Hypotesting, Kap. 9, forts.

Data $X_i \sim f(x;\theta)$ uif $i=1,\cdots,n$

Nullhypotese $H_0: \theta = \theta_0$

Alternativ hypotese $H_a: \theta \neq \theta_0$

$H_a': \theta > \theta_0$

$H_a'': \theta < \theta_0$

Testobservator: $T = T(X_1,\cdots,X_n)$ s.a.

forkaste H_0 hvis $T > t_0$ passende valgt t_0.

Forkastningsområde $R_x = \{(x_1,\cdots,x_n): T(x_1,\cdots,x_n) > t_0\}$

Signifikansnivå α: $P(\text{Forkaste } H_0 \mid H_0 \text{ sann}) = \alpha$

$= P(T > t_0 \mid H_0: \theta = \theta_0)$

sep 19-12:15

Type I feil: Forkaste H_0 hvis H_0 sann

Type II feil: Beholde H_0 når H_a sann (N.B. endret notasjon side igjen)

Styrkefunksjon $\gamma(\theta) = 1 - \beta(\theta) = P(\text{Forkast } H_0 \mid \theta)$

$\beta(\theta) = P(\text{Type II feil med } \theta \neq \theta_0)$

Med ensidige alt. hypoteser settes problemet noen ganger opp som

$H_0: \theta \leq \theta_0$ mot $H_a: \theta > \theta_0$

el $H_0: \theta \geq \theta_0$ mot $H_a: \theta < \theta_0$

Velger først alternativ hypotese H_a som det vi ønsker å konkludere med

Nullhypotese H_0 er da negasjonen til H_a.

sep 19-12:16

Ex) $X_i \sim N(\mu,\sigma^2)$, σ kjent, uavh, $i=1,\cdots,n$

Med $H_0: \mu = \mu_0$ mot $H_a: \mu > \mu_0$

benytte testobs. $Z = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \sim N(0,1)$ under H_0

og vi forkaster hvis $Z > z_\alpha = (1-\alpha)100$ persentil i $N(0,1)$

Med $H_0: \mu = \mu_0$ mot $H_a: \mu < \mu_0$

blir $Z = \frac{\mu_0 - \bar{X}}{\sigma}\sqrt{n}$, forkaster når $Z > z_\alpha$

Med $H_0: \mu = \mu_0$ mot $H_a: \mu \neq \mu_0$

blir testobs. $|Z| = \frac{|\bar{X}-\mu_0|}{\sigma}\sqrt{n}$

og vi forkaster når $|Z| > z_{\alpha/2}$

så tosidige testen tilsvarer ensidige med

$\alpha' = \alpha/2$

Ved ensidig testing: Bestem alternativ før du gjør testen

sep 19-12:16

Anta nå at σ ukjent, men at n er stor

Da er empirisk standardavvik S en presis estimator for σ og $\frac{\bar{X}-\mu}{S}\sqrt{n} \overset{\text{tiln}}{\sim} N(0,1)$

Benytter testene med kjent σ der vi bytter ut σ med S. Får da tilnærmet nivå α.

Så hva hvis n er liten?

Vet at $T = \frac{\bar{X}-\mu_0}{S}\sqrt{n} \sim t_{n-1}$ fordelt under H_0.

Derfor med alternativ $H_a: \mu > \mu_0$ får vi en test med eksakt nivå α ved å forkaste hvis $T > t_\alpha = (1-\alpha)100$ persentil i t_{n-1} ford.

og med alt $H_a': \mu \neq \mu_0$ tilsv. med

$|T| > t_{\alpha/2}$

sep 19-12:39

Ex) X_i = gj.sn. temp år $2000+i$, $i=1,\cdots,16$

Meteorologisk inst. får $\mu_0 = 5.74$ (som normaltemp.

Med $\alpha = 0.05$ har $z_\alpha = 1{,}645$ og $z_{\alpha/2} = 1{,}96$

Med $\alpha = 0{,}05$ og $n=16$ (df = 15) fås $t_\alpha = 1{,}75$ og $t_{\alpha/2} = 2.13$

Finner $\bar{x} = 7.06$, $s = 0.767$ som gir

$t = \frac{\bar{x} - 7.66}{0.767}\sqrt{16} = 6.88 > t_{\alpha/2}$

Forkaster – uansett ensidig el tosidig alt.

Med $\alpha = 0.001$ blir $t_\alpha = 3.73$ og $t_{\alpha/2} = 4.07$

Forkaster fortsatt. Sterkt signifikans.

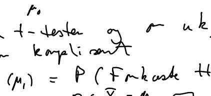
sep 19-12:46

Vi så at med $X_i \sim N(\mu,\sigma^2)$ σ^2 kjent og $H_a: \mu > \mu_0$ og $\alpha = 0{,}05$ ble styrkefunksjonen

$\gamma(\mu_1) = P(\text{Forkaste } H_0 \mid \mu_1) = 1 - \beta(\mu_1)$

$= P\left(\frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} > 1{,}645 \mid \mu_1\right)$

$= \Phi\left(\frac{\mu_1-\mu_0}{\sigma}\sqrt{n} - 1{,}645\right)$ der $\Phi(z) = P(Z \leq z)$, $Z \sim N(0,1)$



For t-testen og σ ukjent – blir styrkefn mer komplisert

$\gamma(\mu_1) = P(\text{Forkaste } H_0 \mid \mu_1)$

$= P\left(\frac{\bar{X}-\mu_0}{S}\sqrt{n} > t_\alpha \mid \mu_1\right)$

$= P\left(\underbrace{\frac{\bar{X}-\mu_1}{S}\sqrt{n}}_{T} > t_\alpha - \frac{\mu_1-\mu_0}{S}\sqrt{n}\right)$

S er stokastisk

Har $T \sim t_{n-1}$ når $\mu = \mu_1$

Men høyre siden i $\gamma(\mu_1)$ er stokastisk, kan ikke bare bruke kumulativ for t_{n-1} ford.

Styrka kan

- avleses fra figur i D&B (Ikke gjør dette)
- Programvare: power.t.test i R
- Simulere! (Særlig i mer kompliserte situasjoner)

sep 19-13:16

Utvalgsstørrelse: Vil ofte kreve n stor nok til at testen kan 80% eller 90%

For Z-testen, kjent σ, finner vi en grei formel for påkrevd n. Med $z_\beta = z_{1-\gamma}$ = (1−β) 100 persentil i N(0,1) har vi

$$\gamma = \gamma(\mu_1) = \Phi\left(\frac{\mu_1-\mu_0}{\sigma}\sqrt{n} - z_\alpha\right) = 1-\beta$$

som er ekvivalent med

$$z_\beta = \frac{\mu_1-\mu_0}{\sigma}\sqrt{n} - z_\alpha$$

$$\Leftrightarrow \quad n = (z_\beta + z_\alpha)^2 \frac{\sigma^2}{(\mu_1-\mu_0)^2}$$

Så med $H_a: \mu > \mu_0$, nivå $\alpha = 0{,}05 \rightarrow z_\alpha = 1{,}645$

og styrke $\gamma = 0{,}8 \rightarrow z_\beta = 0.84$ 80 = 100 − 20 persentil i N(0,1)

trengs $n \geq (0.84 + 1.645)^2 \frac{\sigma^2}{(\mu_1-\mu_0)^2} = 6.18$ når $\sigma = \mu_1 - \mu_0$

altså $n = 7$

sep 19-13:29

9.3 Tester om binomisk p

$X \sim \text{bin}(n, p)$

$H_0: p = p_0$ mot $H_a: p \neq p_0$, $H_a': p > p_0$, $H_a'': p < p_0$

Når n stor ($np_0 > 10$ og $n(1-p_0) > 10$) har vi for $\hat{p} = \frac{X}{n}$ at

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} \overset{\text{tiln}}{\sim} N(0,1) \text{ under } H_0$$

Rimelig å basere testobs. på Z.

f.eks. med |Z| for $H_a: p \neq p_0$

For $H_a': p > p_0$ forkastes H_0 med tilnærmet nivå α

når $Z > z_\alpha \Leftrightarrow \hat{p} > p_0 + z_\alpha \frac{\sqrt{p_0(1-p_0)}}{n}$

og tilsv. for $H_a: p \neq p_0$ når $|Z| > z_{\alpha/2}$

$\Leftrightarrow \hat{p} < p_0 - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$ el. $\hat{p} > p_0 + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}$

Ex) Årstemp. $n = 16$, $X = \#\{\text{År med temp} > 5.74\}$

Kan vi oversette $H_0: \mu > 5.74$ til $H_a: p > \frac{1}{2}$

(hvis symmetrisk temp. ford.)

sep 19-13:38

Obs. $X = 15$ som $\hat{p} = 15/16$ og

$$Z = \frac{\hat{p} - 0.5}{\sqrt{0.5 \cdot 0.5}}\sqrt{16} = 3.5 > 1.96 = z_{0.025}$$

Forkaster med nivå 2.5% f.eks.

og også med nivå 0.1%, $z_\alpha = 3.09$

Her er $np_0 = n(1-p_0) = 16 \cdot 0.5 = 8 < 10$

så testens tiln. til normalfordeling er ikke perfekt.

Men: Vi kan teste mot p_0 uten normal tiln.

sep 19-13:47

Med ensidig alt. $H_a: p > p_0$

forkaster vi når $X \geq c$ for velvalgt c.

Bestemmer c ved at nivået skal være lik eller mindre enn α slik at

$$(*) \quad \alpha \geq P(X \geq c \mid H_0: p = p_0) = \sum_{x=c}^{n} \binom{n}{x} p_0^x (1-p_0)^{n-x}$$

Vil sjelden ha likhet i (*). Velger c som minste verdi slik at ulikheten er oppfylt.

F.eks. med $n = 16$ og $p = \frac{1}{2}$ fås

$$P(X \geq 10) = 0.102 > 0.05 > P(X \geq 11) = 0.032$$

Vi får forkastningsgrense $c = 11$ med nivå $\alpha = 0.05$

og forkastningsområde $\{11, 12, 13, 14, 15, 16\} = R_x$

Observerte $X = 15 \in R_x$, Forkaster H_0

sep 19-13:52