Uitwerkingen Tentamen Mathematische Statistiek (2S990) donderdag 19 januari 2006, 9.00-12.00 uur.

De notatie voor quantielen van de χ^2-verdeling volgt de conventie $P(\chi^2_n < \chi^2_{n;\alpha}) = \alpha$.

---

1. (a) Aangezien de waarnemingen op dezelfde tijdstippen plaats vinden en verder niet duidelijk is of de metingen op de verschillende tijdstippen onder dezelfde weersomstandigheden plaatsvinden, zijn hier alleen gepaarde waarnemingen van toepassing. De hypothesen zijn $H_0 : \mu_d = 0$ en $H_a : \mu_d \neq 0$. Het gemiddelde van de verschillen is 0,3 en de standaardafwijking van de verschillen is 0,59. De toetsingsgrootheid T is gelijk aan $\overline{X}_d/(S_d/\sqrt{6}) = 0,3/(0,59/\sqrt{6}) = 1{,}25$. Het kritieke gebied is $\mid T \mid > t_{5;0{,}025} = 2{,}571$. Conclusie: we verwerpen H_0 niet.

   (b) De standaardafwijking van het verschil is gelijk aan $\sqrt{2}$. Aangezien de variantie nu bekend is, wordt de toetsingsgrootheid nu $V = \overline{X}_d/(\sigma_d/\sqrt{6}) = \overline{X}_d/(2/\sqrt{6})$, hetgeen onder H_0 standaardnormaal verdeeld is. Het kritieke gebied is dus $\mid V \mid > z_{0{,}025} = 1,96$. Als $\mu_d = 0,8$, dan is het onderscheidingsvermogen, d.w.z. de kans om H_0 te verwerpen, gelijk aan $P(V > 1,96) + P(V < -1,96)$. Aangezien

$$P(V > 1,96) = P\left(\frac{\overline{X}_d - 0,8}{\sqrt{2}/\sqrt{6}} > 1,96 - \frac{0,8}{\sqrt{2}/\sqrt{6}}\right) = P(Z > 0,57) = 0,2843$$

   en

$$P(V < -1,96) = P\left(\frac{\overline{X}_d - 0,8}{\sqrt{2}/\sqrt{6}} < -1,96 - \frac{0,8}{\sqrt{2}/\sqrt{6}}\right) = P(Z < -3,35) = 0,0004,$$

   volgt dat het onderscheidingsvermogen gelijk is $0,2843 + 0,0004 = 0,2847$.

   (c) De afwijking is gelijk aan de halve breedte van het betrouwbaarheidsinterval, dus gelijk aan $z_{0{,}025}\sqrt{2}/\sqrt{n} = 2,77/\sqrt{n}$. De minimale steekproefomvang n moet dus voldoen aan $\sqrt{n} \geq 2,77/0,5 \Rightarrow n \geq 31$.

2. (a) De momentgenererende functie van een Gammaverdeling met parameters κ en θ wordt gegeven door $(1 - x\theta)^{-\kappa}$. Dus de momentgenererende functie van $\sum_{i=1}^n X_i$ is gelijk aan $(1 - x\theta)^{-n\kappa}$. Hieruit volgt tenslotte dat de momentgenererende functie van $2\sum_{i=1}^n X_i/\theta$ is gelijk aan $(1 - (2x/\theta)\theta)^{-n\kappa} = (1 - 2x)^{-n\kappa}$, hetgeen de momentgenererende functie is van een χ^2 verdeling met $2n$ vrijheidsgraden.

   (b) Uit a) volgt dat $P\left(\chi^2_{2n;\alpha/2} \leq 2\sum_{i=1}^n X_i/\theta \leq \chi^2_{2n;1-\alpha/2}\right) = 1 - \alpha$. Met $\alpha = 0{,}05$ concluderen we dus dat $\left(\frac{2\sum_{i=1}^n X_i}{\chi^2_{2n;0{,}975}}, \frac{2\sum_{i=1}^n X_i}{\chi^2_{2n;0{,}025}}\right)$ een 95%-betrouwbaarheidsinterval is voor θ.

3. De Cramér-Rao ondergrens met $\tau(p) = (1 - p)^2$ en $\tau' = -2(1 - p)$ is gelijk aan

$$\begin{aligned}
4(1-p)^2\left(-n\mathrm{E}\left[\frac{\partial^2}{\partial p^2}\log\left\{p^X(1-p)^{1-X}\right\}\right]\right)^{-1} &= 4(1-p)^2\left(n\mathrm{E}\left[\frac{X}{p^2} + \frac{1-X}{(1-p)^2}\right]\right)^{-1} \\
&= 4(1-p)^2\left(n\left[\frac{1}{p} + \frac{1}{(1-p)}\right]\right)^{-1} \\
&= 4(1-p)^2\left(p\,(1-p)\right)/n.
\end{aligned}$$

We concluderen dat de Cramér-Rao ondergrens voor $(1-p)^2$ gelijk is aan $4p(1-p)^3/n$.

4. (a) De likelihood is gelijk aan $\theta^n \left(\prod_{i=1}^n x_i\right)^{\theta-1}$. Dit is van de vorm $f(x)g(S,\theta)$ met $S = \prod_{i=1}^n X_i$. Uit de factorisatiestelling volgt nu dat $\prod_{i=1}^n X_i$ een voldoende steekproefgrootheid is voor θ.

   (b) De loglikelihood is gelijk aan $n \log\theta + (\theta-1)\sum_{i=1}^n \log x_i$. Nulstellen van de afgeleide van de loglikelihood levert dat $n/\theta = -\sum_{i=1}^n \log x_i$. Hieruit volgt dat $-n/\sum_{i=1}^n \log X_i$ de Maximum Likelihoodschatter voor θ is.

   (c) Door de dichtheid te integreren van 0 tot x vinden we dat $F_{X_i}(x) = x^\theta$ voor $0 < x < 1$. Aangezien $P(-\log X_i > x) = P(X_i < e^{-x}) = e^{-x\theta}$, zien we dat $-\log X_i$ exponentieel verdeeld is met parameter θ. Hieruit volgt dat $-\sum_{i=1}^n \log X_i$ een Gammaverdeling volgt met parameters n en θ.

   (d) Aangezien de exponentiële verdeling een eindig tweede moment bezit, volgt uit de Centrale Limietstelling dat

$$\frac{\sum_{i=1}^n \log X_i - \sum_{i=1}^n \mathrm{E}\left(\log X_i\right)}{\sqrt{\sum_{i=1}^n \mathrm{Var}\left(\log X_i\right)}} = \frac{nW - n\theta}{\sqrt{n\theta^2}} = \frac{\sqrt{n}\,(W-\theta)}{\theta} \xrightarrow{d} N(0,1)$$

   als $n \to \infty$. We concluderen dus dat $\sqrt{n}\,(W-\theta)$ in verdeling naar een normale verdeling met verwachting 0 en variantie θ^2 convergeert.

5. We gebruiken het lemma van Neyman-Pearson om eerst een UMP toets af te leiden voor de enkelvoudige alternatieve hypothese $H_a : \lambda = \lambda_0$ met $\lambda_0 > 1$. Volgens dit lemma is een toets gebaseerd op likelihood ratio

$$T = \frac{e^{-n}/\prod_{i=1}^n X_i!}{e^{-\lambda_0 n}\lambda_0^{\sum_{i=1}^n X_i}/\prod_{i=1}^n X_i!} = \frac{e^{-n(1-\lambda_0)}}{\lambda_0^{\sum_{i=1}^n X_i}}$$

en een kritiek gebied van de vorm $T < c$ voor geschikte c, het meest onderscheidend. Aangezien $\lambda_0 > 1$ volgt dat

$$P(T < c) = P\left(\lambda_0^{\sum_{i=1}^n X_i} > e^{-n(1-\lambda_0)}/c\right) = P(\sum_{i=1}^n X_i) > \tilde{c}).$$

Hieruit volgt dat voor elke $\lambda_0 > 1$, de toets met toetsingsgrootheid $V = \sum_{i=1}^n X_i$ en kritiek gebied $V > c_2$ waarbij $P(V > c_2) = \alpha$, een UMP toets is voor het toetsen van $H_0 : \lambda = 1$ tegen $H_a : \lambda > 1$ voor steekproeven uit een Poissonverdeling. Merk op dat V onder H_0 zelf ook Poisson verdeeld is met parameter n.

De notatie voor quantielen van de χ^2-verdeling volgt de conventie $P(\chi^2_n < \chi^2_{n;\alpha}) = \alpha$.