1 Statistika

1.1 Úvod

V následujícím textu budeme jako N označovat libovolné přirozené číslo nebo $+\infty$.

Definice 1.1. *Buďte $(X_j)_{j=1}^N$ náhodné veličiny na prostorech $(\Omega_j, \mathcal{A}_j, \mathrm{P}_j)$ s rozděleními $\mathrm{P}^{X_j} = \mathrm{P} \circ X_j^{-1}$. Definujme nyní*

$$\Omega^{(N)} = \times_{j=1}^N \Omega_j$$

$$\mathcal{A}^{(N)} = \bigotimes_{j=1}^N \mathcal{A}_j \stackrel{d}{=} \sigma\left(\times_{k=1}^l \mathcal{A}_{j_k} \ : \ l \in \widehat{N}\right) = \sigma\left(\times_{k=1}^l A_{j_k} \ : \ A_{j_k} \in \mathcal{A}_{j_k}, \ l \in \widehat{N}\right)$$

$$\mathrm{P}^{(N)} = \mathrm{P}_1 \otimes \cdots \otimes \mathrm{P}_N$$

přičemž $\mathrm{P}^{(N)}$ je taková pravděpodobnostní míra na prostoru $\left(\Omega^{(N)}, \mathcal{A}^{(N)}\right)$, pro kterou platí

$$\mathrm{P}^{(N)}\left(\times_{k=1}^l A_{j_k}\right) = \prod_{k=1}^l \mathrm{P}_{j_k}\left(A_{j_k}\right)$$

Víme tedy, že existuje přirozené prodloužení $\left(\widetilde{X}_j\right)_{j=1}^N$ na prostoru $\left(\Omega^{(N)}, \mathcal{A}^{(N)}, \mathrm{P}^{(N)}\right)$. Tudíž $j-$tý prvek zachovává vlastnosti $j-$tého původního prvku a současně $j-$tá pravděpodobnost zachovává vlastnosti $j-$té původní pravděpodobnosti.

Poznámka 1.2. *V předchozí definici nelze $\sigma-$algebru $\mathcal{A}^{(N)}$ nadefinovat jako*

$$\mathcal{A}^{(N)} = \times_{j=1}^N A_j \qquad A_j \in \mathcal{A}_j \ j \in \widehat{N}$$

protože to by nebyla $\sigma-$algebra.

Věta 1.3. *Buďte $(X_j)_{j=1}^N$ náhodné veličiny a $\left(\widetilde{X}_j\right)_{j=1}^N$ buď jejich přirozeným prodloužením. Potom*

1. *$\left(\widetilde{X}_j\right)_{j=1}^N$ jsou nezávislé*
2. *$\mathrm{P}^{\widetilde{X}_j} = \mathrm{P}_j^{X_j}$ pro $\forall j \in \widehat{N}$, tj.*

$$\forall B \in \mathcal{B} \ \forall j \in \widetilde{N} \ \mathrm{P}^{\widetilde{X}_j}(B) = \mathrm{P}^{X_j}(B)$$

Důkaz. Nejdříve dokážeme stejnost rozdělení. Volme $B_j \in \mathcal{B}, \ j \in \widehat{N}$ libovolně. Potom

$$\{X_j \in B_j\} = \mathrm{P}\left(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_{j-1} \times \{X_j \in B_j\} \times \Omega_{j+1} \times \cdots \Omega_N\right) =$$

$$= 1 \cdot 1 \cdots 1 \cdot \mathrm{P}\left(X_j \in B_j\right) \cdot 1 \cdots 1 \cdot 1 = \mathrm{P}_j\left(X_j \in B_j\right) = \mathrm{P}^{X_j}\left(B_j\right)$$

a nyní nezávislost. Chceme ukázat, že sdružená pravděpodobnost

$$\mathrm{P}\left(X_{j_1} \in B_{j_1}, X_{j_2} \in B_{j_2}, \ldots, X_{j_l} \in B_{j_l}\right)$$

je stejná jako součin pravděpodobností, a to lze velice jednodue ukázat přes kartézský součin (stejně jako u stejnosti rozdělení), tj.

$$\mathrm{P}\left(\widetilde{X}_{j_k} \in B_{j_k}\right) = \mathrm{P}^{(N)}\left(\Omega_1 \times \cdots \times \Omega_{j_{k-1}} \times \left\{X_{j_k} \in B_{j_k} \times \Omega_{j_{k+1}}\right\} \times \cdots \times \Omega_{j_l}\right)$$

ale průnik takovýchto kartézských součinů je opět kartézský součin, a tedy

$$\mathrm{P}^{(N)}\left(X_{k=1}^{N}\left\{X_{j_k} \in B_{j_k}\right\}\right) = \prod_{k=1}^{l} \mathrm{P}\left(X_{j_k} \in B_{j_k}\right) = \prod_{k=1}^{l} \mathrm{P}^{(N)}\left(\widetilde{X}_{j_k} \in B_{j_k}\right)$$

□

Poznámka 1.4 (Komentář Kůse)**.** *V předchozí větě jsem někde něco moc urychlil, takže je tam někde nějaký problém.*

Důsledek 1.5.

1. *Pokud označíme* $\mathrm{P}^{X_j} = \mathrm{P}^X$*, potom* $\left(\widetilde{X}_j\right)_{j=1}^{N}$ *jsou nezávislé a mají stejné rozdělení (jsou i.i.d. s* P^X*).*
2. *Pokud* $X_j \sim \mathrm{P}^X$ *pro* $\forall j \in \widehat{N}$ *a současně*

   $$(\Omega_j, \mathcal{A}_j, \mathrm{P}_j) = (\Omega, \mathcal{A}, \mathrm{P})$$

   potom jsou $\left(\widetilde{X}_j\right)_{j=1}^{N}$ *i.i.d. náhodné veličiny na prostoru* $\left(\Omega^{(N)}, \mathcal{A}^{(N)}, \mathrm{P}^{(N)}\right)$*, který budeme v následujícím textu (pro matení čtenáře) značit* $(\Omega, \mathcal{A}, \mathrm{P})$*.*

1.2 Statistika - základní pojmy a definice

Statistika sice využívá modelu z předchozího důsledku, ale jednotlivým částem prostoru $(\Omega, \mathcal{A}, \mathrm{P})$ přiřazuje mírně odlišný význam.

Ω - populace

ω - individuum, element

$X : \Omega \to \mathbb{R}$**, měřitelná** - vlastnost Ω

$\omega^{(N)} \in \Omega^{(N)}$ - výběr individuí z populace Ω

$\mathrm{P}^{(N)}$ - součinová pravděpodobnostní míra

pozorování Buďte $\left(\widetilde{X}_j\right)_1^N$ i.i.d. s rozdělením P^X. Potom $\widetilde{X}_j$ nazýváme pozorováními X na populaci Ω.

Poznámka 1.6. *Budeme značit* $\widetilde{X}_j = X_j$ *na prostoru* $(\Omega, \mathcal{A}, \mathrm{P})$*.*

Úkolem statistiky obecně je na základě realizací $\widetilde{X}_j\left(\omega^{(N)}\right) = \widetilde{X}_j$ „odhadnout“ tvar P^X. Je to vlastně postup obrácený vzhledem k počtu pravděpodobnosti. Statistické úlohy bychom dále mohli rozdělit na:

(A) Odhad parametrů rozdělení P^X

(A1) Bodový odhad parametrů Nechť θ je nějaký parametr spojený s Ω. Úkolem je najít funkci $\widehat{\theta}(\mathbf{X})$, která na základě pozorování $\mathbf{X} = (X_j)_{j=1}^n$ odhaduje parametr θ.

(A2) Intervalový odhad parametrů Úkol je obdobný jako v případě bodového odhadu parametrů, nicméně hledáme takové borelovsky měřitelné funkce

$$\underline{\theta}(\mathbf{X})$$

$$\overline{\theta}(\mathbf{X})$$

že $\mathrm{P}\left(\theta \in \left[\underline{\theta}(\mathbf{X}), \overline{\theta}(\mathbf{X})\right]\right) \geq 1 - \alpha \quad \alpha \in (0,1)$

(A2') Konfidenční odhad parametrů Hledáme obecnou množinu $\mathbf{C}(\mathbf{X}) \subset \mathbf{R}^k$ (pro parametr $\theta \in \mathbb{R}^k$ takovou, že

$$\mathrm{P}(\theta \in \mathbf{C}(\mathbf{X})) \geq 1 - \alpha \quad \alpha \in (0,1)$$

(B) Testování hypotéz o rozdělení P^X Předmětem zkoumání je opět parametr, přičemž mohu vyslovit například hypotézu $H_0 : \theta = 5$. Abych ale o takové hypotéze vůbec něco mohl říct, potom si musím sehnat pozorování $\mathbf{X} = (X_1, \ldots, X_n)$ a na jeho základě se mohu pokusit zjistit zda

$$\mathrm{P}(H_0 \text{ platí}) \geq 1 - \alpha$$

Pokud tento vztah platí, potom hypotézu přijmeme, jinak ji zamítneme.

1.3 Bodový odhad parametrů

Uvažujme následující model: Buďte $X \sim \mathrm{P}^X$ náhodné veličiny na prostoru $(\Omega, \mathcal{A}, \mathrm{P})$, a nechť parametr θ, jehož hodnotu se snažíme zjistit, je $\theta \in \mathbf{\Theta} \subset \mathbb{R}^k$. Přitom $\mathbf{\Theta}$ nazýváme **parametrickým prostorem**. Můžeme také odhadovat hodnotu nějaké funkce $\tau(\theta)$. Takovou funkci nazýváme **parametrickou**.

Postup je zhruba následující - vytáhneme si $\omega^{(n)} = (\omega_1, \ldots, \omega_n)$, proměříme je pomocí X (tj. $\widetilde{X}_j\left(\omega^{(n)}\right) = X_j(\omega)$, a nakonec odhadneme $\tau(\theta)$.

Definice 1.7 (Odhad parametrické funkce)**.** *Buď $\mathbf{X} = (X_j)_{j=1}^n$ je pozorování X (náhodné veličiny s rozdělením P^X). Potom libovolnou borelovsky měřitelnou funkci $T(\mathbf{X}) : \Omega \to \mathbb{R}^s$ nazýváme **odhadem parametrické funkce** $\tau(\theta)$ na základě pozorování $\mathbf{X}$. Specielně pro $\tau(\theta) = \theta$ označujeme $T(\mathbf{X}) = \widehat{\theta}(\mathbf{X})$.*

Definice 1.8 (Eficientní odhad)**.** *$T(\mathbf{X})$ je eficientním (vydatným) odhadem parametrické funkce $\tau(\theta)$, pokud*

$$E(T(\mathbf{X}) - \tau(\theta))^2 \leq E\left(\widetilde{T}(\mathbf{X}) - \tau(\theta)\right)^2 \quad \textit{pro } \forall \widetilde{T}(\mathbf{X})$$

Definice 1.9 (Nestranný odhad)**.** *$T(\mathbf{X})$ je nestranným odhadem parametrické funkce $\tau(\theta)$ pokud*

$$E(T(\mathbf{X})) = \tau(\theta) \quad \textit{pro } \forall \theta \in \Theta$$

Definice 1.10 (Asymptoticky nestranná posloupnost odhadů)**.** *Buď $(T_n(\mathbf{X}))_{n=1}^{\infty}$ taková posloupnost odhadů parametrické funkce $\tau(\theta)$, že $T_n(\mathbf{X}) = T_n(X_1, \ldots, X_n)$ (tj. odhady jsou založeny na stále více pozorováních). Říkáme, že odhad $(T_n(\mathbf{X}))$ je asymptoticky nestranný, pokud*

$$\lim_{n\to\infty} E\left(T_n(\mathbf{X})\right) = E\tau(\theta) \quad \textit{pro } \forall\theta \in \Theta$$

Definice 1.11 (Slabá konzistence)**.** *Posloupnost odhadů $(T_n(\mathbf{X}))_{n=1}^{\infty}$ je slabě konzistentním odhadem parametrické funkce $\tau(\theta)$, pokud*

$$T_n(\mathbf{X}) \xrightarrow{\mathrm{P}_\theta} \tau(\theta) \quad \textit{pro } \forall\theta \in \Theta$$

Definice 1.12 (Silná konzistence)**.** *Posloupnost odhadů $(T_n(\mathbf{X}))_{n=1}^{\infty}$ je slabě konzistentním odhadem parametrické funkce $\tau(\theta)$, pokud*

$$T_n(\mathbf{X}) \xrightarrow{s.j.} \tau(\theta) \quad \textit{pro } \forall\theta \in \Theta$$

tj. $\mathrm{P}^\theta\left(|T_n(\mathbf{X}) - \tau(\theta)| < \varepsilon\right) \to 1$.

Věta 1.13. *Buď $(T_n(\mathbf{X}))_{n=1}^{\infty}$ taková posloupnost odhadů parametrické funkce $\tau(\theta)$, že*

1. $ET_n(\mathbf{X}) \to \tau(\theta) \quad \textit{pro } \forall\theta \in \Theta$
2. $DT_n(\mathbf{X}) \to 0$

Potom je $T_n(\mathbf{X})$ slabě konzistentním odhadem.

Definice 1.14 (Asymptoticky normální posloupnost odhadů)**.** *Posloupnost odhadů $(T_n(\mathbf{X}))_{n=1}^{\infty}$ se nazývá asymptoticky normální s kovarianční maticí $\mathbf{C}(\theta)$ pokud*

$$\sqrt{n}\left(T_n(\mathbf{X}) - \tau(\theta)\right) \xrightarrow{\mathcal{D}} N_s\left(0, \mathbf{C}(\theta)\right) \quad \textit{pro } \forall\theta \in \Theta$$

tj. $T_n(\mathbf{X}) \sim AN_s\left(\tau(\theta), \frac{1}{n}\mathbf{C}(\theta)\right)$, a specielně pro $s = 1$ platí $\sqrt{n}\left(T_n(\mathbf{X}) - \tau(\theta)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2(\theta)\right)$, kde $\sigma^2(\theta)$ je asymptotický rozptyl.

Poznámka 1.15. *Nechť $T_n(\mathbf{X}) \sim AN_s\left(\tau(\theta), \frac{1}{n}\mathbf{C}(\theta)\right)$, potom*

$$T_n(\mathbf{X}) \xrightarrow{\mathrm{P}} \tau(\theta)$$

Věta 1.16. *Buď $\mathbf{X} = (X_1, \ldots, X_n)$ pozorování na prostoru $(\Omega, \mathcal{A}, \mathrm{P})$, přičemž $\mathbf{X} \in \mathcal{L}_2$. Potom*

1. $$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

   je konzistentním, nestranným a asymptoticky normálním odhadem EX

2. $$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2$$

   je konzistentním a nestranným odhadem DX

3.

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2$$

je konzistentním a asymptoticky nestranným odhadem DX

Důkaz.

1. **konzistentnost** Ze zákona velkých čísel přímo vyplývá, že $\overline{X}_n \xrightarrow{\mathrm{P}} EX$.

   nestrannost

$$E\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} EX_j = EX$$

2.

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2 \xrightarrow{\mathrm{P}} DX$$

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2 = \frac{n}{n-1}\widehat{\sigma}_n^2 \xrightarrow{\mathrm{P}} DX$$

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \left(\overline{X}_n\right)^2$$

$$E\left(\overline{X}_n\right)^2 = E\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 = \frac{1}{n^2}E\left(\sum_{j=1}^{n} X_j^2 + \sum_{i,j=1,i\neq j}^{n} X_iX_j\right) =$$

$$= \frac{1}{n^2}\left(nEX^2 + n\left(n-1\right)\left(EX\right)^2\right)$$

$$EX^2 - E\left(\overline{X}_n\right)^2 = EX^2 - \frac{EX^2}{n} - \frac{n-1}{n}\left(EX\right)^2 = \frac{n-1}{n}\left(EX^2 - \left(EX\right)^2\right) =$$

$$= \frac{n-1}{n}DX \to DX$$

$$E\left(s_n^2\right) = E\left(\frac{n}{n-1}\widehat{\sigma}_n^2\right) = \frac{n}{n-1}E\left(\widehat{\sigma}_n^2\right) = DX$$

□

Věta 1.17. *Buď $\mathbf{X} = (X_1, \ldots, X_n)$ pozorování na prostoru $(\Omega, \mathcal{A}, \mathrm{P})$, přičemž $\mathbf{X} \in \mathcal{L}_r$ pro $r \geq 2$. Potom*

1.

$$m'_r(\mathbf{X}) = \frac{1}{n}\sum_{j=1}^{n} X_j^r$$

je konzistentním odhadem $\mu'_r(X) = E\left(X^r\right)$

2.

$$m_r(\mathbf{X}) = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \overline{X}_n\right)^2$$

je konzistentním odhadem $\mu_r(X) = E\left(E - EX\right)^r$

Definice 1.18 (Výběrový kvantil)**.** *Buďte* $(X_j)_{j=1}^n$ *pozorování* X*, a označme* $\left(X_{(j)}\right)_{j=1}^n$ *tato pozorování seřazená dle velikosti. Potom*

$$\widehat{X}_{p\in(0,1)} = X_{[np]+1}$$

nazýváme ***výběrovým kvantilem****. Specielně pro* $p = 0.5$ *tento kvantil nazýváme* ***výběrovým mediánem****.*

Poznámka 1.19. *Pro výběrový medián zřejmě platí*

$$\widehat{X}_{0.5} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{pro } n \text{ lichá} \\ \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right) & \text{pro } n \text{ sudá} \end{cases}$$

Definice 1.20 (Výběrové rozpětí)**.** *Buď* $(X_j)_{j=1}^n$ *pozorování* X*. Potom* ***výběrové rozpětí*** *definujeme jako*

$$\max_{j\in\widehat{n}} X_j - \min_{j\in\widehat{n}} X_j$$

Definice 1.21 (Empirická distribuční funkce)**.** *Buď* $(X_j)_{j=1}^n$ *pozorování* X*.* ***Empirickou distribuční funci*** *poté definujeme jako*

$$F_n\left(\mathbf{X}, x\right) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{(-\infty,X_j]}(\mathbf{X})$$

kde $\mathbf{1}_{(-\infty,X_j]}(X)$ *je indikátor jevu* $X_j \in (-\infty, X_j]$ *(viz* **??***).*

1.4 Nestranné odhady s minimálním rozptylem - UMVUE

V tomto paragrafu budeme obecně hledat dolní mez střední kvadratické chyby, tj. výrazu

$$DT(\mathbf{X}) = E\left(T(\mathbf{X}) - \tau(\theta)\right)^2$$

Uvažujme dva nestranné odhady $T_n^{(1)}(\mathbf{X})$ a $T_n^{(2)}(\mathbf{X})$. Pokud tyto odhady budeme chtít srovnat, můžeme za „lepší“ považovat například ten s menším rozptylem, tj. ten pro který je menší pravděpodobnost že mi „uletí“ od $ET(\mathbf{X}) = \tau(\theta)$. Pokud totiž dosadíme do Čebyševovy nerovnosti, potom

$$\mathrm{P}\left(|T(\mathbf{X}) - ET(\mathbf{X})| \geq \varepsilon\right) = \mathrm{P}\left(|T(\mathbf{X}) - \tau(\theta))| \geq \varepsilon\right) \leq \frac{DT(\mathbf{X})}{\varepsilon^2}$$

Otázkou však zůstává, jak malého rozptylu lze vůbec u nestranných odhadů dosáhnout.

Definice 1.22 (Regulární systém hustot v $\mathbb{R}^1$)**.** *Buď* $\Theta \subset \mathbb{R}$*. Potom systém hustot*

$$\mathcal{F} = \{f(x,\theta) \mid \theta \in \Theta\}$$

nazveme ***regulárním systémem hustot****, pokud platí*

1. $\mathrm{supp} f = \{x \mid f(x,\theta) \geq 0\}$ *nezávisí na* θ.

2. *Parciální derivace*
$$\frac{\partial f(x,\theta)}{\partial \theta}$$
existuje a je konečná pro všechna θ *a skoro všechna* x.

3. *Střední hodnota*
$$E\left(\frac{\partial \ln f(X,\theta)}{\partial \theta}\right) = 0$$
pro všechna θ.

4. *Fisherova míra informace*
$$\mathcal{I}(\theta) = E\left(\frac{\partial \ln f(X,\theta)}{\partial \theta}\right)^2 > 0$$
pro všechna θ.

Poznámka 1.23.

1. *Podmínka nulovosti střední hodnoty v předchozí větě je obecně splněna právě když lze* $\int f$ *derivovat za integrálem, protože*
$$E\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right) = \int \frac{\partial \ln f(x,\theta)}{\partial \theta} f(x,\theta)dx = \int \frac{f'(x,\theta)}{f(x,\theta)} f(x,\theta)dx = \int f'(x,\theta)dx =$$
$$= \frac{d}{d\theta}\int f(x,\theta)dx = 0$$

2. *Díky předchozímu bodu ale také platí, že*
$$\mathcal{I}(\theta) = D\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right)$$

Věta 1.24. *Buďte* $\left(\widetilde{X}_j\right)_{j=1}^n$ *nezávislé s ASR, a nechť jim odpovídají systémy hustot*
$$\mathcal{F}_j = \left\{f_{X_j}(x_j,\theta) \mid \theta \in \Theta\right\}$$
regulární pro všechna $j \in \widehat{n}$. *Potom platí*
$$\mathcal{I}_{X_1,X_2}(\theta) = \sum_{j=1}^n \mathcal{I}_{X_j}(\theta)$$

Důkaz. Důkaz provedeme matematickou indukcí. Provedeme pouze první krok (pro $n = 2$), druhý krok je zřejmý.
$$f_{X_1,X_2}(x_1,x_2,\theta) = f_{X_1}(x_1,\theta) f_{X_2}(x_2,\theta)$$
$$\mathcal{I}_{X_1,X_2}(\theta) = E\left(\frac{\partial \ln f_{X_1,X_2}(x_1,x_2,\theta)}{\partial \theta}\right)^2 = \iint \left(\frac{\partial \ln f_{X_1,X_2}}{\partial \theta}\right)^2 f_{X_1,X_2} dx_1 dx_2 =$$

$$= \iint \left(\frac{\partial f_{X_1}}{\partial \theta} \right) f_{X_1,X_2} dx_1 dx_2 + 2 \iint \frac{\partial \ln f_{X_1}}{\partial \theta} \frac{\partial \ln f_{X_2}}{\partial \theta} f_{X_1} f_{X_2} dx_1 dx_2 +$$
$$+ \iint \left(\frac{\partial \ln f_{X_2}}{\partial \theta} \right) f_{X_1,X_2} dx_1 dx_2 = \mathcal{I}_{X_1}(\theta) + \mathcal{I}_{X_2}(\theta)$$

□

Důsledek 1.25. *Buďte $(X_j)_{j=1}^n$ pozorování na X (tj. i.i.d). Potom zřejmě*

$$\mathcal{I}_{X_1,\ldots,X_n}(\theta) = n\mathcal{I}_{X_1}(\theta)$$

Věta 1.26. *Buď $\{f(x,\theta) \mid \theta \in \Theta\}$ regulární systém hustot, a nechť $\int f$ lze dvakrát derivovat za integrálem. Potom platí*

$$\mathcal{I}(\theta) = -E\left(\frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} \right)$$

Věta 1.27 (Rao - Cramerova nerovnost)**.** *Buď $\theta \in \Theta \subset \mathbb{R}$, $\mathcal{F} = \{f(\mathbf{x},\theta) \mid \theta \in \Theta\}$ nechť je regulární systém hustot a parametrická funkce $\tau(\theta)$ nechť je diferencovatelná. $T(\mathbf{X})$ nechť je nějaký nestranný odhad $\tau(\theta)$ takový, že $E(T(\mathbf{X}))$ lze derivovat pod znakem E pro $\forall \theta \in \Theta$. Potom*

$$D(T(\mathbf{X})) \geq \frac{[\tau'(\theta)]^2}{\mathcal{I}_{\mathbf{X}}(\theta)} \tag{1}$$

Přitom rovnost nastává právě když existuje $K = K(\theta, n)$ taková, že skoro jistě (s pravděpodobností rovnou jedné) platí

$$\frac{\partial \ln f}{\partial \theta}(\mathbf{X},\theta) = K(T(\mathbf{X}) - \tau(\theta)) \tag{2}$$

Důkaz.

$$\tau'(\theta) = [ET(\mathbf{X})]' = \frac{d}{d\theta} \int T(\mathbf{x}) f(\mathbf{x},\theta) d\mathbf{x} = \int T(\mathbf{x}) \frac{\partial f}{\partial \theta}(\mathbf{x},\theta) d\mathbf{x} =$$
$$= \int T(\mathbf{x}) \frac{\partial \ln f}{\partial \theta}(\mathbf{x},\theta) f(\mathbf{x},\theta) d\mathbf{x} = E\left(T(\mathbf{X}) \frac{\partial \ln f}{\partial \theta} \right) = \mathrm{Cov}\left(T(\mathbf{X}), \frac{\partial \ln f}{\partial \theta}(\mathbf{X},\theta) \right)$$

Ze Schwarzovy nerovnosti potom vyplývá

$$\left| \mathrm{Cov}\left(T, \frac{\partial \ln f}{\partial \theta} \right) \right|^2 \leq DT(\mathbf{X}) D\left(\frac{\partial \ln f}{\partial \theta} \right)$$

a tedy

$$\left[\tau'(\theta)\right]^2 \leq D(T(\mathbf{X}))\, \mathcal{I}(\theta)$$

a rovnost ve Schwarzově nerovnosti nastává právě když platí

$$\frac{\partial \ln f}{\partial \theta}(\mathbf{x},\theta) - E\left(\frac{\partial \ln f}{\partial \theta}(\mathbf{X},\theta) \right) = K(T(\mathbf{X}) - ET(\mathbf{X}))$$

odkud již plyne tvrzení věty, protože

$$E\left(\frac{\partial \ln f}{\partial \theta}(\mathbf{X},\theta) \right) = 0$$

□

Věta 1.28. *Buď $T(\mathbf{X})$ nestranný odhad $\tau(\theta)$. Potom $D\left(T(\mathbf{X})\right) = RCLB(\theta)$ právě když $f_{\mathbf{X}}(\mathbf{x},\theta)$ tvoří jednoparametrickou exponenciální třídu hustot tvaru*

$$f_{\mathbf{X}}(\mathbf{x},\theta) = h(\mathbf{x})c(\theta)\exp\left\{Q(\theta)T(\mathbf{X})\right\}$$

a parametrická funkce $\tau(\theta)$ je tvaru

$$\tau(\theta) = -\frac{1}{c(\theta)}\frac{c'(\theta)}{Q'(\theta)}$$

Definice 1.29. *Buď $\mathbf{\Theta} \subset \mathbb{R}^k$ otevřená množina. Potom říkáme že systém hustot*

$$\mathcal{F} = \left\{ f_X(x,\theta) \mid \theta \in \mathbf{\Theta} \subset \mathbb{R}^k \right\}$$

je regulární, pokud jsou splněny následující podmínky:

1. $\mathrm{supp} f_X$ *nezávisí na volbě θ*
2. *Parciální derivace*
$$\frac{\partial f_X}{\partial \theta_j}$$
existuje pro všechna $j \in \widehat{k}$, pro všechna θ a pro skoro všechna x.
3. *Pro všechna $j \in \widehat{k}$ a pro pro všechna $\theta \in \mathbf{\Theta}$ platí*
$$E\left[\frac{\partial \ln f}{\partial \theta_j}(\mathbf{X},\theta)\right] = 0$$
4. *Fischerovská informační matice, definovaná jako*
$$\mathbb{I}_{i,j}(\theta) = E\left[\frac{\partial \ln f}{\partial \theta_i}, \frac{\partial \ln f}{\partial \theta_j}\right] = \mathrm{Cov}\left[\frac{\partial \ln f}{\partial \theta_i}, \frac{\partial \ln f}{\partial \theta_j}\right] = -E\left[\frac{\partial^2 \ln f}{\partial \theta_i \partial \theta_j}\right]$$
je konečná a regulární

Poznámka 1.30. *Nechť jsou $(X_j)_{j=1}^n$ nezávislé. Potom platí*

$$\mathbb{I}_{\mathbf{X}}(\theta) = \sum_{j=1}^{n} \mathbb{I}_{X_j}(\theta)$$

Věta 1.31 (Rao - Cramerova nerovnost)**.** *Buď $\left\{ f_{\mathbf{X}}(\mathbf{x},\theta) \mid \theta \in \mathbf{\Theta} \subset \mathbb{R}^k \right\}$ regulární systém hustot a nechť $\tau(\theta)$ je taková funkce do $\mathbb{R}$, že parciální derivace*

$$\frac{\partial \tau(\theta)}{\partial \theta_j}$$

existují pro všechna $j \in \widehat{k}$. Dále nechť $T(\mathbf{X})$ je takový nestranný odhad $\tau(\theta)$, že $ET(\mathbf{X})$ lze pro každé $j \in \widehat{k}$ derivovat pod znakem E. Potom

$$D\left(T(\mathbf{X})\right) \geq \tau'(\theta)\mathbb{I}^{-1}\left(\tau'(\theta)\right)^T \tag{3}$$

Věta 1.32 (Bhattacharya)**.** *Buď $\theta \in \Theta \subset \mathbb{R}$, nechť platí stejné podmínky jako v případě Rao - Cramerovy nerovnosti, a navíc nechť pro všechna $j \in \widehat{m}, m \geq 1$ existují parciální derivace*

$$\frac{\partial^j ET(\mathbf{X})}{\partial \theta^j}$$

Potom platí

$$D\left(T(\mathbf{X})\right) \geq \widetilde{\tau}'(\theta)\widetilde{\mathbb{I}}_{\mathbf{X}}(\theta)\left(\widetilde{\tau}'(\theta)\right)^T \quad \textit{pro } \forall\theta \in \Theta \tag{4}$$

kde

$$\widetilde{\tau}'(\theta) = \left(\tau^{(1)}(\theta), \tau^{(2)}(\theta), \ldots, \tau^{(m)}(\theta)\right)$$

$$\widetilde{\mathbb{J}}_{i,j} = E\left[\frac{\frac{\partial^i \ln f_{\mathbf{X}}(\mathbf{x},\theta)}{\partial\theta^i}\frac{\partial^j f_{\mathbf{X}}(\mathbf{x},\theta)}{\partial\theta^j}}{f_{\mathbf{X}}^2}\right]$$

přičemž matice $\widetilde{\mathbb{J}}$ je konečná a nesingulární.

Důkaz. Důkaz je stejný jako v případě $RCLB(\theta),\ \theta \in \boldsymbol{\Theta} \subset \mathbb{R}^k$, pouze matice $\mathbb{D}$ je definována jako

$$\mathbb{D} = \begin{pmatrix} D\left(T(\mathbf{X})\right) & \left(\widetilde{\tau}'\right)^T \\ \widetilde{\tau}' & \widetilde{\mathbb{J}} \end{pmatrix}$$

□

Definice 1.33 (Asymptoticky eficientní odhad)**.** *Buď $T_n(\mathbf{X})$ posloupnost nestranných odhadů $\tau(\theta)$. Potom říkáme že $T_n(\mathbf{X})$ je **asymptoticky eficientní** pokud platí*

$$\frac{RCLB_n(\theta)}{D\left(T_n(\mathbf{X})\right)} \to 1$$

Poznámka 1.34 (Nedostatky UMVUE)**.**

1. *Předpoklady jsou velice restriktivní.*
2. *Existují odhady, které sice nejsou nestranné, ale přitom mají lepší rozptyl než UMVUE.*
3. *Nestranný odhad nemusí existovat.*
4. *Pokud nestranný odhad existuje, nemusí být v praxi použitelný.*

Věta 1.35. *Buďte $(X_j)_{j=1}^n$ i.i.d. náhodné veličiny s rozdělením P^X a nechť je systém hustot*

$$\mathcal{F} = \left\{f_X(x,\theta) \mid \theta \in \mathbb{R}^1\right\}$$

regulární. Nechť je navíc platí

1. *Parciální derivace*

$$\frac{\partial^2 f}{\partial\theta^2}$$

   je spojitá v θ.

2. Existuje taková $M(X)$, $EM(X) < \infty$ pro kterou

$$\left| \frac{\partial^2 \ln f}{\partial \theta^2} \right|$$

3. $T_n(\mathbf{X})$ je nej. AN $\left(\tau(\theta), \frac{\sigma^2(\theta)}{n}\right)$, tj.

$$\sqrt{n}\left(T_n - \tau(\theta)\right) \xrightarrow{\mathcal{D}} N\left(0, \sigma^2(\theta)\right)$$

Potom

$$\sigma^2(\theta) \geq \frac{[\tau'(\theta)]}{\mathcal{I}_{X_1}(\theta)}$$

až na množinu míry nula.

Definice 1.36 (Asymptoticky eficientní odhady)**.** *Nechť $T_n(\mathbf{X})$ je asymptoticky normální odhad AN $\left(\tau(\theta), \frac{\sigma^2(\theta)}{n}\right)$. Říkáme, že $T_n(\mathbf{X})$ je (asymptoticky) eficientní, pokud*

$$\sigma^2(\theta) = \frac{[\tau'(\theta)]^2}{\mathcal{I}_{X_1}(\theta)}$$

Definice 1.37 (Asymptoticky relativní eficience)**.** *Buďte $T_n^{(1)}(\mathbf{X})$ a $T_n^{(2)}(\mathbf{X})$ dva asymptoticky normální odhady s asymptotickými rozptyly $\sigma^2_{(1)}(\theta)$, $\sigma^2_{(2)}(\theta)$. Potom* ***asymptoticky relativní eficienci*** *(ARE) definujeme jako*

$$e_{T^{(1)}T^{(2)}} = \frac{\sigma^2_{(1)}(\theta)}{\sigma^2_{(2)}(\theta)}$$

1.5 Metoda momentů

Uvažujme prostor $(\Omega, \mathcal{A}, \mathrm{P})$, X buď náhodná veličina a $\mathbf{X} = (X_1, \ldots, X_n)$ její pozorování. Nechť $\theta \in \mathbf{\Theta} \subset \mathbb{R}^k$ a $\tau(\theta)$ buď odhadovaná parametrická funkce. Pro $j \in \widehat{k}$ nechť existuje $EX^j = \mu'_j(\theta)$, a označme

$$\mu'(\theta) = \left(\mu'_1(\theta), \mu'_2(\theta), \ldots, \mu'_k(\theta)\right)$$

takže $\mu'(\theta) : \mathbb{R}^k \to \mathbb{R}^k$ a předpokládejme že existuje funkce inverzní, tj. $(\mu'(\theta))^{-1}$.

Definice 1.38 (Momentový odhad)**.** *Označme*

$$m'_j(\mathbf{X}) = \frac{1}{n}\sum_{k=1}^{n} X_k^j$$

Potom ***momentový odhad*** $\widehat{\mathbf{\Theta}}_M(\mathbf{X})$ *parametru θ definujeme jako*

$$\widehat{\mathbf{\Theta}}_M(\mathbf{X}) = \left(\mu'\right)^{-1}\left(m'_1(\mathbf{X}), \ldots, m'_k(\mathbf{X})\right)$$

a momentový odhat $T_M(\mathbf{X})$ parametrické funkce $\tau(\theta)$ definujeme jako

$$T_M(\mathbf{X}) = \tau\left(\widehat{\mathbf{\Theta}}_M(\mathbf{X})\right)$$

Poznámka 1.39.

1. *Momentový odhad $\widehat{\mathbf{\Theta}}_M(\mathbf{X})$ je řešením soustavy k nelineárních rovnic*

$$\mu'_j(\theta) = m'_j(\mathbf{X}) \quad j \in \widehat{k}$$

   (tzv. soustava momentových rovnic)

2. *Odhad $\widehat{\mathbf{\Theta}}_M(\mathbf{X})$ nemusí být určen jednoznačně.*

3. *Místo μ'_j a m'_j lze použít centrální momenty μ_j a m_j.*

Věta 1.40. *Pokud je $(\mu')^{-1}$ funkce spojitá, potom je $\widehat{\mathbf{\Theta}}_M(\mathbf{X})$ konzistentním odhadem parametru θ. Pokud je navíc $\tau(\theta)$ spojitá funkce, potom je $T_M(\mathbf{X}) \overset{\text{def.}}{=} \tau(\widehat{\mathbf{\Theta}}_M)$ jejím konzistentním odhadem.*

Poznámka 1.41. *Odhady získané metodou momentů jsou sice konzistentní, ale nejsou eficientní. Problémy jsou také se splněním předpokladů (existence momentů a spojitost).*

1.6 Metoda maximální věrohodnosti

Definice 1.42 (Věrohodnostní funkce)**.** *Buďte $\mathbf{X} = (X_1, \ldots, X_n)$ nezávislá pozorování X, tj. X_k, $k \in \widehat{n}$ i.i.d. s rozdělením P^X. Potom libovolnou funkci tvaru*

$$L(\theta|\mathbf{x}) = c(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}, \theta)$$

*nazýváme **věrohodnostní funkcí**, a funkci tvaru*

$$l(\theta|\mathbf{x}) = \ln L(\theta|\mathbf{x})$$

*nazýváme **logaritmickou věrohodnostní funkcí**.*

Definice 1.43 (Maximálně věrohodný odhad - MLE)**.** *Buď $\widehat{\mathbf{\Theta}}_{ML}(\mathbf{X})$ taková borelovsky měřitelná funkce na Ω, že platí*

$$L\left(\widehat{\mathbf{\Theta}}_{ML}(\mathbf{X})|\mathbf{X}\right) = \sup_{\theta \in \mathbf{\Theta}} L(\theta, \mathbf{X})$$

*Pokud $\widehat{\mathbf{\Theta}}_{ML}(\mathbf{X})$ závisí na $\mathbf{X}$ a pokud je určena jednoznačně, potom je nazývána **maximálně věrohodným odhadem** parametru θ a $T_{ML}(\mathbf{X}) = \tau\left(\widehat{\mathbf{\Theta}}_{ML}(\mathbf{X})\right)$ nazýváme **maximálně věrohodným odhadem** parametrické funkce $\tau(\theta)$.*

Lemma 1.44 (Jensenova nerovnost)**.** *Buď X náhodná veličina na prostoru $(\Omega, \mathcal{A}, \mathrm{P})$, a nechť $X \in \mathcal{L}_1$. Dále nechť $\Phi(t)$ je konvexní (resp. konkávní) funkce. Potom*

$$\Phi(EX) \leq E\Phi(X)$$

$$\textit{resp.} \quad \Phi(EX) \geq E\Phi(X)$$

Věta 1.45. *Buď $X \sim f(x, \theta)$, $\theta \in \mathbf{\Theta} \subset \mathbb{R}^k$, nechť $\mathrm{supp} f$ nezávisí na θ a $E|\ln f| < \infty$. Buďte $\mathbf{X} = (X_1, \ldots, X_n)$ pozorovánína X. Potom pro všechna $\theta \neq \theta_0$ platí*

$$\lim_{n \to \infty} \mathrm{P}\left(L\left(\theta_0|\mathbf{x}\right) > L\left(\theta|\mathbf{x}\right)\right) = 1$$

kde θ_0 je skutečná hodnota parametru, a $\theta \neq \theta_0$. je libovolný bod z $\mathbf{\Theta}$.

Důkaz.

$$\{L(\theta_0|\mathbf{X}) > L(\theta|\mathbf{X})\} = \left\{\frac{L(\theta|\mathbf{X})}{L(\theta_0|\mathbf{X})} < 1\right\} = \left\{\prod_{j=1}^{n} \frac{f_{X_j}(x_j,\theta)}{f_{X_j}(x_j,\theta_0)} < 1\right\} =$$

$$= \left\{\frac{1}{n}\sum_{j=1}^{n} \ln\left(\frac{f_{X_j}(x_j,\theta)}{f_{X_j}(x_j,\theta_0)}\right) < 0\right\}$$

přitom platí

$$\frac{1}{n}\sum_{j=1}^{n} \ln\left(\frac{f_{X_j}(x_j,\theta)}{f_{X_j}(x_j,\theta_0)}\right) \xrightarrow{\mathrm{P}} E\left(\ln\frac{f_{X_j}(X_j,\theta)}{f_{X_j}(X_j,\theta_0)}\right)$$

a tedy

$$E_{\theta_0}\left(\ln\frac{f_{X_j}(X_j,\theta)}{f_{X_j}(X_j,\theta_0)}\right) < \ln E\left(\frac{f_{X_j}(X_j,\theta)}{f_{X_j}(X_j,\theta_0)}\right) = \ln\int_{\mathbb{R}} \frac{f_{X_j}(x_j,\theta)}{f_{X_j}(x_j,\theta_0)} f_{X_j}(x_j,\theta_0)dx = \ln 1 = 0$$

□

Poznámka 1.46. *Pokud je* $\boldsymbol{\Theta} \subset \mathbb{R}^k$ *otevřená množina,* supp$f$ *nezávisí na* θ *a parciální derivace* $\frac{\partial L}{\partial \theta_j}$ *existují pro všechna* $j \in \widehat{k}$*, potom je odhad* $\widetilde{\boldsymbol{\Theta}}_{ML}(\mathbf{X})$ *řešením soustavy věrohodnostních rovnic*

$$\frac{\partial L(\theta|\mathbf{x})}{\partial \theta_j} = 0 \qquad j \in \widehat{k}$$

Věta 1.47. *Buďte* $X_1, \ldots, X_n$ *i.i.d. náhodné veličiny s hustotou* $f(\mathbf{x},\theta)$*,* $\theta \in \Theta \subset \mathbb{R}$*, přičemž* Θ *je otevřená množina a nechť* θ_0 *je skutečná hodnota parametru. Dále nechť existuje* $\delta' > 0$ *takové, že na intervalu* $(\theta_0 - \delta', \theta_0 + \delta')$ *existuje parciální derivace*

$$\frac{\partial l(\theta,\mathbf{x})}{\partial \theta}$$

Potom s pravděpodobností jdoucí k jedné (při $n \to \infty$*) existuje takové řešení věrohodnostních rovnice, které je konsistentním odhadem* θ_0*.*

Důkaz. Volme $\delta < \delta'$ a označme $l_n = \ln L_n$. Potom

$$\mathrm{P}\left(l_n(\theta_0) - l_n(\theta_0 - \delta) > 0\right) \to 1$$

$$\mathrm{P}\left(l_n(\theta_0) - l_n(\theta_0 + \delta) > 0\right) \to 1$$

takže pro $\forall \delta < \delta'$ musí (spojitá) funkce l_n nabývat maxima. Takže věrohodnostní rovnice

$$\frac{\partial l}{\partial \theta} = 0$$

má na intervalu $(\theta_0 - \delta', \theta_0 + \delta')$ řešení s pravděpodobností jdoucí k 1. A to jsme chtěli dokázat. □

Věta 1.48. *Buďte $X_1, \ldots, X_n$ i.i.d. náhodné veličiny s hustotou pravděpodobnosti $f_X(x,\theta)$, $\theta \in \Theta \subset \mathbb{R}$ kde Θ je otevřená množina. Buď*

$$\mathcal{F} = \{f_X(x,\theta) \ : \ \theta \in \Theta\}$$

regulární systém hustot a nechť platí

$$\left|\frac{\partial^3 \ln f_X}{\partial \theta^3}(x,\theta)\right| \leq M(x)$$

kde $EM(X) < \infty$. Potom pro každé konzistentní řešení $\widehat{\theta}_{ML}(\mathbf{X})$ věrohodnostní rovnice platí

$$\sqrt{n}\left(\widehat{\theta}_{ML}(\mathbf{X}) - \theta_0\right) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{\mathcal{I}_X(\theta_0)}\right)$$

tj. $\widehat{\theta}_{ML}(\mathbf{X}) \sim AN\left(\theta_0, \frac{1}{n\mathcal{I}_X(\theta_0)}\right)$.

Důkaz. Odhad $\widehat{\theta}_{ML}(\mathbf{X})$ řeší věrohodnostní rovnici

$$\frac{\partial l}{\partial \theta}(\theta, x) = 0$$

takže $l'\left(\widehat{\theta}_{ML}(\mathbf{X})\right) = 0$. Provedeme Taylorův rozvoj v bodě θ_0, tj.

$$0 = l'\left(\widehat{\theta}_{ML}(\mathbf{X})\right) = l'_n(\theta_0) + l''_n(\theta_0)\left(\left(\widehat{\theta}_{ML}(\mathbf{X})\right) - \theta_0\right) + \frac{l'''_n(\theta_n^*)}{2}\left(\left(\widehat{\theta}_{ML}(\mathbf{X})\right) - \theta_0\right)^2 + \cdots$$

přičemž $\theta_n^* \in (\theta_0, \widehat{\theta}_n)$ nebo $\theta_n^* \in (\widehat{\theta}_n, \theta_0)$. Potom

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) = -\frac{\sqrt{n}\ln'(\theta_0)}{\left[\ln''(\theta_0) + \frac{\ln'''(\widehat{\theta}_n^*)}{2n}(\widehat{\theta}_n - \theta_0)\right]}$$

přičemž z konzistence řešení $\widehat{\theta}_n$ vyplývá

$$\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{\mathrm{P}} 0$$

Ukažme nyní omezenost vztahu

$$\frac{\ln'''(\widehat{\theta}_n^*)}{n} = \frac{1}{n}\sum_1^n \left|l'''_1\left(\widehat{\theta}_n^*\right)\right|$$

a protože dle předpokladu je

$$\left|\frac{\partial^3 \ln f_X}{\partial \theta^3}(x,\theta)\right| leq M(x)$$

potom dle zákona velkých čísel (Kolmogorov) platí

$$\frac{\ln'''\left(\widehat{\theta}_n^*\right)}{n} \leq \frac{1}{n}\sum_1^n M_n \to EM(X) < \infty$$

Odtud však vyplývá, že

$$\mathrm{P}\left(\left|\frac{1}{n}\sum_{1}^{n} M_n(X)\right| \leq K\right) \to 1$$

a celkem tedy

$$\frac{\ln'''(\widehat{\theta}_n^*)}{2n}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathrm{P}} 0$$

Podívejme se nyní na další člen

$$\frac{1}{n}\ln''(\theta_0) = \frac{1}{n}\sum_{1}^{n} l''_1(\theta_0) \overset{s.j.}{\to} E\left(l''_1(\theta_0)\right) = E\left[\frac{\partial^2 f_x}{\partial^2}(\theta_0)\right] = -\mathcal{I}(\theta_0)$$

a pro další člen platí

$$\frac{1}{\sqrt{n}}\ln'(\theta_0) = \sqrt{n}\left(\frac{1}{n}\ln'(\theta_0)\right) = \sqrt{n}\left(\frac{1}{n}\sum_{1}^{n} l'_1(\theta_0) - El'_1(\theta_0)\right) \overset{CLT}{\to} N\left(0, D\left(l'_1(\theta_0)\right)\right)$$

Ale současně

$$D\left(l'_1(\theta_0)\right) = E\left(\frac{\partial \log f_X(\theta_0)}{\partial\theta}\right) = \mathcal{I}(\theta_0)$$

přičemž dle Slutskyho lemmatu platí

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{\mathcal{D}} \frac{1}{\mathcal{I}(\theta_0)}Y$$

protože $Y \sim N\left(0, \mathcal{I}(\theta_0)\right)$.

□