3. Úkol na předmět Mnohorozměrná analýza

Jaromír Macoun

Březen 2022

3. úloha

Mějme náhodný vektor $\mathbf{X} \sim N_2(\mu, \Sigma)$, kde $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ a $\mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

Uvažujme náhodné veličiny vzniklé z náhodného vektoru $\mathbf{X}$ transformací $\mathbb{A}\mathbf{X} = (1,1)\mathbf{X}$ a $\mathbb{B}\mathbf{X} = (1,-1)\mathbf{X}$. Tyto veličiny vznikly lineární transformací z mnohorozměrného normálního rozdělení, proto jsou také normální.

Pro normální náhodné veličiny se nezávislost ověřuje jednoduše, stačí zjistit hodnotu kovariance.

$$cov(\mathbb{A}\mathbf{X}, \mathbb{B}\mathbf{X}) = \mathbb{A}cov(\mathbf{X}, \mathbf{X})\mathbb{B}^T = \mathbb{A}\mathbb{B}^T = (1,1)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0.$$

Tímto je ukázáno, že se opravdu jedná o nezávislé náhodné veličiny.

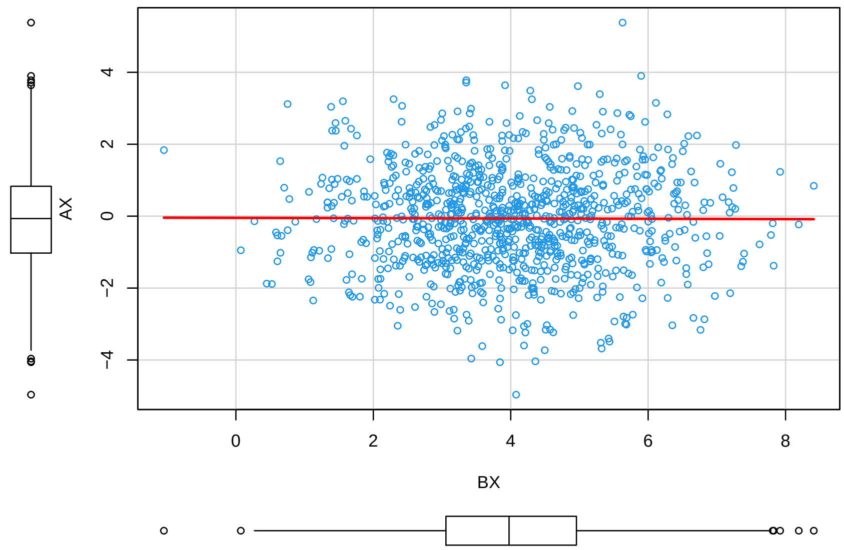
Mějme náhodnou realizaci vektoru $\mathbf{X}$ o 1000 pozorováních. Nezávislost náhodných veličin je za předpokladu normality ekvivalentní nulové korelaci. Pokud uvažujeme lineární model, kde jako odezvu bereme $Z_1 = \mathbb{B}\mathbf{X}$ a jako vysvětlující proměnnou $Z_2 = \mathbb{A}\mathbf{X}$. Nulová korelace je ekvivalentní nulovosti parametru směrnice u regresoru Z_2. V obrázku 1 je vykreslena regresní přímka společně s pozorováními a lze vidět, že přímka je téměř vodorovná. To je ve shodě s výsledkem teoretické části úlohy.

Marginální normální rozdělení

Mějme náhodnou veličinu $\mathbf{X}$ s normovaným normálním rozdělením. Dále uvažujme náhodnou veličinu $\mathbf{Z}$, která se získá z $\mathbf{X}$ následující transformací

$$\mathbf{Z} = \Phi^{-1}(F(\mathbf{Z}^2)),$$

kde F je distribuční funkce chí kvadrát rozdělení o jednom stupni volnosti a Φ^{-1} je kvantilová funkce normovaného normálního rozdělení. $\mathbf{Z}$ má tedy také normované normální rozdělení, neboť aplikací distribuční funkce na náhodnou veličinu s touto distribuční funkcí dostáváme veličinu U s rovnoměrným rozdělením na intervalu $(0, 1)$. Dále transformací náhodné veličiny U kvantilovou funkcí Φ^{-1} dostáváme normované normální rozdělení.



Obrázek 1: Klasický bodový graf s regresní přímkou

Veličiny $\mathbf{Z}$ a $\mathbf{X}$ jsou nekorelované neboť platí

$$cov(\mathbf{X}, \mathbf{Z}) = \mathbb{E}\mathbf{XZ} = 0$$

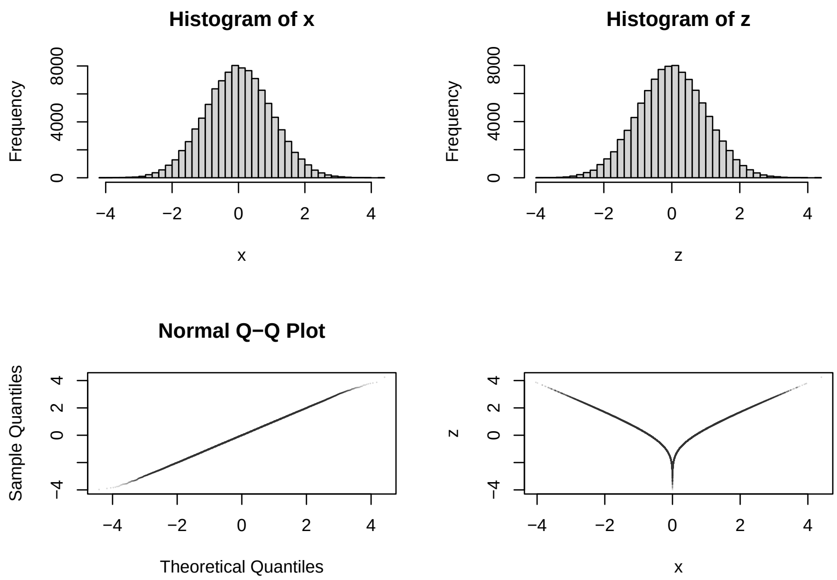
Poslední rovnost plyne z toho, že $\mathbf{Z}$ je symetrické okolo nuly, tedy výraz $\mathbf{XZ}$ (resp. funkce) je lichý.

Máme veličiny, které jsou nekorelované, ale jsou zřejmě závislé. Z toho vyplývá, že $\mathbf{X}$ a $\mathbf{Z}$ nemůžou mít sdružené normální rozdělení.

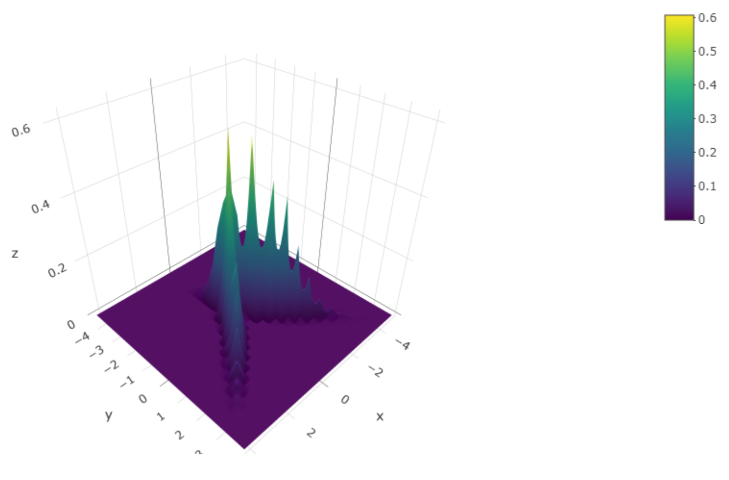
V simulaci jsem uvažoval 100000 pozorování náhodného výběru z normovaného normálního rozdělení. Jak je vidět z histogramu náhodné veličiny Z(obrázek č 2, zdá se že Z má opravdu normované normální rozdělení(pro srovnání je i uveden histogram z náhodného výběru veličiny X). QQplot se zdá být také normální. Nicméně na bodovém grafu(obrázek č.2 vpravo dole) je vidět silná závislost mezi $\mathbf{X}$ a $\mathbf{Z}$.

Dále je vidět porušení sdružené normality na obrázku č3.Tento obrázek zobrazuje odhad hustoty. Pro 3D zobrazení lze prohlédnout zde:

`https://labk10.karlin.mff.cuni.cz/~macounj1/obrazekweb.html`



Obrázek 2: Souhrn vysledků



Obrázek 3: Odhad 3D hustoty