

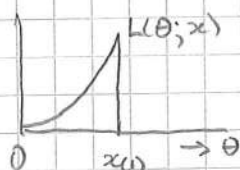
Tentamen Algemene Statistiek 17-12-2014

1. a) $p = \sup_{\theta \in \Theta} P(T \geq t)$ als éénzijdig, of $p = 2 \min(p_r, p_e)$ tweezijdig.
In woorden: de p-waarde is het supremum over alle mogelijkheden onder H_0 op de geobserveerde uitkomst of nog extremer.
Gebruik: als p heel klein is, $p < \alpha_0$, dan wordt H_0 verworpen.
- b) Nee. Als $\pi(\theta; K_1) > \pi(\theta; K_2)$ voor alle θ , dan heeft K_1 een hogere onbetrouwbaarheid, $\sup_{\theta \in \Theta_0} \pi(\theta; K_1)$. Dat is onwenselijk.
Voor $\theta \in \Theta_1$ is de ongelijkheid wel in het voordeel van K_1 .
- c) $\lambda(X) = \frac{\sup_{\theta \in \Theta} P_{\theta}(X)}{\sup_{\theta \in \Theta_0} P_{\theta}(X)} = \frac{P_{\hat{\theta}_{ML}}(X)}{P_{\hat{\theta}_0}(X)}$. Als $\lambda(X)$ groot is duidt dat erop dat in Θ_1 een hogere likelihoodwaarde mogelijk is. Alleen verworpen voor grote waarden. Kleine waarden van $\lambda(X)$ geven aan dat de likelihood niet veel hoger kan dan in Θ_0 .

2. $X_1, \dots, X_n \sim P_{\theta}(x) = \frac{3\theta^3}{x^4} \mathbf{1}_{x \geq \theta}, \theta > 0.$

a) $L(\theta; x) = \prod_{i=1}^n \frac{3\theta^3}{x_i^4} \mathbf{1}_{x_i \geq \theta} = 3^n \theta^{3n} \frac{1}{\prod x_i^4} \mathbf{1}_{\theta \leq x_{(1)}}$

schets:



Dus een randmaximum in

$$\hat{\theta}_{ML} = X_{(1)}.$$

b) $MSE(\theta; \bar{X}) = \text{Var}_{\theta}(\bar{X}) + (E\bar{X} - \theta)^2 = \frac{1}{n} \text{Var}_{\theta} X_1 + (EX_1 - \theta)^2$

$$EX_1 = \int_{\theta}^{\infty} \frac{3\theta^3}{x^4} x dx = -\frac{3\theta^3}{2} x^{-2} \Big|_{\theta}^{\infty} = \frac{3}{2} \theta$$

$$EX_1^2 = \int_{\theta}^{\infty} \frac{3\theta^3}{x^4} x^2 dx = -3\theta^3 x^{-1} \Big|_{\theta}^{\infty} = 3\theta^2$$

$$\begin{aligned} MSE(\theta; \bar{X}) &= \frac{1}{n} (EX_1^2 - (EX_1)^2) + (EX_1 - \theta)^2 = \frac{1}{n} (3\theta^2 - \frac{9}{4}\theta^2) + \left(\frac{\theta}{2}\right)^2 \\ &= \frac{3\theta^2}{4n} + \frac{\theta^2}{4} \end{aligned}$$

$$\begin{aligned} MSE(\theta; \frac{2}{3}\bar{X}) &= \text{Var}_{\theta}(\frac{2}{3}\bar{X}) + (E(\frac{2}{3}\bar{X}) - \theta)^2 = \frac{4}{9} \text{Var}_{\theta} \bar{X} + 0 \\ &= \frac{4}{9} \frac{3\theta^2}{4n} = \frac{\theta^2}{3n} \end{aligned}$$

(2)

2.

c) De schatter met de kleinste m.s.E voor alle θ verdient de voorkeur. Hier geldt $\frac{\theta^2}{3n} < \frac{3\theta^2}{4n} + \frac{\theta^2}{4}$ voor alle θ , dus $\frac{2}{3}\bar{X}$ is beter.

3.

$X_1, \dots, X_n \sim p_\theta(x) = \frac{2x}{\theta^2} \mathbb{1}_{[0, \theta]}(x)$

$$\begin{aligned} \text{a.) } P(X_{(n)} \leq u) &= P(X_1 \leq u, X_2 \leq u, \dots, X_n \leq u) = (P(X_1 \leq u))^n \\ &= \left(\int_0^u \frac{2x}{\theta^2} dx \right)^n = \left(\frac{x^2}{\theta^2} \Big|_0^u \right)^n = \left(\frac{u^2}{\theta^2} \right)^n = \frac{u^{2n}}{\theta^{2n}}. \end{aligned}$$

b.) $T = T(X, \theta)$ is een pivot als de verdeling van T bekend is, en dus niet van θ afhangt.

$$\text{c.) } T = \frac{X_{(n)}}{\theta}. \quad P(T \leq t) = P\left(\frac{X_{(n)}}{\theta} \leq t\right) = P(X_{(n)} \leq t\theta) = \frac{(t\theta)^{2n}}{\theta^{2n}} = t^{2n}, \quad t \in [0, 1].$$

Hangt dus niet af van θ . Dus $X_{(n)}/\theta$ is een pivot.

d.) Kies een a en b zodat $P(a \leq T \leq b) = 0.95$.

$$\text{ofwel } b^{2n} - a^{2n} = 0.95 \text{ met } a, b \in (0, 1]$$

$$\begin{aligned} \text{Dan omschrijven } 0.95 &= P(a \leq \frac{X_{(n)}}{\theta} \leq b) = P\left(\frac{1}{b} \leq \frac{\theta}{X_{(n)}} \leq \frac{1}{a}\right) \\ &= P\left(\frac{X_{(n)}}{b} \leq \theta \leq \frac{X_{(n)}}{a}\right). \text{ Neem } a, b \neq 0. \end{aligned}$$

4.

$X_1, \dots, X_n \sim p_\theta(x) = \theta(1-\theta)^{x-1}, \quad \theta \in (0, 1), \quad x \in \{1, 2, \dots\} \quad E_\theta X_1 = \frac{1}{\theta}, \quad \text{Var}_\theta X_1 = \frac{1-\theta}{\theta^2}$

a.) Mbt factorisatie in exponentiële familie:

$$p_\theta(x_1, \dots, x_n) = \prod_{i=1}^n \theta(1-\theta)^{x_i-1} = \theta^n (1-\theta)^{-n} (1-\theta)^{\sum x_i} = c(\theta) e^{\sum x_i \log(1-\theta)}$$

Definieer $V(X) = \sum X_i$ en $Q(\theta) = \log(1-\theta)$.

De verzameling $\{\log(1-\theta); \theta \in (0, 1)\} = (-\infty, 0) \subset \mathbb{R}$

heeft een inwendig punt binnen $\mathbb{R}$. Dus $V(X)$ is v.d.t.v.

b.) Een umvz-schatter baseren op $V(X)$.

$$\begin{aligned} E_\theta(V(X)) &= E_\theta(\sum X_i) = n E_\theta X_1 = \frac{n}{\theta}. \text{ Dus } \frac{1}{n} \sum X_i = \bar{X} \text{ is zuiver} \\ &\text{voor } \frac{1}{\theta} \text{ en gebaseerd op v.d.t.v. statistiek, dus } \bar{X} \text{ is umvz.} \end{aligned}$$

$$\begin{aligned} \text{c.) } i_\theta &= -E_\theta(\ddot{\ell}_\theta(X_1)) = -E_\theta \frac{\partial^2}{\partial \theta^2} \log p_\theta(X_1) = -E_\theta \frac{\partial^2}{\partial \theta^2} [\log \theta + (X_1 - 1) \log(1-\theta)] \\ &= -E_\theta \frac{\partial}{\partial \theta} \left[\frac{1}{\theta} - \frac{X_1 - 1}{1-\theta} \right] = -E_\theta \left[-\frac{1}{\theta^2} - \frac{X_1 - 1}{(1-\theta)^2} \right] = \frac{1}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} \\ &= \frac{1}{\theta^2} + \frac{1-\theta}{\theta(1-\theta)^2} = \frac{1}{\theta^2} + \frac{1}{\theta(1-\theta)} = \frac{1-\theta}{\theta^2(1-\theta)} + \frac{\theta}{\theta^2(1-\theta)} = \frac{1}{\theta^2(1-\theta)}. \end{aligned}$$

(3)

$$d.) \text{ CRB: } \text{Var}_\theta T \geq \frac{(g'(\theta))^2}{n i_\theta} = \frac{(-1/\theta^2)^2}{n/\theta^2(1-\theta)} = \frac{\theta^2(1-\theta)}{n\theta^4} = \frac{1-\theta}{n\theta^2}.$$

De variantie van $\bar{X}$ is $\text{var}_\theta \bar{X} = \frac{1}{n} \text{Var}_\theta X_1 = \frac{1-\theta}{n\theta^2}$. Dus de CRB wordt gehaald en is scherp.

5. a) $X_i = \text{uitgave } i^{\text{e}} \text{ persoon vorig jaar}$ $X_i \sim N(\mu, \sigma^2)$ $\mu \in \mathbb{R}, \sigma^2 > 0$
 $Y_j = \text{ " " } j^{\text{e}} \text{ " dit jaar}$ $Y_j \sim N(\nu, \tau^2)$ $\nu \in \mathbb{R}, \tau^2 > 0$

$$H_0: \mu \geq \nu \text{ vs } H_1: \mu < \nu$$

b) $T = \frac{\bar{X} - \bar{Y}}{S_{X,Y} \sqrt{\frac{1}{m} + \frac{1}{n}}} \stackrel{\mu=\nu}{\sim} t_{m+n-2}$ als $\sigma^2 = \tau^2$. Aanname: $\tau^2 = \sigma^2$

$$m+n-2 = 17+15-2 = 30.$$

$$S_{X,Y}^2 = (14 \cdot (3.26)^2 + 16 \cdot (3.34)^2) / 30 = (3.30)^2$$

$$K_T = (-\infty, t_{30;0.05}] = (-\infty, -1.7]$$

Uitrekenen: $T = \frac{22.73 - 24.58}{3.30 \sqrt{\frac{1}{15} + \frac{1}{17}}} = -1.58$

T ligt niet in K_T , dus niet genoeg bewijs om H_0 te verwerpen.

Conclusie: we kunnen niet concluderen dat Nederlanders dit jaar significant meer uitgeven aan kerstcadeaus.