

Ekstraoppgave 2

I denne oppgaven skal vi studere en metode, kalt *Bonferronis metode*, som kan brukes til å bestemme simultane konfidensintervall. For enveis variansanalyse gir Bonferronis metode et alternativ til Tukeys metode som kan brukes også når det er forskjellig antall observasjoner i gruppene.

Vi vil først vise Bonferronis ulikhet (som også blir kalt Booles ulikhet). Denne ulikheten sier at hvis A_1, A_2, \dots, A_k er begivenheter (som ikke trenger å være disjunkte), så har vi at

$$P\left(\bigcup_{i=1}^k A_i\right) \leq \sum_{i=1}^k P(A_i) \quad (1)$$

- Vis at (1) gjelder for $k = 2$. *Vink:* Bruk addisjonssetningen gitt nederst på side 60 i Devore & Berk (D&B).
- Bruk induksjon til å vise at (1) gjelder for vilkårlig verdi av k .

En versjon av Bonferronis ulikhet er

$$P\left(\bigcap_{i=1}^k A_i\right) \geq 1 - \sum_{i=1}^k \{1 - P(A_i)\} \quad (2)$$

- Vis ulikheten (2). *Vink:* Bruk De Morgans lov $(\bigcap_{i=1}^k A_i)' = \bigcup_{i=1}^k A_i'$, der A_i' er den komplementære hendelse til A_i .

Anta nå at vi har konfidensintervall for k parametere $\theta_1, \theta_2, \dots, \theta_k$. Mer presist lar vi (L_i, U_i) være et $100(1 - \alpha_i)\%$ konfidensintervall for θ_i ; $i = 1, 2, \dots, k$. Vi har da at $P(L_i < \theta_i < U_i) = 1 - \alpha_i$.

- Vis at

$$P(L_i < \theta_i < U_i \text{ for } i = 1, 2, \dots, k) \geq 1 - \sum_{i=1}^k \alpha_i \quad (3)$$

Vink: Bruk (2) med $A_i = \{L_i < \theta_i < U_i\}$.

Merk at hvis vi velger $\alpha_i = \alpha/k$ så gir (3) at konfidensintervallene (L_i, U_i) ; $i = 1, 2, \dots, k$; har *simultan* konfidenskoeffisient som minst er lik $100(1 - \alpha)\%$. Konfidensintervall lagd på denne måten kalles Bonferroni konfidensintervall.

Vi ser nå på situasjonen i avsnittene 11.1 – 11.3 i D&B og vi bruker notasjonen derfra. Vi vil imidlertid ikke kreve at det er like mange observasjoner i de I gruppene, og vi lar J_i være antall observasjoner i gruppe i (jf. side 576 i D&B). Vi er interessert i å lage konfidensintervall for $\mu_i - \mu_j$ for alle $i < j$.

e) Vis at Bonferroni konfidensintervallene

$$\bar{X}_i. - \bar{X}_j. \pm t_{\alpha/(2k), n-I} \sqrt{MSE \left(\frac{1}{J_i} + \frac{1}{J_j} \right)} \quad (4)$$

med $k = I(I - 1)/2$ har simultan konfidenskoeffisient som minst er lik $100(1 - \alpha)\%$.

Når $J_i = J$ for $i = 1, 2, \dots, k$ gjelder Tukeys simultane konfidensintervall, jf. side 566 i D&B. For denne situasjonen er både Bonferroni konfidensintervallene og Tukey konfidensintervallene av formen

$$\bar{X}_i. - \bar{X}_j. \pm c\sqrt{MSE/J}, \quad (5)$$

men konstanten c er ikke den samme for de to metodene.

d) Bruk R til å bestemme c i (5) for Bonferronis og Tukeys metoder når $I = 4, 6, 10$ og $J = 4, 6, 10$. Kommenter de resultatene du finner. *Vink:* Du kan bruke kommandoene `qt` og `qtukey` til å finne persentiler for t -fordelingen og for den studentifiserte variasjonsbredden.