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 n	Treatment 1 X	Treatment 1 s	Treatment 2 n	Treatment 2 X	Treatment 2 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.

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

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

(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.

In all the above four parts, show the detail of your computation.

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 TSS* is defined the same as TSS but with $\boldsymbol{C}$ replaced by $\boldsymbol{C}^*$, show that TSS* = 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.)

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

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

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

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:

Stratum		Treatment 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.

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

(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:

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

(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.

Note: for those who are required to submit the assignment, please submit the solutions to Question 1 and 2 only.