TMA4240 H2004: Oversikt i statistisk inferens

Situasjon - observator - fordeling

	Situasjon	Vil estimere	Hva vi har	Observator	Fordeling
1	Ett utvalg	μ	kjent σ^2	$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	normal
2	Ett utvalg	μ	ukjent σ^2	$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$	t med $n-1$ frihetsgr.
3	To uavhengige utvalg	$\mu_1 - \mu_2$	kjent $\sigma_1^2 = \sigma_2^2 = \sigma^2$	$Z = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sigma\sqrt{1/n_1+1/n_2}}$	normal
4	To uavhengige utvalg	$\mu_1 - \mu_2$	ukjent $\sigma_1^2 = \sigma_2^2 = \sigma^2$	$T = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{1/n_1+1/n_2}}$	t med $(n_1 + n_2 - 2)$ frihetsgr.
5	To uavhengige utvalg	$\mu_1 - \mu_2$	kjente $\sigma_1^2 \neq \sigma_2^2$	$Z = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$	normal
6	To uavhengige utvalg	$\mu_1 - \mu_2$	ukjente $\sigma_1^2 \neq \sigma_2^2$	$T = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_1^2/n_1+S_2^2/n_2}}$	tilnærmet t-fordelt med ν^* frihetsgr.
7	To par-utvalg	differansen $\mu_d = \mu_1 - \mu_2$	Se kommentar	$T = \frac{\overline{D}-\mu_d}{S_d/\sqrt{n}}$	t med $n-1$ frihetsgr.
8	Binomisk	Andel p	Se kommentar	$Z = \frac{\widehat{P}-p}{\sqrt{p(1-p)/n}}$	tilnærmet normal
9	Binomisk	To andeler $p_1 - p_2$	Se kommentar	$Z = \frac{(\widehat{P}_1-\widehat{P}_2)-(p_1-p_2)}{\sqrt{p_1(1-p_1)/n_1+p_2(1-p_2)/n_2}}$	tilnærmet normal
10	Ett utvalg	σ^2	-	$V = \frac{(n-1)S^2}{\sigma^2}$	Kji-kvadrat med $n-1$ frihetsgr.

Kommentarer til tabellen

- **Til pkt. 1+2+10**: " Ett utvalg, normalfordelt". Vi har følgende situasjon: $X_1, X_2, ..., X_n$ u.i.f. normal(μ,σ). Videre er

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{og} \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

- **Til pkt. 3+4+5+6+7**: " To utvalg, normalfordelt". De to utvalgene er trukket uavhengige av hverandre, og vi har:
  $X_1^{(1)}, X_2^{(1)}, ..., X_{n_1}^{(1)}$ u.i.f. normal(μ_1,σ_1), med estimatorer

$$\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i \qquad \text{og} \qquad S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i^{(1)} - \overline{X}_1)^2$$

  For det andre utvalget: $X_1^{(2)}, X_2^{(2)}, ..., X_{n_2}^{(2)}$ u.i.f. normal(μ_2,σ_2), med estimatorer

$$\overline{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_j \qquad \text{og} \qquad S_2^2 = \frac{1}{n_2-1}\sum_{j=1}^{n_2}(X_j^{(2)} - \overline{X}_2)^2$$

- **Til pkt. 4:** "Pooled" varians:

$$S_p^2 = \frac{\sum_{i=1}^{n_1}(X_i^{(1)} - \overline{X}_1)^2 + \sum_{j=1}^{n_2}(X_j^{(2)} - \overline{X}_2)^2}{n_1 + n_2 - 2} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}.$$

- **Til pkt. 6:** Antall frihetsgrader for to-utvalg med forskjellige ukjente varianser:

$$\nu^* = \frac{(s_1^2/n_1)^2 + (s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}.$$

  Denne må avrundes til nærmeste heltall (skal ikke huske denne formelen).

- **Til pkt. 7:** Vi vil teste forskjellen mellom to utvalg, $\mu_1 - \mu_2 = \mu_d$. Testene kan ikke gjøres uten at andre faktorer spiller inn, derfor settes hver test opp til å sammenlikne to individ (enheter), ett fra hvert utvalg. Her er n lik antall slike parvise sammenligninger ($n = n_1 = n_2$), dvs halvparten av antall elementer totalt i de to utvalgene.

- **Til pkt. 8:** Når vi lager konfidensintervall for andelen p, bruker vi $\widehat{P} = \frac{X}{n}$ som estimator for p i variansen under rot-tegnet i nevneren av observatoren. Når vi tester hypoteser om p så setter vi inn verdien under nullhypotesen, p_0, i nevneren når vi skal ha fordelingen til testobservatoren under nullhypotesen. Når n er liten, eller p er nær 0 eller 1, så kan vi ikke bruke normaltilnærmingen. En tommelfinger-regel er at $np \geq 5$, $n(1-p) \geq 5$. For hypotesetesting så har vi jobbet med en andel og funnet forkastningsområde og/eller p-verdi basert på binomisk fordeling. For intervallestimering har vi bare sett på normaltilnærmingen.

- **Til pkt. 9:** Når vi lager konfidensintervall for differensen mellom to andeler, bruker vi $\widehat{P}_1 = \frac{X_1}{n_1}$ som estimator for p_1 og $\widehat{P}_2 = \frac{X_2}{n_2}$ som estimator for p_2 i variansen under rot-tegnet i nevneren av observatoren. Når vi tester hypoteser om $p_1 - p_2$ så ser vi gjerne på tilfellet der $p_1 - p_2 = 0$ som nullhypotese, og da kan vi lage en *pooled* estimator for p_1 og p_2 til å bruke i nevneren for testobservatoren under nullhypotesen,

$$\hat{P}_p = \frac{X_1 + X_2}{n_1 + n_2}$$

  Vi setter da inn $\hat{P}_p$ for både p_1 og p_2 i variansen under rot-tegnet i nevneren av observatoren. Samme reservasjon om når normaltilnærmingen kan brukes som for pkt. 8.

Konfidensintervall

i) Normalfordeling : Konfidensintervall for normalfordeling beregnes fra:

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

ii) t-fordeling : Konfidensintervall for t-fordeling beregnes fra:

$$P(-t_{\alpha/2,\nu} < T < t_{\alpha/2,\nu}) = 1 - \alpha$$

iii) Kji-kvadratfordeling : Konfidensintervall for kji-kvadratfordeling beregnes fra:

$$P(\chi^2_{1-\alpha/2,\nu} < V < \chi^2_{\alpha/2,\nu}) = 1 - \alpha$$

Hypotesetesting

(se også s 313 i læreboka)

1) Null- og alternativ hypotese: Ut fra situasjonen som er beskrevet setter vi først opp en nullhypotese, H_0, som er en "konservativ" fremstilling av verden (som vi vil ha sterke bevis for å forkaste), og en alternativ hypotese H_1. Hypotesetesten kan være ensidig (enten $\leq$ vs. $>$, eller $\geq$ vs. $<$) eller tosidig ($=$ vs. $\neq$).

2) Signifikansnivå: Vi velger så signifikansnivå, α, som angir hvor stor sannsynlighet vi godtar for å begå en "Type I" feil, dvs. å forkaste H_0 når den en sann. Vi sammenligner dette med en tiltalt som er uskyldig til han er bevist skyldig, og vi vil ha liten sannsynlighet for å begå justismord (forkaste nullhypotesen nå den er sann). Dette er den type feil vi er mest bekymret for å gjøre.

	H_0 sann	H_0 falsk
Aksepter H_0	Korrekt	Type-II feil
Forkast H_0	Type-I feil	Korrekt

3) Testobservator og forkastningsområde: Med utgangspunkt i situasjonen, så velger vi en observator, fra tabellen, og lager en testobsevator ved å sette inn parameterverdien under nullhypotesen. Forkastningsområdet beregnes fra
P(begå type I feil)=P(forkaste $H_0|H_0$ sann)$\leq \alpha$. Vi lar k være en "kritisk verdi" som vi bruker til å definere forkastningsområdet.

a) Tosidig test: H_0 (med $=$) mot H_1 (med $\neq$)

i) Normalfordeling: Løser $P(|Z| \geq k|H_0 \text{ sann}) \leq \alpha$, som gir $k = z_{\alpha/2}$. Dvs. forkast H_0 hvis $|Z| \geq z_{\alpha/2}$.

ii) t-fordeling: Løser $P(|T| \geq k|H_0 \text{ sann}) \leq \alpha$, som gir $k = t_{\alpha/2,\nu}$ med antallet frihetsgrader som testobservatoren har. Dvs. forkast H_0 hvis $|T| \geq t_{\alpha/2,\nu}$.

iii) Kji-kvadrat fordeling: Løser $P(k_1 \leq V \leq k_2|H_0 \text{ sann}) \leq \alpha$, som gir $k_1 = \chi^2_{1-\alpha/2,\nu}$ og $k_1 = \chi^2_{\alpha/2,\nu}$. Dvs. forkast H_0 hvis $V \leq \chi^2_{1-\alpha/2,\nu}$ eller $V \geq \chi^2_{\alpha/2,\nu}$.

b) Ensidig test: H_0 (med $\leq$) mot H_1 (med $>$), dvs. forkaster H_0 for store verdier av testobservatoren.

i) Normalfordeling: Løser $P(Z \geq k|H_0 \text{ sann}) \leq \alpha$, som gir $k = z_\alpha$. Dvs. forkast H_0 hvis $Z \geq z_\alpha$.

ii) t-fordeling: Løser $P(T \geq k|H_0 \text{ sann}) \leq \alpha$, som gir $k = t_{\alpha,\nu}$ med antallet frihetsgrader som testobservatoren har. Dvs. forkast H_0 hvis $T \geq t_{\alpha,\nu}$.

iii) Kji-kvadrat fordeling: Løser $P(V \geq k|H_0 \text{ sann}) \leq \alpha$, som gir $k = \chi^2_{\alpha,\nu}$. Dvs. forkast H_0 hvis $V \geq \chi^2_{\alpha,\nu}$.

c) Ensidig test: H_0 (med $\geq$) mot H_1 (med $<$), dvs. forkaster H_0 for små verdier av testobservatoren Som over, men forkaster når testobservatoren er liten (mindre eller lik gitte grense), slik at nå byttes z_α med $-z_\alpha$, $t_{\alpha,\nu}$ med $-t_{\alpha,\nu}$ og $\chi^2_{\alpha,\nu}$ med $\chi^2_{1-\alpha,\nu}$.

4) Beregninger: Sett inn verdi av observasjoner og beregn verdi for testobservatoren. Se om denne ligger i forkastningsområdet definert over. Ligger observert verdi av testobservatoren i forkastningsområdet så forkaster du nullhypotesen og konkluderer med at den alternative hypotesen gjelder. Ligger observert verdi av testobservatoren utenfor forkastningsområdet så finnes det ikke tilstrekkelig bevis til å forkaste nullhypotesen.

5) Andre størrelser :

- Det er også mulig å bytte ut beregningen av forkastningsområde med en p-verdi. P-verdi regnes ut som
  P(det du har observert eller noe verre$|H_0$ er sann).
- Det kan også være av interesse å se på styrken til testen. Da ser i på
  P(forkaste $H_0|$ ny verdi av parameteren)=$1 - \beta$, der β er sannsynligheten for å begå en "Type II feil". Denne størrelsen er avhengig av hvilken verdi vi velger for parameteren (dvs. en funksjon av parameterens sanne verdi).
- Vi har også sett på hvor mange observasjoner vi må ha for å oppdage et gitt avvik fra nullhypotesen med en gitt styrke for situasjonen i pkt. 1 i tabellen.