

Tõenäosusteooria ja statistika II

Loeng 16

Tõepärasuhte test

Märt Möls

Tõepärasuhte test

	parameetrid	tõepära	tõepärafunktsiooni maksimum
H_0 : võime kasutada lihtsamat mudelit (vähem hinnatavaid parameetreid)	θ_0	$L(\theta_0)$	$L_0 := L(\hat{\theta}_0)$
H_1 : peame kasutama suuremat mudelit – rohkem hinnatavaid parameetreid	θ	$L(\theta)$	$L := L(\hat{\theta})$

$$-2 \ln \frac{L_0}{L} \underset{n \rightarrow \infty}{\overset{H_0}{\sim}} \chi_{df}^2$$

- * Kui fikseeritud parameetrid ei asu parameetri võimalike väärtuste ruumi piiril
- * suure valimi korral

Tõepärasuhte test - eeldused

$$Y \sim N(\mu; \sigma^2)$$

Võime kasutada testimaks hüpoteese

$$H_0 : \mu = 0 \quad H_0 : \sigma^2 = 10$$

Ei saa kasutada testimaks hüpoteesi

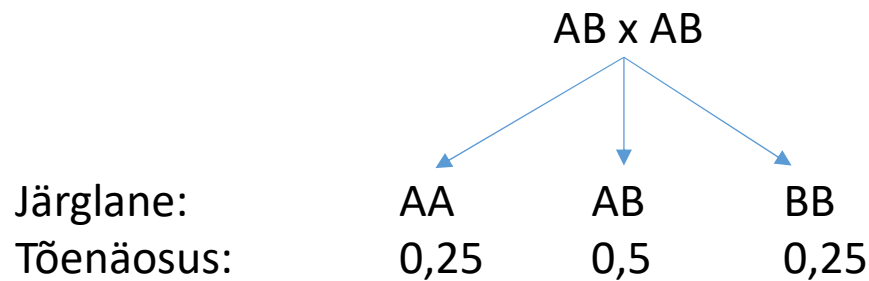
$$H_0 : \sigma^2 = 0$$

$$-2 \ln \frac{L_0}{L} \underset{n \rightarrow \infty}{\overset{H_0}{\sim}} \chi_{df}^2$$

- * Kui fikseeritud parameetrid ei asu parameetri võimalike väärtuste ruumi piiril
- * suure valimi korral

Tõepärasuhte test - näide

H_0 : Teise põlvkonna hübriidide jaoks kehtib Mendeli seadus:



$$\hat{\theta}_0 = \begin{pmatrix} 0.25 \\ 0.5 \\ 0.25 \end{pmatrix}$$

$$\hat{\theta} = \begin{pmatrix} n_{AA}/n \\ n_{AB}/n \\ n_{BB}/n \end{pmatrix}$$

$$L_0 = 0,25^{n_{AA}} \cdot 0,5^{n_{AB}} \cdot 0,25^{n_{BB}}$$

$$L = (n_{AA}/n)^{n_{AA}} \cdot (n_{AB}/n)^{n_{AB}} \cdot (n_{BB}/n)^{n_{BB}}$$

$$L_0/L = (0,25n/n_{AA})^{n_{AA}} (0,5n/n_{AB})^{n_{AB}} (0,25n/n_{BB})^{n_{BB}}$$

$$-2 \ln(L_0/L) = 2 \sum_{i=1}^k O_i \ln(O_i/E_i)$$

Tuntakse ka G-testi nime all, sageli soovitatakse hii-ruut testi asemel.

$$-2 \ln \frac{L_0}{L} \underset{n \rightarrow \infty}{\overset{H_0}{\sim}} \chi_{df}^2$$

Tõepärasuhte test – alternatiivsed esitused

$$-2 \ln \frac{L_0}{L} \stackrel{H_0}{\sim} \chi_{df}^2$$

$$-2(l_0 - l) \stackrel{H_0}{\sim} \chi_{df}^2$$

$$2(l - l_0) \stackrel{H_0}{\sim} \chi_{df}^2$$

Tõepärasuhte test – tõestuse idee

Vaatame, kuidas saame tõestada, et tõepärasuhte testi teststatistik on asümptootiliselt hii-ruut jaotusega siis, kui tarvis on hinnata vaid ühte parameetrit (ja nullhüpoteesi kehtides polegi hinnatavaid parameetreid)

Arendame log-tõepära $l(\theta_0)$ Tayloriga punkti $\hat{\theta}$ ümbruses:

$$l(\theta_0) \approx l(\hat{\theta}) + l'(\hat{\theta})(\theta_0 - \hat{\theta}) + \frac{l''(\hat{\theta})}{2}(\theta_0 - \hat{\theta})^2$$

$=0$

$$l(\theta_0) - l(\hat{\theta}) \approx \frac{l''(\hat{\theta})}{2}(\theta_0 - \hat{\theta})^2$$

$$-2(l(\theta_0) - l(\hat{\theta})) \approx \underbrace{\frac{l''(\hat{\theta})}{\sigma(\hat{\theta})^2}}_1 \left(\underbrace{\frac{\theta_0 - \hat{\theta}}{\sigma(\hat{\theta})}}_{\substack{\text{kui } H_0 \text{ siis} \\ \sim N(0; 1)}} \right)^2 \underset{H_0}{\sim} \chi_{df=1}^2$$