# Uitwerking Toets 2 Statistiek - 24 oktober 2014

**Opgave 1** We berekenen eerst de gemiddelden en standaardafwijkingen voor de $x_i$'s, de $y_i$'s en de verschillen $z_i$'s (*bij voorkeur invoeren in rekenmachine en* $\overline{x}$ *en* $s$ *aflezen*!)

| $i$ | 1 | 2 | 3 | 4 | Gemiddelde | Stand. afw. | Variantie |
|---|---|---|---|---|---|---|---|
| $x_i$ | 3.09 | 4.67 | 7.72 | 6.89 | 5.5925 | 2.11 | 4.44 |
| $y_i$ | 4.56 | 4.44 | 9.29 | 8.01 | 6.575 | 2.45 | 6.02 |
| $z_i = y_i - x_i$ | 1.47 | -0.23 | 1.57 | 1.12 | 0.9825 | 0.831 | 0.691 |

**a.** Toets op $H_0$: $\mu_X - \mu_Y = 0$ tegen $H_1$: $\mu_X - \mu_Y \neq 0$: één steekproef $t$-toets op de verschillen:
$T = \frac{\overline{Z}-0}{S_Z/\sqrt{4}} \overset{H_0}{\cong} T(3)$ en $t = \frac{0.9825}{0.831/2} \approx 2.36$
p-waarde = $2 \cdot P(T \geq 2.36|H_0)$. Omdat $P(T(3) \geq 2.35) = 0.05$, is de p-waarde (iets)kleiner dan 10% = $\alpha_0$, dus $H_0$ verwerpen.

Twee **onafhankelijke aselecte steekproeven** uit normale verdelingen:

**b.** We toetsen $H_0$: $\sigma_X^2 = \sigma_Y^2$ tegen $H_1$: $\sigma_X^2 \neq \sigma_Y^2$ met $\alpha_0 = 0.10$
Toetsingsgrootheid $F = \frac{S_X^2}{S_Y^2}$ heeft onder $H_0$ een F(3,3)-verdeling en als waarde $\left(\frac{2.11}{2.45}\right)^2 \approx 0.74$
Verwerp $H_0$ als $F \leq c_1$ of $F \geq c_2$.
Uit de F(3,3)-tabel met staartkans 5% blijkt $c_2 = 9.28$ en $c_1 = \frac{1}{9.28}$
$F = 0.74$ ligt niet in het Kritieke Gebied, dus $H_0$ niet verwerpen: de varianties kunnen gelijk verondersteld worden.

**c.** Toets $H_0$: $\mu_X = \mu_Y$ tegen $H_1$: $\mu_X < \mu_Y$, aangenomen dat $\sigma_X^2 = \sigma_Y^2$: 2 steekproeven $t$ −toets:
Toetsingsgrootheid $T = \frac{\overline{X}-\overline{Y}}{\sqrt{S^2\left(\frac{1}{4}+\frac{1}{4}\right)}}$ met $S^2 = \frac{S_X^2+S_Y^2}{2}$ en $T$ is onder $H_0$ $T(4+4-2)$-verdeeld
waargenomen: $t = \frac{5.5925-6.575}{\sqrt{\frac{2.11^2+2.45^2}{2} \cdot \left(\frac{1}{4}+\frac{1}{4}\right)}} \approx -0.61$ $\quad (s^2 \approx 5.23)$
De linker overschrijdingskans $P(T(6) < -0.61) > 10\% = \alpha_0$, dus $H_0$ niet verwerpen.

**d.** Rangsomtoets van Wilcoxon:
We toetsen $H_0$: $F_Y = F_X$ tegen $H_1$: $F_Y \leq F_X$ met $F_Y \neq F_X$
W = som van de rangnummers van de $Y$-waarden = $\sum_{i=1}^{4} R(Y_i)$

| Rangnummer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
|---|---|---|---|---|---|---|---|---|---|
| Geordende waarnemingen | 3.09 | 4.44 | 4.56 | 4.67 | 6.89 | 7.72 | 8.01 | 9.29 | som |
| $R(Y_i)$ | | 2 | 3 | | | | 7 | 8 | W = 20 |

Het Kritieke Gebied: elke combinatie van 4 rangnummers heeft kans $\frac{1}{\binom{8}{4}} = \frac{1}{70}$
Wat is de minimale rangsom $c$, zodat $P(W \geq c|H_0) \leq \alpha_0 = 5\%$?

- Er is één combinatie met de grootste rangsom $5+6+7+8 = 26$.
- $25 = 4+6+7+8$, dus $P(W \geq 24|H_0) = \frac{2}{70} = 2.9\%$
- $24 = 3+6+7+8 = 4+5+7+8$, dus $P(W \geq 24|H_0) = \frac{4}{70} \approx 5.7\%$

Conclusie: het Kritieke gebied wordt gegeven door $W \geq 25$.
Daar $W = 20$ niet in dit kritieke gebied ligt, kunnen we $H_0$ niet verwerpen.

**Opgave 2**

Het tweezijdig 90%-betrouwbaarheidsinterval voor $\sigma_X^2$ wordt gegeven door
$\left(\frac{(n_1-1)S_X^2}{c_2}, \frac{(n_1-1)S_X^2}{c_1}\right)$ en het naar beneden begrensde interval dus door $\left(\frac{(n_1-1)S_X^2}{c_2}, \infty\right)$
$n_1$= 4, $s_X^2 = 4.44$ en $c_2$= 6.25 zodat $P(W(3) \geq c_2) = \alpha = 0.10$
Dus het naar beneden begrensde 90%-BI $(\sigma_X^2) \approx (2.13, \infty)$.

**Opgave 3**.

**a.** Een benaderend tweezijdig 95%-BI($p$) = $\left(\hat{p} - c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\ , \hat{p} + c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

waarin $n = 125, \hat{p} = \frac{55}{125}$ en $c = 1.96$, zodat $\Phi(c) = 0.975$

95%-BI($p$) = $\left(\frac{55}{125} - 1.96 \cdot \sqrt{\frac{55 \cdot 70}{125^3}}\ , 0.440 + 0.087\right) \approx (0.353, 0.527)$

**b.** Een exact naar boven begrensd 90%-betrouwbaarheidsinterval voor $p$:
Als $H_0 : p = p_0$ getoetst wordt tegen $H_1 : p < p_0$ met $\alpha_0 = 0.10$, is
95%-BI($p$) = $\{p_0 | P(X \leq 2 | H_0) > 0.10\}$
$P(X \leq 2 | p = 0.45) = 0.1495$ en $P(X \leq 2 | p = 0.5) = 0.0898$
Lineaire interpolatie: $P\left(X \leq 2 \middle| p = 0.5 - \frac{0.0102}{0.0597} \cdot 0.05\right) \approx 0.10$
90%-BI($p$) $\approx$ [0, 0.491) (of [0, 0.49))

**Opgave 4**

$$\Lambda = \frac{\sup\limits_{\theta \in \Theta_0} \prod_{i=1}^n g(x_i, \theta)}{\sup\limits_{\theta \in \Theta} \prod_{i=1}^n g(x_i, \theta)} = \frac{\prod_{i=1}^n e^{-x_i}}{\sup\limits_{\mu \geq 1} \prod_{i=1}^n \frac{1}{\mu} e^{-\frac{x_i}{\mu}}} = \frac{e^{-\sum x_i}}{\sup\limits_{\mu \geq 1} \frac{e^{-\sum \frac{x_i}{\mu}}}{\mu^n}} = \frac{e^{-n\overline{x}}}{\frac{e^{-\sum \frac{x_i}{\overline{x}}}}{\overline{x}^n}} = \overline{x}^n e^{-n\overline{x}+n}, \text{voor } \overline{x} \geq 1$$

Anders $(0 \leq \overline{x} < 1)$ is $\Lambda = 1$
Verwerpen voor kleine $\Lambda$ komt overeen met verwerpen voor grote $\overline{x}$, dus KG is $\overline{X} \geq c$.
Voor $n = 100$ en $\alpha_0 = 0.05$ geldt $P(\overline{X} \geq c | H_0) \leq 0.05$.
Omdat $\overline{X}$ onder $H_0$ bij benadering $N\left(1, \frac{1}{100}\right)$-verdeeld is vinden we uit $\Phi\left(\frac{c-1}{0.1}\right) = 0.95$
$\frac{c-1}{0.1} = 1.645$ ofwel $c = 1.1645$.
(Je kunt ook uitgaan van parameter $\lambda$ en, wegens $\mu = 1/\lambda$, $H_0 : \lambda = 1$ toetsen tegen $H_1 : \lambda < 1$ )

**Opgave 5**

de p-waarde is $P(\overline{X} > \overline{x} | H_0) = 1 - \Phi\left(\frac{\overline{x}-0}{\sqrt{1/n}}\right)$, dus $p(X_1, .., X_n) = 1 - \Phi\left(\frac{\overline{X}-0}{\sqrt{1/n}}\right)$

$$P\left(1 - \Phi\left(\frac{\overline{X} - 0}{\sqrt{\frac{1}{n}}}\right) \leq \ \alpha_0 \ \middle| \ \mu = 0\right) = P\left(\frac{\overline{X} - 0}{\sqrt{\frac{1}{n}}} \geq \Phi^{-1}(1 - \alpha_0) | \mu = 0\right)$$
$$= 1 - \Phi\left(\Phi^{-1}(1 - \alpha_0)\right) = 1 - (1 - \alpha_0) = \alpha_0$$