ST4241: Design and Analysis of Clinical Trials

2009/2010: Semester I

Tutorial 6

1. In Example 3 of Lecture notes 5, a study comparing two methods for treating children with moderate neurological problems is considered. The children were stratified according to their sex and their parent's social class. The response is a measure of neurological functioning obtained eight weeks after the start of the treatment. The summary data is in the following table:

	Social class	Sex	Treatment 1			Treatment 2		
			n	X	s	n	X	s
1	L	F	41	1.38	0.22	40	1.36	0.28
2	L	M	41	1.26	0.25	38	1.28	0.19
3	M	F	33	1.51	0.31	35	1.41	0.27
4	M	M	45	1.46	0.28	46	1.39	0.33
5	H	F	18	1.61	0.34	20	1.51	0.41
6	H	M	23	1.59	0.46	23	1.44	0.30

(i) Compute the pooled s^2.

The s^2 is computed by the following R-code:

```
s = c(0.22,0.25,0.31,0.28,0.34,0.46,0.28,0.19,0.27,0.33,0.41,0.30)
n = c(41,41,33,45,18,23,40,38,35,46,20,23)
s2 = sum(s^2 *(n-1))/(sum(n)-12)
```

$s^2 = 0.08830563$.

(ii) Compute $\bar{d}$, the weighted average of the mean response differences between the two treatments across the 6 strata.

$\bar{d}$ is computed by the following R-code:

```
X1 = c(1.38,1.26,1.51,1.46,1.61,1.59)
X2 = c(1.36,1.28,1.41,1.39,1.51,1.44)
n1 = c(41,41,33,45,18,23)
n2 = c(40,38,35,46,20,23)
d = X1-X2
w = (n1*n2)/(n1+n2)
dbar = sum(w*d)/sum(w)
```

$\bar{d} = 0.05933681$.

(iii) Compute ISS and the F-statistics for testing the significance of interaction effect. Give the degrees of freedom of the F-statistic.

ISS *and the corresponding* F *are computed by:*

```
iss = sum(w*(d-dbar)^2)
F.iss = iss/5/s2
```

ISS $= 0.2963276$, $F_{\text{ISS}} = 0.6711409$.

The df's for the F *are 5 and 391.*

(iv) Compute the F-statistics for testing the significance of the overall mean difference of the two treatments. Give the degrees of freedom of the F-statistic.

It is computed by

```
F.tss = dbar^2*sum(w)/s2
```

$F_{\text{TSS}} = 4.014026$.

The df's for this F *are 1 and 391.*

2. For the comparison of more than two treatments in a stratified design, the treatment sum of squares is defined as TSS $= s^2\bar{\boldsymbol{d}}'\boldsymbol{S}^{-1}\bar{\boldsymbol{d}}$, where

$$\boldsymbol{S} = s^2\left[\sum_{a=1}^{A}\boldsymbol{W}_a\right]^{-1}, \quad \bar{\boldsymbol{d}} = \left[\sum_{a=1}^{A}\boldsymbol{W}_a\right]^{-1}\sum_{a=1}^{A}\boldsymbol{W}_a\boldsymbol{d}_a.$$

Here $\boldsymbol{d}_a = \boldsymbol{C}\boldsymbol{x}_a$, where $\boldsymbol{x}_a$ is the vector of treatment means in stratum a and $\boldsymbol{C}$ is a $g-1$ by g matrix whose rows are independent contrast vectors, and $\boldsymbol{W}_a = s^2[\text{Var}(\boldsymbol{d}_a)]^{-1}$.

(i) If $\boldsymbol{C}^*$ is another matrix of independent contrast vectors and $\textsc{tss}^*$ is defined the same as $\textsc{tss}$ but with $\boldsymbol{C}$ replaced by $\boldsymbol{C}^*$, show that $\textsc{tss}^* = \textsc{tss}$, i.e., the definition of the treatment sum of square does not depend on any particular matrix $\boldsymbol{C}$. (Hint: $\boldsymbol{C}^*$ can be expressed as $B\boldsymbol{C}$ where B is a $g-1$ by $g-1$ matrix of full rank.)

Since the rows of C are linearly independent, they in fact constitute a basis for the $g-1$ dimensional linear space of all contrast vectors. Hence, we have $\boldsymbol{C}^* = B\boldsymbol{C}$ for some B. Since the rows of $\boldsymbol{C}^*$ are also linearly independent, B is of full rank, i.e., it is invertible. Then we have

$$\begin{aligned}
\boldsymbol{d}_a^* &= \boldsymbol{C}^*\boldsymbol{x}_a = B\boldsymbol{C}\boldsymbol{x}_a = B\boldsymbol{d}_a,\\
\boldsymbol{W}_a^* &= s^2[\,Var(\boldsymbol{d}_a^*)]^{-1} = s^2[B\,Var(\boldsymbol{d}_a)B']^{-1} = B'^{-1}s^2[\,Var(\boldsymbol{d}_a)]^{-1}B^{-1}\\
&= B'^{-1}\boldsymbol{W}_aB^{-1},\\
\sum \boldsymbol{W}_a^* &= B'^{-1}\sum \boldsymbol{W}_aB^{-1}.
\end{aligned}$$

Thus

$$\begin{aligned}
\boldsymbol{S}^* &= s^2\left[\sum \boldsymbol{W}_a^*\right]^{-1} = s^2[B'^{-1}\sum \boldsymbol{W}_aB^{-1}]^{-1}\\
&= Bs^2[\sum \boldsymbol{W}_a]^{-1}B' = B\boldsymbol{S}B';\\
\bar{\boldsymbol{d}}^* &= \left[\sum \boldsymbol{W}_a^*\right]^{-1}\sum_{a=1}^{A}\boldsymbol{W}_a^*\boldsymbol{d}_a^* = B\left[\sum \boldsymbol{W}_a\right]^{-1}B'B'^{-1}\sum_{a=1}^{A}\boldsymbol{W}_aB^{-1}B\boldsymbol{d}_a\\
&= B\left[\sum_{a=1}^{A}\boldsymbol{W}_a\right]^{-1}\sum_{a=1}^{A}\boldsymbol{W}_a\boldsymbol{d}_a = B\bar{\boldsymbol{d}};\\
\textsc{tss}^* &= \bar{\boldsymbol{d}}^{*\prime}\boldsymbol{S}^{*-1}\bar{\boldsymbol{d}}^*\\
&= \bar{\boldsymbol{d}}'B'B'^{-1}\boldsymbol{S}^{-1}B^{-1}B\bar{\boldsymbol{d}} = \textsc{tss}.
\end{aligned}$$

(ii) Show that the interaction sum of squares defined by

$$\textsc{iss} = \sum_{a=1}^{A}(\boldsymbol{d}_a - \bar{\boldsymbol{d}})'\boldsymbol{W}_a(\boldsymbol{d}_a - \bar{\boldsymbol{d}})$$

does not depend on $\boldsymbol{C}$ either.

$$\begin{aligned}
\text{ISS}^* &= \sum_{a=1}^{A} (\boldsymbol{d}_a^* - \bar{\boldsymbol{d}}^*)' \boldsymbol{W}_a^* (\boldsymbol{d}_a^* - \bar{\boldsymbol{d}}^*) \\
&= \sum_{a=1}^{A} (\boldsymbol{d}_a - \bar{\boldsymbol{d}})' B' B'^{-1} \boldsymbol{W}_a B^{-1} B (\boldsymbol{d}_a - \bar{\boldsymbol{d}}) \\
&= \sum_{a=1}^{A} (\boldsymbol{d}_a - \bar{\boldsymbol{d}})' \boldsymbol{W}_a (\boldsymbol{d}_a - \bar{\boldsymbol{d}}) = \text{ISS}.
\end{aligned}$$

3. In the example of Lecture notes 5 for studying the changes in systolic blood pressure associated with four treatments in a stratified design, we have the summary data:

		Treatment			
Stratum		1	2	3	4
1	n	6	5	3	5
	Mean	29.33	28.00	16.33	13.60
	sd	13.02	10.98	14.19	10.55
2	n	4	4	5	6
	Mean	28.25	33.50	4.40	12.83
	sd	5.85	2.08	6.91	10.34
3	n	5	6	4	5
	Mean	20.40	18.17	8.50	14.20
	sd	13.37	12.53	9.00	8.93

The following R-codes are used for the computation of the desired results:

```
x=c(29.33333,28.25,20.4,
     28,33.5,18.166666667,
     16.333333333,4.4,8.5,
     13.6,12.83333333,14.2)
n=c(6,4,5,5,4,6,3,5,4,5,6,5)
S=factor(rep(c(1,2,3),4))
T=factor(c(rep(1,3),rep(2,3),rep(3,3),rep(4,3)))
options(contrasts=c("contr.treatment","contr.poly"))
lm.fit=lm(x~S+T, x=T, weights=n)
```

From the fitted object `lm.fit` the following anova table and variance-covariance matrix are extracted:

```
          Df  Sum Sq Mean Sq F value  Pr(>F)
S          2  488.64  244.32  2.0727 0.20685
T          3 3063.43 1021.14  8.6627 0.01338 *
Residuals  6  707.27  117.88
```

```
            (Intercept)      S2      S3      T2      T3      T4
(Intercept)     11.2265 -5.4680 -5.7295 -7.4765 -7.0383 -7.3855
S2              -5.4680 12.6511  6.2831 -0.4189 -1.8977 -1.2396
S3              -5.7295  6.2831 12.1619 -0.8108 -0.9425 -0.4273
T2              -7.4765 -0.4189 -0.8108 15.7711  7.9213  7.8870
T3              -7.0383 -1.8977 -0.9425  7.9213 17.9663  8.0445
T4              -7.3855 -1.2396 -0.4273  7.8870  8.0445 15.3513
```

(i) Compute the pooled s^2.

The s^2 is computed as follows:

```
sd =c(13.02,5.85,13.37,10.98,2.08,12.53,
      14.19,6.91,9,10.55,10.34,8.93)
s2 = sum(sd^2*(n-1))/(sum(n)-12)
```

$s^2 = 110.4564.$

(ii) Describe the linear model behind the R-codes given above.

The model is as follows:

$$X = \mu_0 + \sum_{i=1}^{2} \alpha_i s_i + \sum_{j=1}^{3} \beta_j t_j + \epsilon,$$

where s_i and t_j are defined as follows:

$$s_i = \begin{cases} 1, & \textit{if stratum } i+1, \\ 0, & \textit{otherwise}, \ i = 1, 2; \end{cases} \quad t_j = \begin{cases} 1, & \textit{if treatment } j+1, \\ 0, & \textit{otherwise}, \ j = 1, 2, 3; \end{cases}$$

(iii) Let μ_j denote the mean response of treatment $j, j = 1, 2, 3, 4$. Express the following three contrasts in terms of the parameters of the linear model:

$$C_1 = \mu_1 - \mu_4, \ C_2 = \mu_2 - \mu_4, \ C_3 = \mu_3 - \mu_4.$$

In terms of the parameters of the linear model, the three contrasts are equivalent to

$$\tilde{C}_1 = \beta_3, \; \tilde{C}_2 = \beta_1 - \beta_3, \; \tilde{C}_3 = \beta_2 - \beta_3.$$

(iv) Test the significance of the three contrasts in (iii) using an appropriate multiple comparison procedure to control the overall error rate at 0.05.

The test statistics for the three contrasts are computed using the R code below. The estimates of β_1, β_2 and β_3 and their variance-covariance matrix are extracted from the fitted object `lm.fit` *first. Note that the covariance matrix is adjusted by dividing by 117.88 and multiplying by s^2. (Think why). Then the test statistics are constructed in the usual way.*

```
V=vcov(lm.fit)
V.b = V[4:6,4:6]/117.88*s2
b = lm.fit$coef[4:6]
c1 = c(0,0,1)
c2 = c(1,0,-1)
c3 = c(0,1,-1)
L1 = t(c1)%*%b / sqrt(  t(c1)%*%V.b%*%c1 )
L2 = t(c2)%*%b / sqrt(  t(c2)%*%V.b%*%c2 )
L3 = t(c3)%*%b / sqrt(  t(c3)%*%V.b%*%c3 )
```

$L1 = -3.287631, L2 = 3.260416, L3 = -1.126660.$

The Dunnett's criterion should be used. The dfs for Dunnett's criterion are $p = 3$ and $\nu = 46$. From Table A.6 of Fleiss, we found the following critical values at level 0.05 for two-sided test: $d_{3,40,0.05} = 2.44$, $d_{3,60,0.05} = 2.41$. By interpolaton, we obtain

$$d_{3,46,0.05} = (46 - 40)\frac{2.41 - 2.44}{60 - 40} + 2.44 = 2.431.$$

According to Dunnett's criterion, the first two contrasts are significant, but not the third one.