0}$, $h(\mathbf{A}) = r \geq 1$, tak faktorizujeme $\mathbf{A} = \mathbf{C}_{n,r} \mathbf{C}'$, $h(\mathbf{C}) = r$. Náhodný vektor

$$\begin{pmatrix} \mathbf{C}' \mathbf{X} \\ \mathbf{B} \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{C}' \\ \mathbf{B} \end{pmatrix} \mathbf{X} \sim N \left(\begin{pmatrix} \mathbf{C}' \\ \mathbf{B} \end{pmatrix} \boldsymbol{\mu}, \begin{pmatrix} \mathbf{C}' \\ \mathbf{B} \end{pmatrix} \boldsymbol{\Sigma} (\mathbf{C} : \mathbf{B}') = \begin{pmatrix} \mathbf{C}' \boldsymbol{\Sigma} \mathbf{C} & \mathbf{C}' \boldsymbol{\Sigma} \mathbf{B}' \\ \mathbf{B} \boldsymbol{\Sigma} \mathbf{C} & \mathbf{B} \boldsymbol{\Sigma} \mathbf{B}' \end{pmatrix} \right).$$

Z predpokladu $\mathbf{B} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{0}$ postupne dostávame

$$\mathbf{B} \boldsymbol{\Sigma} \mathbf{C}' = \mathbf{0}$$

$$\mathbf{B} \boldsymbol{\Sigma} \mathbf{C}' (\mathbf{C}')^- = \mathbf{0}$$

a podľa Vety 3.10 $\mathbf{B} \boldsymbol{\Sigma} \mathbf{C} = \mathbf{0}$. Pretože $\text{cov}(\mathbf{B} \mathbf{X}, \mathbf{C}' \mathbf{X}) = \mathbf{B} \boldsymbol{\Sigma} \mathbf{C} = \mathbf{0}$ a $\begin{pmatrix} \mathbf{C}' \mathbf{X} \\ \mathbf{B} \mathbf{X} \end{pmatrix}$ je normálne rozdelený, sú vektory $\mathbf{B} \mathbf{X}$ a $\mathbf{C}' \mathbf{X}$ nezávislé. Čiže aj pre ľubovoľné vektory $\mathbf{b} \in \mathcal{R}^m$ a $\mathbf{a} \in \mathcal{R}^n$ sú $\mathbf{b} + \mathbf{B} \mathbf{X}$ a $\mathbf{C}'(\mathbf{X} - \mathbf{a})$ nezávislé, ale aj (ich funkcie) $\mathbf{Y} = \mathbf{B} \mathbf{X} + \mathbf{b}$ a $(\mathbf{C}'(\mathbf{X} - \mathbf{a}))'(\mathbf{C}'(\mathbf{X} - \mathbf{a})) = (\mathbf{X} - \mathbf{a})' \mathbf{A} (\mathbf{X} - \mathbf{a}) = V$. Q.E.D.

Veta 3.16: Nech $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a $\mathbf{A}_{n,n}$, $\mathbf{B}_{n,n}$ sú reálne symetrické pozitívne semidefinitné matice také, že $\mathbf{B} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{0}$. Potom náhodné veličiny $Y_1 = (\mathbf{X} - \mathbf{a})' \mathbf{A} (\mathbf{X} - \mathbf{a})$ a $Y_2 = (\mathbf{X} - \mathbf{b})' \mathbf{B} (\mathbf{X} - \mathbf{b})$ sú nezávislé pre ľubovoľné vektory $\mathbf{a} \in \mathcal{R}^n$ a $\mathbf{b} \in \mathcal{R}^n$.

Dôkaz: Stačí uvažovať prípad $h(\mathbf{A}) = r \geq 1$, $h(\mathbf{B}) = s \geq 1$. Faktorizujeme \mathbf{A} aj \mathbf{B} , teda $\mathbf{A} = \mathbf{C}_{n,r} \mathbf{C}'$, $h(\mathbf{C}) = r$ a $\mathbf{B} = \mathbf{G}_{n,s} \mathbf{G}'$, $h(\mathbf{G}) = s$. Z predpokladu $\mathbf{B} \boldsymbol{\Sigma} \mathbf{A} = \mathbf{0}$ dostávame $\mathbf{G} \mathbf{G}' \boldsymbol{\Sigma} \mathbf{C} \mathbf{C}' = \mathbf{0}$ a $\mathbf{G}^- \mathbf{G} \mathbf{G}' \boldsymbol{\Sigma} \mathbf{C} \mathbf{C}' (\mathbf{C}')^- = \mathbf{0}$, čiže (použijúc Vetu 3.9 a Vetu 3.10) $\mathbf{G}' \boldsymbol{\Sigma} \mathbf{C} = \mathbf{0}$. Teda $\mathbf{G}'(\mathbf{X} - \mathbf{a})$ a $\mathbf{C}'(\mathbf{X} - \mathbf{a})$ sú pre ľubovoľné vektory $\mathbf{a} \in \mathcal{R}^n$ a $\mathbf{b} \in \mathcal{R}^n$ nekorelované a v tomto prípade (normality) aj nezávislé. Ale potom aj ich funkcie $(\mathbf{C}'(\mathbf{X} - \mathbf{a}))'(\mathbf{C}'(\mathbf{X} - \mathbf{a})) = (\mathbf{X} - \mathbf{a})' \mathbf{C} \mathbf{C}' (\mathbf{X} - \mathbf{a}) = (\mathbf{X} - \mathbf{a})' \mathbf{A} (\mathbf{X} - \mathbf{a}) = Y_1$ a $(\mathbf{G}'(\mathbf{X} - \mathbf{b}))'(\mathbf{G}'(\mathbf{X} - \mathbf{b})) = (\mathbf{X} - \mathbf{b})' \mathbf{G} \mathbf{G}' (\mathbf{X} - \mathbf{b}) = (\mathbf{X} - \mathbf{b})' \mathbf{B} (\mathbf{X} - \mathbf{b}) = Y_2$ sú pre ľubovoľné vektory $\mathbf{a} \in \mathcal{R}^n$ a $\mathbf{b} \in \mathcal{R}^n$ nezávislé. Q.E.D.

Teraz uveďme jednu veľmi dôležitú aplikáciu predchádzajúcich viet v štatistike.

Veta 3.17: Nech X_1, X_2, \dots, X_n je náhodný výber z $N(\mu, \sigma^2)$. $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ je výberový priemer a $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ je výberový rozptyl. Potom platí

- (i) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$,
- (ii) pre $\sigma^2 > 0$, $n \geq 2$ je $\frac{n-1}{\sigma^2} S^2 \sim \chi_{n-1}^2$,
- (iii) pre $n \geq 2$ sú \bar{X} a S^2 nezávislé.

Dôkaz: Podľa Príkladu 1.6 keď označíme $\mathbf{X} = (X_1, \dots, X_n)'$, tak platí $\mathbf{X} \sim N_n(\mathbf{1}\mu, \sigma^2 \mathbf{I}_{n,n})$, kde $\mathbf{1}$ je n -rozmerný vektor, ktorého každá zložka je rovná 1. Potom ale

$$\bar{X} = \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n} \right) \mathbf{X} = \frac{1}{n} \mathbf{1}' \mathbf{X} \quad (8)$$

a podľa Príkladu 1.6 pre $n \geq 2$ je

$$S^2 = \mathbf{X}' \left\{ \frac{1}{n-1} \mathbf{I} - \frac{1}{n(n-1)} \mathbf{1} \mathbf{1}' \right\} \mathbf{X} = \mathbf{X}' \left\{ \frac{1}{n-1} \mathbf{I} - \frac{1}{n(n-1)} \mathbf{E} \right\} \mathbf{X} \quad (9)$$

($n \times n$ matica \mathbf{E} má všetky prvky rovné 1).

$$(i) \text{ Podľa Vety 2.6 dostávame } \bar{X} \sim N_1\left(\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)\mathbf{1}\mu = \mu, \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)\sigma^2\mathbf{I} \begin{pmatrix} \frac{1}{n} \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix} = \frac{\sigma^2}{n}\right).$$

(ii) Z (9) dostávame pre $n \geq 2$ $\frac{n-1}{\sigma^2}S^2 = \mathbf{X}'\left\{\frac{1}{\sigma^2}\mathbf{I} - \frac{1}{n\sigma^2}\mathbf{1}\mathbf{1}'\right\}\mathbf{X} = \mathbf{X}'\mathbf{A}\mathbf{X}$, pričom $\mathbf{A}\mathbf{1} = \frac{1}{\sigma^2}\mathbf{1} - \frac{1}{n\sigma^2}\mathbf{1}n = \mathbf{0}$.
Preto

$$\frac{n-1}{\sigma^2}S^2 = (\mathbf{X} - \mathbf{1}\mu)' \left\{ \frac{1}{\sigma^2}\mathbf{I} - \frac{1}{n\sigma^2}\mathbf{1}\mathbf{1}' \right\} (\mathbf{X} - \mathbf{1}\mu).$$

Ukážeme, že \mathbf{A} spĺňa predpoklady Vety 3.14.

Platí $\mathbf{A} = \mathbf{A}'$ (\mathbf{A} je symetrická), pre každý vektor $\mathbf{y} \in \mathcal{R}^n$ je $\mathbf{y}'\mathbf{A}\mathbf{y} = \frac{1}{\sigma^2}\mathbf{y}'(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{y} = \frac{1}{\sigma^2}\mathbf{y}'(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}')(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{y} \geq 0$ (\mathbf{A} je pozitívne semidefinitná), $\mathbf{A}\sigma^2\mathbf{I} \neq \mathbf{0}$ a $\mathbf{A}\sigma^2\mathbf{I}\mathbf{A}\sigma^2\mathbf{I} = \{\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\}\{\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\} = \{\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\} = \mathbf{A}\sigma^2\mathbf{I}$. Pretože $Tr\mathbf{A}\sigma^2\mathbf{I} = Tr\{\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\} = n - \frac{1}{n}Tr\mathbf{1}\mathbf{1}' = n - 1$, priamo z Vety 3.14 dostávame, že $\frac{n-1}{\sigma^2}S^2 \sim \chi_{n-1}^2$.

(iii) Pretože $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})\sigma^2\mathbf{I}\mathbf{A} = \frac{1}{n}\mathbf{1}'\{\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\} = \frac{1}{n}\mathbf{1}' - \frac{1}{n^2}\mathbf{1}'\mathbf{1}\mathbf{1}' = \mathbf{0}$, podľa Vety 3.15 sú \bar{X} a S^2 nezávislé. Q.E.D.

4 Teoretické základy lineárnej regresie a korelácie

Y, X_1, X_2, \dots, X_k nech sú náhodné veličiny na tom istom pravdepodobnostnom priestore (Ω, \mathcal{A}, P) s konečnými druhými momentami. Naším cieľom je predikovať Y pomocou \mathbf{X} . Predikciou rozumieme (vhodnú) náhodnú veličinu $\hat{Y} = g(X_1, \dots, X_n)$, kde $g: \mathcal{R}^k \rightarrow \mathcal{R}^1$ je merateľné zobrazenie. My sa budeme zaoberať lineárnu predikciou

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k = \beta_0 + \beta'\mathbf{X},$$

teda ak $g(x_1, \dots, x_k) = \beta_0 + \sum_{i=1}^k \beta_i x_i$. Kvalitu predikcie budme posudzovať *strednou kvadratickou chybou*

$$\mathcal{E}(Y - \hat{Y})^2.$$

Veta 4.1: Nech Y, X_1, X_2, \dots, X_k sú náhodné veličiny na tom istom pravdepodobnostnom priestore (Ω, \mathcal{A}, P) s konečnými druhými momentami a kovariančná matica náhodného vektora $\mathbf{X} = (X_1, \dots, X_n)'$ je pozitívne definitná. Pre lineárnu predikciu $\hat{Y} = \beta_0 + \beta'\mathbf{X}$ platí

$$\mathcal{E}(Y - \hat{Y})^2 = \mathcal{E}(Y - \beta_0 - \beta'\mathbf{X})^2 \geq \mathcal{D}(Y) - cov(Y, \mathbf{X})[cov(\mathbf{X})]^{-1}cov(\mathbf{X}, Y).$$

Rovnosť sa dosiahne práve ak

$$\beta_0 = \mathcal{E}(Y) - \beta'\mathcal{E}(\mathbf{X}) \quad \text{a} \quad \beta = [cov(\mathbf{X})]^{-1}cov(\mathbf{X}, Y). \quad (10)$$

Dôkaz: Pre ľubovoľnú náhodnú veličinu ξ s konečným druhým momentom platí $\mathcal{D}(\xi) = \mathcal{E}(\xi^2) - \mathcal{E}^2(\xi)$, preto

$$\mathcal{E}(Y - \hat{Y})^2 = \mathcal{D}(Y - \hat{Y}) + \mathcal{E}^2(Y - \hat{Y}) \geq \mathcal{D}(Y - \hat{Y})$$

(rovnosť nastane práve ak $\mathcal{E}(Y - \hat{Y}) = 0$, teda práve ak $\beta_0 = \mathcal{E}(Y) - \beta'\mathcal{E}(\mathbf{X})$)

$$= \mathcal{D}(Y - \beta_0 - \beta'\mathbf{X}) = \mathcal{D}(Y - \beta'\mathbf{X}) = \mathcal{D}\left(\begin{pmatrix} 1 \\ \vdots \\ -\beta' \end{pmatrix} \begin{pmatrix} Y \\ \mathbf{X} \end{pmatrix}\right) = \begin{pmatrix} 1 & -\beta' \end{pmatrix} \begin{pmatrix} \mathcal{D}(Y) & cov(Y, \mathbf{X}) \\ cov(\mathbf{X}, Y) & cov(\mathbf{X}) \end{pmatrix} \begin{pmatrix} 1 \\ -\beta \end{pmatrix} =$$

$$\begin{aligned}
&= \left(\mathcal{D}(Y) - \boldsymbol{\beta}' \text{cov}(\mathbf{X}, Y) : \text{cov}(Y, \mathbf{X}) - \boldsymbol{\beta}' \text{cov}(\mathbf{X}) \right) \begin{pmatrix} 1 \\ -\boldsymbol{\beta} \end{pmatrix} = \mathcal{D}(Y) - \boldsymbol{\beta}' \text{cov}(\mathbf{X}, Y) - \text{cov}(Y, \mathbf{X})\boldsymbol{\beta} + \boldsymbol{\beta}' \text{cov}(\mathbf{X})\boldsymbol{\beta} = \\
&= \mathcal{D}(Y) + [\boldsymbol{\beta} - (\text{cov}(\mathbf{X}))^{-1} \text{cov}(\mathbf{X}, Y)]' \text{cov}(\mathbf{X}) [\boldsymbol{\beta} - (\text{cov}(\mathbf{X}))^{-1} \text{cov}(\mathbf{X}, Y)] - \text{cov}(Y, \mathbf{X}) (\text{cov}(\mathbf{X}))^{-1} \text{cov}(\mathbf{X}, Y) \geq \\
&(\text{rovnosť nastane práve ak } \boldsymbol{\beta} = [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)) \\
&\geq \mathcal{D}(Y) - \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y). \quad \text{Q.E.D.}
\end{aligned}$$

Poznámka 4.2: Optimálna lineárna predikcia, t.j. lineárna predikcia s minimálnu strednekvadratickou chybou je

$$\hat{Y} = \mathcal{E}(Y) - \boldsymbol{\beta}' \mathcal{E}(\mathbf{X}) + \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} \mathbf{X} = \mathcal{E}(Y) + \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} (\mathbf{X} - \mathcal{E}(\mathbf{X})).$$

Pre túto predikciu platí

$$\begin{aligned}
\sigma_{Y, \mathbf{X}}^2 &= \mathcal{E} \left(Y - (\mathcal{E}(Y) - \boldsymbol{\beta}' \mathcal{E}(\mathbf{X}) + \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} \mathbf{X}) \right)^2 = \mathcal{D}(Y) - \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y) = \\
&= \mathcal{D}(Y) - \boldsymbol{\beta}' \text{cov}(\mathbf{X}) \boldsymbol{\beta}. \tag{11}
\end{aligned}$$

$\sigma_{Y, \mathbf{X}}^2$ sa volá *reziduálny rozptyl*. Funkcia

$$\hat{Y} = \mathcal{E}(Y) + \text{cov}(Y, \mathbf{X}) [\text{cov}(\mathbf{X})]^{-1} (\mathbf{X} - \mathcal{E}(\mathbf{X}))$$

sa volá *lineárna regresná funkcia*.

V prípade $k = 1$ je

$$\hat{Y} = \mathcal{E}(Y) - \frac{\text{cov}(X, Y)}{\mathcal{D}(X)} \mathcal{E}(X) + \frac{\text{cov}(X, Y)}{\mathcal{D}(X)} X = \mathcal{E}(Y) + \varrho_{X, Y} \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}} (X - \mathcal{E}(X)).$$

Regresná priamka preto je

$$\hat{y} = \mathcal{E}(Y) + \varrho_{X, Y} \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}} (x - \mathcal{E}(X))$$

(x je realizácia X a \hat{y} je realizácia \hat{Y}). Obyčajne teoretické (skutočné) hodnoty $\mathcal{E}(X)$, $\mathcal{E}(Y)$, $\mathcal{D}(X)$, $\mathcal{D}(Y)$, $\varrho_{X, Y}$ nepoznáme, ale použijeme ich odhady $\widehat{\mathcal{E}(X)} = \bar{X}$, $\widehat{\mathcal{E}(Y)} = \bar{Y}$, $\widehat{\mathcal{D}(X)} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, $\widehat{\mathcal{D}(Y)} = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$, $\widehat{\varrho_{X, Y}} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{j=1}^n (Y_j - \bar{Y})^2}}$.

Veta 4.3: Platí

$$0 \leq \sigma_{Y, \mathbf{X}}^2 \leq \mathcal{D}(Y).$$

Dôkaz: Obidve nerovnosti sú zrejmé zo vzťahu (11), prvá nerovnosť vyplýva z toho, že $\sigma_{Y, \mathbf{X}}^2$ je stredná hodnota nezápornej náhodnej veličiny. Q.E.D.

Poznámka 4.4: Ak sú Y a \mathbf{X} nekorelované, tak $\text{cov}(Y, \mathbf{X}) = \mathbf{0}$, teda z (11) je $\sigma_{Y, \mathbf{X}}^2 = \mathcal{D}(Y)$. Z (11) vyplýva, že vždy je $\sigma_{Y, \mathbf{X}}^2 \leq \mathcal{D}(Y)$, lebo $[\text{cov}(\mathbf{X})]^{-1}$ je pozitívne definitná matica (ľahko sa o tom presvedčíme).

Definícia 4.5: Nech $\text{cov}(\mathbf{X})$ je regulárna matica. *Koeficientom mnohonásobnej korelácie* medzi Y a \mathbf{X} nazývame korelačný koeficient $\varrho(Y, \hat{Y})$ a značíme $\varrho_{Y, \mathbf{X}}$, pričom \hat{Y} je optimálne lineárna predikcia, teda $\hat{Y} = \beta_0 + \boldsymbol{\beta}' \mathbf{X}$, kde $\beta_0 = \mathcal{E}(Y) - \boldsymbol{\beta}' \mathcal{E}(\mathbf{X})$ a $\boldsymbol{\beta} = [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)$. Ak $\mathcal{D}(Y) = 0$ alebo $\mathcal{D}(\hat{Y}) = 0$, tak položíme $\varrho_{Y, \mathbf{X}} = 0$.

Poznámka 4.6: $\mathcal{D}(\hat{Y}) = 0 \iff \beta' \text{cov}(\mathbf{X})\beta = 0 \iff \beta = \mathbf{0}$.

Veta 4.7: Platí

(i) $\varrho_{Y,\mathbf{X}} \geq 0$.

Ak $0 < \mathcal{D}(Y) < \infty$, tak

(ii) $\varrho_{Y,\mathbf{X}}^2 = \frac{\beta' \text{cov}(\mathbf{X})\beta}{\mathcal{D}(Y)}$,

(iii) $\varrho_{Y,\mathbf{X}}^2 = \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Y)$,

(iv) $\varrho_{Y,\mathbf{X}}^2 = 1 - \frac{\sigma_{Y,\mathbf{X}}^2}{\mathcal{D}(Y)}$.

Dôkaz: (i) Ak je $\mathcal{D}(Y) = 0$ & $\beta = \mathbf{0}$ alebo $\mathcal{D}(Y) = 0$ & $\beta \neq \mathbf{0}$, tak priamo z definície $\varrho(Y, \hat{Y}) = 0$. Ak $\mathcal{D}(Y) > 0$ & $\beta = \mathbf{0}$, tak podľa Poznámky 4.6 je $\mathcal{D}(\hat{Y}) = 0$ a opäť z definície je $\varrho(Y, \hat{Y}) = 0$. Preto stačí uvažovať $\mathcal{D}(Y) > 0$, $\beta \neq \mathbf{0}$. V tomto prípade

$$\begin{aligned} \varrho_{Y,\mathbf{X}} = \varrho(Y, \hat{Y}) &= \frac{\text{cov}(Y, \beta_0 + \beta' \mathbf{X})}{\sqrt{\mathcal{D}(Y)\mathcal{D}(\beta_0 + \beta' \mathbf{X})}} = \frac{\text{cov}(Y, \beta' \mathbf{X})}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} = \frac{\text{cov}\left(\begin{pmatrix} 1 & \mathbf{0}' \end{pmatrix} \begin{pmatrix} Y \\ \mathbf{X} \end{pmatrix}, \begin{pmatrix} 0 \\ \beta' \end{pmatrix} \begin{pmatrix} Y \\ \mathbf{X} \end{pmatrix}\right)}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} = \\ &= \frac{\begin{pmatrix} 1 & \mathbf{0}' \end{pmatrix} \begin{pmatrix} \mathcal{D}(Y) & \text{cov}(Y, \mathbf{X}) \\ \text{cov}(\mathbf{X}, Y) & \text{cov}(\mathbf{X}) \end{pmatrix} \begin{pmatrix} 0 \\ \beta \end{pmatrix}}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} = \frac{\begin{pmatrix} \mathcal{D}(Y) & \text{cov}(Y, \mathbf{X}) \end{pmatrix} \begin{pmatrix} 0 \\ \beta \end{pmatrix}}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} = \\ &= \frac{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} = \frac{\beta' \text{cov}(\mathbf{X})\beta}{\sqrt{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta}} > 0. \end{aligned}$$

(ii) Ak $\beta = \mathbf{0}$, tak podľa Poznámky 4.6 je $\mathcal{D}(\hat{Y}) = 0$ a z definície je $\varrho_{Y,\mathbf{X}} = 0$, ale takisto $\beta' \text{cov}(\mathbf{X})\beta = 0$, teda (ii) platí. Ak $\beta \neq \mathbf{0}$, tak (využívajúc odvodzovanie v (i)) dostávame

$$\begin{aligned} \varrho_{Y,\mathbf{X}}^2 = \varrho^2(Y, \hat{Y}) &= \frac{\text{cov}^2(Y, \beta_0 + \beta' \mathbf{X})}{\mathcal{D}(Y)\mathcal{D}(\beta_0 + \beta' \mathbf{X})} = \frac{\{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)\}^2}{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta} = \\ &= \frac{\{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)\}^2}{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta} = \frac{[\beta' \text{cov}(\mathbf{X})\beta]^2}{\mathcal{D}(Y)\beta' \text{cov}(\mathbf{X})\beta} = \frac{\beta' \text{cov}(\mathbf{X})\beta}{\mathcal{D}(Y)}. \end{aligned}$$

(iii) Ak $\beta = \mathbf{0}$, tak $\text{cov}(\mathbf{X}, Y) = \mathbf{0}$ ale aj $\text{cor}(\mathbf{X}, Y) = \mathbf{0}$, teda $\text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Y) = 0$. Na druhej strane v tomto prípade podľa Poznámky 4.6 je $\mathcal{D}(\hat{Y}) = 0$, teda $\varrho_{Y,\mathbf{X}}^2 = 0$. Ak $\beta \neq \mathbf{0}$, tak podľa (ii)

$$\begin{aligned} \varrho_{Y,\mathbf{X}}^2 &= \frac{\beta' \text{cov}(\mathbf{X})\beta}{\mathcal{D}(Y)} = \frac{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)}{\mathcal{D}(Y)} = \frac{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)}{\mathcal{D}(Y)} = \\ &= \frac{1}{\sqrt{\mathcal{D}(Y)}} \left(\text{cov}(Y, X_1) \dots \text{cov}(Y, X_k) \right) \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix} \left[\begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} \right] \times \\ &= \left[\text{cov}(\mathbf{X})^{-1} \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} \right] \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix} \begin{pmatrix} \text{cov}(X_1, Y) \\ \text{cov}(X_2, Y) \\ \vdots \\ \text{cov}(X_k, Y) \end{pmatrix} \frac{1}{\sqrt{\mathcal{D}(Y)}} = \\ &= \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Y) \end{aligned}$$

(využívajúc známy fakt, že pre regulárne matice \mathbf{A}, \mathbf{B} platí $(\mathbf{A}\mathbf{B}\mathbf{A})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$).

(iii) Podľa (11) pre reziduálny rozptyl platí

$$\sigma_{\hat{Y}, \mathbf{X}}^2 = \mathcal{D}(Y) - \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1}\text{cov}(\mathbf{X}, Y),$$

čiže

$$\sigma_{\hat{Y}, \mathbf{X}}^2 = \mathcal{D}(Y) - \boldsymbol{\beta}'\text{cov}(\mathbf{X})\boldsymbol{\beta}.$$

Keď vydělíme túto rovnicu nenulovou hodnotou $\mathcal{D}(Y)$ dostaneme

$$\frac{\sigma_{\hat{Y}, \mathbf{X}}^2}{\mathcal{D}(Y)} = 1 - \frac{\boldsymbol{\beta}'\text{cov}(\mathbf{X})\boldsymbol{\beta}}{\mathcal{D}(Y)}$$

a pomocou (ii) konečne dostávame

$$\varrho_{Y, \mathbf{X}}^2 = 1 - \frac{\sigma_{\hat{Y}, \mathbf{X}}^2}{\mathcal{D}(Y)}. \quad \text{Q.E.D.}$$

Poznámka 4.8: Obyčajne vyjadrujeme $\varrho_{Y, \mathbf{X}}^2$ pomocou prvkov korelačnej matice

$$\text{cor}(Y, \mathbf{X}) = \begin{pmatrix} 1 & \text{cor}(Y, \mathbf{X}) \\ \text{cor}(\mathbf{X}, Y) & \text{cor}(\mathbf{X}) \end{pmatrix}.$$

Poznámka 4.9:

- (i) Ak Y, \mathbf{X} sú nekorelované, tak $\boldsymbol{\beta} = \mathbf{0}$ a $\hat{Y} = \mathcal{E}(Y)$, čiže $\mathcal{D}(\hat{Y}) = 0$ a preto $\varrho_{Y, \mathbf{X}}^2 = 0$, teda $\varrho_{Y, \mathbf{X}} = 0$.
- (ii) Ak \hat{Y} je optimálna lineárna predikcia, teda $\hat{Y} = \mathcal{E}(Y) - \boldsymbol{\beta}'\mathcal{E}(\mathbf{X}) + \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1}\mathbf{X}$, $0 < \mathcal{D}(Y) < \infty$ a $Y = \hat{Y}$, tak $\varrho_{Y, \mathbf{X}} = \frac{\text{cov}(Y, Y)}{\sqrt{\mathcal{D}(Y)\mathcal{D}(Y)}} = 1$.
- (iii) $\varrho_{Y, \mathbf{X}}$ je ukazovateľ (miera) štatistickej väzby (stochastickej väzby) medzi Y a \mathbf{X} .
- (iv) $100\varrho_{Y, \mathbf{X}}^2$ udáva v % variabilitu Y , ktorá sa dá vysvetliť variabilitou \mathbf{X} .

Veta 4.10: Nech $0 < \mathcal{D}(Y) < \infty$. Potom platí

$$\varrho_{Y, \mathbf{X}} = \max_{\substack{d \in \mathcal{R}^1 \\ \mathbf{0} \neq \mathbf{b} \in \mathcal{R}^k}} |\varrho(Y, d + \mathbf{b}'\mathbf{X})|,$$

teda $\varrho_{Y, \mathbf{X}}$ je maximálny korelačný koeficient (v absolútnej hodnote) medzi Y a ľubovoľnou lineárnou kombináciou $\mathbf{d}'\mathbf{X} + d$.

Dôkaz: Nech $d \in \mathcal{R}^1$, $\mathbf{0} \neq \mathbf{b} \in \mathcal{R}^k$.

$$\begin{aligned} |\varrho(Y, d + \mathbf{b}'\mathbf{X})| &= \left| \frac{\text{cov}(1Y, d + \mathbf{b}'\mathbf{X})}{\sqrt{\mathcal{D}(Y)\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right| = \left| \frac{\text{cov}(1Y, \mathbf{b}'\mathbf{X})}{\sqrt{\mathcal{D}(Y)\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right| = \left| \frac{1\text{cov}(Y, \mathbf{X})\mathbf{b}}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right| = \\ &= \left| \frac{\text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1}\text{cov}(\mathbf{X})\mathbf{b}}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right| = \left| \frac{\boldsymbol{\beta}'\text{cov}(\mathbf{X})\mathbf{b}}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right|. \end{aligned}$$

Maticu $\text{cov}(\mathbf{X})$ faktorizujeme a píšeme $\text{cov}(\mathbf{X}) = \mathbf{B}\mathbf{B}'$, teda

$$\left| \frac{\boldsymbol{\beta}'\text{cov}(\mathbf{X})\mathbf{b}}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \right| = \frac{1}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} |(\mathbf{B}'\boldsymbol{\beta})'(\mathbf{B}'\mathbf{b})| \leq$$

(použijeme Schwarzovu nerovnosť, podľa ktorej pre ľubovoľné dva vektory $\mathbf{u}, \mathbf{w} \in \mathcal{R}^k$ platí $|\mathbf{u}'\mathbf{w}| \leq \sqrt{\mathbf{u}'\mathbf{u}}\sqrt{\mathbf{w}'\mathbf{w}}$)

$$\leq \frac{1}{\sqrt{\mathcal{D}(Y)}\sqrt{\mathbf{b}'\text{cov}(\mathbf{X})\mathbf{b}}} \sqrt{\boldsymbol{\beta}'\mathbf{B}\mathbf{B}'\boldsymbol{\beta}}\sqrt{\mathbf{b}'\mathbf{B}\mathbf{B}'\mathbf{b}} = \sqrt{\frac{\boldsymbol{\beta}'\text{cov}(\mathbf{X})\boldsymbol{\beta}}{\mathcal{D}(Y)}} = \sqrt{\varrho_{Y, \mathbf{X}}^2} = \varrho_{Y, \mathbf{X}}$$

(použili sme Vetu 4.7 (ii) a (i)).

Príklad 4.11: Vyjadrite $\varrho_{Z,(X,Y)}^2$ pomocou "obyčajných" korelačných koeficientov..

Riešenie: Podľa Vety 4.7 (iii) platí

$$\begin{aligned}\varrho_{Z,(X,Y)}^2 &= \text{cor}\left(Z, \begin{pmatrix} X \\ Y \end{pmatrix}\right) \left[\text{cor}\left(\begin{pmatrix} X \\ Y \end{pmatrix}\right)\right]^{-1} \text{cor}\left(\begin{pmatrix} X \\ Y \end{pmatrix}, Z\right) = (\varrho_{Z,X}, \varrho_{Z,Y}) \begin{pmatrix} 1 & \varrho_{X,Y} \\ \varrho_{Y,X} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varrho_{X,Z} \\ \varrho_{Y,Z} \end{pmatrix} = \\ &= \frac{1}{1 - \varrho_{X,Y}^2} (\varrho_{Z,X}, \varrho_{Z,Y}) \begin{pmatrix} 1 & -\varrho_{X,Y} \\ -\varrho_{X,Y} & 1 \end{pmatrix} \begin{pmatrix} \varrho_{X,Z} \\ \varrho_{Y,Z} \end{pmatrix} = \frac{\varrho_{Z,X}^2 + \varrho_{Z,Y}^2 - 2\varrho_{Z,X}\varrho_{Z,Y}\varrho_{X,Y}}{1 - \varrho_{X,Y}^2}.\end{aligned}$$

(Použili sme vzorec $\begin{pmatrix} a & b \\ b & c \end{pmatrix}^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$ a ak $\mathbf{X} = X$, tak $\varrho_{Z,\mathbf{X}}^2 = \varrho_{Z,X}^2$.)

Teraz si zavedieme parciálny korelačný koeficient.

Nech Y, Z, X_1, \dots, X_k sú náhodné veličiny na pravdepodobnostnom priestore (Ω, \mathcal{A}, P) , majú konečné druhé momenty, $\mathcal{D}(Y) \neq 0$, $\mathcal{D}(Z) \neq 0$ a $\text{cov}(\mathbf{X})$ je regulárna. Cieľom je získať mieru štatistickej (stochastickej) väzby medzi Y a Z pri eliminácii vplyvu $\mathbf{X} = (X_1, \dots, X_k)'$ ("očistenú závislosť"). V zhode s predchádzajúcim označením označme

\hat{Y} najlepšiu lineárnu predikciu Y pomocou \mathbf{X} ,

\hat{Z} najlepšiu lineárnu predikciu Z pomocou \mathbf{X} ,

teda $\hat{Y} = \beta_0 + \beta'\mathbf{X}$, kde $\beta_0 = \mathcal{E}(Y) - \beta'\mathcal{E}(\mathbf{X})$ a $\beta = [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y)$ a

teda $\hat{Z} = \alpha_0 + \alpha'\mathbf{X}$, kde $\alpha_0 = \mathcal{E}(Z) - \alpha'\mathcal{E}(\mathbf{X})$ a $\alpha = [\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Z)$.

Náhodné veličiny $R_Y = Y - \hat{Y}$, $R_Z = Z - \hat{Z}$ voláme rezíduá.

Definícia Veta 4.11: Nech platia označenia a predpoklady z predchádzajúceho odstavca. Korelačný koeficient $\varrho(R_Y, R_Z)$ nazývame *parciálnym korelačným koeficientom* medzi Y a Z pri danom náhodnom vektore \mathbf{X} (niekedy sa povie "pri eliminácii vplyvu náhodného vektora \mathbf{X} "). Značíme ho $\varrho_{Y,Z,\mathbf{X}}$ alebo $\varrho_{Y,Z,X_1,X_2,\dots,X_k}$. Ak $\mathcal{E}(Y - \hat{Y})^2 = \mathcal{D}(Y - \hat{Y}) = \sigma_{Y,\mathbf{X}}^2 = 0$ alebo $\mathcal{E}(Z - \hat{Z})^2 = \mathcal{D}(Z - \hat{Z}) = \sigma_{Z,\mathbf{X}}^2 = 0$ (teda ak $\varrho_{Y,\mathbf{X}} = 1$ alebo $\varrho_{Z,\mathbf{X}} = 1$), tak kladieme $\varrho_{Y,Z,\mathbf{X}} = 0$.

Veta 4.12: Ak $\mathcal{D}(Y - \hat{Y}) \neq 0$ a $\mathcal{D}(Z - \hat{Z}) \neq 0$, tak

$$\varrho_{Y,Z,\mathbf{X}} = \frac{\varrho_{Y,Z} - \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Z)}{\sqrt{(1 - \varrho_{Y,\mathbf{X}}^2)(1 - \varrho_{Z,\mathbf{X}}^2)}}. \quad (12)$$

Dôkaz: Platí

$$\varrho_{Y,Z,\mathbf{X}} = \frac{\text{cov}(Y - \hat{Y}, Z - \hat{Z})}{\sqrt{\mathcal{D}(Y - \hat{Y})\mathcal{D}(Z - \hat{Z})}}. \quad (13)$$

Počítajme

$$\begin{aligned}\text{cov}(Y - \hat{Y}, Z - \hat{Z}) &= \text{cov}(Y - \beta_0 - \beta'\mathbf{X}, Z - \alpha_0 - \alpha'\mathbf{X}) = \text{cov}(Y - \beta'\mathbf{X}, Z - \alpha'\mathbf{X}) = \\ &= \text{cov}\left((1:0: -\beta') \begin{pmatrix} Y \\ Z \\ \mathbf{X} \end{pmatrix}, (0:1: -\alpha') \begin{pmatrix} Y \\ Z \\ \mathbf{X} \end{pmatrix}\right) = \\ &= (1:0: -\beta') \begin{pmatrix} \mathcal{D}(Y) & \text{cov}(Y, Z) & \text{cov}(Y, \mathbf{X}) \\ \text{cov}(Z, Y) & \mathcal{D}(Z) & \text{cov}(Z, \mathbf{X}) \\ \text{cov}(\mathbf{X}, Y) & \text{cov}(\mathbf{X}, Z) & \text{cov}(\mathbf{X}) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -\alpha \end{pmatrix} =\end{aligned}$$

$$\begin{aligned}
&= \left(\mathcal{D}(Y) - \beta' \text{cov}(\mathbf{X}, Y) : \text{cov}(Y, Z) - \beta' \text{cov}(\mathbf{X}, Z) : \text{cov}(Y, \mathbf{X}) - \beta' \text{cov}(\mathbf{X}) \right) \begin{pmatrix} 0 \\ 1 \\ -\alpha \end{pmatrix} = \\
&= \text{cov}(Y, Z) - \beta' \text{cov}(\mathbf{X}, Z) - \text{cov}(Y, \mathbf{X})\alpha + \beta' \text{cov}(\mathbf{X})\alpha = \\
&= \text{cov}(Y, Z) - \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Z) - \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Z) + \\
&\quad + \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Z) = \\
&= \sqrt{\mathcal{D}(Y)\mathcal{D}(Z)} \left\{ \varrho_{Y,Z} - \frac{1}{\sqrt{\mathcal{D}(Y)}} \left(\text{cov}(Y, X_1) : \dots : \text{cov}(Y, X_k) \right) \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix} \right\} \times \\
&\quad \left[\begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} [\text{cov}(\mathbf{X})]^{-1} \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} \right] \times \\
&\quad \left. \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix} \begin{pmatrix} \text{cov}(X_1, Z) \\ \text{cov}(X_2, Z) \\ \vdots \\ \text{cov}(X_k, Z) \end{pmatrix} \frac{1}{\sqrt{\mathcal{D}(Z)}} \right\} = \\
&= \sqrt{\mathcal{D}(Y)\mathcal{D}(Z)} \{ \varrho_{Y,Z} - \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Z) \} = \text{cov}(Y - \hat{Y}, Z - \hat{Z}). \tag{14}
\end{aligned}$$

Ďalej platí

$$\begin{aligned}
\mathcal{D}(Y - \hat{Y}) &= \mathcal{D}(Y - \beta_0 - \beta' \mathbf{X}) = \mathcal{D}(Y - \beta' \mathbf{X}) = \mathcal{D} \left((1 : -\beta') \begin{pmatrix} Y \\ \mathbf{X} \end{pmatrix} \right) = \\
&= (1 : -\beta') \begin{pmatrix} \mathcal{D}(Y) & \text{cov}(Y, \mathbf{X}) \\ \text{cov}(\mathbf{X}, Y) & \text{cov}(\mathbf{X}) \end{pmatrix} \begin{pmatrix} 1 \\ -\beta \end{pmatrix} = \\
&= \left(\mathcal{D}(Y) - \beta' \text{cov}(\mathbf{X}, Y) : \text{cov}(Y, \mathbf{X}) - \beta' \text{cov}(\mathbf{X}) \right) \begin{pmatrix} 1 \\ -\beta \end{pmatrix} = \mathcal{D}(Y) - \beta' \text{cov}(\mathbf{X}, Y) - \text{cov}(Y, \mathbf{X})\beta + \beta' \text{cov}(\mathbf{X})\beta = \\
&= \mathcal{D}(Y) - \text{cov}(Y, \mathbf{X})[\text{cov}(\mathbf{X})]^{-1} \text{cov}(\mathbf{X}, Y) = \\
&= \mathcal{D}(Y) \left\{ 1 - \frac{1}{\sqrt{\mathcal{D}(Y)}} \left(\text{cov}(Y, X_1) : \dots : \text{cov}(Y, X_k) \right) \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix} \right\} \times \\
&\quad \left[\begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} [\text{cov}(\mathbf{X})]^{-1} \begin{pmatrix} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{pmatrix}^{-1} \right] \times
\end{aligned}$$

$$\left(\begin{array}{cccc} \frac{1}{\sqrt{\mathcal{D}(X_1)}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\mathcal{D}(X_2)}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{\mathcal{D}(X_n)}} \end{array} \right) \left(\begin{array}{c} \text{cov}(X_1, Y) \\ \text{cov}(X_2, Y) \\ \vdots \\ \text{cov}(X_k, Y) \end{array} \right) \frac{1}{\sqrt{\mathcal{D}(Y)}} \Bigg\} =$$

$$= \mathcal{D}(Y) \{1 - \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Y)\} = \mathcal{D}(Y)(1 - \varrho_{Y, \mathbf{X}}^2) = \mathcal{D}(Y - \hat{Y}). \quad (15)$$

(podľa Vety 4.7 (iii)) Úplne analogicky dostaneme

$$\mathcal{D}(Z - \hat{Z}) = \mathcal{D}(Z)(1 - \varrho_{Z, \mathbf{X}}^2). \quad (16)$$

Dosadením (14), (15) a (16) do (13) ľahko dostaneme

$$\varrho_{Y, Z, \mathbf{X}} = \frac{\varrho_{Y, Z} - \text{cor}(Y, \mathbf{X})[\text{cor}(\mathbf{X})]^{-1} \text{cor}(\mathbf{X}, Z)}{\sqrt{(1 - \varrho_{Y, \mathbf{X}}^2)(1 - \varrho_{Z, \mathbf{X}}^2)}}. \quad \text{Q.E.D.}$$

Poznámka 4.13: K výpčtu $\varrho_{Y, Z, \mathbf{X}}$ potrebujeme vedieť $\varrho_{Y, Z}$, ϱ_{Y, X_j} , ϱ_{X_i, X_j} , ϱ_{Z, X_j} , $i = 1, 2, \dots, k$, $j = 1, 2, \dots, k$.

Poznámka 4.14: Medzi $\varrho_{Y, Z, \mathbf{X}}$ a $\varrho_{Y, Z}$ nie je žiaden (všeobecný) vzťah.

Príklad 4.15: Vyjadrite $\varrho_{Z, (X, Y)}$ pomocou "obyčajných" korelačných koeficientov.

Riešenie: Podľa (12) platí

$$\varrho_{Y, Z, X} = \frac{\varrho_{Y, Z} - \varrho_{Y, X} \varrho_{Z, X}}{\sqrt{(1 - \varrho_{Y, X}^2)(1 - \varrho_{Z, X}^2)}}.$$

5 Lineárny regresný model

Príklad 5.1: Merajme neznámu dĺžku stola β n -krát nezávisle s meradlami "rovnakej kvality", teda rovnakej (neznámej) štandardnej neistoty σ (smerodajná odchýlka meradla). Merania modelujeme náhodnými veličinami Y_1, Y_2, \dots, Y_n , $\mathcal{E}(Y_i) = \beta$ - meracie prístroje sú bez systematickej chyby. To znamená, skutočne namerané hodnoty (čísla) y_1, y_2, \dots, y_n sú realizáciami náhodných veličín Y_1, Y_2, \dots, Y_n . "Celé" meranie modelujeme observačným (pozorovaným) náhodným vektorom (vektorom meraní) $\mathbf{Y} = (Y_1, \dots, Y_n)'$, ktorého stredná hodnota je $\mathcal{E}(\mathbf{Y}) = \mathbf{1}\beta$ a kovariančná matica je $\sigma^2 \mathbf{I}$, teda modelujeme ho "trojicou" $(\mathbf{Y}, \mathbf{1}\beta, \sigma^2 \mathbf{I})$.

Poznámka 5.2: V skutočnosti meráme tým istým meracím prístrojom, čo spôsobuje "závislosť" medzi meraniami. Štandardnú neistotu σ^2 meracieho prístroja môžeme poznať (napríklad z certifikátu prístroja), ale nemusíme poznať. Reálny prístroj nie je bez systematickej chyby. Náš jednoduchý model merania nie je úplne dokonalý (je to len určité priblíženie reality).

Príklad 5.3: (Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 104.) Merajme medenú trubku (nominálnej dĺžky $L_0 = 1000\text{mm}$ pri 20°C) postupne pri 30°C , 40°C , 50°C , 60°C , 70°C , 80°C . Vysledky meraní sú

Δt (zmena teploty)	10°C	20°C	30°C	40°C	50°C	60°C
predĺženie ΔL [mm]	0,18	0,35	0,48	0,65	0,84	0,97

Zákon rozťažnosti (z fyziky) tvrdí, že $\Delta L = L_0\alpha\Delta t$, kde α je koeficient tepelnej rozťažnosti (pre daný materiál), teda

$$\begin{aligned} Y_1 &= L_0\alpha 10 + \varepsilon_1 \\ Y_2 &= L_0\alpha 20 + \varepsilon_2 \\ &\vdots \\ Y_6 &= L_0\alpha 60 + \varepsilon_6, \end{aligned}$$

pričom predpokladáme, že $\varepsilon_1, \dots, \varepsilon_n$ sú nezávislé, $\mathcal{E}(\varepsilon_i) = 0$, $i = 1, 2, \dots, 6$ a $\mathcal{D}(\varepsilon_i) = \sigma^2$, $i = 1, 2, \dots, 6$. Vektor

observácií $\mathbf{Y}_{6,1}$ má strednú hodnotu $\begin{pmatrix} L_0 10 \\ L_0 20 \\ \vdots \\ L_0 60 \end{pmatrix} \alpha = \mathbf{X}\alpha$ (\mathbf{X} je známa matica a α neznámy parameter,

ktorého hodnota nás zaujíma). Kovariančná matica $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}_{6,6}$. Teda "celé" meranie modelujeme "trojicou" ($\mathbf{Y}, \mathbf{X}\alpha, \sigma^2 \mathbf{I}$).

Príklad 5.4: (Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 111.) U automobilu Trabant sa merala spotreba paliva (v litroch/100 km) v závislosti na jeho rýchlosti (pri stále zaradenom 4. rýchlostnom stupni, aby boli rovnaké podmienky jazdy). Rýchlosť považujeme za bezchybne určenú. Konkrétne namerané spotreby y_1, \dots, y_7 považujeme za realizácie náhodných veličín Y_1, \dots, Y_7 , pričom spotreba je (podľa vyjadrenia odborníkov) kvadratickou funkciou rýchlosti. Merania spotreby modelujeme ako

$$Y_i = a + bx_i + cx_i^2 + \varepsilon_i, \quad i = 1, 2, \dots, 7,$$

kde x_i je rýchlosť pri ktorej sa namerala spotreba y_i – realizácia náhodnej veličiny Y_i . Keby sme merali bezchybne, spotreba pri rýchlosti x_i by bola vždy $a + bx_i + cx_i^2$. Náhodné veličiny ε_i sú náhodné chyby. O nich predpokladáme, že sú nezávislé, majú nulovú strednú hodnotu a rovnakú disperziu σ^2 . Namerané hodnoty sú

rýchlosť (km/hod)	40	50	60	70	80	90	100
spotreba [l]	6,1	5,8	6,0	6,5	6,8	8,1	10,0

Observačný vektor (vektor meraní) $\mathbf{Y}_{7,1} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_7 \end{pmatrix}$ má strednú hodnotu $\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_7 & x_7^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{X}\beta$,

pričom \mathbf{X} je známa (pevná) matica a β je vektor neznámych parametrov, ktoré nás zaujímajú. Kovariančná matica observačného vektora je $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$, teda "celé" meranie zase modelujeme "trojicou" ($\mathbf{Y}, \mathbf{X}\beta, \sigma^2 \mathbf{I}$).

Príklad 5.5: (Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 113.) Majme dvojrozmernú náhodnú premennú $\mathbf{X} = (X, Y)'$, kde X –počet detí v rodine, Y –výdavky na stravu v rodine. Namerané hodnoty sú

počet detí	2	0	2	3	1	2
výdavky na stravu v tis.	4	3	4	6	4	5

Môžeme predpokladať (napríklad z "vynesenia" bodov (x_i, y_i)), že náhodná veličina $Y/X = x$ má strednú hodnotu $\mathcal{E}(Y/X = x) = a + bx$. Teda hodnotu 4 tis. môžeme považovať za realizáciu náhodnej premennej $Y/X = 2$, atď. V takomto prípade môžeme napísať model merania ako

$$Y/X = x_i = a + bx_i + \varepsilon_i, \quad i = 1, 2, \dots, 6,$$

kde nezávislé chyby $\varepsilon_1, \dots, \varepsilon_6$ majú nulové stredné hodnoty a (predpokladajme) rovnaké disperzie σ^2 . Ob-

servačný vektor (vektor meraní) $\mathbf{Y}_{7,1} = \begin{pmatrix} Y_1 = Y/X = x_1 \\ Y_2 = Y/X = x_2 \\ \vdots \\ Y_7 = Y/X = x_7 \end{pmatrix}$ má strednú hodnotu $\begin{pmatrix} 1 & 2 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} =$

$\mathbf{X}\boldsymbol{\beta}$, pričom zase \mathbf{X} je známa (pevná) matica a $\boldsymbol{\beta}$ je vektor neznámych parametrov, ktoré nás zaujímajú. Kovariančná matica observačného vektora je $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$, teda "celé" meranie opäť modelujeme "trojicou" $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$.

Poznámka 5.6: Ak môžeme považovať v predchádzajúcom príklade náhodný vektor $(X, Y)'$ za normálne rozdelený, teda $\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \mathbf{V} \right)$, \mathbf{V} je regulárna, tak priamo z teórie vychádza $Y/X = x \sim N(a + bx, \sigma^2)$ (pozri (6) a odvodte, že $a = \mu_Y - \mu_X \varrho \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}}$, $b = \varrho \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}}$, $\sigma^2 = \mathcal{D}(Y)(1 - \varrho^2)$).

Príklad 5.7: Majme body $A(\beta_1, \beta_2)$, $B(0, 0; 0, 0)$, $C(2365, 22; 0, 0)$, $D(3603, 67; 823, 35)$ v rovine. Meriame vzdialenosti AB , AC , AD a chceme zistiť (odhadnúť) súradnice β_1 a β_2 bodu A . Popíšte model merania.

Platí

$$\begin{aligned} AB &= \sqrt{(\beta_1 - 0)^2 + (\beta_2 - 0)^2} \\ AC &= \sqrt{(\beta_1 - 2365, 22)^2 + (\beta_2 - 0)^2} \\ AD &= \sqrt{(\beta_1 - 3603, 67)^2 + (\beta_2 - 823, 35)^2}. \end{aligned}$$

Odmeriame AB , teda realizujeme Y_1 a nameráme $y_1 = 1980, 102$; odmeriame AC , teda realizujeme Y_2 a nameráme $y_2 = 2040, 243$; a odmeriame AD , teda realizujeme Y_3 a nameráme $y_3 = 2598, 897$. Z rovníc

$$\begin{aligned} 1980, 102 &= \sqrt{\beta_1^2 + \beta_2^2} \\ 2040, 243 &= \sqrt{(\beta_1 - 2365, 22)^2 + \beta_2^2} \end{aligned}$$

vypočítame "približné" hodnoty $\beta_1^0 = 1131, 5$; $\beta_2^0 = 1625, 0$. (Na výpočet približných hodnôt môžeme použiť ľubovoľné dve rovnice.) Teraz (nelineárne) vzdialenosti linearizujeme okolo približných hodnôt pomocou Taylorovej vety, t.j.

$$\begin{aligned} AB &\approx \sqrt{(\beta_1^0)^2 + (\beta_2^0)^2} + \frac{\partial(AB)}{\partial\beta_1} \Big|_{\beta_1^0, \beta_2^0} \overbrace{(\beta_1 - \beta_1^0)}^{\Delta\beta_1} + \frac{\partial(AB)}{\partial\beta_2} \Big|_{\beta_1^0, \beta_2^0} \overbrace{(\beta_2 - \beta_2^0)}^{\Delta\beta_2} = \\ &= 1980, 131 + \frac{\beta_1^0 - 0}{\sqrt{(\beta_1^0)^2 + (\beta_2^0)^2}} \Delta\beta_1 + \frac{\beta_2^0 - 0}{\sqrt{(\beta_1^0)^2 + (\beta_2^0)^2}} \Delta\beta_2 = \\ &= 1980, 131 + \frac{1131, 5}{1980, 131} \Delta\beta_1 + \frac{1625}{1980, 131} \Delta\beta_2 = 1980, 131 + 0, 571 \Delta\beta_1 + 0, 821 \Delta\beta_2. \end{aligned}$$

Podobne dostaneme

$$\begin{aligned} AC &\approx 2040,267 - 0,605\Delta\beta_1 + 0,796\Delta\beta_2 \\ AD &\approx 2598,897 - 0,951\Delta\beta_1 + 0,308\Delta\beta_2. \end{aligned}$$

Tentokrát máme 3 merania, a síce Y_1 –meranie AB , Y_2 –meranie AC , Y_3 –meranie AD . Náhodný vektor

$\mathbf{W} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1980,131 \\ 2040,267 \\ 2598,897 \end{pmatrix}$ má strednú hodnotu $\begin{pmatrix} 0,571 & 0,821 \\ -0,605 & 0,796 \\ -0,951 & 0,308 \end{pmatrix} \begin{pmatrix} \Delta\beta_1 \\ \Delta\beta_2 \end{pmatrix} = \mathbf{X}\boldsymbol{\delta}$ a kovariančnú maticu $\sigma^2\mathbf{I}$. Modelom "celého" merania je opäť $(\mathbf{W}, \mathbf{X}\boldsymbol{\delta}, \sigma^2\mathbf{I})$.

V každom z predchádzajúcich príkladov sme pri matematicko-štatistickom modelovaní reálnej situácie dostali náhodný vektor (meraní, pozorovaní), ktorého stredná hodnota bola $\mathbf{X}\boldsymbol{\beta}$, pričom \mathbf{X} je známa matica a $\boldsymbol{\beta}$ vektor neznámych parametrov, ktoré nás zaujímali (chceli by sme ich "odhadnúť" (zistiť)). Kovariančná matica náhodného vektora meraní bola známa, alebo v tvare $\sigma^2 \times$ známa matica. Takýto model sa nazýva *lineárny regresný model (LRM)*, alebo aj *lineárny model, regresný model, model lineárnej regresie*. Dospeli sme k nasledujúcej definícii.

Definícia 5.8: Povieme, že náhodný vektor \mathbf{Y} sa riadi *lineárnym rešeným modelom* s maticou plánu \mathbf{X} (známa matica), vektorom chýb $\boldsymbol{\varepsilon}$, vektorom (neznámych) parametrov $\boldsymbol{\beta}$, ak

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

- pričom
1. $\mathcal{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$,
 2. $cov(\mathbf{Y}) = cov(\boldsymbol{\varepsilon}) = \mathbf{V}$, resp. $\sigma^2\mathbf{H}$

(\mathbf{V} resp. \mathbf{H} sú známe pozitívne semidefinitné matice, skalárny faktor σ^2 kovariančnej matice môžeme, ale nemusíme poznať). Neexistuje funkčný vzťah medzi $\boldsymbol{\beta}$ a σ^2 . Model značíme LRM $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \mathbf{V})$.

Ak $h(\mathbf{X}_{n,k}) = k < n$ a \mathbf{V} je pozitívne definitná (regulárna) matica, tak model sa nazýva *regulárny regresný model* alebo *model plnej hodnosti*. Inak to je *model neúplnej hodnosti*.

V tejto celej prednáške (celý semester) budeme uvažovať LRM $(\mathbf{Y}_{n,1}, \mathbf{X}_{n,k}\boldsymbol{\beta}_{k,1}, \sigma^2\mathbf{I}_{n,n})$ plnej hodnosti. Úlohou, cieľom je odhadnúť (určiť) neznáme parametre $\boldsymbol{\beta}$ (strednej hodnoty) modelu. Odhadujeme ich *metódou najmenších štvorcov (metoda najmenších čtverců-MNČ)*. Sú to také $\hat{\beta}_1(\mathbf{Y}), \dots, \hat{\beta}_k(\mathbf{Y})$, ktoré minimalizujú výraz

$$\mathcal{S}(\boldsymbol{\beta}) = \mathcal{S}(\beta_1, \dots, \beta_k) = \sum_{i=1}^n (Y_i - \sum_{j=1}^k x_{ij}\beta_j)^2 = \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2.$$

Teda MNČ odhad parametrov $\boldsymbol{\beta}$ v LRM $(\mathbf{Y}_{n,1}, \mathbf{X}_{n,k}\boldsymbol{\beta}_{k,1}, \sigma^2\mathbf{I}_{n,n})$ plnej hodnosti je

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta} \in \mathcal{R}^k} \mathcal{S}(\boldsymbol{\beta})$$

a preto

$$\mathcal{S}(\hat{\boldsymbol{\beta}}) = \min_{\boldsymbol{\beta} \in \mathcal{R}^k} \mathcal{S}(\boldsymbol{\beta}) = \min_{\boldsymbol{\beta} \in \mathcal{R}^k} \|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2.$$

Veta 5.9: Nech $(\mathbf{Y}_{n,1}, \mathbf{X}_{n,k}\boldsymbol{\beta}_{k,1}, \sigma^2\mathbf{I}_{n,n})$ je LRM plnej hodnosti. Odhad $\hat{\boldsymbol{\beta}}$ parametrov $\boldsymbol{\beta}$ metódou najmenších štvorcov je ekvivalentný riešeniu *normálnych rovníc* $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{Y}$, teda $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

Dôkaz: Podľa Vety 3.6 je $\mathcal{M}(\mathbf{X}') = \mathcal{M}(\mathbf{X}'\mathbf{X})$, teda $k = h(\mathbf{X}) = h(\mathbf{X}') = h(\mathbf{X}'\mathbf{X})$. Pretože matica $\mathbf{X}'\mathbf{X}$ je rozmeru $k \times k$ a aj hodnosť má rovnú k , je to regulárna matica a existuje $(\mathbf{X}'\mathbf{X})^{-1}$. Normálne rovnice majú jediné riešenie. Ak označíme riešenie normálnych rovníc $\hat{\beta}$, tak $\mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \mathbf{0}$ a pre ľubovoľné $\beta \in \mathcal{R}^k$

$$\begin{aligned} \mathcal{S}(\hat{\beta}) &= \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \\ &= (\mathbf{Y} - \mathbf{X}\hat{\beta} + \mathbf{X}\hat{\beta} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\hat{\beta} + \mathbf{X}\hat{\beta} - \mathbf{X}\beta) = \\ &= (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}) + \underbrace{(\mathbf{Y} - \mathbf{X}\hat{\beta})'\mathbf{X}}_{\mathbf{0}'}(\hat{\beta} - \beta) + (\hat{\beta} - \beta)'\underbrace{\mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta})}_{\mathbf{0}} + (\hat{\beta} - \beta)'\underbrace{\mathbf{X}'\mathbf{X}}_{p.d.matica}(\hat{\beta} - \beta) \geq \\ &\geq (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}), \end{aligned}$$

teda

$$\|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 = \min_{\beta \in \mathcal{R}^k} \|\mathbf{Y} - \mathbf{X}\beta\|^2.$$

Naopak, ak hľadáme minimum $\mathcal{S}(\hat{\beta})$, tak v tomto minime nutne $\frac{\partial \mathcal{S}(\beta)}{\partial \beta_m} = 0$, $m = 1, 2, \dots, k$. Pretože

$$\frac{\partial \mathcal{S}(\beta)}{\partial \beta_m} = \frac{\partial}{\partial \beta_m} \sum_{i=1}^n (Y_i - \sum_{j=1}^k x_{ij}\beta_j)^2 \Big|_{\beta=\hat{\beta}} = \sum_{i=1}^n 2(Y_i - \sum_{j=1}^k x_{ij}\hat{\beta}_j)(-x_{im}) \Big|_{\beta=\hat{\beta}} = 0, \quad m = 1, 2, \dots, k,$$

tak nutne

$$\sum_{i=1}^n (Y_i - \sum_{j=1}^k x_{ij}\hat{\beta}_j)(x_{im}) = 0, \quad m = 1, 2, \dots, k,$$

čo môžeme písať

$$(Y_1 - \sum_{j=1}^k x_{1j}\hat{\beta}_j)(x_{1m}) + (Y_2 - \sum_{j=1}^k x_{2j}\hat{\beta}_j)(x_{2m}) + \dots + (Y_n - \sum_{j=1}^k x_{nj}\hat{\beta}_j)(x_{nm}) = 0, \quad m = 1, 2, \dots, k,$$

alebo aj

$$(\mathbf{Y} - \mathbf{X}\hat{\beta})'\{\mathbf{X}\}_{\bullet m} = 0, \quad m = 1, 2, \dots, k.,$$

Dostávame postupne

$$\begin{aligned} (\mathbf{Y} - \mathbf{X}\hat{\beta})'\mathbf{X} &= (0, 0, \dots, 0)_{1,k} = \mathbf{0} \\ \mathbf{Y}'\mathbf{X} - \hat{\beta}'\mathbf{X}'\mathbf{X} &= \mathbf{0} \\ \mathbf{X}'\mathbf{Y} - \mathbf{X}'\mathbf{X}\hat{\beta} &= \mathbf{0} \\ \hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \quad \text{Q.E.D.} \end{aligned}$$

Poznámka 5.10: MNČ odhad $\hat{\beta}$ parametrov β je odhadom lineárnym, lebo jeho zložky sú lineárne funkcie (observačného) náhodného vektora \mathbf{Y} , ktorý máme k dispozícii pre odhadnutie parametrov β .

Poznámka 5.11: Pretože $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{Y} \iff \mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \mathbf{0} \iff \mathbf{Y} - \mathbf{X}\hat{\beta} \perp \mathcal{M}(\mathbf{X})$ t.j. $\mathbf{Y} - \mathbf{X}\hat{\beta} \perp \{\mathbf{X}\}_{\bullet i}$, $i = 1, 2, \dots, k$, dostávame, že

$$\|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2 = \min_{\beta \in \mathcal{R}^k} \|\mathbf{Y} - \mathbf{X}\beta\|^2 \iff \mathbf{Y} - \mathbf{X}\hat{\beta} \perp \mathcal{M}(\mathbf{X}).$$

Geometricky je v \mathcal{R}^n náhodný vektor $\mathbf{X}\hat{\beta}$ ležiaci v $\mathcal{M}(\mathbf{X})$ ortogonálnou projekciou observačného (náhodného) vektora \mathbf{Y} na $\mathcal{M}(\mathbf{X})$, t.j. má od \mathbf{Y} minimálnu vzdialenosť.

Odhad získaný metódou najmenších štvorcov je odhad získaný optimalizačnou (numerickou) metódou. Vôbec sme pri jeho hľadaní neuvažovali o jeho pravdepodobnostných (štatistických) vlastnostiach. Vieme, že ak máme náhodný vektor $\mathbf{Y} = (Y_1, \dots, Y_n)'$, ktorého rozdelenie pravdepodobnosti závisí od (neznámych) parametra $\beta_{k,1}$, tak odhad (bodový) tohto parametra je (ľubovoľné) merateľné zobrazenie $\mathbf{g} : \mathcal{R}^n \rightarrow \mathcal{R}^k$ (ktorého predpis nezávisí od β) také, že náhodný vektor (niekedy sa používa názov *štatistika*) $\hat{\beta} = \mathbf{g}(\mathbf{Y})$ v nejakom "rozumnom zmysle" aproximuje neznámy vektor parametrov $\beta_{k,1}$.

Majme ľubovoľný (pevný) vektor $\mathbf{c} \in \mathcal{R}^k$. $\gamma = \mathbf{c}'\beta$ je (lineárna) parametrická funkcia vektora β . Ak máme m (pevných) vektorov $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m \in \mathcal{R}^k$, tak

$$\mathbf{C}' = \begin{pmatrix} \mathbf{c}'_1 \\ \mathbf{c}'_2 \\ \vdots \\ \mathbf{c}'_m \end{pmatrix}_{m,k} \quad \text{a} \quad \gamma(\beta) = \mathbf{C}'\beta = \begin{pmatrix} \gamma_1(\beta) \\ \gamma_2(\beta) \\ \vdots \\ \gamma_m(\beta) \end{pmatrix} \quad \text{je vektorová parametrická funkcia.}$$

Definícia 5.11: Povieme, že vektorová štatistika (náhodný vektor) $\mathbf{T} = \begin{pmatrix} T_1 \\ \vdots \\ T_m \end{pmatrix} = \mathbf{T}(\mathbf{Y}) = \begin{pmatrix} T_1(\mathbf{Y}) \\ \vdots \\ T_m(\mathbf{Y}) \end{pmatrix}$ je

najlepší nevychýlený (nestranný) lineárny odhad - NNLO vektorovej parametrickej funkcie $\gamma = \begin{pmatrix} \mathbf{c}'_1\beta \\ \vdots \\ \mathbf{c}'_m\beta \end{pmatrix} =$
 $= \gamma(\beta) = \mathbf{C}'\beta$ ak

- (i) $\mathbf{T} = \mathbf{L}_{m,n}\mathbf{Y}$, kde \mathbf{L} je reálna matica (linearita odhadu),
- (ii) $\mathcal{E}_\beta(\mathbf{T}) = \gamma(\beta)$ pre každé $\beta \in \mathcal{R}^k$ (nevychýlenosť odhadu),
- (iii) ak \mathbf{T}^* je iný lineárny nevychýlený odhad parametrickej funkcie γ , tak $\text{cov}(\mathbf{T}^*) - \text{cov}(\mathbf{T})$ je pozitívne semidefinitná matica (niekedy sa píše $\text{cov}(\mathbf{T}^*) - \text{cov}(\mathbf{T}) \geq 0$).

Veta 5.12: Majme LRM $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ plnej hodnosti a $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ je MNČ odhad parametrov β . Potom platí

- (i) $\mathcal{E}_\beta(\hat{\beta}) = \beta$ pre každé $\beta \in \mathcal{R}^k$ (MNČ odhad je nevychýlený)
- (ii) $\text{cov}(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

Dôkaz:

- (i) $\mathcal{E}_\beta(\hat{\beta}) = \mathcal{E}_\beta((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta = \beta$ pre každé $\beta \in \mathcal{R}^k$,
- (ii) $\text{cov}(\hat{\beta}) = \text{cov}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Q.E.D.

Veta 5.13: V LRM $(\mathbf{Y}, \mathbf{X}_{n,k}\beta, \sigma^2\mathbf{I})$ plnej hodnosti je $\hat{\gamma} = \mathbf{C}'_{m,k}\hat{\beta}$ NNLO vektorovej parametrickej funkcie $\gamma = \mathbf{C}'\beta$, pričom $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ je MNČ odhad parametrov β .

Dôkaz:

$$\mathbf{C}'\hat{\beta} = \mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

je lineárny odhad, pričom

$$\mathcal{E}(\mathbf{C}'\hat{\beta}) = \mathcal{E}(\mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) = \mathbf{C}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta = \mathbf{C}'\beta \quad \forall \beta \in \mathcal{R}^k,$$

teda je nevychýlený. Nech $\gamma^* = \mathbf{L}^* \mathbf{Y}$ je ľubovoľný nevychýlený odhad vektorovej parametrickej funkcie γ , tak

$$\mathcal{E}(\mathbf{L}^* \mathbf{Y}) = \mathbf{L}^* \mathbf{X} \boldsymbol{\beta} = \mathbf{C}' \boldsymbol{\beta} \quad \forall \boldsymbol{\beta} \in \mathcal{R}^k \iff \mathbf{L}^* \mathbf{X} = \mathbf{C}'.$$

$$\begin{aligned} \text{cov}(\gamma^*) - \text{cov}(\hat{\gamma}) &= \mathbf{L}^* \sigma^2 \mathbf{I} (\mathbf{L}^*)' - \mathbf{C}' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \sigma^2 \mathbf{I} \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{C} = \sigma^2 \mathbf{L}^* (\mathbf{L}^*)' - \sigma^2 \underbrace{\mathbf{C}'}_{\mathbf{L}^* \mathbf{X}} (\mathbf{X}' \mathbf{X})^{-1} \underbrace{\mathbf{C}}_{\mathbf{X}' (\mathbf{L}^*)'} = \\ &= \sigma^2 \{ \mathbf{L}^* (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') (\mathbf{L}^*)' \} \geq 0, \end{aligned}$$

lebo $\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}$ je symetrická a idempotentná, teda pozitívne semidefinitná matica. Teda $\hat{\gamma} = \mathbf{C}' \hat{\boldsymbol{\beta}}$ je NNLO vektorovej parametrickej funkcie $\gamma = \mathbf{C}' \boldsymbol{\beta}$. Q.E.D.

Poznámka 5.14: Ak v predchádzajúcej vete $\mathbf{C}' = \mathbf{e}'_i$, tak $\hat{\beta}_i = \mathbf{e}'_i (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}$ je NNLO parametra β_i .

Definícia 5.15: *Reziduálny súčet štvorcov* je náhodná veličina

$$S_e(\hat{\boldsymbol{\beta}}) = (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}}).$$

Je to miera kvality odhadu v danom LRM.

Veta 5.16: Platí

$$S_e = \mathbf{Y}' (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \mathbf{Y} = \mathbf{Y}' \mathbf{Y} - \mathbf{Y}' \mathbf{X} \hat{\boldsymbol{\beta}}.$$

Dôkaz:

$$\begin{aligned} S_e &= (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}}) = (\mathbf{Y} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y})' (\mathbf{Y} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}) = \\ &= \mathbf{Y}' (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}')' (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \mathbf{Y} = \mathbf{Y}' (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \mathbf{Y} = \mathbf{Y}' \mathbf{Y} - \mathbf{Y}' \mathbf{X} \hat{\boldsymbol{\beta}}. \quad \text{Q.E.D.} \end{aligned}$$

Veta 5.17: Štatistika

$$s^2 = \frac{1}{n-k} S_e = \frac{1}{n-k} \mathbf{Y}' (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \mathbf{Y}$$

je nevychýleným odhadom σ^2 .

Dôkaz: Podľa Vety 1.5 dostávame

$$\begin{aligned} \mathcal{E}(s^2) &= \mathcal{E} \left(\mathbf{Y}' \left[\frac{1}{n-k} (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \right] \mathbf{Y} \right) = \\ &= \boldsymbol{\beta}' \mathbf{X}' \left[\frac{1}{n-k} (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \right] \mathbf{X} \boldsymbol{\beta} + \text{Tr} \left[\frac{1}{n-k} (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') \right] \sigma^2 \mathbf{I} = \\ &= \frac{\sigma^2}{n-k} \text{Tr} (\mathbf{I} - \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}') = \frac{\sigma^2}{n-k} (n - \text{Tr} \mathbf{X}' \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1}) = \sigma^2. \quad \text{Q.E.D.} \end{aligned}$$

Poznámka 5.18: Náhodný vektor $\hat{\mathbf{Y}} = \mathbf{X} \hat{\boldsymbol{\beta}}$ je "aproximáciou" bezchybných meraní, teda NNLO vektora stredných hodnôt $\mathbf{X} \boldsymbol{\beta}$, čiže $\hat{\mathbf{Y}} = \widehat{\mathbf{X}} \hat{\boldsymbol{\beta}}$. Niekedy sa mu hovorí vektor vyrovnaných hodnôt.

Definícia 5.19: Vektor $\mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{r}$ voláme *vektor rezíduí* alebo *reziduálny vektor*. Jeho i -tu zložku (súradnicu) voláme *i -te rezíduum*.

Poznámka 5.20: Reziduá sú jedným z prostriedkov diagnostikovania modelu, teda posúdenia vhodnosti modelovania nameraných údajov daným modelom. Keď si nakreslíme graf rezíduí, t.j. body (i, r_i) (tu r_i je hodnota (realizácia) i -teho rezídua), tak táto postupnosť nesmie vykazovať pri správnej voľbe modelu žiadnu systematickosť.

Doteraz sme nič nepredpokladali o rozdelení pravdepodobnosti observačného (náhodného) vektora \mathbf{Y} . Pri ďalších štatistických inferenciách (odvodzovaniach) budeme predpokladať, že $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, čo je to isté ako predpoklad $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$.

Veta 5.21: Majme LRM $(\mathbf{Y}, \mathbf{X}_{n,k}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ plnej hodnosti a $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$, $\hat{\boldsymbol{\beta}}$ je MNČ odhad $\boldsymbol{\beta}$. Potom platí

- (i) $\hat{\boldsymbol{\beta}} \sim N_k(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$,
- (ii) $\frac{n-k}{\sigma^2}s^2 = \frac{S_e}{\sigma^2} \sim \chi_{n-k}^2$,
- (iii) $\hat{\boldsymbol{\beta}}$ a s^2 sú nezávislé.

Dôkaz: (i) Z predpokladov platí $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$. Pretože $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, podľa Vety 2.6 je $\hat{\boldsymbol{\beta}} \sim N_k(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$.

(ii) Náhodná veličina

$\frac{n-k}{\sigma^2}s^2 = \mathbf{Y}' \left[\frac{1}{\sigma^2}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \right] \mathbf{Y} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \left[\frac{1}{\sigma^2}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \right] (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) =$
 $= (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{A} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ je kvadratickou formou náhodného vektora $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$ s maticou kvadratickej formy \mathbf{A} . Pretože platí, že $\mathbf{A} = \frac{1}{\sigma^2}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$ je symetrická, pozitívne semidefinitná, $\mathbf{A}\sigma^2\mathbf{I} \neq \mathbf{0}$, $\mathbf{A}\sigma^2\mathbf{I}\mathbf{A}\sigma^2\mathbf{I} = \mathbf{A}\sigma^2\mathbf{I}$, podľa Vety 3.14 $\frac{n-k}{\sigma^2}s^2 \sim \chi_{Tr\mathbf{A}\sigma^2\mathbf{I}}^2$. Ale $Tr\mathbf{A}\sigma^2\mathbf{I} = Tr(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') =$
 $n - k$, teda $\frac{n-k}{\sigma^2}s^2 \sim \chi_{n-k}^2$.

(iii) Platí, že $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{B}\mathbf{Y}$, $s^2 = \mathbf{Y}' \left[\frac{1}{n-k}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \right] \mathbf{Y}$, pričom $\frac{1}{n-k}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$ je symetrická a pozitívne semidefinitná a $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$. Pretože $\mathbf{B}\sigma^2\mathbf{I} \left[\frac{1}{n-k}(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \right] = \frac{\sigma^2}{n-k}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') = \mathbf{0}$, podľa Vety 3.15 sú $\hat{\boldsymbol{\beta}}$ a s^2 nezávislé. Q.E.D.

Nech $\mathbf{c} \in \mathcal{R}^k$ je daný vektor, teda majme parametrickú funkciu $\gamma = \mathbf{c}'\boldsymbol{\beta}$.

Veta 5.22: Majme LRM $(\mathbf{Y}, \mathbf{X}_{n,k}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ plnej hodnosti, $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$ a $\gamma = \mathbf{c}'\boldsymbol{\beta}$ (funkciu parametrov). Nech $\hat{\boldsymbol{\beta}}$ je MNČ odhad vektora $\boldsymbol{\beta}$. Potom

$$T = \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{s\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim t_{n-k}, \quad \text{ak } \mathbf{c} \neq \mathbf{0}.$$

Dôkaz: Pretože $\mathbf{c}'\hat{\boldsymbol{\beta}} \sim N(\mathbf{c}'\boldsymbol{\beta}, \sigma^2\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c})$, je $\frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{\sigma\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim N(0, 1)$. Podľa Vety 5.21 (ii) $\frac{n-k}{\sigma^2}s^2 = \frac{S_e}{\sigma^2} \sim \chi_{n-k}^2$ a podľa Vety 5.21 (iii) sú $\hat{\boldsymbol{\beta}}$ a s^2 sú nezávislé, teda aj $\mathbf{c}'\hat{\boldsymbol{\beta}}$ (ako funkcia $\hat{\boldsymbol{\beta}}$) a s^2 sú nezávislé. Potom ale (priamo z definície Studentovho t -rozdelenia)

$$T = \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{\sigma\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} = \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{s\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim t_{n-k}. \quad \text{Q.E.D.}$$

Z Vety 5.22 vyplýva, že pre dané $\alpha \in (0, 1)$

$$P \left\{ t_{n-k} \left(\frac{\alpha}{2} \right) \leq \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{s\sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \leq t_{n-k} \left(1 - \frac{\alpha}{2} \right) \right\} = 1 - \alpha, \quad (17)$$

kde $t_g(\beta)$ je β -kvantil Studentovho t rozdelenia s g stupňami voľnosti. Teda ak náhodná veličina $T \sim t_g$ (T má Studentovo t rozdelenie pravdepodobnosti s g stupňami voľnosti), tak $t_g(\beta)$ je také číslo, pre ktoré platí $P\{T < t_g(\beta)\} = \beta$. Upozorňujeme len, že v niektorej literatúre (napr. v knižke Anděl, J., *Matematická statistika*, SNTL, Praha, 1985) sa pracuje (na rozdiel od tohto textu) s kritickými hodnotami a nie s kvantilmi.

Zo vzťahu (17) úpravami dostaneme

$$P \left\{ t_{n-k} \left(\frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \leq \mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta} \leq t_{n-k} \left(1 - \frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \right\} = 1 - \alpha,$$

$$P \left\{ \mathbf{c}'\hat{\boldsymbol{\beta}} - t_{n-k} \left(1 - \frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \leq \mathbf{c}'\boldsymbol{\beta} \leq \mathbf{c}'\hat{\boldsymbol{\beta}} + t_{n-k} \left(1 - \frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \right\} = 1 - \alpha$$

(lebo pre kvantily Studentovho rozdelenia platí $t_g(\beta) = -t_g(1 - \beta)$) a teda $100(1 - \alpha)\%$ -ný interval spoľahlivosti (konfidenčný interval) pre $\gamma = \mathbf{c}\boldsymbol{\beta}$ je

$$\left(\mathbf{c}'\hat{\boldsymbol{\beta}} - t_{n-k} \left(1 - \frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}, \mathbf{c}'\hat{\boldsymbol{\beta}} + t_{n-k} \left(1 - \frac{\alpha}{2} \right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}} \right). \quad (18)$$

Vetu 5.22 použijeme pri testovaní hypotézy o hodnote lineárnej funkcie $\gamma = \mathbf{c}\boldsymbol{\beta}$.

Majme LRM $(\mathbf{Y}, \mathbf{X}_{n,k}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ plnej hodnosti, pričom $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$, (σ^2 nepoznáme), $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ (MNČ odhad). Ďalej majme danú $\gamma = \mathbf{c}'\boldsymbol{\beta}$ (lineárna funkcia parametrov $\boldsymbol{\beta}$).

Test hypotézy

$$H_0 : \mathbf{c}'\boldsymbol{\beta} = \gamma_0 \text{ (dané číslo)} \quad \not\approx \quad H_1 : \mathbf{c}'\boldsymbol{\beta} \neq \gamma_0 \quad (19)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \gamma_0}{s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}} \sim t_{n-k} \quad (\text{za platnosti } H_0).$$

$$\text{Ak } |T_0| \geq t_{n-k} \left(1 - \frac{\alpha}{2} \right) \implies H_0 \text{ zamietame,}$$

$$\text{ak } |T_0| < t_{n-k} \left(1 - \frac{\alpha}{2} \right) \implies H_0 \text{ nezamietame.}$$

Podľa Vety 5.22 má tento test hladinu významnosti α .

Dôležité špeciálne prípady testu (19) sú testy o hodnote jednotlivých zložiek vektora parametrov. Pretože $\beta_j = \mathbf{e}'_j\boldsymbol{\beta}$, $j = 1, 2, \dots, k$, ak v teste (19) za \mathbf{c} vezmeme \mathbf{e}_j a za γ_0 vezmeme β_j^0 (dané číslo), dostávame nasledujúci test.

Test hypotézy

$$H_0 : \beta_j = \beta_j^0 \text{ (dané číslo)} \quad \not\approx \quad H_1 : \beta_j \neq \beta_j^0 \quad (20)$$

realizujeme pomocou testovacej štatistiky

$$T_{0j} = \frac{\hat{\beta}_j - \beta_j^0}{s \sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{jj}}} \sim t_{n-k} \quad (\text{za platnosti } H_0).$$

$$\text{Ak } |T_{0j}| \geq t_{n-k} \left(1 - \frac{\alpha}{2} \right) \implies H_0 \text{ zamietame na hladine významnosti } \alpha,$$

$$\text{ak } |T_{0j}| < t_{n-k} \left(1 - \frac{\alpha}{2} \right) \implies H_0 \text{ nezamietame na hladine významnosti } \alpha.$$

Z (18) (pri voľbe $\mathbf{c} = \mathbf{e}_j$) okamžite dostávame, že $100(1 - \alpha)\%$ -ný interval spoľahlivosti (konfidenčný interval) pre β_j je

$$\left(\hat{\beta}_j - t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{jj}}, \hat{\beta}_j + t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{jj}} \right). \quad (21)$$

Ak $\beta_j = 0$, tak môžeme "vynechať" j -ty stĺpec matice plánu, teda dostaneme jednoduchší model (s menej parametrami). Vektor \mathbf{Y} nezáleží od parametra β_j .

Teraz si odvodíme $(1 - \alpha)$ -predikčný interval pre náhodnú veličinu (meranie) $Y_{\mathbf{c}} = \mathbf{c}'\boldsymbol{\beta} + \varepsilon$, ktorá má (skutočnú) strednú hodnotu (bezchybnú hodnotu) $\mathbf{c}'\boldsymbol{\beta}$, disperziu σ^2 a je nezávislá od Y_1, \dots, Y_n .

$(1 - \alpha)$ -predikčný interval pre $Y_{\mathbf{c}} = \mathbf{c}'\boldsymbol{\beta} + \varepsilon$ je náhodný interval $(D_{\mathbf{c}}, H_{\mathbf{c}})$, pre ktorý platí

$$P\{Y_{\mathbf{c}} \in (D_{\mathbf{c}}, H_{\mathbf{c}})\} = 1 - \alpha.$$

Náhodná veličina

$$\tilde{Y}_{\mathbf{c}} = \mathbf{c}'\hat{\boldsymbol{\beta}} - \varepsilon,$$

pričom MNČ odhad $\hat{\boldsymbol{\beta}}$ a náhodná chyba $\varepsilon \sim N(0, \sigma^2)$ sú nezávislé, má strednú hodnotu $\mathcal{E}(\tilde{Y}_{\mathbf{c}}) = \mathbf{c}'\boldsymbol{\beta}$ a disperziu $\mathcal{D}(\tilde{Y}_{\mathbf{c}}) = \sigma^2 \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + \sigma^2$, čiže

$$\tilde{Y}_{\mathbf{c}} \sim N(\mathbf{c}'\boldsymbol{\beta}, \sigma^2(\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1)).$$

Podľa Vety 5.21 (iii) sú $s^2 = \frac{1}{n-k} \mathbf{Y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}$ a MNČ odhad $\hat{\boldsymbol{\beta}}$ nezávislé a ε je nezávislá s \mathbf{Y} (teda aj s funkciami \mathbf{Y} , čo sú s^2 aj $\hat{\boldsymbol{\beta}}$). Náhodná veličina $\frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \mathbf{c}'\boldsymbol{\beta}}{\sigma \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}}$ má $N(0, 1)$ rozdelenie, a je nezávislá s $\frac{n-k}{\sigma^2} s^2$, ktorá má χ_{n-k}^2 rozdelenie. Potom ale

$$\frac{\frac{\tilde{Y}_{\mathbf{c}} - \mathbf{c}'\boldsymbol{\beta}}{\sigma \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1}}}{\sqrt{\frac{\frac{n-k}{\sigma^2} s^2}{n-k}}} = \frac{\tilde{Y}_{\mathbf{c}} - \mathbf{c}'\boldsymbol{\beta}}{s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1}} = \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \varepsilon - \mathbf{c}'\boldsymbol{\beta}}{s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1}} \sim t_{n-k},$$

čiže

$$P \left\{ -t_{n-k} \left(1 - \frac{\alpha}{2}\right) \leq \frac{\mathbf{c}'\hat{\boldsymbol{\beta}} - \varepsilon - \mathbf{c}'\boldsymbol{\beta}}{s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1}} \leq t_{n-k} \left(1 - \frac{\alpha}{2}\right) \right\} = 1 - \alpha,$$

$$P \left\{ -t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1} - \mathbf{c}'\hat{\boldsymbol{\beta}} \leq -\mathbf{c}'\boldsymbol{\beta} - \varepsilon \leq t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1} - \mathbf{c}'\hat{\boldsymbol{\beta}} \right\} = 1 - \alpha,$$

odkiaľ dostávame

$$P \left\{ \mathbf{c}'\hat{\boldsymbol{\beta}} - t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1} \leq \overbrace{\mathbf{c}'\boldsymbol{\beta} + \varepsilon}^{Y_{\mathbf{c}}} \leq \mathbf{c}'\hat{\boldsymbol{\beta}} + t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1} \right\} = 1 - \alpha.$$

Preto $(1 - \alpha)$ -predikčný interval pre $Y_{\mathbf{c}}$, $\mathbf{c} \in \mathcal{R}^k$ je

$$\left(\mathbf{c}'\hat{\boldsymbol{\beta}} - t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1}, \mathbf{c}'\hat{\boldsymbol{\beta}} + t_{n-k} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} + 1} \right). \quad (22)$$

Poznámka 5.23: Ak za \mathbf{c} zvolíme $(x_{i1}, x_{i2}, \dots, x_{ik})' = \{\mathbf{X}'\}_{i\bullet}$, tak z (22) dostaneme $(1 - \alpha)$ -predikčný interval pre nové (nezávisle zopakované) meranie Y_i .

Ak $m \in \{1, 2, \dots, k - 1\}$, rozdelíme β na 2 časti, a síce na $\beta_1 = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$, $\beta_2 = \begin{pmatrix} \beta_{m+1} \\ \beta_{m+2} \\ \vdots \\ \beta_k \end{pmatrix}$, pričom $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$

a analogicky rozdelíme $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ a $\mathbf{S} = (\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix}$.

Veta 5.24: Majme LRM $(\mathbf{Y}, \mathbf{X}_{n,k}\beta, \sigma^2\mathbf{I})$ plnej hodnosti, $\varepsilon \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$. Potom

$$F = \frac{(\hat{\beta}_2 - \beta_2)' \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2)}{(k - m)s^2} \sim F_{k-m, n-k}.$$

Dôkaz: Podľa Vety 5.21 sú $\hat{\beta}$ a $\frac{n-k}{\sigma^2}s^2$ nezávislé, teda aj $\hat{\beta}_2$ a $\frac{n-k}{\sigma^2}s^2$ sú nezávislé, pričom $\frac{S_e}{\sigma^2} \sim \chi_{n-k}^2$. Pomocou tvrdenia Vety 2.6 a analogickým postupom ako v dôkaze Vety 2.7 ľahko ukážeme, že $\hat{\beta}_2 \sim N_{k-m}(\beta_2, \sigma^2\mathbf{S}_{22})$. Z toho vyplýva podľa Vety 3.13, že $(\hat{\beta}_2 - \beta_2)' \frac{1}{\sigma^2} \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2) \sim \chi_{k-m}^2$. (Upozorňujeme len, že podľa Dôkazu Vety 5.9 existuje $(\mathbf{X}'\mathbf{X})^{-1}$ (samozrejme regulárna) a podľa Lemy 2.14 (i) je \mathbf{S}_{22} tiež regulárna.) Preto

$$F = \frac{\frac{(\hat{\beta}_2 - \beta_2)' \frac{1}{\sigma^2} \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2)}{k-m}}{\frac{(n-k)s^2}{\sigma^2}} = \frac{(\hat{\beta}_2 - \beta_2)' \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2)}{(k-m)s^2} \sim F_{k-m, n-k}. \quad \text{Q.E.D.}$$

Poznámka 5.25: Veta 5.24 platí aj pre $m = 0$, teda pre $\beta_2 = \beta$.

Test hypotézy

$$H_0 : \beta_2 = \beta_2^0 \text{ (daný vektor)} \quad \not\asymp \quad H_1 : \beta_2 \neq \beta_2^0 \quad (23)$$

realizujeme pomocou testovacej štatistiky

$$F_0 = \frac{(\hat{\beta}_2 - \beta_2^0)' \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2^0)}{(k-m)s^2} \sim F_{k-m, n-k} \quad (\text{za platnosti } H_0).$$

Ak $F_0 \geq s^2(k-m)F_{k-m, n-k}(1-\alpha) \implies H_0$ zamietame,

ak $F_0 < s^2(k-m)F_{k-m, n-k}(1-\alpha) \implies H_0$ nezamietame,

pričom $F_{k-m, n-k}(1-\alpha)$ je $(1-\alpha)$ kvantil Fisherovho-Snedecorovho F rozdelenia s $k-m$ a $n-k$ stupňami voľnosti. Podľa Vety 5.24 má tento test hladinu významnosti α .

Poznámka 5.26: Z Vety 5.24 dostávame, že

$$P \left\{ \frac{(\hat{\beta}_2 - \beta_2)' \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2)}{(k-m)s^2} \leq F_{k-m, n-k}(1-\alpha) \right\} = 1-\alpha,$$

teda

$$P \left\{ (\hat{\beta}_2 - \beta_2)' \mathbf{S}_{22}^{-1} (\hat{\beta}_2 - \beta_2) \leq s^2(k-m)F_{k-m, n-k}(1-\alpha) \right\} = 1-\alpha. \quad (24)$$

Vzťahom (24) je určená $(1 - \alpha)100\%$ -ná konfidenčná oblasť (oblasť spoľahlivosti, konfidenčný elipsoid), ktorý s pravdepodobnosťou $1 - \alpha$ pokrýva (neznámy) vektor β_2 .

Poznámka 5.27: Nie je podstatné delenie β na $\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, môžeme vziať $\beta_1 = \begin{pmatrix} \beta_{i_1} \\ \beta_{i_2} \\ \vdots \\ \beta_{i_m} \end{pmatrix}, \beta_2 = \begin{pmatrix} \beta_{j_1} \\ \beta_{j_2} \\ \vdots \\ \beta_{j_{k-m}} \end{pmatrix}$,

aby $\{1, 2, \dots, k\} = \{i_1, i_2, \dots, i_m\} \cup \{j_1, j_2, \dots, j_{k-m}\}$.

Špeciálne regresné modely

a) Jednovýberový t -test.

Majme náhodný výber Y_1, \dots, Y_n z $N(\mu, \sigma^2)$ rozdelenia, μ ani σ^2 nepoznáme. LRM je

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \mathbf{1}\mu + \boldsymbol{\varepsilon}, \quad \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}.$$

MNČ odhad $\hat{\mu} = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{Y} = \bar{Y}$, $s^2 = \frac{1}{n-1}\mathbf{Y}'(\mathbf{I} - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}')\mathbf{Y} = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y})^2$ (dokážte ako cvičenie). Keď aplikujeme Vetu 5.22 a zvolíme $\mathbf{c} = 1, \gamma_0 = \mu_0$ (dané číslo), dostávame:

Test hypotézy

$$H_0 : \mu = \mu_0 \text{ (dané číslo)} \quad \not\asymp \quad H_1 : \mu \neq \mu_0 \quad (25)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{1\hat{\mu} - 1\mu_0}{s\sqrt{\mathbf{1}'\frac{1}{n}\mathbf{1}}} = \frac{\bar{Y} - \mu_0}{s}\sqrt{n} \sim t_{n-1} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n-1}\left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n-1}\left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Podľa Vety 5.22 má tento test hladinu významnosti α .

b) Dvojjvýberový t -test.

Majme

náhodný výber $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1n_1})'$ z $N(\mu_1, \sigma^2)$ rozdelenia a nezávislý s ním

náhodný výber $\mathbf{Y}_2 = (Y_{21}, \dots, Y_{2n_2})'$ z $N(\mu_2, \sigma^2)$ rozdelenia. LRM je

$$\mathbf{Y}_{n_1+n_2,1} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{1}_{n_1,1} & \mathbf{0}_{n_1,1} \\ \mathbf{0}_{n_2,1} & \mathbf{1}_{n_2,1} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}, \quad \text{cov}(\mathbf{Y}) = \text{cov}\left(\begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}\right) = \sigma^2\mathbf{I}_{n_1+n_2, n_1+n_2}.$$

Matica plánu $\mathbf{X} = \begin{pmatrix} \mathbf{1}_{n_1,1} & \mathbf{0}_{n_1,1} \\ \mathbf{0}_{n_2,1} & \mathbf{1}_{n_2,1} \end{pmatrix}$. Presvedčte sa, že platí $\mathbf{X}'\mathbf{X} = \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix}$, $(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{pmatrix}$,
 $\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum_{i=1}^{n_1} Y_{1i} \\ \sum_{i=1}^{n_2} Y_{2i} \end{pmatrix}$, $\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} = \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{pmatrix}$, $s^2 = \frac{1}{n_1 + n_2 - 2} \left(\sum_{i=1}^{n_1} Y_{1i}^2 - n_1 \bar{Y}_1^2 + \sum_{j=1}^{n_2} Y_{2j}^2 - n_2 \bar{Y}_2^2 \right)$. Keď aplikujeme Vetu 5.22 a zvolíme $\mathbf{c} = (1, -1)'$ a $\gamma_0 = 0$, dostávame:

Test hypotézy

$$H_0 : \mu_1 = \mu_2 \quad \not\approx \quad H_1 : \mu_1 \neq \mu_2 \quad (26)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{\bar{Y}_1 - \bar{Y}_2}{s \sqrt{(1, -1) \begin{pmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}}} = \frac{\bar{Y}_1 - \bar{Y}_2}{s \sqrt{\frac{n_1 + n_2}{n_1 n_2}}} \sim t_{n_1 + n_2 - 2} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n_1 + n_2 - 2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n_1 + n_2 - 2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Podľa Vety 5.22 má tento test hladinu významnosti α .

c) Zovšeobecnenie na k výberov

Majme

náhodný výber $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1n_1})'$ z $N(\mu_1, \sigma^2)$ rozdelenia,

náhodný výber $\mathbf{Y}_2 = (Y_{21}, \dots, Y_{2n_2})'$ z $N(\mu_2, \sigma^2)$ rozdelenia,

⋮

náhodný výber $\mathbf{Y}_k = (Y_{k1}, \dots, Y_{kn_k})'$ z $N(\mu_k, \sigma^2)$ rozdelenia.

Všetky výbery sú nezávislé. Testujeme $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$. Je to úloha analýzy rozptylu, budeme sa ňou zaoberať v prednáške LSM 2.

d) Regresná priamka.

Majme nezávislé náhodné veličiny (merania) Y_1, \dots, Y_n , pre ktoré platí

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad n \geq 3.$$

(Bezchybné merania ležia na priamke $y = \beta_0 + \beta_1 x$, hodnoty x_i , $i = 1, 2, \dots, n$ poznáme bezchybne (úplne presne)). LRM je

$$\mathbf{Y}_{n,1} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \boldsymbol{\varepsilon}, \quad \text{cov}(\mathbf{Y}) = \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n,n}.$$

Opäť sa presvedčte, že platí

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}, \quad (\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix},$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{1}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n x_i Y_i \end{pmatrix},$$

preto

$$\hat{\beta}_0 = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i Y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (27)$$

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n x_i Y_i - \sum_{i=1}^n x_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (28)$$

Pretože platí (tiež sa presvedčte výpočtom)

$$\hat{\beta}_0 + \hat{\beta}_1 \bar{x} = \bar{Y}, \quad (29)$$

obyčajne sa najprv spočíta $\hat{\beta}_1$ a potom $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} = \frac{1}{n} \{ \sum_{i=1}^n Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i \}$. Ešte potrebujeme

$$s^2 = \frac{1}{n-2} (\mathbf{Y}'\mathbf{Y} - \hat{\beta}'\mathbf{X}'\mathbf{Y}) = \frac{1}{n-2} \left(\sum_{i=1}^n Y_i^2 - \hat{\beta}_0 \sum_{i=1}^n Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i Y_i \right). \quad (30)$$

Keď aplikujeme Vetu 5.22 a zvolíme $\mathbf{c} = (0, 1)'$, $\gamma_0 = \mu_0$ (dané číslo), dostávame:

Test hypotézy

$$H_0 : \beta_1 = 0 \text{ (dané číslo)} \quad \approx \quad H_1 : \beta_1 \neq 0 \quad (31)$$

realizujeme pomocou testovacej štatistiky

$$T = \frac{\hat{\beta}_1}{s} \sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \sim t_{n-2} \quad (\text{za platnosti } H_0).$$

Ak $|T| \geq t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T| < t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Podľa Vety 5.22 má tento test hladinu významnosti α .

Overte výpočtom, že platí

$$(0, 1)(\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{n}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{1}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}. \quad (32)$$

Z (18) dostaneme $100(1 - \alpha)\%$ -ný interval spoľahlivosti pre β_1

$$\left(\hat{\beta}_1 - t_{n-2} \left(1 - \frac{\alpha}{2}\right) \frac{s}{\sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}, \hat{\beta}_1 + t_{n-2} \left(1 - \frac{\alpha}{2}\right) \frac{s}{\sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} \right). \quad (33)$$

Analogicky odvoďte test hypotézy

$$H_0 : \beta_0 = 0 \text{ (dané číslo)} \quad \approx \quad H_1 : \beta_0 \neq 0. \quad (34)$$

Keď aplikujeme Vetu 5.22 a zvolíme $\mathbf{c} = (1, x)'$, γ_0 (x dané reálne číslo, γ_0 je hypotetická hodnota $\beta_0 + \beta_1 x$), dostávame:

Test hypotézy

$$H_0 : \beta_0 + \beta_1 x = \gamma_0 \text{ (dané číslo)} \quad \asymp \quad H_1 : \beta_0 + \beta_1 x \neq \gamma_0 \quad (35)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{\hat{\beta}_0 + \hat{\beta}_1 x - \gamma_0}{s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \sim t_{n-2} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Podľa Vety 5.22 má tento test hladinu významnosti α .

Overte výpočtom, že platí

$$(1, x)(\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} 1 \\ x \end{pmatrix} = \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}. \quad (36)$$

Z (18) dostaneme $100(1 - \alpha)\%$ -ný interval spoľahlivosti pre $\beta_0 + \beta_1 x$

$$\left(\hat{\beta}_0 + \hat{\beta}_1 x - t_{n-2} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}, \hat{\beta}_0 + \hat{\beta}_1 x + t_{n-2} \left(1 - \frac{\alpha}{2}\right) s \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} \right). \quad (37)$$

Poznámka 5.28: Ak vynesieme (37) pre každé $x \in \mathcal{R}^1$, dostaneme $100(1 - \alpha)\%$ -ný pás spoľahlivosti okolo regresnej priamky, ktorý pre každé x (zvlášť) pokrýva skutočnú (bezchybnú) hodnotu $\beta_0 + \beta_1 x$ s pravdepodobnosťou $1 - \alpha$. Najužší je pre $x = \bar{x}$. Jeho šírka sa dá ovplyvniť výberom bodov x_1, \dots, x_n , t.j. dizajnom experimentu.

$100(1 - \alpha)\%$ -ný pás spoľahlivosti pre celú regresnú priamku je

$$\left(\hat{\beta}_0 + \hat{\beta}_1 x - s \sqrt{2F_{2, n-2}(1 - \alpha) \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \right)}, \right. \\ \left. \hat{\beta}_0 + \hat{\beta}_1 x + s \sqrt{2F_{2, n-2}(1 - \alpha) \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \right)} \right). \quad (38)$$

Pokrýva s pravdepodobnosťou $1 - \alpha$ celú priamku $\beta_0 + \beta_1 x$ (celú teoretickú regresnú preiamku). Je širší ako pás spoľahlivosti okolo regresnej priamky. Odvodíme si ho neskôr. Pozri Obr. 2, str. 106 v knihe Anděl, J., *Matematická statistika*, SNTL, Praha, 1985.

Poznámka 5.29: Ak vynesieme (22) pre $\mathbf{c} = (1, x)'$ pre každé $x \in \mathcal{R}^1$, dostaneme $100(1 - \alpha)\%$ -ný predikčný pás (pás spoľahlivosti pre jednotlivé merania), ktorý pre každé x (zvlášť) obsahuje meranie v bode x s pravdepodobnosťou $1 - \alpha$:

$$\left(\hat{\beta}_0 + \hat{\beta}_1 x - t_{n-2} \left(1 - \frac{\alpha}{2}\right) s \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}, \right. \\ \left. \hat{\beta}_0 + \hat{\beta}_1 x + t_{n-2} \left(1 - \frac{\alpha}{2}\right) s \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} \right). \quad (39)$$

e) Dvojica regresných priamok.

Majme skupinu náhodných veličín (meraní) Y_1, \dots, Y_n , pre ktoré platí

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad n \geq 3$$

a od nich nezávislú inú skupinu náhodných veličín (meraní) $Y_1^*, \dots, Y_{n^*}^*$, pre ktoré platí

$$Y_i^* = \beta_0^* + \beta_1^* x_i^* + \varepsilon_i^*, \quad i = 1, 2, \dots, n^*, \quad n^* \geq 3$$

LRM pre prvú skupinu meraní je $(\mathbf{Y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ a pre druhú skupinu meraní je $(\mathbf{Y}^*, \mathbf{X}^*\boldsymbol{\beta}^*, \sigma^2\mathbf{I})$, (σ^2 je pre obe skupiny meraní rovnaká), teda

$$\mathbf{Y}_{n,1} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \boldsymbol{\varepsilon}, \quad \text{cov}(\mathbf{Y}) = \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}_{n,n},$$

$$\mathbf{Y}_{n^*,1}^* = \begin{pmatrix} 1 & x_1^* \\ 1 & x_2^* \\ \vdots & \vdots \\ 1 & x_{n^*}^* \end{pmatrix} \begin{pmatrix} \beta_0^* \\ \beta_1^* \end{pmatrix} + \boldsymbol{\varepsilon}_{n^*,1}^* = \mathbf{X}^* \begin{pmatrix} \beta_0^* \\ \beta_1^* \end{pmatrix} + \boldsymbol{\varepsilon}^*, \quad \text{cov}(\mathbf{Y}^*) = \text{cov}(\boldsymbol{\varepsilon}^*) = \sigma^2\mathbf{I}_{n^*,n^*},$$

pričom $\text{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}^*) = \mathbf{0}$. Označme MNČ odhady v jednotlivých LRM $\hat{\boldsymbol{\beta}}$ a $\hat{\boldsymbol{\beta}}^*$, ďalej (podľa (30))

$$s^2 = \frac{1}{n-2}(\mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}) = \frac{1}{n-2} \left(\sum_{i=1}^n Y_i^2 - \hat{\beta}_0 \sum_{i=1}^n Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i Y_i \right), \quad (40)$$

$$s^{*2} = \frac{1}{n^*-2}(\mathbf{Y}^{*\prime}\mathbf{Y}^* - (\hat{\boldsymbol{\beta}}^*)'\mathbf{X}^{*\prime}\mathbf{Y}^*) = \frac{1}{n^*-2} \left(\sum_{i=1}^{n^*} Y_i^{*2} - \hat{\beta}_0^* \sum_{i=1}^{n^*} Y_i^* - \hat{\beta}_1^* \sum_{i=1}^{n^*} x_i^* Y_i^* \right). \quad (41)$$

Podľa Vety 5.21 je $\frac{n-2}{\sigma^2}s^2 \sim \chi_{n-2}^2$ a $\frac{n^*-2}{\sigma^2}s^{*2} \sim \chi_{n^*-2}^2$, sú nezávislé a preto (podľa Vety 3.5)

$$\frac{n-2}{\sigma^2}s^2 + \frac{n^*-2}{\sigma^2}s^{*2} = \frac{1}{\sigma^2} \left[(n-2)s^2 + (n^*-2)s^{*2} \right] \sim \chi_{n+n^*-4}^2. \quad (42)$$

Podľa Vety 5.21 (iii) $\frac{n-2}{\sigma^2}s^2 + \frac{n^*-2}{\sigma^2}s^{*2}$ nezávisí od $\hat{\boldsymbol{\beta}}$ ani od $\hat{\boldsymbol{\beta}}^*$ (ktoré sú tiež medzi sebou nezávislé) a (pomocou Vety 5.21) dostávame

$$\hat{\gamma} = \hat{\beta}_1 - \hat{\beta}_1^* \sim N(\beta_1 - \beta_1^*, \sigma^2\{(\mathbf{X}'\mathbf{X})^{-1}\}_{22} + \sigma^2\{(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\}_{22}),$$

čiže

$$\frac{\hat{\beta}_1 - \hat{\beta}_1^* - (\beta_1 - \beta_1^*)}{\sqrt{\sigma^2 \{(\mathbf{X}'\mathbf{X})^{-1}\}_{22} + \{(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\}_{22}}} = \frac{\hat{\beta}_1 - \hat{\beta}_1^* - (\beta_1 - \beta_1^*)}{\sqrt{\sigma^2 \left[\frac{1}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} + \frac{1}{\sum_{i=1}^{n^*} (x_i^*)^2 - n^*\bar{x}^{*2}} \right]}} \sim N(0,1). \quad (43)$$

Zo vzťahov (42) a (43) dostávame, že štatistika

$$\begin{aligned} T &= \frac{\hat{\beta}_1 - \hat{\beta}_1^* - (\beta_1 - \beta_1^*)}{\sqrt{\sigma^2 \left[\frac{1}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} + \frac{1}{\sum_{i=1}^{n^*} (x_i^*)^2 - n^*\bar{x}^{*2}} \right]}}} = \\ &= \frac{[\hat{\beta}_1 - \hat{\beta}_1^* - (\beta_1 - \beta_1^*)] \sqrt{n+n^*-4}}{\sqrt{(n-2)s^2 + (n^*-2)s^{*2}} \sqrt{\frac{1}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} + \frac{1}{\sum_{i=1}^{n^*} (x_i^*)^2 - n^*\bar{x}^{*2}}}}} \sim t_{n+n^*-4} \end{aligned} \quad (44)$$

Test hypotézy

$$H_0 : \beta_1 = \beta_1^* \quad \not\approx \quad H_1 : \beta_1 \neq \beta_1^* \quad (45)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{(\hat{\beta}_1 - \hat{\beta}_1^*) \sqrt{n+n^*-4}}{\sqrt{(n-2)s^2 + (n^*-2)s^{*2}} \sqrt{\frac{1}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} + \frac{1}{\sum_{i=1}^{n^*} (x_i^*)^2 - n^*\bar{x}^{*2}}}}} \sim t_{n+n^*-4} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n+n^*-4} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n+n^*-4} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Test má hladinu významnosti α .

Podľa Vety 5.21 (i) $\hat{\beta} \sim N(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$ a $\hat{\beta}^* \sim N(\beta^*, \sigma^2(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1})$. Tieto odhady sú nezávislé a preto

$$\hat{\beta} - \hat{\beta}^* \sim N(\beta - \beta^*, \sigma^2[(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}])$$

(k dôkazu stačí napr. vhodne použiť Vetu 2.6). Podľa Vety 3.13 je

$$[\hat{\beta} - \hat{\beta}^* - (\beta - \beta^*)]' \frac{1}{\sigma^2} [(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}]^{-1} [\hat{\beta} - \hat{\beta}^* - (\beta - \beta^*)] \sim \chi_2^2$$

a samozrejme je táto náhodná veličina nezávislá s

$$\frac{1}{\sigma^2} [(n-2)s^2 + (n^*-2)s^{*2}] \sim \chi_{n+n^*-4}^2.$$

Ľahko dostávame, že štatistika

$$F = \frac{[\hat{\beta} - \hat{\beta}^* - (\beta - \beta^*)]' [(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}]^{-1} [\hat{\beta} - \hat{\beta}^* - (\beta - \beta^*)]}{(n-2)s^2 + (n^*-2)s^{*2}} \frac{n+n^*-4}{2} \sim F_{2, n+n^*-4}.$$

Test hypotézy (o totožnosti (celých) teoretických regresných priamok)

$$H_0 : \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} \beta_0^* \\ \beta_1^* \end{pmatrix} \quad \not\approx \quad H_1 : \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \neq \begin{pmatrix} \beta_0^* \\ \beta_1^* \end{pmatrix} \quad (46)$$

realizujeme pomocou testovacej štatistiky

$$F_0 = \frac{(\hat{\beta} - \hat{\beta}^*)'[(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}]^{-1}(\hat{\beta} - \hat{\beta}^*)}{(n-2)s^2 + (n^*-2)s^{*2}} \frac{n+n^*-4}{2} \sim F_{2, n+n^*-4}. \quad (\text{za platnosti } H_0).$$

Ak $F_0 \geq F_{2, n+n^*-4}(1-\alpha) \implies H_0$ zamietame,

ak $F_0 < F_{2, n+n^*-4}(1-\alpha) \implies H_0$ nezamietame.

Test má hladinu významnosti α .

Nech v "nehviezdičkovom" LRM je disperzia náhodných chýb σ^2 a v "hviezdičkovanom" σ^{*2} . Podľa Vety 5.21 je $\frac{n-2}{\sigma^2}s^2 \sim \chi_{n-2}^2$ a $\frac{n^*-2}{\sigma^{*2}}s^{*2} \sim \chi_{n^*-2}^2$. s^2 a s^{*2} , sú nezávislé, preto

$$F = \frac{\frac{s^2}{\sigma^2}}{\frac{s^{*2}}{\sigma^{*2}}} \sim F_{n-2, n^*-2}.$$

Test hypotézy

$$H_0 : \sigma^2 = \sigma^{*2} \quad \not\approx \quad H_1 : \sigma^2 \neq \sigma^{*2} \quad (47)$$

realizujeme pomocou testovacej štatistiky

$$F_0 = \frac{s^2}{s^{*2}} \sim F_{n-2, n^*-2} \quad (\text{za platnosti } H_0).$$

Ak $F_0^{real} \geq F_{n-2, n^*-2}(1-\frac{\alpha}{2})$ alebo $F_0^{real} \leq F_{n-2, n^*-2}(\frac{\alpha}{2}) \implies H_0$ zamietame,

inak H_0 nezamietame. Test má hladinu významnosti α .

Poznámka 5.30: Obidva LRM ("nehviezdičkovany" a "hviezdičkovany") sa dajú modelovať jediným LRM, a síce

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{Y}^* \end{pmatrix} = \begin{pmatrix} \mathbf{1}_{n,1} & \mathbf{x} & \mathbf{0}_{n,1} & \mathbf{0}_{n,1} \\ \mathbf{0}_{n^*,1} & \mathbf{0}_{n^*,1} & \mathbf{1}_{n^*,1} & \mathbf{x}^* \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_0^* \\ \beta_1^* \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}^* \end{pmatrix}, \quad cov \left(\begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}^* \end{pmatrix} \right) = \sigma^2 \mathbf{I}_{n+n^*, n+n^*},$$

kde $\mathbf{x} = (x_1, \dots, x_n)'$ a $\mathbf{x}^* = (x_1^*, \dots, x_{n^*}^*)'$. Všetky testy uvedené v bode e) sa dajú odvodiť v tomto modeli.

f) Regresná parabola (kvadratická regresia).

Majme nezávislé náhodné veličiny (merania) Y_1, \dots, Y_n , pre ktoré platí

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad n \geq 4.$$

(Bezchybné merania ležia na parabole $y = \beta_0 + \beta_1 x + \beta_2 x^2$, hodnoty x_i , $i = 1, 2, \dots, n$ poznáme bezchybne (úplne presne)). LRM je

$$\mathbf{Y}_{n,1} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \boldsymbol{\varepsilon}, \quad \text{cov}(\mathbf{Y}) = \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n,n}.$$

Samozrejme $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ a $s^2 = \frac{1}{n-3}(\sum_{i=1}^n Y_i^2 - \hat{\beta}_0 \sum_{i=1}^n Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i Y_i - \hat{\beta}_2 \sum_{i=1}^n x_i^2 Y_i)$.

Test hypotézy

$$H_0 : \beta_2 = 0 \quad \not\approx \quad H_1 : \beta_2 \neq 0 \quad (48)$$

realizujeme pomocou testovacej štatistiky

$$T_2 = \frac{\hat{\beta}_2}{s \sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{33}}} \sim t_{n-3} \quad (\text{za platnosti } H_0).$$

Test hypotézy

$$H_0 : \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \not\approx \quad H_1 : \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (49)$$

realizujeme pomocou testovacej štatistiky

$$F_0 = \frac{1}{2s^2}(\hat{\beta}_1, \hat{\beta}_2) \begin{pmatrix} \{(\mathbf{X}'\mathbf{X})^{-1}\}_{22} & \{(\mathbf{X}'\mathbf{X})^{-1}\}_{23} \\ \{(\mathbf{X}'\mathbf{X})^{-1}\}_{32} & \{(\mathbf{X}'\mathbf{X})^{-1}\}_{33} \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \sim F_{2,n-3} \quad (\text{za platnosti } H_0).$$

V tomto prípade testujeme, či $Y_i = \beta_0 + \varepsilon_i$ (teda, či Y_i nezávisia od x_i) oproti alternatíve, že závisia lineárne alebo kvadraticky.

g) Polynomická regresia.

Majme nezávislé náhodné veličiny (merania) Y_1, \dots, Y_n , pre ktoré platí

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_p x_i^p + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad n \geq p + 2.$$

(Bezchybné merania ležia na polynome p -teho stupňa $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p$, hodnoty x_i , $i = 1, 2, \dots, n$ poznáme bezchybne (úplne presne)). LRM je

$$\mathbf{Y}_{n,1} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^p \\ 1 & x_2 & x_2^2 & \dots & x_2^p \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^p \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \boldsymbol{\varepsilon}, \quad \text{cov}(\mathbf{Y}) = \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n,n}.$$

Testy dostaneme analogicky ako v prípade regresnej priamky alebo paraboly.

h) Model s dvomi vysvetľujúcimi premennými.

Majme nezávislé náhodné veličiny (merania) Y_1, \dots, Y_n , pre ktoré platí

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad n \geq 4.$$

teda bezchybné merania lineárne závisia od dvoch (vysvetľujúcich) premenných x a z . Hodnoty x_i a z_i , $i = 1, 2, \dots, n$ poznáme úplne bezchybne. LRM je

$$\mathbf{Y}_{n,1} = \begin{pmatrix} 1 & x_1 & z_1 \\ 1 & x_2 & z_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & z_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \boldsymbol{\varepsilon}_{n,1} = \mathbf{X} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \boldsymbol{\varepsilon}, \quad \text{cov}(\mathbf{Y}) = \text{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n,n}.$$

Odhady $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ a $s^2 = \frac{1}{n-3}(\sum_{i=1}^n Y_i^2 - \hat{\beta}_0 \sum_{i=1}^n Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i Y_i - \hat{\beta}_2 \sum_{i=1}^n z_i Y_i)$.

Test hypotézy

$$H_0 : \beta_2 = 0 \quad \not\approx \quad H_1 : \beta_2 \neq 0 \quad (50)$$

realizujeme pomocou testovacej štatistiky

$$T_2 = \frac{\hat{\beta}_2}{s\sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{33}}} \sim t_{n-3} \quad (\text{za platnosti } H_0).$$

Test hypotézy

$$H_0 : \beta_1 = 0 \quad \not\approx \quad H_1 : \beta_1 \neq 0 \quad (51)$$

realizujeme pomocou testovacej štatistiky

$$T_1 = \frac{\hat{\beta}_1}{s\sqrt{\{(\mathbf{X}'\mathbf{X})^{-1}\}_{22}}} \sim t_{n-3} \quad (\text{za platnosti } H_0).$$

Test hypotézy

$$H_0 : \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \not\approx \quad H_1 : \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (52)$$

realizujeme pomocou testovacej štatistiky

$$F_0 = \frac{1}{2s^2}(\hat{\beta}_1, \hat{\beta}_2) \begin{pmatrix} \{(\mathbf{X}'\mathbf{X})^{-1}\}_{22} & \{(\mathbf{X}'\mathbf{X})^{-1}\}_{23} \\ \{(\mathbf{X}'\mathbf{X})^{-1}\}_{32} & \{(\mathbf{X}'\mathbf{X})^{-1}\}_{33} \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \sim F_{2,n-3} \quad (\text{za platnosti } H_0).$$

6 Výberový korelačný koeficient

Majme náhodný výber $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ z dvojrozmerného rozdelenia s distribučnou funkciou $F(x, y; \boldsymbol{\theta})$.

Definícia 6.1: Nech $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ sú výberové priemery. Štatistiku

$$S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

nazývame výberovou kovarianciou,

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right), \quad S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n Y_i^2 - n\bar{Y}^2 \right)$$

sú výberové rozptyly. Výberový korelačný koeficient je

$$r_{XY} = r = \frac{S_{XY}}{S_X S_Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{j=1}^n (Y_j - \bar{Y})^2}} = \frac{\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) \left(\sum_{j=1}^n Y_j^2 - n\bar{Y}^2\right)}}.$$

Poznámka 6.2: Dá sa ukázať (pozri napr. Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str.116), že ak máme náhodný výber rozsahu aspoň 2 z absolútne spojitého rozdelenia, tak r_{XY} je definovaný s pravdepodobnosťou 1.

Veta 6.3: Nech $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ je náhodný výber z regulárneho $N_2 \left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \mathcal{D}(X) & \varrho\sqrt{\mathcal{D}(X)\mathcal{D}(Y)} \\ \varrho\sqrt{\mathcal{D}(X)\mathcal{D}(Y)} & \mathcal{D}(Y) \end{pmatrix} \right)$ rozdelenia, $n > 2$, $0 < \mathcal{D}(X) < \infty$, $0 < \mathcal{D}(Y) < \infty$, $\varrho^2 \neq 1$. Ak $\varrho = 0$ tak

$$T = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \sim t_{n-2}.$$

Dôkaz: Uvažujme podmienené rozdelenie $Y_i/X_i = x_i \stackrel{\text{ozn.}}{=} (x_i)Y_i$, $i = 1, 2, \dots, n$.

Podľa (6) $(x_i)Y_i \sim N \left(\mu_Y + \varrho \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}} (x_i - \mu_X), \mathcal{D}(Y)(1 - \varrho^2) \right)$. Ak označíme

$$\beta_0 = \mu_Y - \mu_X \varrho \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}} \quad \text{a} \quad \beta_1 = \varrho \sqrt{\frac{\mathcal{D}(Y)}{\mathcal{D}(X)}},$$

tak dostávame, že $(x_i)Y_i \sim N(\beta_0 + \beta_1 x_i, \mathcal{D}(Y)(1 - \varrho^2))$, $i = 1, 2, \dots, n$, pričom $(x_1)Y_1, (x_2)Y_2, \dots, (x_n)Y_n$ sú nezávislé. Máme teda LRM

$$\begin{pmatrix} (x_1)Y_1 \\ (x_2)Y_2 \\ \vdots \\ (x_n)Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + {}_{(\mathbf{x})}\varepsilon, \quad {}_{(\mathbf{x})}\varepsilon \sim N_n(\mathbf{0}, \mathcal{D}(Y)(1 - \varrho^2)\mathbf{I}_{n,n}).$$

V tomto LRM platí $\varrho = 0 \iff \beta_1 = 0$. Z (28), (29) a (30) dostávame

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i - {}_{(\mathbf{x})}\bar{Y}}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{\beta}_0 = {}_{(\mathbf{x})}\bar{Y} - \hat{\beta}_1 \bar{x}, \quad {}_{(\mathbf{x})}\bar{Y} = \frac{1}{n} \sum_{i=1}^n (x_i)Y_i,$$

$${}_{(\mathbf{x})}s^2 = \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i^2 - \hat{\beta}_0 \sum_{i=1}^n (x_i)Y_i - \hat{\beta}_1 \sum_{i=1}^n x_i (x_i)Y_i \right).$$

Za platnosti $\varrho = 0$ (teda $\beta_1 = 0$)

$${}_{(\mathbf{x})}T = \frac{\hat{\beta}_1}{{}_{(\mathbf{x})}s} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\hat{\beta}_1}{{}_{(\mathbf{x})}s} \sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \sim t_{n-2} \quad (53)$$

(pozri (31) a nižšie).

Ešte upravme výraz pre $_{(\mathbf{x})}s^2$:

$$\begin{aligned}
{(\mathbf{x})}s^2 &= \frac{1}{n-2} \left(\sum{i=1}^n (x_i)Y_i^2 -_{(\mathbf{x})}\hat{\beta}_0 \sum_{i=1}^n (x_i)Y_i -_{(\mathbf{x})}\hat{\beta}_1 \sum_{i=1}^n x_i (x_i)Y_i \right) = \\
&= \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i^2 - \left(_{(\mathbf{x})}\bar{Y} -_{(\mathbf{x})}\hat{\beta}_1\bar{x}\right) \sum_{i=1}^n (x_i)Y_i -_{(\mathbf{x})}\hat{\beta}_1 \sum_{i=1}^n x_i (x_i)Y_i \right) = \\
&= \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i^2 - \left(_{(\mathbf{x})}\bar{Y} -_{(\mathbf{x})}\hat{\beta}_1\bar{x}\right) n_{(\mathbf{x})}\bar{Y} -_{(\mathbf{x})}\hat{\beta}_1 \sum_{i=1}^n x_i (x_i)Y_i \right) = \\
&= \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i^2 - n_{(\mathbf{x})}\bar{Y}^2 +_{(\mathbf{x})}\hat{\beta}_1 \left[n\bar{x}_{(\mathbf{x})}\bar{Y} - \sum_{i=1}^n x_i (x_i)Y_i \right] \right) = \\
&= \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2 - \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y}}{\sum_{i=1}^n (x_i - \bar{x})^2} \left[\sum_{i=1}^n x_i (x_i)Y_i - n\bar{x}_{(\mathbf{x})}\bar{Y} \right] \right) = \\
&= \frac{1}{n-2} \left(\sum_{i=1}^n (x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2 - \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y}}{\sum_{i=1}^n (x_i - \bar{x})^2} \left[\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right] \right) = \\
&= \frac{1}{n-2} \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^n (x_j)Y_j -_{(\mathbf{x})}\bar{Y} - \left(\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.
\end{aligned}$$

Dosadením do (53) dostávame

$$\begin{aligned}
{(\mathbf{x})}T &= \frac{\frac{\sum{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y}}{\sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{\frac{1}{n-2} \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^n (x_j)Y_j -_{(\mathbf{x})}\bar{Y} - \left(\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = \\
&= \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y}}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^n (x_j)Y_j -_{(\mathbf{x})}\bar{Y} - \left(\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2}} \sqrt{n-2} = \\
&= \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y}}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^n (x_j)Y_j -_{(\mathbf{x})}\bar{Y} - \left(\sum_{i=1}^n (x_i - \bar{x})(x_i)Y_i -_{(\mathbf{x})}\bar{Y} \right)^2}} \sqrt{n-2} \sim t_{n-2}.
\end{aligned}$$

Teda $_{(\mathbf{x})}T \sim t_{n-2}$ pre ľubovoľné $\mathbf{x} = (x_1, \dots, x_n)'$. To je ale to isté, ako tvrdenie $T = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \sim t_{n-2}$ ($_{(\mathbf{x})}T$ je podmienené T za podmienky $\mathbf{X} = \mathbf{x}$, ale nezáleží na podmienke, teda "nepodmienené" T má rovnaké rozdelenie ako podmienené $_{(\mathbf{x})}T$). Q.E.D.

Majme náhodný výber $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ z dvojrozmerného regulárneho normálneho rozdelenia. Pomocou Vety 6.3 testujeme hypotézu o nezávislosti X a Y .

Test hypotézy

$$H_0 : \varrho = 0 \quad \not\approx \quad H_1 : \varrho \neq 0 \quad (54)$$

realizujeme pomocou testovacej štatistiky

$$T_0 = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \sim t_{n-2} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame.

Tento test má hladinu významnosti α .

Poznámka 6.4: Zobrazenie $z : (-1, 1) \rightarrow \mathcal{R} : z(x) = \frac{1}{2} \ln \frac{1+x}{1-x}$ sa nazýva *Fisherova Z-transformácia*. Bez dôkazu si uvedieme nasledujúce tvrdenie:

Veta 6.5: Majme náhodný výber $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ z dvojrozmerného regulárneho normálneho rozdelenia s korelačným koeficientom ϱ a výberovým korelačným koeficientom r . Platí

$$Z = \frac{1}{2} \ln \frac{1+r}{1-r} \approx N \left(\frac{1}{2} \ln \frac{1+\varrho}{1-\varrho} + \frac{\varrho}{2(n-1)}, \frac{1}{n-3} \right).$$

Aproximácia je požitelná pre $n \geq 10$ a ϱ nie blízke 1 alebo -1 .

Veta 6.5 sa aplikuje v nasledujúcich prípadoch

a) Test hypotézy

$$H_0 : \varrho = \varrho_0 \quad \not\approx \quad H_1 : \varrho \neq \varrho_0 \quad (55)$$

realizujeme pomocou testovacej štatistiky

$$U_0 = \frac{Z - \frac{1}{2} \ln \frac{1+\varrho_0}{1-\varrho_0} - \frac{\varrho_0}{2(n-1)}}{\sqrt{\frac{1}{n-3}}} \sim N(0, 1) \quad (\text{za platnosti } H_0).$$

Ak $|U_0| \geq u(1 - \frac{\alpha}{2}) \implies H_0$ zamietame,

ak $|U_0| < u(1 - \frac{\alpha}{2}) \implies H_0$ nezamietame,

prícom $u(1 - \frac{\alpha}{2})$ je $(1 - \frac{\alpha}{2})$ - kvantil $N(0, 1)$ rozdelenia. Test má hladinu významnosti približne rovnú α .

Pre veľké n je $Z \approx N \left(\frac{1}{2} \ln \frac{1+\varrho}{1-\varrho}, \frac{1}{n-3} \right)$.

b) Majme dva nezávislé výbery, každý z dvojrozmerného regulárneho normálneho rozdelenia. Ich rozsahy n_1 a n_2 sú aspoň 30. Korelačný koeficient u prvého rozdelenia je ϱ_1 , u druhého je ϱ_2 . Výberové korelačné koeficienty sú r_1, r_2 a $Z_1 = \frac{1}{2} \ln \frac{1+r_1}{1-r_1}, Z_2 = \frac{1}{2} \ln \frac{1+r_2}{1-r_2}$.

Test hypotézy

$$H_0 : \varrho_1 = \varrho_2 \quad \not\approx \quad H_1 : \varrho_1 \neq \varrho_2 \quad (56)$$

realizujeme pomocou testovacej štatistiky

$$U_0 = \frac{Z_1 - Z_2}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \sim N(0, 1) \quad (\text{za platnosti } H_0).$$

Ak $|U_0| \geq u(1 - \frac{\alpha}{2}) \implies H_0$ zamietame,

ak $|U_0| < u(1 - \frac{\alpha}{2}) \implies H_0$ nezamietame,

Test má hladinu významnosti približne rovnú α .

Definícia 6.6: Majme náhodný výber

$$\mathbf{X}_1 = \begin{pmatrix} X_{1,1} \\ X_{1,2} \\ \vdots \\ X_{1,p} \end{pmatrix}, \quad \mathbf{X}_2 = \begin{pmatrix} X_{2,1} \\ X_{2,2} \\ \vdots \\ X_{2,p} \end{pmatrix}, \quad \dots \quad \mathbf{X}_n = \begin{pmatrix} X_{n,1} \\ X_{n,2} \\ \vdots \\ X_{n,p} \end{pmatrix}$$

z p -rozmerného rozdelenia so strednou hodnotou $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ a kovariančnou maticou $\boldsymbol{\Sigma}$.

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i$$

je výberový priemer,

$$\mathbf{S}_{\mathbf{X}} = \mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$$

je výberová kovariančná matica a

$$\mathbf{R}_{\mathbf{X}} = \mathbf{R} = (\text{diag} \mathbf{S})^{-\frac{1}{2}} \mathbf{S} (\text{diag} \mathbf{S})^{-\frac{1}{2}}$$

je výberová korelačná matica, pričom $(\text{diag} \mathbf{S})^{-\frac{1}{2}} = \begin{pmatrix} \{\mathbf{S}\}_{1,1}^{-\frac{1}{2}} & 0 & \dots & 0 \\ 0 & \{\mathbf{S}\}_{2,2}^{-\frac{1}{2}} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \{\mathbf{S}\}_{p,p}^{-\frac{1}{2}} \end{pmatrix}$.

Poznámka 6.7:

(i) Namiesto \mathbf{S} sa niekedy používa $\mathbf{M} = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' = \frac{n-1}{n} \mathbf{S}$.

(ii) $\{\mathbf{S}\}_{j,k}$ je výberová kovariancia $i \neq j$, $\{\mathbf{S}\}_{j,k} = \frac{1}{n-1} \sum_{i=1}^n (X_{i,j} - \bar{X}_j)(X_{i,k} - \bar{X}_k)$, $\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{i,j}$.

(iii) $\{\mathbf{S}\}_{j,j}$ je výberový rozptyl $\{\mathbf{S}\}_{j,j} = \frac{1}{n-1} \sum_{i=1}^n (X_{i,j} - \bar{X}_j)^2$.

(iv) $\mathbf{D} = (\mathbf{X}_1 : \mathbf{X}_2 : \dots : \mathbf{X}_n)$ je dátová matica (matica dát) typu $p \times n$.

Veta 6.8: Nech $\mathbf{X}_1, \dots, \mathbf{X}_n$ je náhodný výber z p -rozmerného rozdelenia so strednou hodnotou $\boldsymbol{\mu}$ a kovariančnou maticou $\boldsymbol{\Sigma}$. Potom

(i) $\bar{\mathbf{X}}$ je nestranný odhad $\boldsymbol{\mu}$;

(ii) $\text{cov}(\bar{\mathbf{X}}) = \frac{1}{n} \boldsymbol{\Sigma}$;

(iii) \mathbf{S} je nestranný odhad $\boldsymbol{\Sigma}$.

Dôkaz: (i) $\mathcal{E}(\bar{\mathbf{X}}) = \mathcal{E}\left(\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i\right) = \frac{1}{n} \sum_{i=1}^n \mathcal{E}(\mathbf{X}_i) = \frac{1}{n} \sum_{i=1}^n \boldsymbol{\mu} = \boldsymbol{\mu}$;

(ii) $\text{cov}(\bar{\mathbf{X}}) = \mathcal{E}((\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})') =$

$$= \mathcal{E}\left(\frac{1}{n}(\mathbf{X}_1 - \boldsymbol{\mu} + \mathbf{X}_2 - \boldsymbol{\mu} + \dots + \mathbf{X}_n - \boldsymbol{\mu}) \frac{1}{n}(\mathbf{X}_1 - \boldsymbol{\mu} + \mathbf{X}_2 - \boldsymbol{\mu} + \dots + \mathbf{X}_n - \boldsymbol{\mu})'\right) =$$

$$= \frac{1}{n^2} \mathcal{E} ((\mathbf{X}_1 - \boldsymbol{\mu})(\mathbf{X}_1 - \boldsymbol{\mu})' + (\mathbf{X}_2 - \boldsymbol{\mu})(\mathbf{X}_2 - \boldsymbol{\mu})' + \dots + (\mathbf{X}_n - \boldsymbol{\mu})(\mathbf{X}_n - \boldsymbol{\mu})') =$$

(lebo $\mathcal{E}(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' = \mathbf{0}$ pre $i \neq j$)

$$= \frac{1}{n^2} \sum_{i=1}^n \mathcal{E}(\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' = \frac{1}{n^2} (n\boldsymbol{\Sigma}) = \frac{1}{n} \boldsymbol{\Sigma}.$$

$$(iii) \mathcal{E}((n-1)\mathbf{S}) = \mathcal{E}(\sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})') = \mathcal{E}(\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu} - (\bar{\mathbf{X}} - \boldsymbol{\mu}))(\mathbf{X}_i - \boldsymbol{\mu} - (\bar{\mathbf{X}} - \boldsymbol{\mu}))') =$$

$$= \mathcal{E} \left\{ \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' - \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' - \sum_{i=1}^n (\bar{\mathbf{X}} - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' \right\} =$$

$$= \mathcal{E} \left\{ \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' - n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' - n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})' \right\} =$$

(lebo $\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) = n(\bar{\mathbf{X}} - \boldsymbol{\mu})$)

$$= \mathcal{E} \left(\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})' \right) - n\mathcal{E}((\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})') = n\boldsymbol{\Sigma} - n \text{cov}(\bar{\mathbf{X}}) = (n-1)\boldsymbol{\Sigma}. \quad \text{Q.E.D.}$$

Poznámka 6.9: Z Vety 6.8 vyplýva, že $\mathcal{E}\{\mathbf{S}\}_{i,j} = \{\boldsymbol{\Sigma}\}_{i,j}$

Nech $\mathbf{X}_1 = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Z}_1 \end{pmatrix}, \dots, \mathbf{X}_n = \begin{pmatrix} \mathbf{Y}_n \\ \mathbf{Z}_n \end{pmatrix}$, kde $\mathbf{Y}_i \in \mathcal{R}^k$, $\mathbf{Z}_i \in \mathcal{R}^{p-k}$, $i = 1, 2, \dots, n$, je náhodný výber z

p -rozmerného rozdelenia so strednou hodnotou $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_Y \\ \boldsymbol{\mu}_Z \end{pmatrix}$, pričom $\boldsymbol{\mu}_Y \in \mathcal{R}^k$, $\boldsymbol{\mu}_Z \in \mathcal{R}^{p-k}$ a kovariančnou

maticou $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{Y,Y} & \boldsymbol{\Sigma}_{Y,Z} \\ \boldsymbol{\Sigma}_{Z,Y} & \boldsymbol{\Sigma}_{Z,Z} \end{pmatrix}$, kde $\boldsymbol{\Sigma}_{Y,Y}$ je $k \times k$ matica, $\boldsymbol{\Sigma}_{Y,Z}$ je $k \times (p-k)$ matica, $\boldsymbol{\Sigma}_{Z,Y}$ je $(p-k) \times k$ matica a $\boldsymbol{\Sigma}_{Z,Z}$ je $(p-k) \times (p-k)$ matica.

Úplne rovnako rozdelíme výberovú kovariančnú maticu $\mathbf{S} = \begin{pmatrix} \mathbf{S}_{Y,Y} & \mathbf{S}_{Y,Z} \\ \mathbf{S}_{Z,Y} & \mathbf{S}_{Z,Z} \end{pmatrix}$ tak, že $\mathbf{S}_{Y,Y}$ je $k \times k$ matica, $\mathbf{S}_{Y,Z}$ je $k \times (p-k)$ matica, $\mathbf{S}_{Z,Y}$ je $(p-k) \times k$ matica a $\mathbf{S}_{Z,Z}$ je $(p-k) \times (p-k)$ matica a tiež výberovú korelačnú maticu $\mathbf{R} = \begin{pmatrix} \mathbf{R}_{Y,Y} & \mathbf{R}_{Y,Z} \\ \mathbf{R}_{Z,Y} & \mathbf{R}_{Z,Z} \end{pmatrix}$ tak, že $\mathbf{R}_{Y,Y}$ je $k \times k$ matica, $\mathbf{R}_{Y,Z}$ je $k \times (p-k)$ matica, $\mathbf{R}_{Z,Y}$ je $(p-k) \times k$ matica a $\mathbf{R}_{Z,Z}$ je $(p-k) \times (p-k)$ matica.

Maticu $\mathbf{S}_{Y,Z}$ nazývame *výberová kovariančná matica náhodných výberov* $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ a $\mathbf{Z}_1, \dots, \mathbf{Z}_n$. Platí

$$\mathbf{S}_{Y,Z} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Z}_i - \bar{\mathbf{Z}})'$$

a podľa Vety 6.8 (iii) je $\mathcal{E}(\mathbf{S}_{Y,Z}) = \boldsymbol{\Sigma}_{Y,Z}$.

Definícia 6.10: Ak $\begin{pmatrix} Y_1 \\ \mathbf{X}_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ \mathbf{X}_n \end{pmatrix}$, kde $\mathbf{Y}_i \in \mathcal{R}$, $\mathbf{X}_i \in \mathcal{R}^p$, $i = 1, 2, \dots, n$, je náhodný výber z $(p+1)$ -rozmerného rozdelenia s regulárnou výberovou korelačnou maticou $\mathbf{R}_{\mathbf{X},\mathbf{X}}$, tak *výberový koeficient mnohnásobnej korelácie* $r_{Y,\mathbf{X}}$ je definovaný ako také nezáporné číslo, pre ktoré platí

$$r_{Y,\mathbf{X}}^2 = \mathbf{R}_{Y,\mathbf{X}} \mathbf{R}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{R}_{\mathbf{X},Y}.$$

Poznámka 6.11: Výberový koeficient mnohnásobnej korelácie $r_{Y,\mathbf{X}}$ je akýsi výberový "proťajšok" teoretického koeficientu mnohnásobnej korelácie $\rho_{Y,\mathbf{X}}$ (pozri Vetu 4.7 (i),(iii)).

Veta 6.12: Nech $\begin{pmatrix} Y_1 \\ \mathbf{X}_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ \mathbf{X}_n \end{pmatrix}$, je náhodný výber z $(p+1)$ -rozmerného regulárneho normálneho rozdelenia s koeficientom mnohonásobnej korelácie $\varrho_{Y,\mathbf{X}} = 0$. Ak $n > p+1$, tak štatistika

$$F = \frac{n-p-1}{p} \frac{r_{Y,\mathbf{X}}^2}{1-r_{Y,\mathbf{X}}^2} \sim F_{p,n-p-1}.$$

Dôkaz: Pomocou vhodného LRM, pozri Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 125.

Test hypotézy

$$H_0 : \varrho_{Y,\mathbf{X}} = 0 \quad \not\approx \quad H_1 : \varrho_{Y,\mathbf{X}} \neq 0 \quad (57)$$

realizujeme pomocou testovacej štatistiky (z Vety 6.12)

$$F_0 = \frac{n-p-1}{p} \frac{r_{Y,\mathbf{X}}^2}{1-r_{Y,\mathbf{X}}^2} \sim F_{p,n-p-1} \quad (\text{za platnosti } H_0).$$

Ak $F_0 \geq F_{p,n-p-1}(1-\alpha) \implies H_0$ zamietame,

ak $F_0 < F_{p,n-p-1}(1-\alpha) \implies H_0$ nezamietame,

Test má hladinu významnosti rovnú α .

Definícia 6.13: Ak $\begin{pmatrix} Y_1 \\ Z_1 \\ \mathbf{X}_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ Z_n \\ \mathbf{X}_n \end{pmatrix}$, kde $Y_i \in \mathcal{R}$, $Z_i \in \mathcal{R}$, $\mathbf{X}_i \in \mathcal{R}^p$, $i = 1, 2, \dots, n$, je náhodný výber z $(p+2)$ -rozmerného rozdelenia s regulárnou výberovou korelačnou maticou

$$\mathbf{R}_{\begin{pmatrix} Y \\ Z \\ \mathbf{X} \end{pmatrix}} = \begin{pmatrix} 1 & r_{Y,Z} & \mathbf{R}_{Y,\mathbf{X}} \\ r_{Z,Y} & 1 & \mathbf{R}_{Z,\mathbf{X}} \\ \mathbf{R}_{\mathbf{X},Y} & \mathbf{R}_{\mathbf{X},Z} & \mathbf{R}_{\mathbf{X},\mathbf{X}} \end{pmatrix},$$

tak výberový koeficient parciálnej korelácie (výberový parciálny korelačný koeficient) je

$$r_{Y,Z,\mathbf{X}} = \frac{r_{Y,Z} - \mathbf{R}_{Y,\mathbf{X}} \mathbf{R}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{R}_{\mathbf{X},Z}}{\sqrt{(1-r_{Y,\mathbf{X}}^2)(1-r_{Z,\mathbf{X}}^2)}},$$

kde $r_{Y,\mathbf{X}}^2 = \mathbf{R}_{Y,\mathbf{X}} \mathbf{R}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{R}_{\mathbf{X},Y}$, $r_{Z,\mathbf{X}}^2 = \mathbf{R}_{Z,\mathbf{X}} \mathbf{R}_{\mathbf{X},\mathbf{X}}^{-1} \mathbf{R}_{\mathbf{X},Z}$, pokiaľ menovateľ nie je rový nule.

Poznámka 6.14: Výberový koeficient parciálnej korelácie $r_{Y,Z,\mathbf{X}}$ je akýsi výberový "proťajšok" teoretického parciálneho korelačného koeficientu $\varrho_{Y,Z,\mathbf{X}}$ (pozri Vetu 4.12).

Veta 6.15: Nech $\begin{pmatrix} Y_1 \\ Z_1 \\ \mathbf{X}_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ Z_n \\ \mathbf{X}_n \end{pmatrix}$ je náhodný výber z $(p+2)$ -rozmerného regulárneho normálneho rozdelenia, ktoré má parciálny korelačný koeficient $\varrho_{Y,Z,\mathbf{X}} = 0$. Ak $n > p+2$, tak štatistika

$$T = \frac{r_{Y,Z,\mathbf{X}}}{\sqrt{1-r_{Y,Z,\mathbf{X}}^2}} \sqrt{n-p-2} \sim t_{n-p-2}.$$

Dôkaz: Pomocou vhodného LRM, pozri Anděl, J., *Matematická statistika*, SNTL, Praha, 1985, str. 128.

Test hypotézy

$$H_0 : \varrho_{Y,Z,\mathbf{X}} = 0 \quad \text{vs} \quad H_1 : \varrho_{Y,Z,\mathbf{X}} \neq 0 \quad (58)$$

realizujeme pomocou testovacej štatistiky (z Vety 6.15)

$$T_0 = \frac{r_{Y,Z,\mathbf{X}}}{\sqrt{1 - r_{Y,Z,\mathbf{X}}^2}} \sqrt{n - p - 2} \sim t_{n-p-2} \quad (\text{za platnosti } H_0).$$

Ak $|T_0| \geq t_{n-p-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ zamietame,

ak $|T_0| < t_{n-p-2} \left(1 - \frac{\alpha}{2}\right) \implies H_0$ nezamietame,

Test má hladinu významnosti rovnú α .