

1. (a) Laat X_i : uitloop hoorcollege i . We nemen aan dat de uitloop normaal verdeeld is. Dan is een kansmodel gegeven door
 $\mathcal{P}_{\mu,\sigma} = \{ \text{kansverdeling van steekproef } X_1, \dots, X_{13} \sim \mathcal{N}(\mu, \sigma^2); \mu \in \mathbb{R}, \sigma > 0 \}$.
 (Merk op: X_i onderling onafhankelijk.)
 Nulhypothese: $H_0 : \mu \leq 0$ tegen alternatieve hypothese $H_1 : \mu > 0$ (of $H_0 : \mu = 0$ tegen $H_1 : \mu \neq 0$).
 (b) Er geldt $\bar{x}_{13} = 2.3946$ en $s_X = 1.7173$. Verder geldt

$$\mathbb{P}_{\mu}\{\sqrt{13}(\bar{X}_{13} - \mu)/S_X \leq t_{12,0.05}\} = 0.95$$

Dit levert als 95% betrouwbaarheidsinterval

$$\left[\bar{x}_{13} - t_{12,0.05} \times \frac{s_X}{\sqrt{13}}, \infty \right)$$

Invullen levert $[2.3946 - 1.78 \times 1.7173/3.6056, \infty) = [1.55, \infty)$

Baseren we het betrouwbaarheidsinterval op $H_0 : \mu = 0$ tegen $H_1 : \mu \neq 0$, dan krijgen we

$$\left[\bar{x}_{13} - t_{12,0.025} \times \frac{s_X}{\sqrt{13}}, \bar{x}_{13} + t_{12,0.025} \times \frac{s_X}{\sqrt{13}} \right] = [1.36, 3.43]$$

- (c) De waarde 0 zit niet in het 95% betrouwbaarheidsinterval, dus we concluderen dat het college gemiddeld uitloopt.
2. (a) Onder de nulhypothese is $X_i \sim \mathcal{N}(1, 1)$.
 De grootheid $\sqrt{n}(\bar{X}_n - 1)$ is de gestandaardiseerde stochast, dus $\mathcal{N}(0, 1)$ -verdeeld. Het kwadraat van deze grootheid is dus chikwadraat verdeeld met één vrijheidsgraad.
- (b) $\bar{X}$ en S_X^2 zijn onderling onafhankelijk. $(n-1)S_X^2$ heeft een chikwadraat verdeling met $n-1$ vrijheidsgraden (de variantie is hier gelijk aan één). T_n is dus de som van twee onderling onafhankelijke stochastische variabelen met, onder de nulhypothese, een chikwadraat verdeling met respectievelijk 1 en $n-1$ vrijheidsgraden. Onder de nulhypothese heeft T_n derhalve een chikwadraat verdeling met n vrijheidsgraden.
- (c) We verwerpen voor grote waarden van T_{30} . Het bovenste 0.90 kwantiel van de χ_{30}^2 verdeling is 40.3, dus het kritieke gebied is $[40.3, \infty)$.

3. (a) Er geldt

$$\int_0^1 x \theta x^{\theta-1} dx = \frac{\theta}{\theta+1}.$$

Gelijkstellen van de verwachting aan het steekproefgemiddelde $\bar{x}_n$ levert

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

- (b) De aannemelijkheidsfunctie op basis van de waarnemingen $x_1, \dots, x_n$ is:

$$\theta \mapsto L(\theta; x_1, \dots, x_n) = \theta^n \prod_{i=1}^n x_i^{\theta-1} 1_{(0,1]}(x_i)$$

Deze functie is differentieerbaar 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 :

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

- (c) We kunnen de kansdichtheid van de steekproef herschrijven als een 1-parameter exponentiële familie:

$$\prod_{i=1}^n f_{\theta}(x_i) = \exp\left\{(\theta - 1) \sum_{i=1}^n \log x_i + n \log \theta\right\} \prod_{i=1}^n 1_{(0,1]}(x_i)$$

Omdat $\{\theta - 1 \mid \theta > 0\}$ een niet-leeg inwendige bevat, is $\sum_{i=1}^n \log X_i$ een volledige en voldoende stochast.

- (d) Een functie van $\sum \log X_i$ die een zuivere schatter is voor $1/\theta$ is de UMVU schatter. Er geldt, met $y = \log x$:

$$\begin{aligned} E(\log X_i) &= \int_0^1 \log x \theta x^{\theta-1} dx \\ &= \int_{-\infty}^0 y \theta \exp\{(\theta - 1)y\} \exp\{y\} dy \\ &= [y \exp\{\theta y\}]_{-\infty}^0 - \int_{-\infty}^0 \exp\{\theta y\} dy = -\frac{1}{\theta} \end{aligned}$$

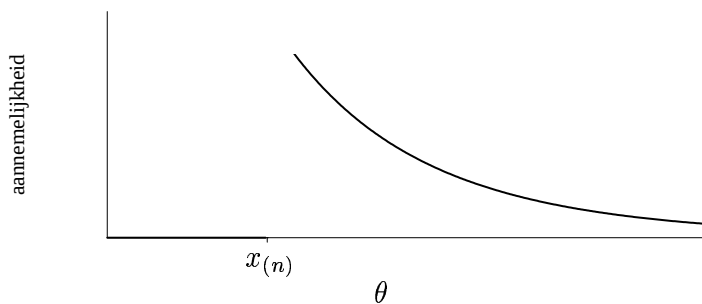
Dus $T = -\frac{1}{n} \sum_{i=1}^n \log X_i$ is zuiver voor $1/\theta$ en derhalve de UMVU schatter.

- (e) Nee, het is mogelijk dat er een niet-zuivere schatter is met een kleinere MSE.

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

$$\theta \mapsto L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n \frac{2x_i}{\theta^2} 1_{[0,\theta]}(x_i) = \frac{2^n}{\theta^{2n}} \prod_{i=1}^n x_i 1_{[x_{(n)}, \infty)}(\theta)$$

In een plaatje geeft dit:



Omdat de functie maximaal is voor $\theta = x_{(n)}$, volgt het beweerde.

(b) Er geldt voor $0 \leq x \leq \theta$:

$$\begin{aligned}\mathbb{P}(X_{(n)} \leq x) &= \prod_{i=1}^n \mathbb{P}(X_i \leq x) \\ &= \left[\int_0^x \frac{2x}{\theta^2} dx \right]^n = \frac{x^{2n}}{\theta^{2n}}\end{aligned}$$

Analoog geldt voor $x < 0$: $\mathbb{P}(X_{(n)} \leq x) = 0$ en voor $x > \theta$: $\mathbb{P}(X_{(n)} \leq x) = 1$.
Differentiëren levert nu het beweerde.

(c) Voor de likelihood ratio toetsingsgrootte (LRT) geldt:

$$\begin{aligned}\text{LRT} &= \frac{\sup_{\theta > 0} \prod_{i=1}^n f_{\theta}(X_i)}{\prod_{i=1}^n f_{\theta=1}(X_i)} \\ &= \frac{\prod_{i=1}^n f_{X_{(n)}}(X_i)}{\prod_{i=1}^n f_1(X_i)} \\ &= \frac{1}{(X_{(n)})^{2n}}\end{aligned}$$

(d) We verdelen de onbetrouwbaarheidsdrempel over het linker en rechter kritieke gebied.

$\mathbb{P}_{\theta=1}(X_{(n)} \leq k_1) = k_1^{20} = 0.025$ levert $k_1 = 0.8315665$

en $\mathbb{P}_{\theta=1}(X_{(n)} \leq k_2) = k_2^{20} = 0.975$ levert $k_2 = 0.9987349$.

Dus het kritieke gebied heeft de vorm $K = (-\infty, 0.832] \cup [0.999, \infty)$.

(e) Het onderscheidend vermogen $\omega(\theta)$ wordt gegeven door $\omega(\theta) = \mathbb{P}_{\theta}(X_{(n)} \in K)$.

Er geldt

$$\mathbb{P}_{\theta}(X_{(n)} \in K) = \begin{cases} 1 & \text{als } \theta \in (0, k_1] \\ (k_1/\theta)^{20} = 0.025 \theta^{-20} & \text{als } \theta \in (k_1, k_2] \\ (k_1/\theta)^{20} + 1 - (k_2/\theta)^{20} = 1 - 0.95 \theta^{-20} & \text{als } \theta \in (k_2, \infty) \end{cases}$$

De grafiek van het onderscheidend vermogen komt er als volgt uit te zien:

