Uitwerking Tentamen Wiskundige Statistiek voor TW (191530382)
Vrijdag 1 november 2013, 13.45-16.45 uur

Opgave 1

a. $E(T_1) = E\left(\frac{\bar{X}}{\sqrt{2}} - \frac{\bar{Y}}{\sqrt{3}}\right) = \frac{E(\bar{X})}{\sqrt{2}} - \frac{E(\bar{Y})}{\sqrt{3}} = \frac{\mu_X}{\sqrt{2}} - \frac{\mu_Y}{\sqrt{3}}$, dus $T_1$is geen zuivere schatter van $\mu_X - \mu_Y$
$E(T_1) = E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$, ongeacht de steekproefomvangen voor X en Y, dus zowel T_2 en T_3 zijn zuivere schatters.

b. Omdat T_2 en T_3zuivere schatters zijn is de verwachte kwadratische fout gelijk aan de variantie:
$var(\bar{X} - \bar{Y}) = var(\bar{X}) + var(\bar{Y}) = \frac{2}{m} + \frac{3}{n}$
$var(T_2) = \frac{2}{10} + \frac{3}{10} = \frac{1}{2}$ en $var(T_3) = \frac{2}{9} + \frac{3}{11} = \frac{49}{99}$.
T_3 heeft de kleinste variantie (MSE) en is dus het beste.

c. Als we de 2*n* waarnemingen gelijk over de twee steekproeven verdelen, is volgens a. $\bar{X} - \bar{Y}$ zuiver en volgens b is $MSE(\bar{X} - \bar{Y}) = var(\bar{X} - \bar{Y}) = \frac{2}{n} + \frac{3}{n} = \frac{5}{n}$
$\lim_{n\to\infty} MSE(\bar{X} - \bar{Y}) = \lim_{n\to\infty} \frac{5}{n} = 0$, dus $\bar{X} - \bar{Y}$ is een consistente schatter.

Opgave 2

a. $f_\theta(x) = \frac{1}{10-\theta}$ voor $\theta \le x \le 10$ en $f_\theta(x) = 0$ buiten het interval.
$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \left(\frac{1}{10-\theta}\right)^n$, met voor alle $i, \theta \le x_i \le 10$, dus $0 \le \theta \le \min(x_1, \dots, x_n) \le 10$
$L(\theta)$ is een stijgende functie ($L'(\theta) = n(10-\theta)^{-n-1} > 0$ voor alle mogelijke θ) en dus maximaal voor de grootst mogelijke waarde van θ, $\min(x_1, \dots, x_n)$.
$\theta^* = \min(X_1, \dots, X_n)$ is de m.a. schatter van θ.

b. Model: $X_1, \dots, X_{15}$ zijn o.o. en alle U$(\theta, 10)$-verdeeld
Toets H_0: $\theta \ge 5.5$ tegen H_1: $\theta < 5.5$ met $\alpha = 0.05$.
We verwerpen voor kleine $\Lambda = \frac{\sup_{\theta\in\Theta_0} \prod_{i=1}^{15} g(x_i,\theta)}{\sup_{\theta\in\Theta} \prod_{i=1}^{15} g(x_i,\theta)}$
Voor de noemer hebben we het maximum in a. bepaald: $\sup_{0\le\theta\le10}(10-\theta)^{-15} = (10-\theta^*)^{-15}$
Voor de teller geldt:

$\sup_{\theta\ge5.5} \prod_{i=1}^{15} g(x_i,\theta) = \begin{cases} 0 & \text{als} \min(x_1, \dots, x_{15}) \le 5.5 \ (\,bijv.\ g(x, 5.5) = 0 \ als\ x < 5.5) \\ (10-\theta^*)^{-15} & \text{als} \min(x_1, \dots, x_{15}) > 5.5 \end{cases}$

Dus $\Lambda = \begin{cases} 0 & \text{als } \min(X, \dots, X_{15}) \le 5.5 \\ 1 & \text{als } \min(X, \dots, X_{15}) \le 5.5 \end{cases}$ $\rightarrow \Lambda$ is klein als $\Lambda = 0$, ofwel als $\min(X, \dots, X_{15}) \le 5.5$
Dus als $\min(X_1, \dots X_n) \le 5.5$ als kritiek gebied wordt gekozen, is de onbetrouwbaarheid $P(\min(X_1, \dots, X_n) \le 5.5 | H_0) = 0$. Immers: als H_0: $\theta \ge 5.5$ waar is, is de kans 0 op een minimum < 5.5
Daar $\alpha_0 = 0.05$ > 0, ligt de waargenomen $\min(x_1, \dots, x_{15}) = 4.8$ in het KG: H_0 verwerpen.
Bij elke onbetrouwbaarheidsdrempel (dus ook voor $\alpha_0 = 0.05$) is aangetoond dat meedoen niet automatisch slagen is.
(<u>*Opmerking*</u>: *voor* $\alpha_0 = 0.05$ *kunnen we voor deze linkseenzijdige toets eventueel het KG van de vorm* $min(X_1, \dots X_n) \le c$: $P_{H_0}(min(X_1, \dots X_n) \le c) \le 0.05 \rightarrow$
$P_{H_0}(min(X_1, \dots X_n) \ge c) = P_{H_0}(X_1 \ge c)^{15} \ge \left(\frac{10-c}{4.5}\right)^{15} \ge 0.95$, *waaruit* $c \approx 5.52$ *volgt)*

c. Als p de kans op slagen is, toetsen we H_0: $p = 0.9$ tegen H_1: $p < 0.9$ met $\alpha_0 = 0.05$ en toetsingsgrootheid *X* = "aantal voldoendes in de steekproef van *n* =15 studenten", die onder H_0 $B(15, 0.9) - verdeeld\ is$.
X = 13, dus verwerp H_0 als de overschrijdingskans
$P(X \le 13 | p = 0.9) = 1 - P(X = 14\ of\ X = 15 | p = 0.9) = 1 - 15 \cdot 0.9^{14} 0.1 - 0.9^{15} = \ < 0.05$
Dit is niet het geval dus H_0 niet verwerpen: bij onbetrouwbaarheidsdrempel 5% is niet aangetoond dat minder dan 90% slaagt.

Opgave 3

a. $\bar{x} = 23.557$ (en s ≈ 2.803), met de rekenmachine

b. $95\% - BI(\mu) = \left(\bar{x} - c\frac{s}{\sqrt{n}}, \bar{x} + c\frac{s}{\sqrt{n}}\right) = (23.557 - 2.26 \times \frac{2.803}{\sqrt{10}}, 23.557 + 2.26 \times \frac{2.803}{\sqrt{10}})$

$$\approx (21.6, 25.6)$$

waarin $P(T(10-1) \le c) = 0.975$, dus $c = 2.26$

c. In het interval in b. liggen 6 van de 10 waarnemingen, dus niet 0 zoals de opgave suggereert: 60% is minder dan 95% omdat het interval betrekking heeft op de verwachte (gemiddelde) afstand.
(Bij grotere aantallen waarnemingen wordt het interval steeds smaller en zal dus een steeds kleiner percentage waarnemingen bevatten, daar de spreiding in de waarnemingen niet kleiner wordt)

Opgave 4

a. We toetsen $H_0: \mu = 3\ tegen\ H_1: \mu < 3$. Onder H_0 is μ = 3 en $\sigma^2 = \frac{1}{\lambda^2} = 9$

De toetsingsgrootheid $\overline{X}$ is onder H_0 volgens de CLS (grote n = 2000) bij benadering $N\left(3, \frac{9}{2000}\right)$.

Dus verwerp H_0 als $\overline{X} \le c, met$

$$P(\overline{X} \le c | \mu = 3) = P\left(\frac{\overline{X}-3}{\sqrt{\frac{9}{2000}}} \le \frac{c-3}{\sqrt{\frac{9}{2000}}} \middle| \mu = 3\right) = \Phi\left(\frac{c-3}{\sqrt{\frac{9}{2000}}}\right) \le 0.001, dus\ \frac{c-3}{\sqrt{\frac{9}{2000}}} = -3.10$$

Ofwel: $c = 3 - 3.10\sqrt{\frac{9}{2000}} \approx 2.79$

We nemen waar: $\overline{x} = 2.74$, dus H_0 verwerpen: bij een onbetrouwbaarheid van 0.1% is aangetoond dat de levensduur van de vissen korter is dan 3 jaar.

b. De toetsingsgrootheid M_2 is onder H_0 volgens de CLS (grote *n* = 2000) bij benadering $N\left(18, \frac{20\times 81}{2000}\right)$.

Omdat $onder\ H_0\ E(X^2) = \frac{2}{\lambda^2} = 18$ en $var(X^2) = \frac{20}{\lambda^2} = 20 \times 81$

Dus verwerp H_0 als $M_2 \le c_2, met$

$$P(M_2 \le c_2 | \mu = 3) = P\left(\frac{M_2-18}{\sqrt{\frac{81}{100}}} \le \frac{c_2-18}{\sqrt{\frac{81}{100}}} \middle| \mu = 3\right) = \Phi\left(\frac{c_2-18}{\sqrt{\frac{81}{100}}}\right) \le 0.001, \quad \text{dus} \quad \frac{c_2-18}{\sqrt{\frac{81}{100}}} = 3.10$$

Ofwel: $\boldsymbol{c_2} = 18 - 3.10 \times 0.9 \approx 15.21$

We nemen waar: $m_2 = 15.33$, dus H_0 **niet** verwerpen: bij een onbetrouwbaarheid van 0.1% is niet aangetoond dat de levensduur van de vissen korter is dan 3 jaar.

c. Als $\mu = 2.5$, is $\lambda = \frac{1}{2.5}$,

dus var(*X*) $= \frac{2}{\lambda^2} = 6.25$, $E(X^2) = \frac{2}{\lambda^2} = 12.5$ en $var(X^2) = \frac{20}{\lambda^4} = 20 \times 2.5^4 = 781.25$

Dus de toetsingsgrootheid $\overline{X}$ is dan volgens de CLS (grote *n* = 2000) bij benadering $N\left(2.5, \frac{6.25}{2000}\right)$

Het onderscheidend vermogen is dan

$$P(\overline{X} \le c | \mu = 2.5) = P\left(\frac{\overline{X}-2.5}{\sqrt{\frac{6.25}{2000}}} \le \frac{2.79-2.5}{\sqrt{\frac{6.25}{2000}}} \middle| \mu = 2.5\right) = \Phi(5.19) \approx 100.00\%$$

De toetsingsgrootheid M_2 is, indien $\mu = 2.5$, volgens de CLS bij benadering $N\left(12.5, \frac{781.25}{2000}\right)$.

$$P(M_2 \le c_2 | \mu = 2.5) = P\left(\frac{M_2-12.5}{\sqrt{\frac{781.25}{2000}}} \le \frac{15.21-12.5}{\sqrt{\frac{781.25}{2000}}} \middle| \mu = 2.5\right) = \Phi(4.37) \approx 100\%$$

Dus de toets met $\overline{X}$ in a. heeft het laagste onderscheidend vermogen: $\Phi(5.19) < \Phi(4.37)$