2 sezóny: 1 sezóna: A vyhrál 3 zápasy ze 4
(tj úspěšnost 75%) B 2/3 (úsp. 66,6%)

2. sezóna A 4/11 (36,4%) B 1/3 (33,3%)

obě sezóny A: 7/15 (47,%) B: 3/6 (50%)

Bodový odhad parametru:

Fce na náh. výběru $\hat{\Theta}(X_1, \dots, X_n)$, požadujeme po ní, aby měla „hezké" vlastnosti.

a) nestrannost: $\mathbb{E}\hat{\Theta}(X_1, \dots, X_n) = \Theta$

b) konzistence

c) efficience

ada: Př.: Pozorujeme n dob čekání na bus, který má intervaly mezi příjezdy b minut.

Odhadněte b.

X doba čekání na bus $\sim Ro(0, b)$

$X_1, \dots, X_n$ ... doby čekání při 1, ..., n. dě jízdě,

tj. $X_1, \dots, X_n$ nezávisle a stejně rozdělené jako

$X \leftarrow$ model pro X_i, $i = 1, \dots, n$

Vaše návrhy: 1) $\hat{b} = 2 \cdot \overline{X_n}$ $\left(\in \mathbb{E}X = \frac{b}{2}\right) \Rightarrow \mathbb{E}\hat{b} = b$ ⇒ nestranný

2) $\hat{b} = \max_{i=1,\dots,n} X_i \Rightarrow \mathbb{E}\hat{b} = \frac{n}{n+1} b \Rightarrow$ není nestranný

3) $\hat{b} = \frac{n+1}{n} \max X_i \Rightarrow \mathbb{E}\hat{b} = \frac{n+1}{n} = \frac{n+1}{n} b$ ⇒ nestranný

ad b) ad 1.) $\operatorname{var}\hat{b} = \operatorname{var} 2\bar{X}_n = \operatorname{var}\frac{2}{n}\cdot\sum_{i=1}^{n} X_i =$

$\frac{4}{n^2}\cdot\operatorname{var}\sum_{i=1}^{n} X_i = \frac{4}{n^2}\cdot\sum_{i=1}^{n}\overbrace{\operatorname{var} X_i}^{b^2/12} = \frac{4}{n^2}\cdot n\cdot\frac{b^2}{12} = \frac{b^2}{3n} \to$

$\xrightarrow[n\to\infty]{} 0 \Rightarrow$ konzistentní

ad 2.) $\operatorname{var}\hat{b} = \operatorname{var}\max_{i=1,\dots,n} X_i = \left(\mathbb{E}\left(\max X_i\right)^2\right) - \left(\mathbb{E}\left(\max X_i\right)\right)^2 = (*)$

$$F_{\max X_i}(x) = \begin{cases} 0 & \text{pro } x \le 0 \\ \left(\frac{x}{b}\right)^n & \text{pro } x \in (0,b) \\ 1 & \text{pro } x \ge b \end{cases} \Rightarrow$$

$$\Rightarrow f_{\max X_i}(x) \begin{cases} \frac{n x^{n-1}}{b^n} & \text{pro } x \in (0,b) \\ 0 & \text{pro } x \notin (0,b) \end{cases}$$

$$(*) = \int_a^b \frac{n x^{n+1}}{b^n}\,dx = \frac{n}{b^n}\cdot\left[\frac{x^{n+2}}{n+2}\right]_0^b =$$

$$= \frac{n}{n+2}\cdot\frac{b^{n+2}}{b^n} = \frac{n}{n+2}b^2$$

$$\Rightarrow \frac{n}{n+2}b^2 - \frac{n^2}{(n+1)^2}b^2 = b^2\cdot\frac{\cancel{n^3}+2n^2+n-\cancel{n^3}-2n^2}{(n+2)(n+1)^2} = \frac{n}{(n+2)(n+1)^2}b^2 =$$

ad 3) $\operatorname{var}\hat{b} = \operatorname{var}\left(\frac{n-1}{n}\max X_i\right) = \left(\frac{n+1}{n}\right)^2 \operatorname{var}(\max X_i) =$

$= \frac{(n+1)^2}{n^2}\,\frac{n}{(n+2)(n+1)^2}\,h^2 = \frac{h^2}{n(n+2)} \xrightarrow[n\to\infty]{} 0 \Rightarrow$ konzistentní

Metoda momentů:

Př. V 5 posluchárnách se píše test, v každé posluchárně je n studentů, psát, že student projde testem je p.

Pozorovány byly následující počty studentů v jednotlivých posluchárnách: 18, 19, 23, 18 a 22.

Určete model pro náh. veličinu X popisující -11- a odhadněte jeho parametry metodou momentů.

$X \sim \text{Binom}(n, p)$ $\quad EX = np$, $\operatorname{var}X = np(1-p)$

odhad $EX \to \bar{X} = \frac{18+19+\ldots+22}{5} = 20$

-11- $\operatorname{var}X \to S_5^2 = \frac{1}{4}\cdot\left((18-20)^2+(19-20)^2+\ldots+(22-20)^2\right) =$

$\left[\frac{1}{n-1}\right] = \frac{4+1+9+4+4}{4} = \frac{11}{2}$

$\left.\begin{aligned}\hat{n}\hat{p} &= 20\\ \hat{n}\hat{p}(1-\hat{p}) &= \tfrac{11}{2}\end{aligned}\right\}\ (1-\hat{p}) = \frac{\frac{11}{2}}{20} = \frac{11}{40} \Rightarrow \hat{p} = \frac{29}{40} \Rightarrow \hat{n}\cdot\frac{29}{40} = 20$

$\Rightarrow 29\hat{n} = 800 \Rightarrow \hat{n} = \frac{800}{29} \doteq 27{,}5$

(tj. $\hat{n} = 27$ nebo $\hat{n} = 28$)

Metoda maximální věrohodnosti (MMV nebo MLE)

Základní myšlenka: „To co jsme naměřili (tj. data) $\to x_1, \dots, x_n$, je nejpravděpodobnější možné." (diskrétní případ)

tj. $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ je max

$$P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_n = x_n) =$$

$$P(X = x_1) \cdot P(X = x_2) \cdots P(X = x_n) = \prod_{i=1}^{n} P(X = x_i)$$

X ... model pro x_i, $i = 1, \dots, n$

Cíl: najít parametr, který maximalizuje $\prod_{i=1}^{n} P(X = x_i)$

resp $\prod_{i=1}^{n} f(x_i)$ pro spojitý model

$$L(\theta) \to L'(\theta) = 0$$

nezikmek $\ell(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln P(X = x_i)$, resp $\sum_{i=1}^{n} \ln f(x_i)$

(můžeme být $\log_a$ pro $a > 0$)

$\ell'(\theta) = 0 \to$ řešením je $\hat{\theta}$

Př. V krabici máme černé a bílé kuličky. 50x jsme vytáhli kuličku, kterou jsme pokaždé vrátili, a pozorovali jsme 30x černou a 20x bílou kuličku.
Odhadněte podíl černých kuliček v krabici metodou max. věrohodnosti.

Ozn. $X=1$, pokud v tahu vytáhneme černou kuličku
$X=1$, — " — bílou
$\Rightarrow X \sim Alt(p)$, tj. $P(X=1)=p$ a $P(X=0)=1-p$

Pozorování: $x_1=1, x_2=1, \dots, x_{30}=1, x_{31}=0, \dots, x_{50}=0$

Postup: 1.) $L(p) = p\,p\cdots p\,(1-p)\cdots(1-p) = p^{30}(1-p)^{20}$

2.) $l(p) = 30 \ln p + 20 \ln(1-p)$

3.) $l'(p) = \frac{30}{p} - \frac{20}{1-p}$

4.) $\frac{30}{\hat{p}} - \frac{20}{1-\hat{p}} = 0$

$30 - 30\hat{p} = 20\hat{p}$

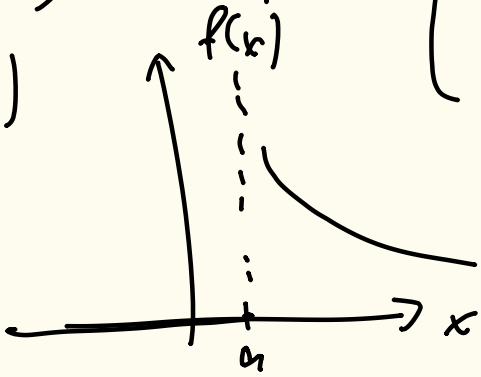
$30 = 50\hat{p}$

$\hat{p} = 0{,}6$

Př. pozorování 1, 2, 2, 2, 3, 3, 4

z rozdělení s hustotou $f(x) = \begin{cases} e^{a-x} & \text{pro } x > a \\ 0 & x \leq a \end{cases}$

Pozn.: $e^{a-x} = e^{-(x-a)}$



odhadněte parametr a metodou momentů a MMV.

$$EX = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{a}^{\infty} x e^{a-x}\,dx = \ldots = a+1$$

$$\bar{x} = \frac{1+2+2+\ldots+4}{7} = 2{,}43 \Rightarrow \hat{a}+1 = 2{,}43 \Rightarrow \hat{a} = 1{,}43 \Rightarrow$$

nemá řešení (na $(-\infty, a)$)

→ ale 1 vyhovuje

MMV: 1.) $L(a) = e^{a-1} \cdot e^{a-2} \cdot e^{a-2} \cdots e^{a-4} = e^{7a-17}$

2.) $l(a) = \ln L(a) = 7a - 17$

3.) $l'(a) = 7$

$4)\ 7 = 0$? →