Matemaattinen tilastotiede 2
Harjoitukset 12, 13.4.2010

1.

$$X_1, \dots, X_m \sim Po(\mu) \quad \rightarrow \quad \sum_{i=1}^{m} X_i \sim Po(m \cdot \mu)$$
$$Y_1, \dots, Y_n \sim Po(\lambda) \quad \rightarrow \quad \sum_{j=1}^{n} Y_j \sim Po(n \cdot \lambda)$$
$$\rightarrow \quad \sum_{i=1}^{m} X_i + \sum_{j=1}^{n} Y_j \sim Po(m\mu + n\lambda)$$

$$P(\Sigma X_i = k \mid \Sigma X_i + \Sigma Y_j = u)$$

$$= \frac{P(\Sigma X_i = k \text{ ja } \Sigma X_i + \Sigma Y_j = u)}{P(\Sigma X_i + \Sigma Y_j = u)} = \frac{P(\Sigma X_i = k \text{ ja } \Sigma Y_j = u-k)}{P(\Sigma X_i + \Sigma Y_j = u)}$$

$$\overset{\Sigma X_i \perp\!\!\!\perp \Sigma Y_j}{=} \frac{P(\Sigma X_i = k) \cdot P(\Sigma Y_j = u-k)}{P(\Sigma X_i + \Sigma Y_j = u)}$$

$$= \frac{\frac{(m\mu)^k}{k!} \cancel{e^{-m\mu}} \cdot \frac{(n\lambda)^{u-k}}{(u-k)!} \cancel{e^{-n\lambda}}}{\frac{(m\mu + n\lambda)^u}{u!} \cancel{e^{-(m\mu+n\lambda)}}}$$

$$= \frac{u!}{k!(u-k)!} \cdot \frac{(m\mu)^k (n\lambda)^{u-k}}{(m\mu + n\lambda)^{u-k+k}} \qquad \Big| u = k + (u-k)$$

$$= \binom{u}{k} \left(\frac{m\mu}{m\mu + n\lambda}\right)^k \left(\frac{n\lambda}{m\mu + n\lambda}\right)^{u-k}$$

$k = 0, \dots, u$
$u = 0, 1, \dots$

2.

$X \sim Bin(m, p_1)$, $Y \sim Bin(n, p_2)$, $X \perp\!\!\!\perp Y$

$H_0: p_1 = p_2$, $H_1: p_1 > p_2$

uskottavuus:

$$L(p_1, p_2) = \binom{m}{x} p_1^x (1-p_1)^{m-x} \cdot \binom{n}{y} p_2^y (1-p_2)^{n-y} = \binom{m}{x}\binom{n}{y} e^{x \log p_1 + y \log p_2 + (m-x)\log(1-p_1) + (n-y)\log(1-p_2)}$$

$$= \binom{m}{x}\binom{n}{y} \exp\left\{ x \log \frac{p_1}{1-p_1} + y \log \frac{p_2}{1-p_2} + m \log(1-p_1) + n \log(1-p_2) \right\}$$

$$\theta_1 = \log \frac{p_1}{1-p_1} - \log \frac{p_2}{1-p_2} \qquad H_0: \theta_1 = 0$$

$$\theta_2 = \log \frac{p_2}{1-p_2} \qquad H_1: \theta_1 > 0$$

$\left(\log \frac{x}{1-x} \text{ kasvava} \right)$

2. jatkuu...

$$L(\theta_1,\theta_2)=\binom{m}{x}\binom{n}{y}\exp\Big\{\underbrace{x\Big(\log\frac{p_1}{1-p_1}-\log\frac{p_2}{1-p_2}\Big)}_{=\theta_1}+\underbrace{x\log\frac{p_2}{1-p_2}+y\log\frac{p_2}{1-p_2}}_{=(x+y)\theta_2}+\underbrace{m\log(1-p_1)+n\log(1-p_2)}_{f(\theta_1,\theta_2)}\Big\}$$

Testi hylätään, jos

$$X \geq c(x+y)$$

missä kriittinen arvo saadaan

$$P_{H_0}\big(X \geq c(u) \mid X+Y=u\big)=\alpha$$

Tehtävässä 3 osoitetaan, että kun $H_0: \theta_1=0$ (eli $p_1=p_2$) pätee, niin $P(X=k \mid X+Y=u)$ saadaan hypergeometrisesta jakaumasta.

3. $p_1 = p_2 = p$

$X \sim Bin(m,p)$ $\quad \overset{X \perp\!\!\!\perp Y}{\Rightarrow} X+Y \sim Bin(n+m,p)$

$Y \sim Bin(n,p)$

$$P(X=k \mid X+Y=u) = \frac{P(X=k \text{ ja } X+Y=u)}{P(X+Y=u)} = \frac{P(X=k, Y=u-k)}{P(X+Y=u)}$$

$$= \frac{P(X=k)\cdot P(Y=u-k)}{P(X+Y=u)} = \frac{\binom{m}{k} p^k (1-p)^{m-k} \cdot \binom{n}{u-k} p^{u-k} (1-p)^{n-(u-k)}}{\binom{n+m}{u} p^u (1-p)^{n+m-u}}$$

$$= \frac{\binom{m}{k}\binom{n}{u-k}}{\binom{n+m}{u}} \cdot \underbrace{p^{k+u-k-u} (1-p)^{\cancel{m}-\cancel{k}+\cancel{n}-u+\cancel{k}-\cancel{n}-\cancel{m}+u}}_{=1} = \frac{\binom{m}{k}\binom{n}{u-k}}{\binom{n+m}{u}}$$

(4.) $Y_1, \dots, Y_n \sim Gamma(\alpha, \beta)$

$H_0: (\alpha,\beta) = (1,1) \qquad H_1: (\alpha,\beta) \neq (1,1)$

$$f(y;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}, \quad L(\alpha,\beta) = \exp\left\{n\alpha\log\beta - n\log\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^n \log y_i - \beta\sum_{i=1}^n y_i\right\}$$

(usk.yhtälöt

$$\begin{cases} \sum_{i=1}^n \log y_i = n(\Psi(\alpha) - \log\beta) \\ \frac{1}{n}\sum_{i=1}^n y_i = \frac{\alpha}{\beta} \end{cases} \rightarrow (\hat\alpha, \hat\beta)$$
)

LR:

$$2\log\frac{L(\hat\alpha,\hat\beta)}{L(1,1)} = 2\log\frac{e^{n\hat\alpha\log\hat\beta - n\log\Gamma(\hat\alpha) + (\hat\alpha-1)\sum\log y_i - \hat\beta\sum y_i}}{e^{n\cdot 1\cdot\underbrace{\log 1}_{=0} - n\underbrace{\log\Gamma(1)}_{=0} + 0\cdot\sum\log y_i - 1\cdot\sum y_i}}$$

$$= 2\left[n\hat\alpha\log\hat\beta - n\log\Gamma(\hat\alpha) + (\hat\alpha-1)\sum_{i=1}^n\log y_i - (\hat\beta-1)\sum_{i=1}^n y_i\right] \xrightarrow{D} \chi^2(2)$$

Waldin testisuure on $n(\hat\theta-\theta_0)' B(\hat\theta)(\hat\theta-\theta_0)$ (yleisessä muodossa)

Nyt $B(\theta) = \begin{bmatrix} \Psi'(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{bmatrix}$ (Demot 11/1)

Testisuure on nyt

$$n\begin{pmatrix}\hat\alpha-1 & \hat\beta-1\end{pmatrix}\begin{pmatrix}\Psi'(\hat\alpha) & -\frac{1}{\hat\beta} \\ -\frac{1}{\hat\beta} & \frac{\hat\alpha}{\hat\beta^2}\end{pmatrix}\begin{pmatrix}\hat\alpha-1 \\ \hat\beta-1\end{pmatrix}$$

$$= n\begin{pmatrix}(\hat\alpha-1)\Psi'(\hat\alpha) - (\hat\beta-1)\frac{1}{\hat\beta} & -(\hat\alpha-1)\cdot\frac{1}{\hat\beta} + (\hat\beta-1)\frac{\hat\alpha}{\hat\beta^2}\end{pmatrix}\begin{pmatrix}\hat\alpha-1 \\ \hat\beta-1\end{pmatrix}$$

$$= n\left((\hat\alpha-1)^2\Psi'(\hat\alpha) - (\hat\alpha-1)(\hat\beta-1)\frac{1}{\hat\beta} - (\hat\alpha-1)(\hat\beta-1)\cdot\frac{1}{\hat\beta} + (\hat\beta-1)^2\cdot\frac{\hat\alpha}{\hat\beta^2}\right)$$

$$= n\left((\hat\alpha-1)^2\Psi'(\hat\alpha) - 2(\hat\alpha-1)(\hat\beta-1)\cdot\frac{1}{\hat\beta} + (\hat\beta-1)^2\cdot\frac{\hat\alpha}{\hat\beta^2}\right) \xrightarrow{D} \chi^2(2)$$

Raon testisuure on $n^{-1}d(\theta_0)B(\theta_0)^{-1}d(\theta_0)$ (yleisesti). Nyt

score vector $d(\theta) = \left(\frac{\partial\log L(\theta)}{\partial\alpha} \quad \frac{\partial\log L(\theta)}{\partial\beta}\right) = \left(n\log\beta - n\underbrace{\frac{d}{d\alpha}\log\Gamma(\alpha)}_{=\Psi'(\alpha)} + \sum_{i=1}^n\log y_i \quad \frac{n\alpha}{\beta} - \sum_{i=1}^n y_i\right)$

$$d(\theta_0) = \left(n\underbrace{\log 1}_{=0} - n\Psi'(1) + \sum\log y_i \quad \frac{n\cdot 1}{1} - \sum y_i\right) = \left(\sum_i\log y_i - n\Psi'(1) \quad n - \sum y_i\right)$$

$$B(\theta_0)^{-1} = \frac{1}{1\cdot\Psi'(1)-1}\begin{bmatrix}1 & 1 \\ 1 & \Psi'(1)\end{bmatrix} = \frac{1}{\Psi'(1)-1}\begin{bmatrix}1 & 1 \\ 1 & \Psi'(1)\end{bmatrix}$$

4. jatkuu

testisuure

$$\frac{1}{n} d(\theta_0)\, B(\theta_0)^{-1} d(\theta_0) = \left(\frac{1}{n}\sum \log y_i - \psi'(1) \quad 1-\frac{1}{n}\sum y_i\right)\begin{bmatrix} 1 & 1 \\ 1 & \psi'(1)\end{bmatrix}\begin{pmatrix}\sum \log y_i - n\psi'(1) \\ n-\sum y_i\end{pmatrix}\cdot\frac{1}{\psi'(1)-1}$$

$$= \left(\frac{1}{n}\sum \log y_i - \psi'(1) + 1 - \frac{1}{n}\sum y_i \quad \frac{1}{n}\sum \log y_i \cancel{-\psi'(1)+\psi'(1)} - \frac{\psi'(1)}{n}\sum y_i\right)\begin{pmatrix}\sum \log y_i - n\psi'(1) \\ n-\sum y_i\end{pmatrix}\cdot\frac{1}{\psi'(1)-1}$$

$$= \left(\frac{1}{n}\left(\sum \log y_i - \psi'(1)\right)^2 + \frac{1}{n}\left(\sum \log y_i - \psi'(1)\right)\left(n-\sum y_i\right) + \frac{1}{n}\left(\sum \log y_i - \psi'(1)\sum y_i\right)\left(n-\sum y_i\right)\right)\cdot\frac{1}{\psi'(1)-1}$$

$$= \frac{1}{n(\psi'(1)-1)}\left(\left(\sum_i \log y_i - \psi'(1)\right)^2 + \left(\sum_i \log y_i - \psi'(1)\right)\left(n-\sum_i y_i\right) + \left(\sum_i \log y_i - \psi'(1)\sum_i y_i\right)\left(n-\sum_i y_i\right)\right)$$

⑤ $Y_1, \dots, Y_n \sim N(\mu, \sigma^2)$

$H_0: \sigma^2 = \sigma_0^2 \qquad H_A: \sigma^2 \neq \sigma_0^2$

LR-testi H_0:lle vastaan H_A, kriittinen alue on $2\log \dfrac{L(\hat{\mu}, \hat{\sigma}^2)}{L(\tilde{\mu}, \sigma_0^2)} > c$

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i-\mu)^2} = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}$$

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2$$

$$\frac{\partial \ell(\mu,\sigma^2)}{\partial\mu} = \cancel{-}\frac{1}{\cancel{2}\sigma^2}\sum \cancel{2}(y_i-\mu)\cdot\cancel{(-1)} = \frac{1}{\sigma^2}\left(\sum y_i - n\mu\right) =: 0 \Rightarrow \hat{\mu} = \frac{\sum y_i}{n} = \bar{y}, \quad \tilde{\mu} = \bar{y}$$

$$\frac{\partial \ell(\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum(y_i-\mu)^2}{2(\sigma^2)^2} =: 0 \Leftrightarrow -n + \frac{\sum(y_i-\mu)^2}{\sigma^2} = 0 \Leftrightarrow \sigma^2 = \frac{\sum(y_i-\mu)^2}{n}$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum(y_i-\bar{y})^2}{n}$$

$$\left(\tilde{\sigma}^2 = \frac{\sum(y_i-\mu_0)^2}{n}\right)$$

$$2\log\frac{L(\hat{\mu},\hat{\sigma}^2)}{L(\tilde{\mu},\sigma_0^2)} = 2\log\frac{\cancel{(2\pi)^{-\frac{n}{2}}}(\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{1}{2\hat{\sigma}^2}\sum(y_i-\bar{y})^2}}{\cancel{(2\pi)^{-\frac{n}{2}}}(\sigma_0^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma_0^2}\sum(y_i-\bar{y})^2}}$$

$$= 2\left[\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{-\frac{n}{2}} - \frac{1}{2}\sum(y_i-\bar{y})^2\left(\frac{1}{\hat{\sigma}^2} - \frac{1}{\sigma_0^2}\right)\right]$$

$$= -n\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) + \left(\frac{1}{\sigma_0^2} - \frac{1}{\hat{\sigma}^2}\right)\underbrace{\sum_{i=1}^{n}(y_i-\bar{y})^2}_{=n\cdot\hat{\sigma}^2}$$

$$= -n\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) + n\frac{\hat{\sigma}^2}{\sigma_0^2} - n$$

$$= n\left(-\log\left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right) + \frac{\hat{\sigma}^2}{\sigma_0^2} - 1\right) \xrightarrow{D} \chi^2(1)$$

6. $(Y_1, Y_2, Y_3)' \sim \text{Multinom}(n: p_1, p_2, p_3)$

$H_0: \ p_1 = \eta^2, \ p_2 = 2\eta(1-\eta), \ p_3 = (1-\eta)^2, \quad 0 < \eta < 1$

Määritelmä: $W = 2 \log \frac{f(y; \hat{\theta})}{f(y; \tilde{\theta})}$

$$f(y: p_1, p_2, p_3) = \frac{n!}{y_1! y_2! y_3!} \cdot p_1^{y_1} p_2^{y_2} p_3^{y_3} \qquad \hat{p}_i = \frac{y_i}{n}, \ i = 1,2,3$$

$$f(y; \hat{p}_1, \hat{p}_2, \hat{p}_3) = \frac{n!}{y_1! y_2! y_3!} \hat{p}_1^{y_1} \hat{p}_2^{y_2} \hat{p}_3^{y_3}$$

$$f(y:\eta) = \frac{n!}{y_1! y_2! y_3!} \underbrace{(\eta^2)^{y_1} (2\eta(1-\eta))^{y_2} ((1-\eta)^2)^{y_3}}_{= 2^{y_2} \eta^{2y_1+y_2} (1-\eta)^{y_2+2y_3}}$$

$$\log f(y:\eta) = \log\left(\frac{n!}{y_1! y_2! y_3!} \cdot 2^{y_2}\right) + (2y_1+y_2)\log\eta + (y_2+2y_3)\log(1-\eta)$$

$$\frac{\partial \log f(y;\eta)}{\partial \eta} = \frac{2y_1+y_2}{\eta} - \frac{y_2+2y_3}{1-\eta} =: 0$$

$$2y_1 + y_2 - (2y_1+y_2)\eta = (y_2+2y_3)\eta$$

$$\hat{\eta} = \frac{2y_1+y_2}{2y_1+y_2+y_2+2y_3} = \frac{2y_1+y_2}{2(y_1+y_2+y_3)} = \frac{2y_1+y_2}{2n}$$

$$W = 2\log \frac{f(y; \hat{p}_1, \hat{p}_2, \hat{p}_3)}{f(y;\hat{\eta})}$$

$$= 2\log\left(\left(\frac{\hat{p}_1}{\hat{\eta}^2}\right)^{y_1} \left(\frac{\hat{p}_2}{2\hat{\eta}(1-\hat{\eta})}\right)^{y_2} \left(\frac{\hat{p}_3}{(1-\hat{\eta})^2}\right)^{y_3} \right)$$

$$= 2\left(y_1 \log\left(\frac{y_1/n}{\hat{\eta}^2}\right) + y_2 \log\left(\frac{y_2/n}{2\hat{\eta}(1-\hat{\eta})}\right) + y_3 \log\left(\frac{y_3/n}{(1-\hat{\eta})^2}\right)\right) \xrightarrow{D} \chi^2(\underset{2-1}{1})$$