12.4 Testing av hypotesar med kontrastar

La oss tenkje oss følgjande tabell.

μ_1	μ_2	μ_3	μ_4
1	-1	0	0
0	0	1	-1
1	1	-1	-1

Denne beskriv 3 kontrastar

1 $\mu_1 - \mu_2 = 0$

2 $\mu_3 - \mu_4 = 0$

3. $\mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$

Definisjon 12.4.1

La $\mu_1, \mu_2, \dots, \mu_k$ vere somme forventingsverdiar til k nivå av ein faktor. Ein lineær kombinasjon, $C = \sum_{j=1}^{k} c_j \mu_j$ av $\mu_1, \dots, \mu_k$ seiast å vere ein kontrast dersom $\sum_{j=1}^{k} c_j = 0$.

For 1 er $c_1 = 1,\ c_2 = -1,\ c_3 = 0,\ c_4 = 0$

For 2 er $c_1 = 0,\ c_2 = 0,\ c_3 = 1,\ c_4 = -1$

For 3 er $c_1 = 1,\ c_2 = 1,\ c_3 = -1,\ c_4 = -1$

Med $k = 5$ kan ein aktuell kontrast vere

$$C = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5)$$

Kontrastar bør brukast for å undersøke planlagte samanlikningar.

1. Velg kontrastane før du ser på resultatet
2. Utfør ikkje testar på kontrastane dersom null hypotesa $\mu_1 = \mu_2 = \cdots = \mu_k$ ikkje er forkasta
3. Bruk ortogonale kontrastar (kan vere vanskeleg).

La $C = \sum_{j=1}^{k} c_j \mu_j$, $\hat{C} = \sum_{j=1}^{k} c_j \bar{Y}_{\cdot j}$

$$E[\hat{C}] = C, \quad Var[\hat{C}] = \sum_{j=1}^{k} c_j^2 \cdot \frac{\sigma^2}{n_j} = \sigma^2 \sum_{j=1}^{k} \frac{c_j^2}{n_j}$$

$$\frac{\hat{C} - C}{\sqrt{Var(\hat{C})}} \sim N(0,1) \Rightarrow \frac{(\hat{C} - C)^2}{Var(\hat{C})} \sim \chi^2(1)$$

σ^2 oftast ukjend og ein estimator for $Var[\hat{C}] = MSE \sum_{j=1}^{k} \frac{c_j^2}{n_j}$

$SSE = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{\cdot j})^2$ er uavh av alle $\bar{Y}_{\cdot j}$

Teorem 12.4.2.

Vi får

$$\frac{\dfrac{\hat{C}^2}{\sigma^2 \sum_{j=1}^{k} \frac{c_j^2}{n_j}}}{\dfrac{SSE}{\sigma^2 (n-k)}} = \frac{\hat{C}^2}{\frac{SSE}{n-k} \sum_{j=1}^{k} \frac{c_j^2}{n_j}} = F \sim F_{1,\, n-k}$$

dersom $H_0: C = 0$ er oppfylt og vi forkastar H_0 dersom $F \geq F_{1-\alpha, 1, n-k}$.

Definisjon 12.4.2

La $C = \sum_{j=1}^{k} c_j \mu_j$. Kvadratsummen assosiert med C er

gitt ved $SS_C = \dfrac{\hat{C}^2}{\sum_{j=1}^{k} \frac{c_j^2}{n_j}}$ der $\hat{C}^2 = \sum_{j=1}^{k} c_j \bar{Y}_{\cdot j}$

La $C_1 = \sum_{j=1}^{k} c_{1j} \mu_j$ og $C_2 = \sum_{j=1}^{k} c_{2j} \mu_j$ vere 2 kontrastar

C_1 og C_2 seiast å vere ortogonale dersom

$$\sum_{j=1}^{k} \frac{c_{1j} c_{2j}}{n_j} = 0$$

Eit eksempel

Anova med $k = 3$, $n = k \cdot r$ som betyr at det er like mange observasjonar i kvar gruppe.

La

$$C_1 = \mu_1 - \mu_2 + 0\mu_3$$

$$C_2 = \mu_1 + \mu_2 - 2\mu_3$$

$\sum_{j=1}^{3} c_{1j} c_{2j} = 0$ slik at C_1 og C_2 er ortogonale.

$$SS_{C_1} = \frac{r(\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 2})^2}{2} \qquad SS_{C_2} = \frac{r(\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} - 2\bar{Y}_{\cdot 3})^2}{6}$$

$$SS_{C_1} + SS_{C_2} = r\left[\frac{3(\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 2})^2 + (\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} - 2\bar{Y}_{\cdot 3})^2}{6}\right]$$

$$\frac{SS_{C_1} + SS_{C_2}}{\sigma^2} \sim \chi^2(2) \quad \text{under } H_0: C_1 = C_2 = 0$$

$$SSTR = \sum_{j=1}^{3} n(\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2 = n\left[\left(\bar{Y}_{\cdot 1} - \frac{\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} + \bar{Y}_{\cdot 3}}{3}\right)^2\right.$$

$$\left. + \left(\bar{Y}_{\cdot 2} - \frac{\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} + \bar{Y}_{\cdot 3}}{3}\right)^2 + \left(\bar{Y}_{\cdot 3} - \frac{\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} + \bar{Y}_{\cdot 3}}{3}\right)^2\right]$$

$$= n\left[\left(\frac{2\bar{Y}_{\cdot 1} - (\bar{Y}_{\cdot 2} + \bar{Y}_{\cdot 3})}{3}\right)^2 + \left(\frac{2\bar{Y}_{\cdot 2} - (\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 3})}{3}\right)^2 + \left(\frac{2\bar{Y}_{\cdot 3} - (\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2})}{3}\right)^2\right]$$

Samanliknar $SS_{C_1} + SS_{C_2}$ med $SSTR$

	$SS_{C_1} + SS_{C_2}$	$SSTR$
$\bar{Y}_{\cdot 1}$	$\frac{2n}{3}\bar{Y}_{\cdot 1}^2$	$\frac{2}{3}n\bar{Y}_{\cdot 1}^2$
$\vdots$	$\vdots$	$\vdots$

Vi får $SSTR = SS_{C_1} + SS_{C_2}$. Under H_0 er

$\frac{SSTR}{\sigma^2} \sim \chi^2(2)$ og $\frac{SS_{C_1} + SS_{C_2}}{\sigma^2} \sim \chi^2(2)$

Teorem 12.4.1

La $\{C_i\}_{i=1}^{k-1}$ vere ei mengde med $k-1$ ortogonale kontrastar

og la $\hat{C}_i = \sum_{j=1}^{k} c_{ij}\bar{Y}_{\cdot j}$, $i = 1, 2, \dots, k-1$ vere estimatorane.

Då er $SSTR = SS_{C_1} + \dots + SS_{C_{(k-1)}}$