Uitwerking Toets 1 Statistiek - 26 september 2014 13:45 - 15:30

1. Als X chikwadraat-verdeling met m vrijheidsgraden, dan is $\mu = E(X) = m$ en $\sigma^2 = 2m$

a. $var(\overline{X}_n) = \frac{\sigma^2}{n} = \frac{2m}{n}$

b. $E(S^2) = \sigma^2 = 2m$

2. X is t-verdeeld met n vrijheidsgraden, dus X^2 is $F(1,n)$-verdeeld en $\frac{1}{X^2} \cong F(n,1)$

(eventueel met de definitie $X = \frac{Z}{\sqrt{\frac{W(n)}{n}}}$, dus $X^2 = \frac{Z^2}{\frac{W(n)}{n}} \cong F(1,n)$, omdat $Z \cong N(0,1)$ en Z en W(n) o.o. zijn en dus ook $Z^2 \cong W(1)$.)

3. Gegeven is dat $E(X) = 0$ en $var(X) = \frac{m}{m-2}$, dus $\frac{m}{m-2} = E(X^2) - (EX)^2 = E(X^2) = \mu_2$

Dus $m = \mu_2(m-2)$, ofwel $2\mu_2 = \mu_2 m - m$, Dus $m = \frac{2\mu_2}{\mu_2 - 1}$

$M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ is de momentenschatter van $E(X^2) = \mu_2$, dus $\frac{2M_2}{M_2-1}$ is de momentenschatter van m

4. **a**. We bepalen eerst de verwachting en variantie van M = $\max(X_1, \dots, X_n)$:

$F_M(x) = P(\max(X_1, \dots, X_n) \leq x) = P(X_1 \leq x) \cdot \ldots \cdot P(X_n \leq x) = \left(\frac{x}{\theta}\right)^n, 0 \leq x \leq \theta$

$f_M(x) = \frac{d}{dx} F_M(x) = \frac{nx^{n-1}}{\theta^n}, 0 \leq x \leq \theta$

$E(M) = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n} dx = \left[\frac{nx^{n+1}}{(n+1)\theta^n}\right]_0^\theta = \frac{n}{n+1}\theta$

$E(M^2) = \int_0^\theta x^2 \cdot \frac{nx^{n-1}}{\theta^n} dx = \left[\frac{nx^{n+2}}{(n+2)\theta^n}\right]_0^\theta = \frac{n}{n+2}\theta^2,$

dus $var(M) = E(M^2) - (EM)^2 = \frac{n}{(n+2)(n+1)^2}\theta^2$

De onzuiverheid van $\hat{\theta}(a) = a \cdot M$ is $b\left(\hat{\theta}(a)\right) = E\left(\hat{\theta}(a)\right) - \theta = E(aM) - \theta = \frac{na}{n+1}\theta - \theta$

En $var\left(\hat{\theta}(a)\right) = var(aM) = a^2 var(M) = \frac{na^2}{(n+2)(n-1)^2}\theta^2$

b. $MSE = b\left(\hat{\theta}(a)\right)^2 + var\left(\hat{\theta}(a)\right) = \left[\frac{(na-n-1)^2}{(n+1)^2} + \frac{na^2}{(n+2)(n+1)^2}\right]\theta^2 = f(a)\theta^2$ is minimaal

voor minimale $f(a)$: $f'(a) = \frac{2n(na-n-1)}{(n+1)^2} + \frac{2na}{(n+2)(n+1)^2} = 0$

$na - n - 1 + \frac{a}{n+2} = 0$, ofwel $\left(n - \frac{1}{n+2}\right)a = n+1$, dus $a = \frac{(n+1)(n+2)}{n^2+2n+1}$

voor deze waarde is MSE minimaal omdat $f''(a) > 0$

5.

a. De onbetrouwbaarheid van deze toets is

$\alpha = P(\sum_{i=1}^{n} X_i = 0 | \mu = 1) = \prod_i^n P_{\mu=1}(X_i = 0) = (e^{-1})^n = e^{-n}$

de onbetrouwbaarheid lager dan $\alpha_0 = 0.05$ als $e^{-n} \leq 0.05$, dus als $n \geq -\ln(0.05) = 2.996$, dus voor $n \geq 3$ (gegeven).

b. Het onderscheidend vermogen van deze toets is $\beta(\mu) = P(\sum_{i=1}^{n} X_i = 0 | \mu)$, $\mu \in \{0,1\}$.

$\beta(1) = e^{-n}$ en

$\beta(0) = 1$ (omdat voor $\mu = 0$, X ontaard is in 0: $P(X = 0) = 1$)

c. De toets is zuiver als $\alpha \leq \beta(\mu)$ voor elke μ uit $\Theta_1 = \{0\}$?
Dus moet gelden $\beta(1) \leq \beta(0)$, hetgeen het geval is $\beta(1) = e^{-n} < 1 = \beta(0)$ voor alle n

d. De toets is consistent als $\lim_{n\to\infty} \beta(0) = 1$, Daar $\beta(0) = 1$, is de toets consistent.

e. Voor $n \geq 3$ is $\alpha < \alpha_0$, zie a.
Deze toets is MP ("*Most Powerful*") voor onbetrouwbaarheidsdrempel 0.05, omdat het onderscheidend vermogen $\beta(0) = 1$ minstens zo groot is als het onderscheidend vermogen van een alternatieve toets.
Opmerking: het afleiden van de MP-toets m.b.v. het lemma van NP is in dit geval moeizaam: de teller is 0 als $\sum_{i=1}^{n} X_i \neq 0$.

6. X uniform (continu) verdeeld op $[\theta^{-1}, \theta]$, dus $f(x) = \frac{1}{\theta - \frac{1}{\theta}} = \frac{\theta}{\theta^2 - 1}$, voor $\frac{1}{\theta} \leq x \leq \theta$.

Omgekeerd betekent de voorwaarde voor θ, dat $\theta \geq x$ èn $\theta \geq \frac{1}{x}$, en dus $\theta \geq \max\left(x, \frac{1}{x}\right)$

Aannemelijkheidsfunctie $L(\theta) = \prod_i f(x_i) = \left(\frac{\theta}{\theta^2-1}\right)^n$, met $\theta \geq \max\left(x_1, \dots, x_n, \frac{1}{x_1}, \dots, \frac{1}{x_n}\right)$

Op dit definitiegebied is $L(\theta)$ een dalende functie ($L'(\theta) = n\left(\frac{\theta}{\theta^2-1}\right)^{n-1} \cdot -\frac{\theta^2+1}{\theta^2-1} < 0$)

Dus de m.a. schatting van θ is het randpunt van het definitiegebied (zie grafiek):

$\max\left(x_1, \dots, x_n, \frac{1}{x_1}, \dots, \frac{1}{x_n}\right)$

$\theta^* = \max\left(X_1, \dots, X_n, \frac{1}{X_1}, \dots, \frac{1}{X_n}\right)$.