

1. (a) Voor zuiverheid moet gelden $E_\theta cX_{(n)} = \theta$ voor alle $\theta > 0$.

Verder geldt

$$E_\theta X_{(n)} = \int_0^\theta x \frac{2nx^{2n-1}}{\theta^{2n}} dx = \frac{2n}{2n+1} \theta$$

Dus $c = \frac{2n+1}{2n}$.

- (b) Wegens $E\bar{X}_n = \frac{1}{n} \sum_{i=1}^n EX_i = EX_1$ en $EX_1 = \int_0^\theta x \frac{2x}{\theta^2} dx = \frac{2}{3}\theta$ geldt

$$E_\theta \frac{3}{2} \bar{X}_n = \theta \quad \text{voor alle } \theta > 0$$

- (c) Als één van beide schatters voor iedere $\theta > 0$ een kleinere (of gelijke) mean square error heeft, dan heeft deze schatter onze voorkeur.

Omdat beide schatters zuiver zijn, is het voldoende de variantie uit te rekenen. Er geldt natuurlijk $\text{var}_\theta(T) = E_\theta T^2 - (E_\theta T)^2$.

T_1 : Bereken $E_\theta T_1^2$:

$$E_\theta \left[\frac{2n+1}{2n} X_{(n)} \right]^2 = \left(\frac{2n+1}{2n} \right)^2 \frac{2n}{2n+2} \theta^2$$

Invullen geeft

$$\text{mse}_\theta(T_1) = \frac{1}{4n^2 + 4n} \theta^2$$

T_2 :

$$\text{mse}_\theta(T_1) = \text{var}_\theta\left(\frac{3}{2} \bar{X}_n\right) = \frac{9}{4} \frac{1}{n} \text{var}_\theta(X_1) = \frac{9}{4n} \frac{1}{18} \theta^2 = \frac{1}{8n} \theta^2$$

$4n^2 + 4n \geq 8n$ met gelijkheid als $n = 1$, dus T_1 heeft de voorkeur

2. (a) De aannemelijkheidsfunctie op basis van de waarnemingen $x_1, \dots, x_n$ is:

$$\theta \mapsto L(\theta; x_1, \dots, x_n) = \theta^n \prod_{i=1}^n x_i^{\theta-1}$$

Dit is een differentieerbare functie in θ . De afgeleide van de log-aannemelijkheidsfunctie is:

$$\frac{n}{\theta} + \sum_{i=1}^n \log x_i$$

Afgeleide nul stellen en controleren of extreem een maximum is geeft als meest aannemelijke schatter voor θ :

$$T_1 = \hat{\theta} = -\frac{n}{\sum_{i=1}^n \log X_i} = -\frac{1}{\overline{\log X_i}}$$

- (b) De momentenschatting wordt verkregen door oplossen van de volgende vergelijking in θ

$$E_{\theta} X_i = \bar{X}_n$$

Uitrekenen geeft $E_{\theta} X_i = \frac{\theta}{\theta+1}$, dus $\frac{T_2}{T_2+1} = \bar{X}_n$, oftewel

$$T_2 = \frac{\bar{X}_n}{1 - \bar{X}_n}$$

- (c) Stel $X = (X_1, \dots, X_n)$. Θ heeft (a-priori) dichtheid $\pi(\theta) = e^{-\theta} 1_{\{\theta > 0\}}(\theta)$. De Bayes-schatting wordt verkregen via

$$T_3 = E(\bar{\Theta} | X = x) = \frac{\int_0^{\infty} \theta \prod_{i=1}^n f_{\theta}(x_i) \pi(\theta) d\theta}{\int_0^{\infty} \prod_{i=1}^n f_{\theta}(x_i) \pi(\theta) d\theta} = \frac{\int_0^{\infty} \theta \theta^n \prod_{i=1}^n x_i^{\theta-1} e^{-\theta} d\theta}{\int_0^{\infty} \theta^n \prod_{i=1}^n x_i^{\theta-1} e^{-\theta} d\theta}$$

Uitrekenen van de teller geeft

$$\frac{(n+1)!}{(1 - \sum_{i=1}^n \log x_i)^{(n+2)} (\prod_{i=1}^n x_i)}$$

en van de noemer geeft

$$\frac{(n)!}{(1 - \sum_{i=1}^n \log x_i)^{(n+1)} (\prod_{i=1}^n x_i)}$$

Dus

$$T_3 = \frac{n+1}{1 - \sum_{i=1}^n \log X_i}$$

3. $X_1, \dots, X_{16} \sim N(\mu, \sigma^2)$. We willen toetsen:

$H_0 : \mu = 12$ tegen $H_1 : \mu \neq 12$

Als toetsingsgrootte nemen we

$$T = 4 \frac{\bar{X}_{16} - 12}{S_X}$$

Merk op dat de variantie na de reorganisatie veranderd kan zijn.

Onder H_0 heeft T een Student-verdeling met 15 vrijheidsgraden.

Het kritieke gebied is: $K = (-\infty, -t_{15;0.025}] \cup [t_{15;0.025}, \infty) = (-\infty; -2,13] \cup [2,13; \infty)$

De waarde van de toetsingsgrootte is:

$$t = 4 \frac{11,92 - 12}{\sqrt{\frac{1}{15} 3,78}} = -0,638$$

Dit zit niet in K , dus we verwerpen H_0 niet.

4. (a) Een werkbaar model is:

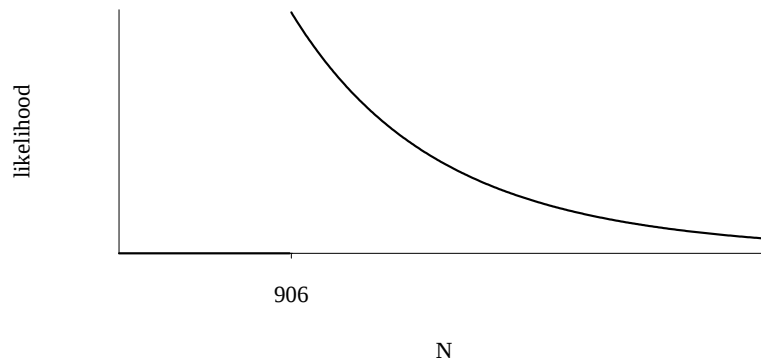
$\mathcal{P} = \{P_N : P_N \text{ kansverdeling van het 10 keer trekken met teruglegging van een getal uit } 1, \dots, N\}$

We nemen hierbij aan dat de waarnemingen van de taxinummers onderling onafhankelijk zijn. Dit is mogelijk niet geheel realistisch, omdat een taxi die zojuist is gepasseerd, even verderop een passagier kan hebben laten uitstappen of kan hebben opgepikt, waarna hij met grotere kans weer langs komt rijden.

(b)

$$L(N; \text{data}) = P_N(234, 60, 10, 125, 502, 906, 568, 422, 588, 357) = \begin{cases} 0 & \text{als } N \leq 905 \\ \left(\frac{1}{N}\right)^{10} & \text{als } N \geq 906 \end{cases}$$

In een plaatje geeft dit:



Dus de meest aannemelijke schatting is 906.

(c) Omdat we H_0 nooit kunnen accepteren, zetten we datgene wat we willen aantonen in de alternatieve hypothese. Dus:

$$H_0 : N \geq 1000 \text{ en } H_1 : N < 1000.$$

(d) $X_{(10)}$ klein geeft indicatie voor klein aantal taxi's, dus we hebben een linker kritiek gebied. Er moet gelden

$$P_N(X_{(10)} \leq k) \leq 0,05 \quad \text{voor alle } N \geq 1000$$

Omdat

$$P_N(X_{(10)} \leq k) = \begin{cases} \left(\frac{k}{N}\right)^{10} & \text{voor } k = 1, 2, \dots, N \\ 0 & \text{elders} \end{cases}$$

zien we dat moet gelden

$$\left(\frac{k}{N}\right)^{10} \leq 0,05 \quad \text{voor alle } N \geq 1000$$

Het is voldoende om naar $N = 1000$ te kijken. We vinden dan $k \leq \sqrt[10]{0,05} N = 741,1$, m.a.w.

$$K = \{1, 2, \dots, 741\}$$

Omdat $906 \notin K$, verwerpen we H_0 niet.

(e) Op dezelfde wijze vinden we met X_{10} als toetsingsgrootheid het kritieke gebied

$$K^* = \{1, 2, \dots, 50\}$$

Voor het onderscheidend vermogen vinden we nu:

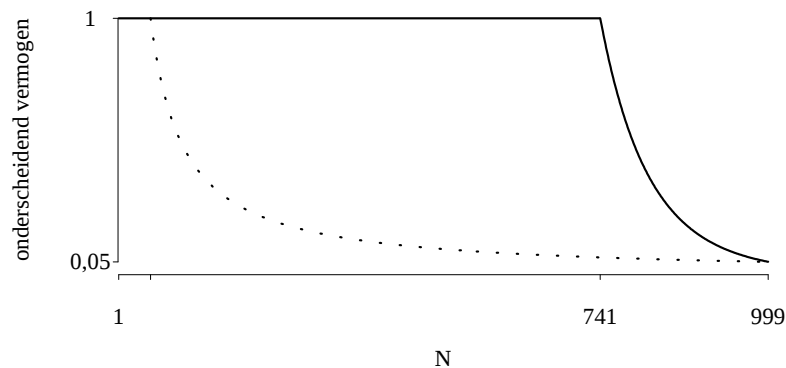
Voor $T = X_{(10)}$:

$$\pi(N; K) = P_N(X_{(10)} \leq 741) = \begin{cases} \left(\frac{741}{N}\right)^{10} & \text{voor } N \geq 741 \\ 1 & \text{voor } N \leq 741 \end{cases}$$

Voor $T = X_{10}$:

$$\pi(N; K^*) = P_N(X_{10} \leq 50) = \begin{cases} \left(\frac{50}{N}\right) & \text{voor } N \geq 50 \\ 1 & \text{voor } N \leq 50 \end{cases}$$

In een plaatje geeft dit:



De gestreepte lijn is voor $T = X_{10}$, de andere lijn voor $T = X_{(10)}$. Omdat de toets op basis van $X_{(10)}$ voor iedere $N \leq 999$ een onderscheidend vermogen heeft dat groter is dan of gelijk is aan het onderscheidend vermogen voor de toets op basis van X_{10} , geven we de voorkeur aan $X_{(10)}$ als toetsingsgrootheid.

(f) Overschrijdingskans:

$$\sup_{N \geq 1000} P_N(X_{(10)} \leq 1091) = P_{1000}(X_{(10)} \leq 1091) = 1$$

Normering:

1(a): 5	2(a): 9	3:12	2(a): 4
1(b): 4	2(b): 9		2(b): 7
1(c):10	2(c): 9		2(c): 3
			2(d): 8
			2(e): 8
			2(f): 2

Eindcijfer = (totaal + 10)/10