Uitwerking Oefententamen Wiskundige Statistiek (153038)-Oktober 2007

Opgave 1

a. 90%-BI(σ^2) = $\left(\frac{(n-1)s^2}{c_2}, \frac{(n-1)s^2}{c_1}\right) = \left(\frac{440}{17.5}, \frac{440}{2.18}\right) \approx (28.4,\ 161.2)$

waarin $(n\text{-}1)s^2 = \sum x_i^2 - n\bar{x}^2 = 440 + 0 = 440$ en

$P(W(8) \le c_1) = P(W(8) \ge c_2) = ½\,(1-\gamma) = 0.05$ dus $c_1 = 2.73$ en $c_2 = 15.5$

(opmerking: het 90%-BI(σ) vinden we door de wortel van de intervalgrenzen te nemen: (5.3, 12.7))

b. Ga met een geschikte toets na of de varianties gelijk verondersteld kunnen worden.

1. Model is gegeven: 2 onafhankelijke steekproeven uit normale verdelingen
2. We toetsen H_0: $\sigma^2_1 = \sigma^2_2$ tegen H_1: $\sigma^2_1 \ne \sigma^2_2$ met $\alpha_0 = 0.05$
3. Toetsingsgrootheid T = S_X^2 / S_Y^2
4. $T \cong F(10\text{-}1, 16\text{-}1)$ onder H_0.
5. $s_X^2 = 440/(9-1) = 55$ en $s_Y^2 = \frac{1}{16-1}(\sum y_i^2 - n\bar{y}^2) = \frac{1}{15}(799 - 16 \times 1.75^2) = 50$
   t = 55/50 = 1.1
6. Tweezijdig kritiek gebied: als $T \le c_1$ of $T \ge c_2$, dan H_0 verwerpen.
   $P(F(9, 15) \le c_1) \le ½\,\alpha_0 = 0.025$, ofwel $P(F(15, 9) \ge 1/c_1) \le 0.025$ dus $c_1 = 1/4.10 = 0.24$
   $P(F(9, 15) \ge c_2) \le ½\,\alpha_0 = 0.025$, dus $c_2 = 3.20$
7. t = 1.1 ligt tussen c_1 en c_2, dus H_0 niet verwerpen
8. Met onbetrouwbaarheid 5% kan gesteld worden dat de varianties niet significant verschillen.

Opgave 2

a. Er is hier sprake van paarsgewijs afhankelijke waarnemingen we bepalen dus de (wel o.o.) verschillen $x_1, \ldots, x_{10}$ per paar: -1, +6, -2, +10, +8, +8, +9, +8, -3, 17, zodat $\bar{x} = 6$ en $s^2 = 39\ 1/9$

1. Model: de verschillen (voor – na) $X_1, \ldots, X_{10}$ zijn o.o. en $N(\mu, \sigma^2)$
2. We toetsen H_0: $\mu = 0$ tegen H_1: $\mu > 0$ met $\alpha_0 = 0.05$
3. Toetsingsgrootheid T = $\frac{\bar{X}}{S/\sqrt{n}}$
4. $T \cong T(10\text{-}1)$ onder H_0 .
5. t = $\frac{6}{6.254/\sqrt{10}} \approx 3.03$
6. Als $T \ge c$, dan H_0 verwerpen. $P(T(9) \ge c) = 0.05$ dus c = 1.83
7. t = 3.03 > 1.83 = c, dan H_0 verwerpen.
8. Conclusie: gewichtsverlies is aangetoond met onbetrouwbaarheid van 5%.

b. Een naar beneden begrensd betrouwbaarheidsinterval voor het verwachte gewichtsverlies met $\gamma = 0.95$ bevat die waarden μ_0 waarvoor H_0: $\mu = \mu_0$ niet verworpen wordt tegen H_1: $\mu > \mu_0$ met $\alpha_0 =$ 1- 0.95. M.a.w. als T niet in het kritieke gebied $[c, \infty)$ ligt dus als $\frac{\bar{X} - \mu_0}{S/\sqrt{10}} < 1.83 (= c)$.
Hieruit volgt $\mu_0 > \bar{x} - 1.83s/\sqrt{10} \approx 2.38$
Met betrouwbaarheid van 95% is het gewichtsverlies dus minstens 2.38 kg.

c. N^+ = het aantal positieve verschillen = 7 (van de 10 niet-nul verschillen)
We toetsen H_0: p = ½ tegen H_1: p > ½ met p = de kans op een positief verschil
N^+ is onder H_0 B(10, ½)
Overschrijdingskans $P(N^+ \ge 7 \mid p = ½) = 1 - P(N^+ \le 6 \mid p = ½) = 1 - 8281 = 0.1719 > \alpha_0 = 0.10$, dus H_0 verwerpen.

Opgave 3

a. T is een zuivere schatter van θ als $E(T) = aE(T_1) + b\,E(T_2) + c\,E(T_3) = (a + b + c)\,\theta = \theta$
dus $a + b + c = 1$ ofwel $c = 1 - a - b$

b. Var (T)= var ($aT, + bT_2 + cT_3$)= a^2 var (T_1) + b^2 var(T_2) + c^2 var(T_3) = $[a^2 + 2b^2 + 3(1\text{-}a\text{-}b)^2]\ \sigma^2/n$
De variantie is minimaal als f(a,b) = $a^2 + 2b^2 + 3(1\text{-}a\text{-}b)^2$ minimaal is. Partiele afgeleiden 0 stellen:
2a -6(1-a-b) =0 en 4b -6(1-a-b) = 0. Dit is het geval als a = , b = en c =
Omdat f een som van kwadraten is en lim f(a,b) $\to\infty$ voor a of b$\to\infty$, is dit extremum een minimum.

c. Uit b volgt dat $\lim_{n\to\infty} var(T) = 0$, voor alle waarden van a, b en c, dus T convergeert in kans naar zijn verwachtingswaarde $ET = (a + b + c)\,\theta$, dus $\lim_{n\to\infty} P(|T - ET| > \varepsilon) = 0,\ voor\ elke\ \varepsilon > 0$.

T (afhankelijk van het aantal waarnemingen n) is een consistente schatter van θ als $\lim_{n\to\infty} P(|T - \theta| > \varepsilon) = 0,\ voor\ elke\ \varepsilon > 0$.

Dit geldt alleen als θ = E(T), dus als $c = 1 - a - b$. Dus T is consistent als $a + b + c = 1$.

(*Opmerking1: De eigenschap*

$\lim_{n\to\infty} E(T) = \theta$ *èn* $\lim_{n\to\infty} var(T) = 0$ => $\lim_{n\to\infty} MSE(T) = 0$ => *T is consistent*

betekent niet dat uit consistentie asymptotische zuiverheid of naar 0 naderende MSE volgt: het zijn voldoende, geen noodzakelijke voorwaarden!

Opmerking 2: $\lim_{n\to\infty} P(|T - ET| > \varepsilon) = 0,\ voor\ elke\ \varepsilon > 0$ *volgt ook uit de ongelijkheid van Chebyshev:*

$P(|T - ET| > \varepsilon) \le \dfrac{var(T)}{\varepsilon^2}$ *voor elke* $\varepsilon > 0$, *tezamen met* $\lim_{n\to\infty} var(T) = 0$)

Opgave 4

a. $L(\lambda) = f(x_1) \times \ldots \times f(x_1) = \lambda^n e^{-\lambda \sum x_i}$

$L^*(\lambda) = \ln L(\lambda) = n \times \ln(\lambda) - \lambda \sum x_i$ Maximum L* bepalen door de afgeleide o te stellen:

$(L^*)' = n/\lambda - \sum x_i = 0$, als $\lambda = \hat{\lambda} = n/\sum x_i = 1/\bar{x}$

$(L^*)'' = -n/\lambda^2 < 0$, dus $\hat{\lambda}$ is de m.a. schatting van λ *en* $\hat{\lambda} = 1/\overline{X}$ de m.a. schatter.

b. $$\Lambda(x_1,\ldots,x_n) = \frac{\sup_{\theta\in\Theta_0} \prod g(x_i,\theta)}{\sup_{\theta\in\Theta} \prod g(x_i,\theta)} = \frac{L(1)}{L(\frac{1}{\bar{x}})} = \frac{e^{-\sum x_i}}{(\bar{x})^{-n} e^{-\sum x_i/\bar{x}}} = \bar{x}^n e^{-n(\bar{x}-1)}$$

de functie $f(y) = y^n \exp[-n(y-1)]$ is positief voor y > 0 en daalt naar 0 als y daalt naar 0 en $y \to \infty$, dus $\Lambda(X_1,\ldots,X_n)$ neemt kleine waarden aan voor kleine en voor grote waarden van $\overline{X}$: indien de kritieke waarden voor $\overline{X}$ zo worden gekozen dat $P_0(\Lambda \le \lambda) = \alpha_0$, is de gevonden toets dus een MP-toets.

c. We bepalen (benaderen met de CLS) voor $n = 25$ het gebruikelijke kritiek gebied voor $\overline{X}$:

$P_0(\overline{X} \le c_1) = P_0(\overline{X} \ge c_2) = 0.025$

$\overline{X}$ is bij benadering $N(\frac{1}{\lambda}, \frac{1/\lambda^2}{n})$- , dus is $\overline{X}$ onder H_0 N(1, 1/n)-verdeeld:

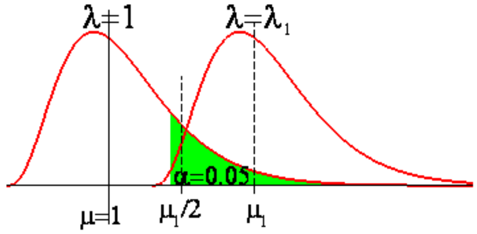
$P(|\overline{X} - 1| \times \sqrt{n} \ge c) \le 0.05$, dus c =1.96. Ofwel: $\overline{X} \le 1 - \frac{1.96}{\sqrt{n}}$ of $\overline{X} \ge 1 + \frac{1.96}{\sqrt{n}}$

Kritiek gebied Z = $\{(x_1,\ldots,x_n) \mid \bar{x} \le 1 - \frac{1.96}{\sqrt{n}}$ of $\bar{x} \ge 1 + \frac{1.96}{\sqrt{n}}\}$.

d. De toets is zuiver als $\alpha \le \beta(\lambda)$ voor elke $\lambda \ne 1$

De toets is consistent als $\lim_{n\to\infty} \beta(\lambda) = 1$ voor elke $\lambda \ne 1$. Om dat aan te tonen kunnen we gebruikmaken van de sterke wet van de grote aantallen: $\overline{X}$ convergeert bijna zeker, dus ook in kans, naar $\mu = \frac{1}{\lambda}$: $\lim_{n\to\infty} P_\mu(|\overline{X} - \mu| > \varepsilon) = 0$ voor elke $\varepsilon > 0$.

Deze limiet impliceert dat voor elke $\varepsilon > 0$ er een n_1 is zodat voor alle $n \ge n_1$: $P_{\mu=1}(|\overline{X} - 1| > \varepsilon) \le \alpha_0$



Voor elke $\lambda_1 < 1$ en dus $\mu_1 = \frac{1}{\lambda_1} > 1$ kunnen we bijvoorbeeld $\varepsilon = ½\times(\mu_1 - 1)$ kiezen: het kritieke gebied omvat dan de waarden waarvoor $\overline{X} \geq ½\times\mu_1$

Dan is $\beta(\lambda_1) \geq P_{\lambda_1}(½\times\mu_1 \leq \overline{X} \leq 1½\times\mu_1) = P_{\lambda_1}(|\overline{X} - \mu_1| \leq ½\times\mu_1)$.

Omdat $\overline{X} \xrightarrow{P} \mu$, gaat deze laatste kans naar 1 voor $n\to\infty$. Daar $\beta(\lambda_1) \leq 1$, geldt voor $\lambda > 1$ook $\lim_{n\to\infty} \beta(\lambda) = 1$, voor $\lambda < 1$ verloopt het bewijs analoog.

Opgave 5

1. Model: X = "het aantal rode balletjes in een aselecte steekproef van 150 balletjes": als p = de fractie rode balletjes, dan is X is B(150, p) (bij trekken met terugleggen, bij benadering bij trekken zonder terugleggen).
2. We toetsen H_0: $m \geq 5000$ tegen H_1: $m < 5000$ met $\alpha_0 = 0.05$, hetgeen equivalent is met:
   H_0: $p \leq 200/5000$ tegen H_1: $p > 200/5000$ met $\alpha_0 = 0.05$,
3. toetsingsgrootheid X
4. X is, onder H_0, B(150, 0.04). We kunnen deze verdeling benaderen met de Poisson-verdeling met $\mu = 150\times0.04 = 6$. *(de normale benadering met N(np, np(1-p)) is in dit geval van kleine p onnauwkeuriger)*
5. Waargenomen X = 10.
6. de rechteroverschrijdingskans is $P(X \geq 10 | H_0) = 1 - P(X \leq 9 | H_0) = 1 - P(\text{Poisson}(6) \leq 9) \approx 8.4\%$
   (- normale benadering met cont. corr. levert 7.2% als uitkomst. De GR berekent de binomiale kans exact
   - de kritieke waarde is c =11 bij een onbetrouwbaarheid van 4.26%.)
7. overschrijdingskans = 8.4% > 5% = α_0, dus H_0 niet verwerpen.
8. Er is niet aangetoond dat het aantal balletjes in de bak kleiner is dan 5000 (met onbetrouwbaarheidsdrempel 5%)