TEORIE ODHADU

Vlastnosti „dobrého“ bodového odhadu

Dobrý“ (věrohodný) odhad musí splňovat určité vlastnosti. Mezi základní vlastnosti věrohodných odhadů patří:

- nestrannost (nevychýlenost, nezkreslenost)
- efektivita (vydatnost,eficience), nestranný odhad, jehož rozptyl je nejmenší mezi všemi nestrannými odhady příslušného parametru, se nazývá **nejlepší nestranný odhad**
- konzistence

Postačující statistika pro parametr Θ

Důkaz toho, zda je určitý odhad efektivní (nejlepší nestranný), není vždy jednoduchý. Abychom našli odhad, který má nejmenší rozptyl, je vhodné nahrazení celého výběru jednou statistikou, a to takovou, která bude obsahovat „veškerou“ informaci o parametru Θ.

Pokud je možné pomocí nějaké statistiky (může se jednat o vícerozměrnou statistiku) odhadnout neznámé parametry souvisejícího rozdělení, hovoříme o **postačující statistice**. Nejjednodušší postačující statistikou je podle definice samotný vektor náhodného výběru $\underline{X} = (X_1, \dots, X_n)$, taková postačující statistika však není příliš užitečná. Smysl má hledat takové postačující statistiky, které mají rozměr menší než n.

Definice:

Reálnou funkci $T_n(\underline{X})$ nazveme **postačující statistikou pro parametr θ**, jestliže sdružené rozdělení náhodného výběru $\underline{X} = (X_1, \dots, X_n)$ podmíněné jevem $T(\underline{X})=t$ není pro žádné t závislé na Θ.

Jestliže pro parametrickou funkci $\tau(\Theta)$ existují nestranné odhady, pak nejlepší z nich (ve smyslu minimálního rozptylu) je funkcí postačujících statistik a je určen jednoznačně.

Jednoduchý postup při hledání postačujících statistik nabízí věta o faktorizaci. Tato věta zároveň umožňuje rychle rozhodnout o tom, zda je určitá statistika postačující.

Věta o faktorizaci:

Nechť $\underline{X} = (X_1, \ldots, X_n)$ je náhodný výběr z rozdělení $f(x;\Theta)$. $T(X)$ je postačující statistikou pro parametr Θ právě tehdy, jestliže sdružené rozdělení náhodného výběru je součinem dvou faktorů:

$$f(\underline{x},\Theta) = g\{T(\underline{x}),\Theta\}\cdot h(\underline{x})$$

DK:

$$T(\underline{X}) \text{ je postačující pro } \theta \iff f(\underline{x},\theta) = g(T(\underline{x}),\theta)\cdot h(\underline{x})$$

(dostatečnost):

$$\Rightarrow \quad P(\underline{X}=\underline{x}) = P(\underline{X}=\underline{x} \mid T(\underline{X})=t)\cdot P(T(\underline{X})=t) = h(\underline{x})\cdot g(t,\theta) \qquad \text{c.b.d}$$

= předpokládám, že $P(T(\underline{X})=t) > 0$

$$P(\underline{X}=\underline{x} \mid T(\underline{X})=t) = \frac{P(\underline{X}=\underline{x} \cap T(\underline{X})=t)}{P(T(\underline{X})=t)} = <$$

1) vezmu takové $\underline{x}$, že $T(\underline{x})=t$

2) vezmu takové $\underline{x}$, že $T(\underline{x})\neq t$

$$1) \quad = \frac{P(\underline{X}=\underline{x})}{P(T(\underline{X})=t)} = \frac{h(\underline{x})\cdot g(t,\theta)}{\sum\limits_{\underline{x}:T(\underline{x})=t} P(\underline{X}=\underline{x})} = \frac{h(\underline{x})\cdot \cancel{g(t,\theta)}}{\cancel{g(t,\theta)}\cdot\sum\limits_{\underline{x}:T(\underline{x})=t} h(\underline{x})} = \frac{h(\underline{x})}{\sum\limits_{\underline{x}:T(\underline{x})=t} h(\underline{x})}$$

$$2) \quad = \frac{P(\underline{X}=\underline{x} \cap T(\underline{X})=t)}{P(T(\underline{X})=t)} = 0$$

neboť $T(\underline{x})=t$ je nemožný jev : $P(A \cap \{\emptyset\}) = 0$

jmenovatel > 0 (předp

Řešený příklad

Nechť $\underline{X} = (X_1, \ldots, X_n)$ je náhodný výběr z Poissonova rozdělení. Dokažme, že $T(X) = \sum_{i=1}^{n} X_i$ je postačující statistikou pro parametr Θ Poissonova rozdělení $(\Theta = \lambda t)$.

$$f(x,\Theta) = \frac{\Theta^x}{x!} \cdot e^{-\Theta} \qquad x = 0,1,2,\ldots, \Theta > 0$$

Sdružené rozdělení výběru má tvar:

$$f(\underline{x},\Theta) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i) = \prod_{i=1}^{n} \frac{\Theta^{x_i}}{x_i!} \cdot e^{-\Theta} =$$

$$= e^{-n\cdot\Theta} \cdot \frac{\Theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!},$$

kde $x_i = 0,1,\ldots,\quad i = 1,2,\ldots,$

Sdružené rozdělení výběru můžeme faktorizovat, tj. můžeme jej zapsat jako součin dvou faktorů:

$$f\left(\underline{x},\Theta\right) = g\left(t,\Theta\right)\cdot h\left(x\right)$$

$$f\left(\underline{x},\Theta\right) = e^{-n\cdot\Theta}\cdot\Theta^{t}\cdot\frac{1}{\prod_{i=1}^{n} x_i!},$$

kde $g\left(t,\Theta\right) = e^{-n\cdot\Theta}\cdot\Theta^{t};\qquad h\left(x\right) = \frac{1}{\prod_{i=1}^{n} x_i!}$

$T(X) = \sum_{i=1}^{n} X_i$ *je tedy postačující statistikou pro parametr Θ.*

Sdružená pravděpodobnostní funkce u některých rozdělení:

Poissonovo rozdělení : $P(x;\lambda) = exp(ln\,\lambda \sum x_i - n\lambda - \sum ln(x_i!))$

a postačující statistikou pro parametr λ je výběrový úhrn $\sum x_i$.

Exponenciálního rozdělení je $f(\mathbf{x};\delta) = \exp\left[-n\ln\delta - \frac{1}{\delta}\sum x_i\right]$

a opět postačující statistikou pro parametr δ je výběrový úhrn $\sum x_i$

Normálního rozdělení $N(0,\sigma^2)$, které má hustotu

$$f(x_i;0;\sigma^2) = \exp\left(-\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \frac{x_i^2}{2\sigma^2}\right)$$

a sdružená hustota $f(\mathrm{x};0;\sigma^2) = \exp\left(-\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right)$

postačující statistikou pro parametr σ^2 je tedy $\sum x_i^2$

Konstrukce efektivních odhadů

Praktický význam postačujících statistik pro výpočet efektivních (nejlepších nestranných) odhadů ukazuje následující věta:

Rao-Blackwellova věta

Nechť $\underline{X} = (X_1, \ldots, X_n)$ je náhodný výběr z rozdělení $f(\underline{x};\Theta)$. Nechť existuje postačující statistika $T(\underline{X})$ pro parametr Θ. Nechť $\tau(\Theta)$ je reálná funkce parametru Θ a $T^*(\underline{X})$ je nestranný odhad této charakteristiky. Potom platí:

1. Pro funkci $\tau(\Theta)$ existuje nestranný odhad $\tilde{T}(X) = \tilde{T}(T(X))$, který je funkcí postačující statistiky $T(\underline{X})$.

2. Nestranný odhad $\tilde{T}(\underline{X})$ má rozptyl menší nebo roven rozptylu odhadu $T^*(\underline{X})$:

$$D\left(\tilde{T}(\underline{X})\right) \leq D\left(T^*(\underline{X})\right) \quad \textit{pro všechna } \Theta$$

3. $D\left(\tilde{T}(\underline{X})\right) = D\left(T^*(\underline{X})\right) \quad \Leftrightarrow \quad P\left(\tilde{T}(\underline{X}) = T^*(\underline{X})\right) = 1 \quad \textit{pro všechna } \Theta$

Pomocné tvrzení

Je-li *(X, Y)* spojitý náhodný vektor se sdruženou hustotou f(x,y), definujeme podmíněnou střední hodnotu takto:

$$E\left(X \middle| Y = y\right) = \int_{-\infty}^{\infty} x \cdot f\left(x \middle| y\right) dx\ ,$$

kde $f\left(x \middle| y\right) = \dfrac{f(x,y)}{f_Y(y)} = \dfrac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dx}$

Důležitou vlastností podmíněné střední hodnoty je, že:

$$E_Y\left[E\left(X \middle| Y\right)\right] = \int_{-\infty}^{\infty} E\left(X \middle| Y\right) \cdot f_Y(y)\,dy = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} x \cdot f\left(x \middle| y\right) dx\,dy =$$

$$= \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} x \cdot \frac{f(x,y)}{f_Y(y)}\,dx\,dy = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx = E(X),$$

kde E_Y je střední hodnota vzhledem k náhodné veličině Y.

Důkaz Rao-Blackwellovy věty:

ad1)

Nechť $T^*(\underline{X})$ je libovolný nestranný odhad parametrické funkce $\tau(\Theta)$ a $T(\underline{X})$ je postačující statistika pro parametr θ.

Položme:

$$\tilde{T}(t) = E\{T^*(X) | T(X) = t\}$$

Protože $T(\underline{X})$ je postačující statistikou, funkce $\tilde{T}(t)$ není funkcí θ, tedy $\tilde{T}(T(\underline{X}))$ je statistika.

Dokážeme, že $\tilde{T}(T(\underline{X}))$ je nestranný odhad parametrické funkce $\tau(\Theta)$.

Pro každé Θ platí:

$$E_T(\tilde{T}) = E_T\{E[T^*(\underline{X}) | T(\underline{X}) = t]\} = E(T^*(\underline{X})) = \tau(\Theta)$$

c.b.d.

(viz pomocné tvrzení)

ad2)

$$D_T\left(T^*\left(\underline{X}\right)\right)=\mathrm{E}_{\mathrm{T}}\left\{\left[T^*\left(\underline{X}\right)-E\left(T^*\left(\underline{X}\right)\right)\right]^2\right\}=\mathrm{E}_{\mathrm{T}}\left\{\left[T^*\left(\underline{X}\right)-\tau\left(\Theta\right)\right]^2\right\}=$$

$$=E_T\left\{\left[\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]+\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]\right]^2\right\}=$$

$$=E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]^2\right\}+2E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]\cdot\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]\right\}+E_T\left\{\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]^2\right\}$$

Střední hodnotu součinu můžeme vyjádřit jako:

$$\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]\cdot\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]$$

$$E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]\cdot\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]\right\}=E_T\left\{E\left[\left(T^*\left(\underline{X}\right)-\tilde{T}(t)\right).\left(\tilde{T}(t)-\tau\left(\Theta\right)\right)\ \middle|\, t\right]\right\}=$$

$$=E_T\left\{\left(\tilde{T}(t)-\tau\left(\Theta\right)\right)\ .\ E\left[\left(T^*\left(\underline{X}\right)-\tilde{T}(t)\right)\ \middle|\, t\right]\right\}=$$

$$=E_T\left\{\left(\tilde{T}(t)-\tau\left(\Theta\right)\right)\cdot 0\right\}=0$$

protože $E\left\{T^*\left(X\right)\middle|\,T\left(X\right)=t\right\}=\tilde{T}\left(t\right)$

Tedy:

$$D_T\left(T^*\left(\underline{X}\right)\right)=E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]^2\right\}+2E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]\cdot\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]\right\}+E_T\left\{\left[\tilde{T}(t)-\tau\left(\Theta\right)\right]^2\right\}=$$

$$=E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]^2\right\}+0+D_T\left(\tilde{T}(t)\right)$$

$$E_T\left\{\left[T^*\left(\underline{X}\right)-\tilde{T}(t)\right]^2\right\}\geq 0\qquad\Rightarrow\qquad D_T\left(T^*\left(\underline{X}\right)\right)\geq D_T\left(\tilde{T}(t)\right)\quad \text{c.b.d.}$$

ad3)

$$D_T\left(T^*(X)\right) \geq D_T\left(\tilde{T}(t)\right) \quad \Leftrightarrow \quad E_T\left\{\left[T^*(\underline{X}) - \tilde{T}(t)\right]^2\right\} = 0 \quad \Leftrightarrow \quad P\left\{T^*(\underline{X}) = \tilde{T}(t)\right\} = 1$$

Z Rao-Blackwellovy věty vyplývá, že při hledání nejlepších nestranných odhadů se můžeme omezit na odhady, které jsou funkcemi postačujících statistik. Tato věta nám ukazuje, jak v případě, že známe libovolný nestranný odhad, určit nestranný odhad, který je funkcí postačující statistiky.

Řešený příklad

Nechť $\underline{X}$ = (X_1, …, X_n) je náhodný výběr z Poissonova rozdělení:

$$f(x,\Theta) = \frac{\Theta^x}{x!} \cdot e^{-\Theta} \qquad x = 0,1,2,\ldots,\ \Theta > 0$$

Nalezněte nejlepší nestranný odhad parametrické funkce: $\tau(\Theta) = e^{-\Theta}$

tj. pravděpodobnosti toho, že náhodná veličina X s Poissonovým rozdělením nabude hodnoty 0.

Hledáme nejlepší nestranný odhad parametrické funkce: $\tau(\Theta) = e^{-\Theta}$

Vzhledem k tomu, že $\tau(\Theta)$ *je pravděpodobnost toho, že náhodná veličina X s Poissonovým rozdělením nabude hodnoty 0, nabízí se jako vhodný nestranný odhad funkce* $\tau(\Theta)$ *relativní četnost nulových hodnot* X_i *ve výběru, tj.*

$$T^*(\underline{X}) = T(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n} Y_i ,$$

kde:

$$Y_i = 0 \qquad pro \quad X_i \geq 1$$
$$Y_i = 1 \qquad pro \quad X_i = 0, \qquad i = 1,2,\ldots,n$$

Už víme, že pro parametr Θ je $T(\underline{X}) = \sum_{i=1}^{n} X_i$ *postačující statistikou.*

Nejlepší nestranný (efektivní) odhad $\tilde{T}(t)$ ***funkce*** $\tau(\Theta)$ *budeme hledat následujícím způsobem:*

1. *Najdeme střední hodnotu $T^*(\underline{X})$ podmíněnou jevem* $T(\underline{X}) = \sum_{i=1}^{n} X_i = t$

$$\tilde{T}(t) = E\left\{T^*(\underline{X}) \middle| T(\underline{X}) = t\right\} = E\left\{\frac{1}{n}\sum_{i=1}^{n} Y_i \middle| \sum_{i=1}^{n} X_i = t\right\} = \frac{1}{n}\sum_{i=1}^{n} P\left\{Y_i = 1 \middle| \sum_{i=1}^{n} X_i = t\right\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n} P\left\{X_i = 0 \middle| \sum_{i=1}^{n} X_i = t\right\} = P\left\{X_1 = 0 \middle| \sum_{i=1}^{n} X_i = t\right\}$$

2. *Postačující statistiku* $T(\underline{X}) = \sum_{i=1}^{n} X_i$ *můžeme zapsat ve tvaru:*

$$T(\underline{X}) = X_1 + \sum_{i=2}^{n} X_i = X_1 + Z \quad ,$$

kde X_1 a Z jsou nezávislé náhodné veličiny s Poissonovým rozdělením:

$$E_\Theta(X_1) = \Theta \qquad E_\Theta(Z) = (n-1)\cdot\Theta$$

Pak:

$$\tilde{T}(t) = \frac{P_\Theta(X_1 = 0) \cdot P_\Theta(Z = t)}{P_\Theta(X_1 + Z = t)} = \frac{e^{-\theta-\theta(n-1)} \cdot \dfrac{[\theta \cdot (n-1)]^t}{t\,!}}{e^{-n\theta} \cdot \dfrac{(n\theta)^t}{t\,!}} = \left(\frac{n-1}{n}\right)^t$$

Efektivním odhadem parametrické funkce $\tau(\Theta) = e^{-\Theta}$ *je odhad* $\tilde{T}(t) = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^{n} x_i}$.

Fisherova míra informace

Důležitým ukazatelem kvality odhadu je jeho rozptyl (připomeňme, že každý odhad je náhodnou veličinou). Vyhovuje-li rozdělení, jehož parametr odhadujeme, jistým obecným předpokladům, lze ukázat, že není možné zkonstruovat odhad s rozptylem menším než jistá hodnota, tzv. Raova-Cramerova hranice. Mezi odhady s požadovanou vlastností se tedy vždy snažíme nalézt odhad, jehož rozptyl je roven této hodnotě. Pokud se to podaří, hovoříme o **odhadu s minimálním rozptylem**. Ve statistické literatuře se často operuje **s nestranným odhadem s minimálním rozptylem**.

K libovolnému rozdělení $f\left(\underline{x}, \Theta\right)$ a k libovolné parametrické funkci τ(Θ) budeme hledat takovou funkci C(Θ), aby libovolný nestranný odhad $\tilde{T}(\underline{X})$, který splňuje podmínky regularity, měl rozptyl větší než C(Θ). Funkce C(Θ) tedy bude dolní mezí rozptylů pro všechny nestranné odhady parametrické funkce τ(Θ).

Existují nestranné odhady, jejichž rozptyl je roven C(Θ). V některých případech je však tato hranice dosažitelná pouze asymptoticky (pro $n \to \infty$).

Dříve než se pustíme do hledání funkce C(Θ), definujeme si některé pojmy.

Definice:

Předpokládejme, že Θ je jednorozměrný parametr. Říkáme, že **systém hustot**

$$\{f(x,\Theta),\ \Theta \in \mathbf{\Theta}\}$$

je regulární právě když:

1. $\mathbf{\Theta}$ je neprázdná otevřená množina
2. Množina M={x: $f(x,\Theta)>0$} není funkcí Θ
3. Pro každé $\underline{x} \in$ M existuje konečná parciální derivace:

$$f'(x,\Theta) = \frac{\partial f(x,\Theta)}{\partial \Theta}$$

4. Pro každé $\Theta \in$ **$\mathbf{\Theta}$,**

$$Z(X,\Theta) = \frac{\partial \ln f(X,\Theta)}{\partial \Theta} \quad \text{platí:} \qquad EZ = 0$$

$$\int_M \frac{\partial \ln f(x,\Theta)}{\partial \Theta} \cdot f(x,\Theta)\, dx = \int_M \frac{f'(x,\Theta)}{f(x,\Theta)} f(x,\Theta)\, dx = \int_M f'(x,\Theta)\, dx = 0$$

5. $0 < \mathrm{DZ} < \infty$ (konečný, kladný rozptyl)

$$DZ = I(\Theta) = \int_M \left(\frac{\partial \, ln\, f(x,\Theta)}{\partial\Theta} - EZ \right)^2 f(x,\Theta)\, dx = \int_M \left(\frac{\partial \, ln\, f(x,\Theta)}{\partial\Theta} \right)^2 f(x,\Theta)\, dx$$

$$= \int_M \left(\frac{f'(x,\Theta)}{f(x,\Theta)} \right)^2 f(x,\Theta)\, dx = \int_M \left(\frac{f'(x,\Theta)}{f(x,\Theta)} \right)^2 f(x,\Theta)\, dx$$

Zjednodušeně často místo o regulárnosti systému hustot mluvíme o **regulárnosti rozdělení *f(xΘ)*.**

Řešený příklad

Dokažte, že hustota normálního rozdělení $N(\Theta,1)$ je regulární.

Hustota normálního rozdělení $N(\Theta,1)$*:*

$$f(x,\Theta)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(x-\Theta)^2}{2}} \qquad \textit{pro } \Theta\in R$$

ad1) $\Theta\in R$, ***Θ*** *je neprázdná otevřená množina*
*ad2) M=R, Množina M={**x**: f(**x**,Θ)>0} není funkcí Θ*

ad3) $\forall x\in R: f'(x,\Theta)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{(x-\Theta)^2}{2}}\cdot(x-\Theta) \qquad \textit{pro } \Theta\in R$

ad4) Pro každé Θ∈ ***Θ,*** $Z(X,\Theta)=\frac{\partial \ln f(X,\Theta)}{\partial\Theta}$ *platí:* $EZ=0$

$$\int_{-\infty}^{\infty} f'(x,\Theta)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\Theta)^2}{2}} \cdot (x-\Theta)dx = \left| \begin{array}{l} t = x-\Theta \\ dt = dx \end{array} \right| = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \cdot t\,dt =$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \left(\int_{-\infty}^{0} e^{-\frac{t^2}{2}} \cdot t\,dt + \int_{0}^{\infty} e^{-\frac{t^2}{2}} \cdot t\,dt \right) = \frac{1}{\sqrt{2\pi}} \cdot (I_1 + I_2) = \frac{1}{\sqrt{2\pi}} \cdot (-1+1)$$

$$= 0$$

kde:

$$I_1 = \int_{-\infty}^{0} e^{-\frac{t^2}{2}} \cdot t\,dt = \left| \begin{array}{l} u = -\frac{t^2}{2} \\ du = -t\,dt \end{array} \right| = -\int_{-\infty}^{0} e^{u} du = -\lim_{v \to (-\infty)} \left[e^{u} \right]_{v}^{0} = -(1-0) = -1$$

$$I_2 = \int_{0}^{\infty} e^{-\frac{t^2}{2}} \cdot t\,dt = \left| \begin{array}{l} u = \frac{t^2}{2} \\ du = t\,dt \end{array} \right| = \int_{0}^{\infty} e^{-u} du = -\lim_{v \to (\infty)} \left[e^{-u} \right]_{0}^{v} = -(0-1) = 1$$

ad5) Pro každé $\Theta \in$ $\boldsymbol{\Theta}$**,** $Z(X,\Theta) = \frac{\partial \ln f(X,\Theta)}{\partial \Theta}$ *platí:* $(0 < \mathrm{DZ} < \infty)$

$$DZ = I(\Theta) = \int_M \left(\frac{f'(x,\Theta)}{f(x,\Theta)} \right)^2 f(x,\Theta)dx = \int_{-\infty}^{\infty} (x-\Theta)^2 \cdot f(x,\Theta)dx =$$

$$= \int_{-\infty}^{\infty} (x-EX)^2 \cdot f(x,\Theta)dx = DX = 1$$

Hustota normálního rozdělení $N(\Theta,1)$ *je tedy regulární.*

Definice:
Nechť náhodná veličina *X* má regulární rozdělení *f(x,Θ)*. Integrál I(Θ) definovaný v podmínce 5 v definici regulárního systému hustot nazýváme **Fisherovou mírou informace**:

$$I(\Theta) = \int_M \left(\frac{\partial \ln f(x,\Theta)}{\partial \Theta} \right)^2 f(x,\Theta)\, dx = \int_M \left(\frac{f'(x,\Theta)}{f(x,\Theta)} \right)^2 f(x,\Theta)\, dx$$

Fischerova míra informace je tedy střední hodnota náhodné veličiny definované jako:

$$\left(\frac{f'(X,\Theta)}{f(X,\Theta)} \right)^2$$

Můžeme ji tedy zapisovat také jako:

$$I(\Theta) = E\left\{ \left(\frac{f'(X,\Theta)}{f(X,\Theta)} \right)^2 \right\} = E\left\{ \left(\frac{\partial \ln f(X,\Theta)}{\partial \Theta} \right)^2 \right\}$$

Následující vlastnost Fisherovy míry informace je využitelná především při výpočtu informace v praktických příkladech.

Věta:

Nechť f(x,Θ) je regulární hustota. Nechť pro všechna x a pro všechna Θ existuje druhá derivace f(x,Θ):

$$f''(x,\Theta) = \frac{\partial^2 f(x,\Theta)}{\partial \Theta^2}$$

Nechť pro každé $\Theta \in \boldsymbol{\Theta}$ platí:

$$\int_M f''(x,\Theta)\, dx = 0$$

Potom:

$$I(\Theta) = -\int_M \frac{\partial^2 \ln f(x,\Theta)}{\partial \Theta^2} f(x,\Theta)\, dx$$

Dk:

$$\frac{\partial \ln f(x,\Theta)}{\partial \Theta} = \frac{f'(x,\Theta)}{f(x,\Theta)}$$

$$\frac{\partial^2 \ln f(x,\Theta)}{\partial \Theta^2} = \frac{f''(x,\Theta) f(x,\Theta) - f'(x,\Theta) . f'(x,\Theta)}{f^2(x,\Theta)} =$$

$$= \frac{f''(x,\Theta)}{f(x,\Theta)} - \left(\frac{f'(x,\Theta)}{f(x,\Theta)}\right)^2$$

$$\left(\frac{f'(x,\Theta)}{f(x,\Theta)}\right)^2 = \frac{f''(x,\Theta)}{f(x,\Theta)} - \frac{\partial^2 \ln f(x,\Theta)}{\partial \Theta^2}$$

Řešený příklad

Nechť náhodná veličina *X* má normální rozdělení $N\left(\mu,\sigma^2\right)$, kde σ^2 známe. Určete míru informace o parametru μ.

$$f\left(x,\mu\right)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{\left(x-\mu\right)^2}{2\sigma^2}} \qquad pro\ x\in R$$

$$\ln f\left(x,\mu\right)=\ln\frac{1}{\sqrt{2\pi}}+\ln\frac{1}{\sigma}+\ln e^{-\frac{\left(x-\mu\right)^2}{2\sigma^2}}=-\frac{1}{2}\ln\left(2\pi\right)-\ln\sigma-\frac{\left(x-\mu\right)^2}{2\sigma^2}$$

$$\frac{\partial\left(\ln f\left(x,\mu\right)\right)}{\partial\mu}=\frac{\left(x-\mu\right)}{\sigma^2}$$

$$\underline{\underline{I\left(\mu\right)}}=E\left\{\left(\frac{\partial\ln f\left(x,\mu\right)}{\partial\mu}\right)^2\right\}=E\left\{\frac{\left(x-\mu\right)^2}{\sigma^4}\right\}=\frac{1}{\sigma^4}\cdot E\left\{\left(x-\mu\right)^2\right\}=\frac{1}{\sigma^4}\cdot DX=\frac{1}{\sigma^4}\cdot\sigma^2=$$

$$=\underline{\underline{\frac{1}{\sigma^2}}}$$

Řešený příklad

Nechť náhodná veličina *X* má Poissonovo rozdělení $Po(\lambda)$. Určete míru informace o parametru λ.

$$f(k,\lambda)=\frac{\lambda^k}{k!}\cdot e^{-\lambda} \qquad k=0,1,\ldots, \quad \lambda>0$$

$$EX=DX=\lambda$$

Je zřejmé, že $f(k,\lambda)$ *je regulární hustota.*

$$\ln f(k,\lambda)=\ln(\lambda)^k-\ln k!+\ln e^{-\lambda}=k\cdot\ln(\lambda)-\ln k!-\lambda$$

$$\frac{\partial(\ln f(k,\lambda))}{\partial\lambda}=\frac{k}{\lambda}-1=\frac{k-\lambda}{\lambda}$$

$$\underline{\underline{I(\lambda)}}=E\left\{\left(\frac{\partial\ln f(k,\lambda)}{\partial\lambda}\right)^2\right\}=E\left\{\frac{(k-\lambda)^2}{\lambda^2}\right\}=\frac{1}{\lambda^2}\cdot E\{(k-\lambda)^2\}=\frac{1}{\lambda^2}\cdot DX=\frac{1}{\lambda^2}\cdot\lambda=\underline{\underline{\frac{1}{\lambda}}}$$

Rao – Cramerova nerovnost

Popisuje vztah mezi Fisherovou mírou informace a dolní mezí rozptylu C(Θ) nestranných odhadů dané parametrické funkce.

Definice:

Nechť $\underline{X} = (X_1, \ldots, X_n)$ je náhodným výběrem, který má regulární rozdělení $f(\mathbf{x},\Theta)$, jenž je funkcí jednoho reálného parametru. Nechť $\tau(\Theta)$ je daná parametrická funkce taková, že její

$$\tau'(\Theta) = \frac{\partial \tau(\Theta)}{\partial \Theta}$$

existuje pro každé $\Theta \in$ **Θ**. **Nestranný odhad** $T(\underline{X})$ ($E(T) = \tau(\Theta)$) parametrické funkce $\tau(\Theta)$ nazveme **regulárním** právě když pro každé $\Theta \in$ **Θ** platí:

$$\int_M T(\underline{x}) \cdot f'(\underline{x},\Theta)\, d\underline{x} = \frac{\partial}{\partial \Theta} \int_M T(\underline{x}) \cdot f(\underline{x},\Theta)\, d\underline{x} = \frac{\partial}{\partial \Theta} E\left(T(\underline{X})\right) = \frac{\partial \tau(\Theta)}{\partial \Theta} = \tau'(\Theta),$$

tzn. když jeho střední hodnotu $\left(E(T(\underline{X})) = \int_M T(\underline{x}) \cdot f(\underline{x},\Theta) dx \right)$ můžeme derivovat podle Θ.

$$\int_M T(\underline{x})\cdot f'(\underline{x},\Theta)\,dx = \int_M T(\underline{x})\cdot \frac{f'(\underline{x},\Theta)}{f(\underline{x},\Theta)}\cdot f(\underline{x},\Theta)\,d\underline{x} =$$

$$\int_M T(\underline{x})\cdot \frac{\partial\, ln\, f(\underline{x},\Theta)}{\partial\Theta}\, f(\underline{x},\Theta)\,d\underline{x} = \tau'(\Theta)$$

Následující věta pak udává dolní mez rozptylu nestranného odhadu dané parametrické funkce $\tau(\Theta)$.

Rao – Cramerova věta

Nechť $\underline{X} = (X_1, \ldots, X_n)$ je náhodným výběrem, který má regulární rozdělení $f(\mathbf{x},\Theta)$. Nechť $\tau(\Theta)$ je daná parametrická funkce. Potom pro každý regulární nestranný odhad $T(\underline{X})$ funkce $\tau(\Theta)$ platí:

$$D(T(\underline{X})) \geq \frac{\{\tau'(\Theta)\}^2}{n\cdot I(\Theta)} = C(\Theta)$$

Odhad $T(\underline{X})$, jehož rozptyl je roven Rao – Cramerově dolní mezi rozptylu C(Θ), je efektivním odhadem.

Schwarzova nerovnost (Rao 71)

f, g – reálné, def. na množině M, integrovatelné

$$\Rightarrow \left(\int_M f \cdot g \, d\underline{x} \right)^2 \leq \int_M f^2(\underline{x}) \, d\underline{x} \cdot \int_M g^2(\underline{x}) \, d\underline{x}$$

důsledek: $\left[\operatorname{cov}(X, Y) \right]^2 \leq D(X) \cdot D(Y)$

My:

$$\left[\operatorname{cov}\left\{ T(\underline{X}), \frac{\partial \ln f(\underline{X}, \theta)}{\partial \theta} \right\} \right]^2 \leq D\left[T(\underline{X}) \right] \cdot D\left\{ \frac{\partial \ln f(\underline{X}, \theta)}{\partial \theta} \right\}$$

ale $\operatorname{cov}\left\{ T(\underline{X}), \frac{\partial \ln f(\underline{X}, \theta)}{\partial \theta} \right\} = E\left\{ \left[T(\underline{X}) - \tau(\theta) \right] \cdot \left[\frac{\partial}{\partial \theta} \ln f(\underline{X}, \theta) - 0 \right] \right\}$

$$= \int_M T(\underline{x}) \cdot \frac{\partial}{\partial \theta} \ln f(\underline{x}, \theta) \cdot f(\underline{x}, \theta) \, d\underline{x} =$$

regularit. odhadu

$$= \tau'(\theta)$$

$\operatorname{cov}(X, Y) = E\left[(X - EX) \cdot (Y - EY) \right]$

Schwarzova nerovnost (Rao 71)

f, g – reálné, def na množině M, integrovatelné

$$\Rightarrow \left(\int_M f \cdot g \, d\underline{x} \right)^2 \leq \int_M f^2(\underline{x})\, d\underline{x} \cdot \int_M g^2(\underline{x})\, d\underline{x}$$

důsledek: $[\mathrm{cov}(X, Y)]^2 \leq D(X) \cdot D(Y)$

My:

$$\left[\mathrm{cov}\left\{ T(\underline{X}), \frac{\partial \ln f(\underline{X},\theta)}{\partial\theta} \right\} \right]^2 \leq D[T(\underline{X})] \cdot D\left\{ \frac{\partial \ln f(\underline{X},\theta)}{\partial\theta} \right\}$$

ale $\mathrm{cov}\left\{ T(\underline{X}), \frac{\partial \ln f(\underline{X},\theta)}{\partial\theta} \right\} = E\left\{ [T(\underline{X}) - \tau(\theta)] \cdot \left[\frac{\partial}{\partial\theta} \ln f(\underline{X},\theta) - 0 \right] \right\}$

$$= \int_M T(\underline{x}) \cdot \frac{\partial}{\partial\theta} \ln f(\underline{x},\theta) \quad f(\underline{x},\theta)\, d\underline{x} =$$

regularita odhadu

$$= \tau'(\theta)$$

$\mathrm{cov}(X,Y) = E[(X - EX) \cdot (Y - EY)]$

Řešený příklad

Nechť $\underline{X}$ = (X_1, …, X_n) je náhodným výběrem z Poissonova rozdělení $Po(\lambda)$. Určete podle Rao – Cramerovy nerovnosti dolní mez rozptylu odhadu parametru λ.

Z předcházejícího řešeného příkladu víme, že Fisherova míra informace parametru λ je:

$$I(\lambda)=\frac{1}{\lambda}$$

Podle Rao – Cramerovy věty má každý regulární nestranný odhad parametru λ $(\tau(\lambda)=\lambda)$ *dolní mez rozptylu*

$$C(\lambda)=\frac{\{\tau'(\lambda)\}^2}{n\cdot I(\lambda)}=\frac{1}{n\cdot\frac{1}{\lambda}}=\frac{\lambda}{n}$$

Rozptyl nestranného odhadu $T(\underline{X})=\overline{X}=\frac{1}{n}\cdot\sum_{i=1}^{n}X_i$ *je* $D(T(\underline{X}))=D(\overline{X})=\frac{\lambda}{n}$

Pro parametr λ tedy existuje nestranný odhad s rozptylem rovnajícím se Rao – Cramerově dolní mezí.

Řešený příklad

Nechť $\underline{X} = (X_1, \dots, X_n)$ je náhodným výběrem z normálního rozdělení $N(\mu,\sigma^2)$, kde σ^2 známe. Určete pomocí Rao – Cramerovy nerovnosti dolní mez rozptylu odhadu parametru μ.

$$f(x,\mu) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad kde\ \sigma^2\ známe$$

Chceme najít Rao-Cramerovu dolní mez rozptylu: $C(\mu) = \dfrac{\{\tau'(\mu)\}^2}{n \cdot I(\mu)}$

Z předcházejících řešených příkladů víme, že Fisherova míra informace pro parametr μ je:

$$I(\mu) = \frac{1}{\sigma^2}$$

a příslušná parametrická funkce je: $\tau(\mu) = \mu$

Proto: $C(\mu) = \dfrac{\{\tau'(\mu)\}^2}{n \cdot I(\mu)} = \dfrac{1}{n \cdot \frac{1}{\sigma^2}} = \dfrac{\sigma^2}{n}$

Víme, že nejlepším nestranným (efektivním) odhadem střední hodnoty μ *je průměr* $\left(T(\underline{X}) = \overline{X}\right)$. *Měl by tedy mít rozptyl roven Rao – Cramerově dolní mezi rozptylu.*

Podle centrální limitní věty můžeme průměr aproximovat normálním rozdělením $N\left(\mu, \frac{\sigma^2}{n}\right)$*, z čehož je zřejmé, že* $D(T(\underline{X})) = \frac{\sigma^2}{n} = C(\mu)$. *Čímž jsme dokázali,že průměr je efektivním odhadem střední hodnoty normálního rozdělení.*

Rayleighovo rozdělení – parametr Θ^2

Rayleighovo rozdělení

$$F(x) = 1 - e^{-\frac{1}{2}\frac{x^2}{\Theta^2}} \qquad f(x) = \frac{x}{\Theta^2} e^{-\frac{1}{2}\frac{x^2}{\Theta^2}} \qquad x > 0$$

N. výběr: $(\underline{X}) = (X_1, \ldots, X_n)$

simult. hustota: $f(\underline{x}) = \Theta^{-2n} \prod_{i=1}^{n} x_i \cdot e^{-\frac{1}{2}\frac{x_i^2}{\Theta^2}}$

$$\ln f(\underline{x}) = -n \ln \Theta^2 + \sum_{i=1}^{n} \ln x_i - \frac{1}{2}\frac{\sum_{i=1}^{n} x_i^2}{\Theta^2}$$

$$f(\underline{x}) = e^{\ln f(\underline{x})} = e^{-n \ln \Theta^2 + \sum_{i=1}^{n} \ln x_i - \frac{1}{2\Theta^2} \cdot \sum_{i=1}^{n} x_i^2}$$

$$= e^{-n \ln \Theta^2 - \frac{1}{2\Theta^2} \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} \ln x_i}$$

$$= g\left(\sum_{i=1}^{n} x_i^2, \Theta^2\right) \cdot h(\underline{x})$$

$\Rightarrow T(\underline{X}) = \sum_{i=1}^{n} X_i^2$ je postačující st. pro Θ^2

MLE: $\dfrac{\partial \ln f(\underline{x})}{\partial \theta^2} = -n \cdot \dfrac{1}{\theta^2} + \dfrac{1}{2}\dfrac{1}{\theta^4}\sum_{i=1}^{n} x_i^2 = 0$

$$\Rightarrow \hat{\theta}^2 = \frac{1}{2n}\sum_{i=1}^{n} x_i^2$$

Pokud $E\hat{\theta}^2 = \theta^2$ $\Rightarrow$ máme efektivní odhad θ^2

DK: $E\hat{\theta}^2 = \dfrac{1}{2n} E\left(\sum_{i=1}^{n} X_i^2\right) = \dfrac{1}{2n}\sum_{i=1}^{n} EX_i^2 = \dfrac{1}{2n}\cdot n \cdot 2\theta^2 = \theta^2$

$$EX^2 = \int_0^\infty \frac{x^3}{\theta^2} e^{-\frac{1}{2}\frac{x^2}{\theta^2}}\,dx = 2\theta^2$$

Subst. $\boxed{t = \dfrac{x^2}{\theta^2}}$

$$f(x,\theta) = \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}}$$

Exp. rozd.

Sestrojte $J(\theta)$

$$J(\theta) = \int_0^\infty \left(\frac{f'}{f}\right)^2 \cdot f(x)\,dx = \int_0^\infty \frac{(f')^2}{f}\,dx =$$

$$= \frac{1}{\theta^4} \int_0^\infty (x-\theta)^2 \frac{1}{\theta} e^{-\frac{x}{\theta}}\,dx = \frac{1}{\theta^4} DX = \frac{\theta^2}{\theta^4} = \frac{1}{\theta^2}$$

$$f'(x) = e^{-\frac{x}{\theta}} \frac{x-\theta}{\theta^3}$$

Metodou MLE: $\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \overline{X} = T(\underline{X})$

$E\hat{\theta} = \theta \quad \wedge \quad \sum_{i=1}^n x_i$ je postačující

$\Rightarrow$ efektivní !?

Srovnej Rao–Cramer:

$$D\,T(\underline{X}) \geq \frac{1}{n \cdot J(\theta)} = \frac{1}{n \cdot \frac{1}{\theta^2}} = \frac{\theta^2}{n}$$

jelikož $D(T(\underline{X}) = D \frac{\sum_{i=1}^n X_i}{n} = \frac{n \cdot \theta^2}{n^2} = \frac{\theta^2}{n}$

6.1. Exponenciální třída rozdělení pravděpodobnosti. Nejlepší nestranné odhady – pokud vůbec pro danou parametrickou funkci nějaký nestranný odhad existuje – lze nejlépe konstruovat, když pozorovaná veličina X má rozdělení patřící do určité speciální třídy, totiž do tzv. exponenciální třídy rozdělení.

Říkáme, že veličina X má *rozdělení z exponenciální třídy* (nebo *rozdělení exponenciálního typu*), jestliže její hustota pravděpodobnosti $f(x)$ [případně pravděpodobnostní funkce $p(x) = P(X = x)$, jde-li o rozdělení diskrétního typu] má tvar

$$(6.1.1) \qquad f(x;\boldsymbol{\theta}) = \exp\left[\sum_{j=1}^{k} Q_j(\boldsymbol{\theta})\, U_j(x) + R(\boldsymbol{\theta}) + V(x)\right]$$

a splňuje podmínky

$$(6.1.2) \qquad \{x \mid f(x;\boldsymbol{\theta}) > 0\} \quad \text{nezávisí na } \boldsymbol{\theta}\,;$$

(6.1.3) parametrický prostor Ω obsahuje k-rozměrný interval,

Většina rozdělení užitečných pro aplikace patří do exponenciální třídy!

Jednorozměrná hustota z exponenciální třídy:

$$f(x;\theta) = \exp\left[\sum_{j=1}^{k} Q_j(\theta)\, U_j(x) + R(\theta) + V(x)\right]$$

Sdružená (simultánní) hustota z exponenciální třídy:

$\underline{X} = (X_1, \ldots, X_n)$ je náhodný výběr

Sdružená hustota pro N. výběr z exp. třídy

$$f(\underline{x}, \underline{\Theta}) = \exp\left[\sum_{j=1}^{k} Q_j(\underline{\Theta})\, S_j(\underline{x}) + n\, R(\underline{\Theta}) + V(\underline{x})\right]$$

$$S_j = \sum_{i=1}^{n} U_j(x_i) \qquad V(\underline{x}) = \sum_{i=1}^{n} V(x_i)$$

min. postačující stat. (→ viz věta o faktorizaci)

$$f(\underline{x}, \Theta) = g\{T(\underline{x}), \Theta\} \cdot h(\underline{x})$$

6.2. Příklady

6.2.1. Poissonovo rozdělení patří do exponenciální třídy, neboť jeho pravděpodobnostní funkci (zde místo θ označíme parametr rozdělení symbolem λ, jak jsme to činili v [24])

$$p(x;\lambda) = P(X = x) = \frac{\lambda^x}{x!}\exp(-\lambda), \qquad x = 0, 1, \dots,$$

lze zapsat ve tvaru

$$p(x;\lambda) = \exp(x \ln \lambda - \lambda - \ln x!),$$

tj. ve tvaru (6.1.1), kde $k = 1$, $Q_1(\lambda) = \lambda$, $U_1(x) = x$; $\{x \mid p(x;\lambda) > 0\} = \{0, 1, \dots\}$, tedy nezávisí na λ, a $\Omega = (0, \infty)$, tedy Ω obsahuje jednorozměrný interval.

$$f(x;\boldsymbol{\theta}) = \exp\left[\sum_{j=1}^{k} Q_j(\boldsymbol{\theta})\, U_j(x) + R(\boldsymbol{\theta}) + V(x)\right]$$

6.2.2. Logaritmicko-normální rozdělení patří do exponenciální třídy, neboť jeho hustotu

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma x \sqrt{(2\pi)}} \exp\left[-\frac{1}{2\sigma^2} (\ln x - \mu)^2 \right], \qquad x > 0,$$

lze zapsat ve tvaru

$$f(x; \mu, \sigma^2) = \exp\left[-\frac{1}{2\sigma^2} (\ln x)^2 + \frac{\mu}{\sigma^2} \ln x - \frac{\mu^2}{2\sigma^2} - \frac{1}{2} \ln \sigma^2 - \ln x - \ln \sqrt{(2\pi)} \right],$$

$$x > 0,$$

tj. ve tvaru (6.1.1) s

$$k = 2,\ \boldsymbol{\theta} = (\mu, \sigma^2),\ Q_1(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2},\ Q_2(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2},\ U_1 = (\ln x)^2,\ U_2 = \ln x,$$

$$R(\boldsymbol{\theta}) = -\frac{1}{2}\left(\frac{\mu^2}{\sigma^2} + \ln \sigma^2 \right) - \ln \sqrt{(2\pi)},\ V(x) = -\ln x.$$

Parametrický prostor $\Omega = \{(\mu, \sigma^2) \mid -\infty < \mu < \infty,\ \sigma^2 > 0\}$ je polorovina, množina $\{x \mid f(x; \mu, \sigma^2) > 0\} = (0, \infty)$, tedy nezávisí na (μ, σ^2).

$$f(x; \boldsymbol{\theta}) = \exp\left[\sum_{j=1}^{k} Q_j(\boldsymbol{\theta})\, U_j(x) + R(\boldsymbol{\theta}) + V(x) \right]$$

6.2.3. Rovnoměrné rozdělení na intervalu $(0, \theta)$ nepatří do exponenciální třídy; jeho hustota totiž je

$$f(x; \theta) = \frac{1}{\theta}, \qquad 0 < x < \theta,$$

takže množina $\{x \mid f(x; \theta) > 0\}$ závisí na θ, a nesplňuje tedy podmínku (6.1.2).

6.2.4. Cauchyovo rozdělení nepatří do exponenciální třídy, neboť jeho hustota

$$f(x; \theta) = \frac{1}{\pi}[1 + (x - \theta)^2]^{-1} = \exp\{-\ln \pi - \ln [1 + (x - \theta)^2]\}$$

Nemá tvar níže

$$f(x; \boldsymbol{\theta}) = \exp\left[\sum_{j=1}^{k} Q_j(\boldsymbol{\theta})\, U_j(x) + R(\boldsymbol{\theta}) + V(x)\right]$$

Vlastnosti ML odhadů (2 vlastnosti)

6.11. Vlastnosti maximálně věrohodných odhadů. Maximálně věrohodné odhady mají dvě důležité vlastnosti, pro které se jich používá tak často, přestože jejich výpočet je numericky náročný. První z těchto vlastností se týká výběrů jakéhokoliv rozsahu *n* a zní:

6.11.1. Věta. *Jestliže rozdělení pozorované veličiny X patří do exponenciální třídy* (6.1.1), *pak maximálně věrohodné odhady $\hat{\theta}_1, \ldots, \hat{\theta}_k$ parametrů $\theta_1, \ldots, \theta_k$ jsou funkcemi minimálních postačujících statistik* (6.3.2).

D ů k a z. Má-li X hustotu (pravděpodobnostní funkci) tvaru (6.1.1), pak sdružená hustota náhodného výběru $\mathbf{X} = (X_1, \ldots, X_n)'$ je typu (6.3.1), logaritmus funkce věrohodnosti má pak tvar

$$\ln L(\theta) = \sum_{l=1}^{k} Q_l(\theta)\, S_l(\mathbf{x}) + n\, R(\theta) + V(\mathbf{x})\,, \tag{6.11.1}$$

věrohodnostní rovnice jsou

$$\sum_{l=1}^{k} S_l(\mathbf{x}) \frac{\partial Q_l(\theta)}{\partial \theta_j}\bigg|_{\theta=\hat{\theta}} + n \frac{\partial R(\theta)}{\partial \theta_j}\bigg|_{\theta=\hat{\theta}} = 0, \quad j = 1, \ldots, k\,, \tag{6.11.2}$$

a jejich řešení závisí na $\mathbf{x}$ jen prostřednictvím hodnot $S_1(\mathbf{x}), \ldots, S_k(\mathbf{x})$ minimálních postačujících statistik.

Z věty odst. 6.4 a věty odst. 6.11.1 pak přímo plyne důležitý důsledek.

↳ Rao Blackwell

Důsledek 1. vlastnosti

6.11.2. *Jestliže pozorovaná náhodná veličina X má rozdělení z exponenciální třídy* (6.1.1) *a jestliže pro nějakou funkci* $T(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ *maximálně věrohodných odhadů parametrů* $\theta_1, \ldots, \theta_k$ *platí*

$$E\left[T(\hat{\theta}_1, \ldots, \hat{\theta}_k)\right] = \tau(\boldsymbol{\theta}) \quad \text{pro všechna} \quad \boldsymbol{\theta} \in \Omega, \tag{6.11.3}$$

pak $T(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ *je nejlepším nestranným odhadem funkce* $\tau(\boldsymbol{\theta})$ *vektoru parametrů* $\boldsymbol{\theta}$.

Druhá vlastnost následuje:

6.11.5. Věta. *Nechť* $\mathbf{X} = (X_1, \ldots, X_n)'$ *je náhodný výběr z rozdělení s hustotou (pravděpodobnostní funkcí)* $f(x; \theta)$ *závislou na parametru* $\theta \in \Omega \subset \mathbf{E}^1$ *splňující podmínky odst. 6.11.3. Nechť skutečná hodnota* θ_0 *parametru* θ *je vnitřním bodem* Ω *a nechť v okolí bodu* θ_0 *platí*

(6.11.10)
$$-\int_M \frac{\partial^2 \ln f(x, \theta)}{\partial \theta^2} f(x, \theta)\, \mathrm{d}x = \int_M \left[\frac{\partial \ln f(x, \theta)}{\partial \theta}\right]^2 f(x, \theta)\, \mathrm{d}x = J(\theta),$$

resp.

$$-\sum_M \frac{\partial^2 \ln f(x. \theta)}{\partial \theta^2} f(x, \theta) = \sum_M \left[\frac{\partial \ln f(x, \theta)}{\partial \theta}\right]^2 f(x, \theta) = J(\theta).$$

Pak náhodná veličina

(6.11.11)
$$n^{1/2}(\hat{\theta} - \theta_0),$$

kde $\hat{\theta}$ *je kořen věrohodnostní rovnice, konverguje v distribuci (viz* [24], *odst.* 26.2) *k normálnímu rozdělení* $\mathrm{N}[0, 1/J(\theta_0)]$.

Je-li $\tau(\theta)$ *diferencovatelná funkce na* Ω, *pak náhodná veličina*

(6.11.12)
$$n^{1/2}[\tau(\hat{\theta}) - \tau(\theta_0)]$$

konverguje v distribuci k normálnímu rozdělení $\mathrm{N}\{0, [\tau'(\theta_0)]^2/J(\theta_0)\}$.

Důsledek druhé vlastnosti

Prakticky to znamená, že při velkých hodnotách n lze aproximovat rozdělení odhadu $\hat{\theta}$ parametru θ rozdělením $\mathrm{N}\{\theta, 1/[n\,J(\theta)]\}$ a rozdělení odhadu $\tau(\hat{\theta})$ pro funkci $\tau(\theta)$ parametru θ rozdělením $\mathrm{N}\{\tau(\theta), [\tau'(\theta)]^2/[n\,J(\theta)]\}$. Stručně říkáme, že maximálně věrohodný odhad $\hat{\theta}$ je *asymptoticky nestranný* a jeho *asymptotické rozdělení* je *normální s rozptylem* $1/[n\,J(\theta)]$. Protože žádný nestranný odhad parametru θ nemůže mít (za podmínek odst. 6.11.3) rozptyl menší než $1/[n\,J(\theta)]$ (viz [1, 27]), říká se, že maximálně věrohodný odhad je *asymptoticky eficientní*.

Aplikace 1

$$f(x,\theta) = \frac{1}{\theta} \cdot e^{-\frac{x}{\theta}}$$

Exp. rozd.

Sestrojte $J(\theta)$

$$J(\theta) = \int_0^\infty \left(\frac{f'}{f}\right)^2 \cdot f(x)\, dx = \int_0^\infty \frac{(f')^2}{f}\, dx =$$

$$= \frac{1}{\theta^4} \int_0^\infty (x-\theta)^2 \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx = \frac{1}{\theta^4} DX = \frac{\theta^2}{\theta^4} = \frac{1}{\theta^2}$$

$$f'(x) = e^{-\frac{x}{\theta}} \frac{x-\theta}{\theta^3}$$

Metodou MLE: $\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \overline{X} = T(X)$

$E\hat{\theta} = \theta \wedge \sum_{i=1}^n x_i$ je postačující

→ **tedy je efektivní**

Tedy aplikace druhé vlastnosti:

$$\sqrt{n}.(\hat{\Theta} - \Theta) \approx N\left(0, \frac{1}{J(\Theta)}\right) \qquad \text{kde} \qquad \hat{\Theta} = \sum_{i=1}^{n} \frac{X_i}{n}$$

a tudíž

$$\hat{\Theta} \approx N\left(\Theta, \frac{1}{n.J(\Theta)}\right) = N(\Theta, \frac{\Theta^2}{n})$$

Tj. známe rozdělení odhadu, můžeme počítat intervalové odhady a další nejistoty okolo parametru Θ!!

Aplikace 2 - Rayleighovo rozdělení – parametr Θ^2 (námět na projekt?)

Rayleighovo rozdělení

$$F(x) = 1 - e^{-\frac{1}{2}\frac{x^2}{\Theta^2}} \qquad f(x) = \frac{x}{\Theta^2} e^{-\frac{1}{2}\frac{x^2}{\Theta^2}} \qquad x > 0$$

MLE: $$\frac{\partial \ln f(x)}{\partial \Theta^2} = -n \cdot \frac{1}{\Theta^2} + \frac{1}{2}\frac{1}{\Theta^4}\sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \hat{\Theta}^2 = \frac{1}{2n}\sum_{i=1}^{n} x_i^2$$

Pokud $E\hat{\Theta}^2 = \Theta^2 \Rightarrow$ mám efektivní odhad Θ^2

Tedy: $$\sqrt{n}.(\hat{\Theta}^2 - \Theta^2) \approx N\left(0, \frac{1}{J(\Theta^2)}\right)$$ kde $$\hat{\Theta}^2 = \sum_{i=1}^{n} \frac{X_i^2}{2n}$$

a tudíž $$\hat{\Theta}^2 \approx N\left(\Theta^2, \frac{1}{n.J(\Theta^2)}\right)$$ tj. zase můžeme počítat nejistoty!!

Příklad 3

6.2.4. Cauchyovo rozdělení nepatří do exponenciální třídy, neboť jeho hustota

$$f(x;\theta) = \frac{1}{\pi}\left[1 + (x-\theta)^2\right]^{-1} = \exp\left\{-\ln\pi - \ln\left[1 + (x-\theta)^2\right]\right\}$$

Ale je regulární !

Spočtěte $J(\Theta)$

$$\underline{\underline{J(\Theta)}} = E\left\{\left(\frac{\partial \ln f(x,\Theta)}{\partial\Theta}\right)^2\right\} = \int_{-\infty}^{+\infty} \frac{4.(x-\Theta)^2}{\left[1+(x-\Theta)^2\right]^2} \bullet f(x,\Theta) = \frac{4}{\pi}\int_{-\infty}^{+\infty} \frac{x^2}{\left[1+x^2\right]^3}\,dx = \ldots = \underline{\underline{\frac{1}{2}}}$$

$$Z = \frac{\partial \ln f(x,\Theta)}{\partial\Theta} = \frac{2.(x-\Theta)}{1+(x-\Theta)^2}$$

Je-li Θ^* MLE odhad $\rightarrow$ potom $\sqrt{n}.(\Theta^* - \Theta) \approx N(0,2)$

neboli $\Theta^* \approx N\left(\Theta, \frac{2}{n}\right)$ a zase můžeme počítat intervalové odhady!!

ZÁKLADY BAYESOVY INDUKCE

Cíl:

- Metodu maximální věrohodnosti – opakování
- Základy Bayesovy indukce
- Apriorní a aposteriorní rozdělení
- Bayesův bodový a intervalový odhad

Metoda maximální věrohodnosti - opakování

Nechť $X_1, X_2, \ldots, X_n$ je náhodný výběr z rozdělení s distribuční funkci $F(x;\Theta)$, kde tvar distribuční funkce je znám a Θ je neznámý parametr. Obecně může distribuční funkce obsahovat více neznámých parametrů, které můžeme označit vektorově jako $\underline{\Theta}$. Problém bodového odhadu nyní spočívá v nalezení statistiky $\hat{\Theta}(X)$, jakožto funkce $X_1, X_2, \ldots, X_n$, která by mohla být použita jako odhad $\underline{\Theta}$. Tato statistika bývá často označována jako **estimátor** a její realizace, řekněme $\hat{\Theta}(x)$, jako odhad. Nechť $f(\underline{x};\underline{\Theta})$ je hustota pravděpodobnosti nebo pravděpodobnostní funkce náhodného výběru $\underline{X}$ $(X_1, X_2, \ldots, X_n)$.

Definice:

Pokud je hustota pravděpodobnosti nebo pravděpodobnostní funkce $f(\underline{x};\underline{\Theta})$ vyšetřovaná jako funkce $\underline{\Theta}$, nazveme ji **věrohodnostní funkcí** založenou na $\underline{\mathrm{x}} = (\mathrm{x}_1, \ldots, \mathrm{x}_n)$ a označíme ji jako $L(\underline{\Theta};\underline{\mathrm{x}})$.

Jestliže $X_1, X_2, \ldots, X_n$ je množina nezávislých náhodných pozorování z rozdělení s hustotami $f_i(x_i;\underline{\Theta})$, $i = 1, \ldots, n$, pak věrohodnostní funkce může být získána jako:

$$L(\underline{\Theta};\mathrm{x}_1,\ldots,x_n) = f_1(x_1;\underline{\Theta})\cdot\ldots\cdot f_n(x_n;\underline{\Theta})$$

Definice:

Nechť $L(\Theta; \underline{X})$ je věrohodnostní funkce založena na náhodném výběru $\underline{X}$ **=** $(X_1, X_2, \ldots, X_n)$ z rozdělení $F(\underline{x}; \Theta)$, kde $\underline{\boldsymbol{\Theta}}$ je vektor neznámých parametrů, který nabývá hodnot z nějakého parametrického prostoru $\boldsymbol{\Theta}$. Pokud $\hat{\Theta} = \hat{\Theta}(\underline{X})$ je náhodný vektor, který maximalizuje $L(\Theta; \underline{X})$ vzhledem k $\hat{\Theta} \in \Theta$, potom $\hat{\Theta}(\underline{X})$ budeme nazývat **maximálně věrohodný estimátor** Θ.

Pro konkrétní realizaci náhodného výběru $\underline{x} = (x_1, \ldots, x_n)$ budeme $\hat{\Theta}(\underline{x})$ nazývat jako **maximálně věrohodný odhad** $\boldsymbol{\Theta}$. Pro tento odhad budeme používat zkratku MVO.

Řešený příklad

MVO pro parametr exponenciálního rozdělení

Nechť $\underline{X} = (X_1, X_2, \ldots, X_n)$ **je náhodný výběr z exponenciálního rozdělení pravděpodobnosti.**

Víme, že hustota pravděpodobnosti exponenciální náhodné veličiny má tvar:

$$f(x) = \lambda \cdot e^{-\lambda x}, \qquad x > 0,$$

kde λ>0 je neznámý parametr.

Věrohodnostní funkce:

$$L(\lambda; \mathrm{x}) = \prod_{i=1}^{n} f(x_i) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Logaritmická věrohodnostní funkce:

$$L^*(\lambda; \mathrm{x}) = \ln\left(\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}\right) = \ln(\lambda^n) + \ln\left(e^{-\lambda \sum_{i=1}^{n} x_i}\right) = n \cdot \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

Nalezení maxima $L^*(\lambda; \mathrm{x})$*:*

$$\frac{dL^*(\lambda; \mathrm{x})}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i \qquad \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}}$$

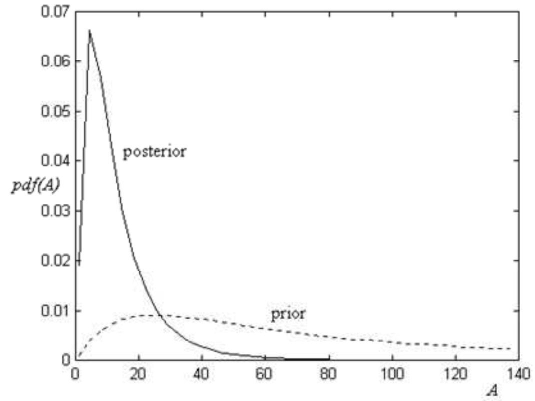
MVO očekávané hodnoty exponenciálního rozdělení EX, označený jako $\mu = \frac{1}{\lambda}$*, může být odvozen pomocí vlastnosti* ***invariance maximálně věrohodných odhadů****. Tato vlastnost říká [Nguyen H. T., Rogers G. S., 1989], že:*

Pokud $\hat{\Theta}$ *je MVO parametru Θ, pak* $g(\hat{\Theta})$ *je MVO funkce parametru* $g(\Theta)$ *za předpokladu, že* $g(.)$ *je prostá funkce.*

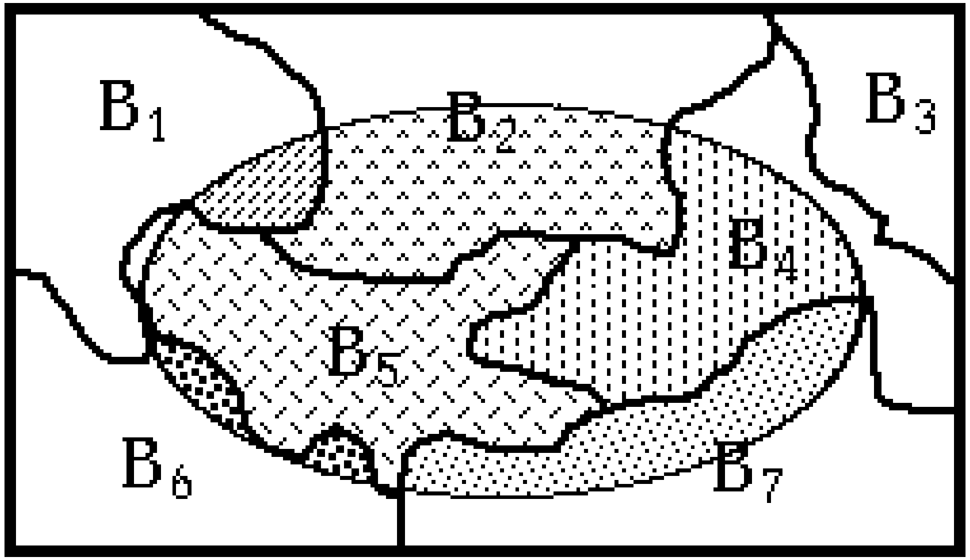
Nyní je jasné, že MVO očekávané hodnoty exponenciálního rozdělení je $\hat{\mu} = \overline{x}$.

Úvod do Bayesovy indukce

Reverend Thomas Bayes v roce 1761 jistě netušil, že jeho teorém najde uplatnění v moderní matematice 21. století



VĚTA BAYESOVA



$$P\{B_k|A\} = \frac{P\{A|B_k\}\,P\{B_k\}}{\sum_{i=1}^{n} P\{A|B_i\}\,P\{B_i\}}$$

$$P\{B_k|A\} = \frac{P\{B_k \cap A\}}{P\{A\}}$$

Tedy srovnej:

$$P\{B_k | A\} = \frac{P\{A|B_k\}\,P\{B_k\}}{\sum_{i=1}^{n} P\{A|B_i\}\,P\{B_i\}} \qquad \textbf{versus} \qquad \pi(\Theta|\underline{x}) = \frac{f(\underline{x}|\Theta)\cdot\pi(\Theta)}{\int_{\Theta} f(\underline{x}|\Theta)\cdot\pi(\Theta)\,d\Theta} \qquad \textit{pro } \Theta \in \Omega$$

V tomto přístupu se provádí statistická indukce o neznámém parametru podmíněně vzhledem k pozorovaným datům. Aby to bylo možné provést, je nezbytné udělit parametru Θ status náhodné veličiny. Klíčová charakteristika v Bayesově přístupu spočívá v tom, že je nezbytné specifikovat pravděpodobnostní rozdělení Θ ještě před pozorováním náhodného výběru. Toto rozdělení se nazývá **apriorní rozdělení pravděpodobnosti**.

Tedy parametr $\underline{\Theta}$, který je předmětem našeho zájmu, je v Bayesově přístupu vyšetřován jako náhodná veličina. Jde-li o parametr náhodné veličiny X s distribuční funkci $F(x;\underline{\Theta})$, pak vždy, když pracujeme s touto funkcí, musíme mít na mysli podmíněnou distribuční funkci $F(x|\underline{\Theta})$.

Apriorní rozdělení

Uvažujme parametr Θ náhodné veličiny X s distribuční funkci $F(x|\Theta)$. Nechť doposud není k dispozici žádné pozorování této náhodné veličiny. Nechť je k dispozici jakási předběžná zkušenost, či apriorní znalost o této populaci nebo jiná informace, která umožní zkonstruovat subjektivní pravděpodobnostní rozdělení pro parametr Θ. Takové rozdělení, které reflektuje nejistotu o hodnotě parametru z hlediska experimentátora ještě před pozorováním aktuálního výběru, se nazývá **apriorní rozdělení** parametru Θ. Dále jej budeme označovat $\pi(\Theta)$.

Úloha najít apriorní rozdělení pro zkoumaný parametr je všeobecně velmi obtížná. V některých případech bývá dokonce velmi výhodné zvolit jako apriorní hustotu takovou funkci, která nemusí být ani integrovatelná, a přesto po implementaci Bayesových metod dává rozumné výsledky (někdy však také vede k nesmyslným výsledkům). Taková hustota bývá označována jako tzv. **nevlastní apriorní hustota**. Bayesovy metody lze použít i v případě, že není dostupná žádná informace o vyšetřovaném parametru. Taková apriorní rozdělení, označována jako **neurčitá**, byla velmi intenzivně studována [Jeffreys, 1961] a jsou základem Bayesových metod vyvinutých autory [Box and Tiao, 1973].

Neexistuje žádný obecný návod, jak by měla být specifikována neurčitá apriorní hustota. Jelikož jsou pod tíhou různých argumentů použity různé definice neurčitých apriorních rozdělení, setkáme se často s různými nevlastními apriorními rozděleními [Zellner, 1977]. Pro účely nalezení neurčitého apriorního rozdělení použijeme jednu z nejpopulárnějších metod, navrženou v [Jeffreys, 1961].

Nechť $f(x|\Theta)$ je hustota pravděpodobnosti nebo pravděpodobnostní funkce pozorované náhodné veličiny X, kde vektor $\underline{\Theta}$ je vektor neznámých parametrů. Za neurčité apriorní rozdělení $\underline{\Theta}$ lze vzít:

$$\pi(\Theta) = \left[\left|I(\Theta)\right|\right]^{1/2} = \sqrt{\det I(\Theta)}\,,$$

Kde $I(\Theta)$ je Fisherova informační matice (regulární rozdělení pro vektorový parametr).

$$I(\Theta) = -E\left(\frac{\partial^2}{\partial\Theta_i \partial\Theta_j} \ln f(X|\Theta)\right) = E\left(\frac{\partial \ln f(X|\Theta)}{\partial\Theta_i} \times \frac{\partial \ln f(X|\Theta)}{\partial\Theta_j}\right) = E(Z_i \times Z_j)$$

Řešený příklad

Uvažujme exponenciální rozdělení s hustotou $f(x|\lambda) = \lambda \cdot e^{-\lambda x}, x > 0$**, kde** $\lambda > 0$ **je parametr, pro který je nutno zkonstruovat apriorní hustotu. Úkolem je nalézt neurčitou apriorní hustotu pro parametr λ a pro jeho převrácenou hodnotu μ** $\left(\mu = \frac{1}{\lambda}\right)$**.**

$$I(\lambda) = -E\left(\frac{\partial^2}{\partial\lambda^2} \ln\left(\lambda \cdot e^{-\lambda x}\right)\right) = -E\left(\frac{\partial^2}{\partial\lambda^2}\left[\ln(\lambda) - \lambda x\right]\right) = -E\left(-\frac{1}{\lambda^2}\right) = \frac{1}{\lambda^2}$$

Odtud vyplývá, že neurčitá (Jeffreysova) apriorní hustota pravděpodobnosti pro λ *je nevlastní apriorní hustota, úměrná* $\frac{1}{\lambda}\left(=\left(\frac{1}{\lambda^2}\right)^{1/2}\right)$*:*

$$\pi(\lambda) = \frac{1}{\lambda}$$

Pokud nás bude zajímat apriorní rozdělení pro převrácenou hodnotu λ, tj. μ, pak Fischerova matice bude:

$$I(\mu) = -E\left(\frac{\partial^2}{\partial\mu^2} ln\left(\frac{1}{\mu}\cdot e^{-\frac{1}{\mu}X}\right)\right) = -E\left(\frac{\partial^2}{\partial\mu^2}\left[ln\left(\frac{1}{\mu}\right) - \frac{1}{\mu}X\right]\right) =$$

$$= -E\left(\frac{\partial}{\partial\mu}\left(-\mu\frac{1}{\mu^2} + \frac{1}{\mu^2}X\right)\right) = -E\left(\frac{\partial}{\partial\mu}\left(-\frac{1}{\mu} + \frac{1}{\mu^2}X\right)\right) = -E\left(\frac{1}{\mu^2} - 2\frac{1}{\mu^3}X\right) =$$

$$= -\frac{1}{\mu^2} + \frac{2}{\mu^3}\cdot EX = -\frac{1}{\mu^2} + \frac{2}{\mu^3}\cdot\mu = \frac{1}{\mu^2}$$

Tedy neurčitá apriorní hustota pro očekávanou hodnotu exponenciálního rozdělení μ je:

$$\pi(\mu) = \left(\frac{1}{\mu^2}\right)^{1/2} = \frac{1}{\mu}$$

Aposteriorní rozdělení

Nechť $\underline{X}$ je náhodný vektor se sdruženou hustotou pravděpodobnosti nebo pravděpodobnostní funkcí $f\left(\underline{x}\middle|\Theta\right)$. Nechť $\pi\left(\underline{\Theta}\right)$ je apriorní rozdělení náhodného vektoru $\underline{\Theta}$. Potom sdružené rozdělení $\underline{X}$ a $\underline{\Theta}$ může být nalezeno jako:

$$h\left(\underline{x},\underline{\Theta}\right)=f\left(\underline{x}\middle|\underline{\Theta}\right)\cdot\pi\left(\underline{\Theta}\right)$$

Za předpokladu, že $\underline{\Theta}$ je spojitý náhodný vektor, marginální rozdělení $\underline{\boldsymbol{X}}$ může být nalezeno jako:

$$g\left(\underline{x}\right)=\int_{\Theta} h\left(\underline{x};\underline{\Theta}\right)d\underline{\Theta}$$

A konečně podmíněné rozdělení $\underline{\Theta}$ při realizaci $\underline{\mathrm{X}}=\underline{\mathrm{x}}$ je:

$$\pi\left(\underline{\Theta}\middle|\underline{x}\right)=\frac{h\left(\underline{x};\underline{\Theta}\right)}{\mathrm{g}\left(\underline{\mathrm{x}}\right)}\qquad pro\ \Theta\in\Omega$$

Toto pravděpodobnostní rozdělení $\underline{\Theta}$ se nazývá **aposteriorní rozdělení $\underline{\Theta}$**. Toto podmíněné rozdělení parametru při daných datech $\underline{x}$ je takto nazváno zejména proto, že odráží představu experimentátora o zkoumaném parametru poté, co byl pozorován náhodný výběr z příslušné populace. Tedy aposteriorní rozdělení kombinuje předběžnou informaci obsaženou v apriorním rozdělení s informací o $\underline{\Theta}$ (obsaženou v datech – věrohodnostní funkce). Pokud budeme ignorovat konstantu úměrnosti, pak aposteriorní rozdělení může být zapsáno následovně:

$$\pi(\Theta|x) \propto f(x|\Theta)\cdot\pi(\Theta) \qquad pro\ \Theta \in \Omega\ ,$$

kde konstanta úměrnosti k může být nalezena tak, aby byla splněna normovací podmínka aposteriorní hustoty.

Pozn.: $\propto$ *označuje přímou úměrnost, tzn.* $\pi(\Theta|x) \propto f(x|\Theta)\cdot\pi(\Theta)$ *je zkráceným zápisem rovnice:* $\pi(\Theta|x) = k\cdot f(x|\Theta)\cdot\pi(\Theta), \qquad k \in R$

Tedy:

aposteriorní rozdělení $\propto$ (věrohodnostní funkce × apriorní rozdělení)

Pokud aposteriorní rozdělení pravděpodobnosti patří do téže třídy rozdělení jako apriorní rozdělení, potom tuto třídu rozdělení nazýváme **přirozený konjugovaný systém rozdělení** pro rozdělení *X*. To znamená, že pokud apriorní rozdělení je konjugované vzhledem k apriornímu rozdělení, pak pro nalezení aposteriorního rozdělení potřebujeme aktualizovat pouze parametry apriorního rozdělení.

Řešený příklad

Nechť $X_1, \ldots, X_n$ je náhodný výběr z exponenciálního rozdělení s parametrem λ. Úkolem je nalézt aposteriorní rozdělení pro λ za předpokladu, že apriorní rozdělení je:

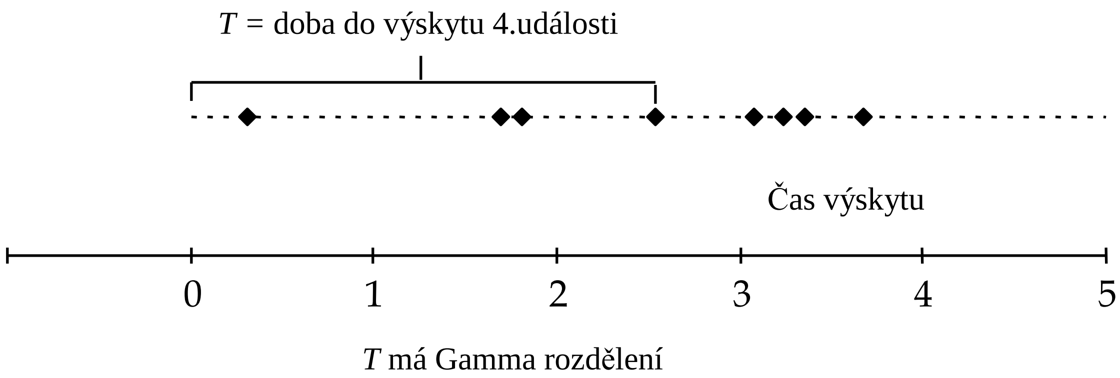
a) $\lambda \to Gamma(a;\mathrm{b})$**, tj.**

$$\pi(\lambda)=\frac{\left(\frac{1}{b}\right)^{a}}{\Gamma(a)}\cdot\lambda^{a-1}\cdot e^{-\frac{\lambda}{b}} \quad , \quad kde\ \Gamma(a)=\int_0^\infty x^{a-1}e^{-x}dx, \quad E\lambda=ab$$

b) $\pi(\lambda)=\frac{1}{\lambda}$

Poznámka – opakování Gamma rozdělení: Ga(k,λ)

- Náhodnou veličinu s Gamma rozdělením si můžeme představit jako součet k nezávislých exponenciálních náhodných veličin ... pomůcka pro výpočet $E(X)$ a $D(X)$.



Distribuční funkce: $F(t) = P(T_k < t) =$

$$P\left(\sum_{i=1}^{k} X_i < t\right) = 1 - P(N_t < k) = 1 - \sum_{j=0}^{k-1} e^{-\lambda t} \cdot \frac{(\lambda t)^j}{j!} \;\; = 1 - e^{-\lambda t}\left[\sum_{j=0}^{k-1} \frac{(\lambda t)^j}{j!}\right]$$

Derivací $F(t)$ dostaneme funkci hustoty pravděpodobnosti:

$$f(t) = \lambda e^{-\lambda t}\left[\sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}\right] - e^{-\lambda t}\left[\sum_{j=1}^{k-1}\frac{\lambda(\lambda t)^{j-1}}{(j-1)!}\right]$$

$$= \lambda^k e^{-\lambda t}\left[\frac{t^{k-1}}{(k-1)!}\right]; \quad t > 0$$

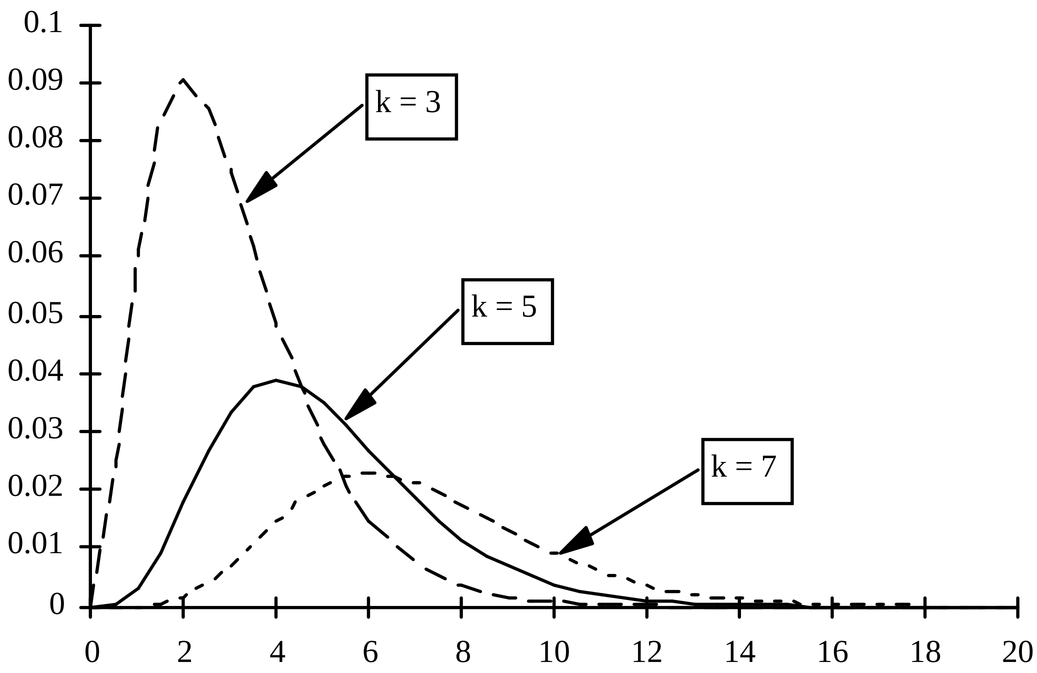
$$T_k = X_1 + X_2 + X_3 + \dots + X_k$$

$$X_i \to E(\lambda)$$

$ET_k = EX_1 + \dots + EX_k = \frac{1}{\lambda} + \dots + \frac{1}{\lambda} = \frac{k}{\lambda}$ ... střední (očekávaná) hodnota

$DT_k = \dots. = \frac{k}{\lambda^2}$ .... rozptyl T_k

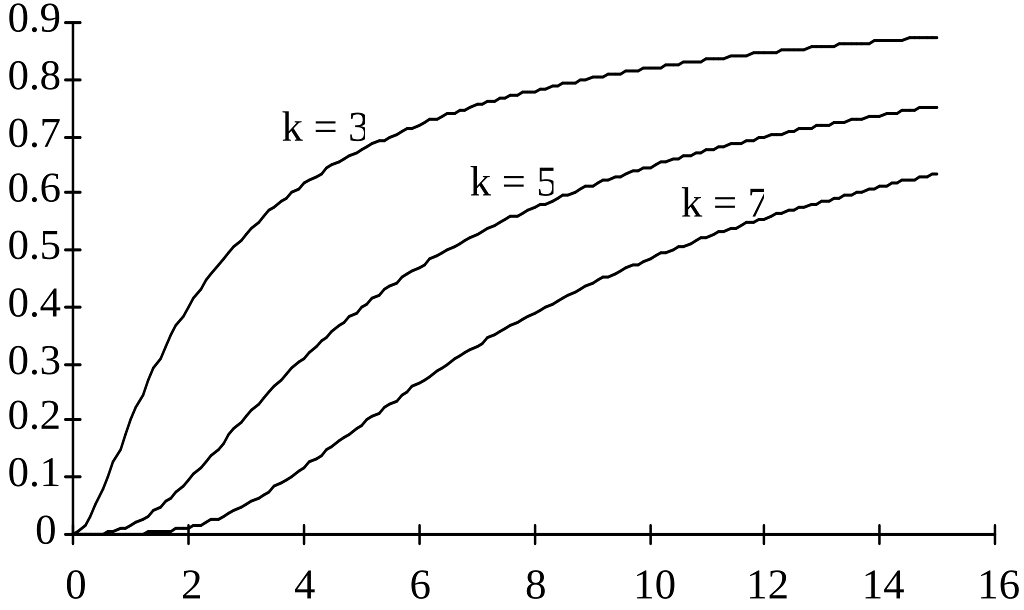
Pokud *k* je celočíselné, toto rozdělení mívá označení jako Erlangovo.



Hustota pravděpodobnosti pro Gamma rozdělení, Ga(k,λ).

Intenzita poruch (hazardní funkce):

$$h(x) = \frac{\mathrm{f(x)}}{1-\mathrm{F(x)}} = \frac{\lambda}{(k-1)!\sum_{j=0}^{k-1} \frac{1}{(k-1-j)!\left(\lambda x\right)^{j}}}$$



Intenzita poruch Gamma rozdělení, λ=1, Ga(k,1).

Ga(k,λ)

$$f(x)=\frac{\lambda^{k}}{\Gamma(k)}\cdot x^{k-1}e^{-\lambda.x}.,\quad \lambda>0$$

reparametrizace: $a=k,\quad b=\frac{1}{\lambda}$

$$f(x)=\frac{\left(\frac{1}{b}\right)^{a}}{\Gamma(a)}\cdot x^{a-1}e^{-\frac{x}{b}}.,\quad \lambda>0$$

je rozdělení $\longrightarrow$ **Ga(a,b)** $\quad a\in R$

A parametr λ bude v pozici náhodné veličiny, tedy jeho apriorní rozdělení bude:

$$f(\lambda)=\frac{\left(\frac{1}{b}\right)^{a}}{\Gamma(a)}\cdot \lambda^{a-1}e^{-\frac{\lambda}{b}}\propto \lambda^{a-1}e^{-\frac{\lambda}{b}}\quad \lambda>0$$

A budeme ho označovat nejčastěji jako: $\pi(\lambda)\propto \lambda^{a-1}e^{-\frac{\lambda}{b}}\quad \lambda>0$

Řešený příklad

Nechť $X_1, \ldots, X_n$ je náhodný výběr z exponenciálního rozdělení s parametrem λ. Úkolem je nalézt aposteriorní rozdělení pro λ za předpokladu, že apriorní rozdělení je:

a) $\lambda \to Gamma(a;\mathrm{b})$**, tj.**

$$\pi(\lambda) = \frac{\left(\frac{1}{b}\right)^a}{\Gamma(a)} \cdot \lambda^{a-1} \cdot e^{-\frac{\lambda}{b}} \quad , \quad kde\ \Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \quad E\lambda = ab$$

b) $$\pi(\lambda) = \frac{1}{\lambda}$$

****

Víme, že hustota pravděpodobnosti exponenciální náhodné veličiny má tvar:

$$f(x) = \lambda \cdot e^{-\lambda x}, \qquad x > 0,$$

kde $\lambda>0$ je neznámý parametr.

Zkonstruujeme nejdříve věrohodnostní funkci:

$$L(\lambda;\mathrm{x}) = \prod_{i=1}^{n} f(x_i) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} = \lambda^n e^{-\lambda n \bar{x}},$$

kde $\bar{x}$ *je výběrový průměr.*

ada) *Pokud ignorujeme konstantu úměrnosti, apriorní rozdělení*

$$\pi(\lambda) \propto \lambda^{a-1} \cdot e^{-\frac{\lambda}{b}}$$

aposteriorní rozdělení $\propto$ *(apriorní rozdělení* × *věrohodnostní funkce):*

$$\pi(\lambda) \propto \lambda^{a-1} \cdot e^{-\frac{\lambda}{b}} \cdot \lambda^n e^{-\lambda n \bar{x}} \Rightarrow \pi(\lambda|x) \propto \lambda^{n+a-1} \cdot e^{-\lambda\left(\frac{1}{b}+n\bar{x}\right)}$$

Lze snadno rozeznat, že se jedná o Gamma rozdělení:

$$(\lambda|x) \rightarrow Gamma\left(n+a; \frac{\mathrm{b}}{\mathrm{nb}\bar{\mathrm{x}}+1}\right)$$

Protože apriorní i aposteriorní rozdělení patří do téže třídy rozdělení, je evidentní, že tato třída (Gamma rozdělení) slouží jako přirozený konjugovaný systém pro parametr exponenciálního rozdělení.

adb) *nevlastní apriorní rozdělení:* $\pi(\lambda) = \dfrac{1}{\lambda}$

aposteriorní rozdělení $\propto \dfrac{1}{\lambda} \cdot \lambda^n e^{-\lambda n \bar{x}} \Rightarrow \pi(\lambda | x) \propto \lambda^{n-1} \cdot e^{-\lambda n \bar{x}}$,

takže:

$$(\lambda | x) \to Gamma\left(n; \frac{1}{n\bar{x}} \right)$$

a skutečnost, že apriorní rozdělení je nevlastní, zde nehraje významnou roli.

Řešený příklad, projekt

Nechť náhodná veličina *X* má Poissonovo rozdělení $Po(\lambda)$**.**

$$f(k,\lambda)=\frac{\lambda^k}{k!}\cdot e^{-\lambda} \qquad\qquad k=0,1,\dots, \quad \lambda>0$$

$$EX=DX=\lambda$$

$$\underline{\underline{I(\lambda)}}=E\left\{\left(\frac{\partial \ln f(k,\lambda)}{\partial\lambda}\right)^2\right\}=E\left\{\frac{(k-\lambda)^2}{\lambda^2}\right\}=\frac{1}{\lambda^2}\cdot E\left\{(k-\lambda)^2\right\}=\frac{1}{\lambda^2}\cdot DX=\frac{1}{\lambda^2}\cdot\lambda=\underline{\underline{\frac{1}{\lambda}}}$$

Jeffreysova apriorní hustota:

$$\pi(\lambda)=\left[\left|I(\lambda)\right|\right]^{1/2}=\frac{1}{\sqrt{\lambda}}=\lambda^{-\frac{1}{2}}$$

Aplikuji věrohodnostní funkci a dostanu aposteriorní hustotu: $Ga(n,\dfrac{1}{\sum_1^n x_i+\frac{1}{2}})$

Námět na projekt: Pokud apriorní je Gamma, jaká je aposteriorní hustota?
Tvoří Gamma rozdělení přirozený konjugovaný systém?

Konjugované apriorní hustoty vybraných jednoparametrických modelů:

Neznámý parametr: θ, ostatní parametry jsou známé

Model	Rozdělení x	Konjug. rozdělení θ
Normální	$\mathcal{N}(\theta, \sigma^2)$	$\mathcal{N}(\mu, \xi^2)$
Normální	$\mathcal{N}(\mu, \frac{1}{\theta})$	$\mathcal{G}(\alpha, \beta)$
Poissonovo	$\mathcal{P}(\theta)$	$\mathcal{G}(\alpha, \beta)$
Gama	$\mathcal{G}(\nu, \theta)$	$\mathcal{G}(\alpha, \beta)$
Binomické	$\mathcal{B}i(n, \theta)$	$\mathcal{B}e(\alpha, \beta)$