McGill University
Faculty of Science

Department of Mathematics and Statistics

April 2008

MATH 204

PRINCIPLES OF STATISTICS II

SOLUTIONS

1. (a) This is a completely randomized design with a single factor, Group, that takes five levels. It is a balanced and complete design, as all treatments (factor levels) have the same non-zero number (five) of observations.

6 MARKS

(b) The completed table is

SOURCE	DF	SS	MS	F
GROUP	4	19.613	4.903	2.136
ERROR	20	45.903	2.295	
TOTAL	24	65.516		

10 MARKS

(c) The ANOVA F-test tests the hypotheses

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 &: \mu_i \neq \mu_j, \text{ for some } i, j \end{aligned}$$

The null distribution is Fisher$(4, 20)$; comparing the F-statistic (2.136) with the 95 % quantile of the Fisher$(4, 20)$ distribution, we see from the table given that the critical value is 2.87, so we do not reject H_0 - there is no evidence that there is any difference in HDRS reduction.

5 MARKS

(d) The p-value in this table reveals that Levene's test for homogeneity of variances does not reject the null hypothesis of equal variances at the $\alpha = 0.05$ significance level, although the result is almost significant. Equality of variances is essential for the ANOVA F-test to give a valid result.

Given the study design, with independent observations being used, the only assumption that needs to be checked is the Normality assumption, that could be checked using boxplots (or residual plots). Also, to verify the results without considering the normality assumptions, could use permutation or randomization tests.

4 MARKS

2. (a) Sample kr experimental units from each of the b subpopulations identified by the levels of the blocking factor, and then allocate r experimental units at random to each of the treatments.

4 MARKS

(b) The four models are

$$1 \qquad \text{TMT} \qquad \text{B} \qquad \text{TMT} + \text{B}$$

which have 1, k, b and $k + b - 1$ parameters in total. No interaction models can be fitted if there are no replicates.

4 MARKS

(c) (i) The ANOVA table indicates that both main effects and interaction are significant at $\alpha = 0.05$. The blocking factor is confirmed as leading to a significantly different response. The interaction term is not significant at $\alpha = 0.01$, indicating that this term is not as important as the other terms in the model.

6 MARKS

(ii) The significant interaction means that the effect of changing treatment level is **different** in the **different** age groups, that is, in a plot demonstrating the group means, lines plotting mean response against age group are **not piecewise parallel**.

6 MARKS

(d) The Age Group 60+ at dose level 4 subgroup corresponds to the baseline group, thus we can get the difference in expected response by adding the main effect for Age Group 1, the main effect for dose level 1, and the interaction for those two groups. This gives

$$(-3.240) + (-4.602) + 5.785 = -2.057$$

5 MARKS

3. (a) This is a factorial design that is complete but not balanced.

3 MARKS

(b) The five models are

Model 1	L + E + A + L.E + L.A + A.E + L.E.A
Model 2	L + E + A + L.E
Model 3	L + E + L.E
Model 4	L + E
Model 5	L

6 MARKS

(c) Extracting the analysis results we have

Comparison	SSE$_C$	SSE$_R$	k	g	F	F_α	Significant at $\alpha = 0.05$
1 vs 2	7	4	16.892	17.253	0.150	3.07	No
2 vs 3	4	3	17.253	18.527	1.772	4.26	No
3 vs 4	3	2	18.527	24.577	8.164	4.24	Yes
4 vs 5	2	1	24.577	34.617	10.621	4.23	Yes

Thus it seems that model 3 is the most appropriate model.

The overall fit is adequate with an R^2 statistic of 0.548.

10 MARKS

(d) Because the study is **not balanced**.

2 MARKS

(e) The simplest way to carry out the procedure would be to form a vector of the response variables, and a list of their corresponding subgroup allocations, and then to cycle through the permutations of the data indices, computing the Fisher-F statistics for the whole table for any hypothesized model. The p-values are computed by enumerating how many times the Fisher-F statistics are more extreme than the observed one.

This may be computationally challenging as the sample size is large, so perhaps sampling permutations would be more amenable.

4 MARKS

4. (a) The five sketches are: (0) Null Model: One horizontal straight line, (1) **Insul**: Two horizontal straight lines, (2) **Temp**: One straight line, (3) **Insul+Temp**: Two parallel straight lines, (4) **Insul*Temp**: Two non-parallel straight lines.

5 MARKS

(b) The full factorial model is the most appropriate; all terms appear to be significant in the ANOVA and in the parameter estimates table. The R^2 value for this model is very high (> 0.9), indicating a very good fit.

6 MARKS

(c) The standardized residual plot against temperature should look like a horizontal band within the bounds ± 2; here, it does, so overall the plot gives no cause for concern - possibly there is an outlier, but this is not certain. The plot allows a check of the necessary assumption that the variance of the residual errors does not depend on the values of the predictors. To check the other assumptions, specifically normality of the residuals, a histogram and P-P plot would be useful.

6 MARKS

(d) We have

$$s^2 = \mathsf{MSE} = \frac{\mathsf{SSE}}{\mathsf{rdf}} = 0.104$$

4 MARKS

(e) The best prediction is

$$((4.724) + (2.130)) + ((-0.278) + (-0.115)) \times 5 = 4.889$$

The answer changes by an amount equal to the coefficient for temperature for uninsulated houses, that is

$$(-0.278) + (-0.115) = -0.393$$

indicating a decrease of 393 cubic feet.

4 MARKS

5. (a) In X^2, the n_{ij} and $\widehat{n}_{ij}$ are the observed and expected frequencies in the cells of the table, where

$$\widehat{n}_{ij} = \frac{n_{i.}n_{.j}}{n} = \frac{i\text{th Row Total} \times j\text{th Column Total}}{\text{Total Sample Size}}$$

If the sample size is large enough, so that the expected counts are at least five, then $X^2 \stackrel{\cdot}{\sim} \chi^2_1$, so we compare the computed test statistic with the 0.95 quantile of this distribution 3.841.

8 MARKS

(b) We have fitted values

	Hay Fever		
Excema	Yes	No	Total
Yes	39.94	521.06	561.00
No	1029.06	13423.94	14453.00
Total	1069.00	13945.00	15014.00

and hence

$$X^2 = 286.00 \qquad \Longrightarrow \qquad \text{Reject } H_0$$

6 MARKS

(c) We carry out a one-way Chi-squared test of the hypothesis

$$H_0 \, : \, p_1 = 0.4, p_2 = 0.3, p_3 = 0.1, p_4 = 0.1, p_5 = 0.1$$

	Ethnic Subpopulation				
	1	2	3	4	5
Expected Proportion (%)	40	30	10	10	10
Expected Count	800	600	200	200	200
Observed Count	822	638	210	157	173
$(O_i - E_i)^2/E_i$	0.60	2.41	0.50	9.24	3.65

and hence

$$X^2 = 16.402 > 13.277 = \text{Chisq}_{0.01}(4) \qquad \Longrightarrow \qquad \text{Reject } H_0$$

so it seems that the remit has not been met.

6 MARKS

(d) We would need a sample size of 50, so that the smallest expected count is 5.

5 MARKS

6. (a) Paired data have dependence between the two samples, either by matching of the experimental units, or because of repeated sampling of the same individuals. Unpaired data are completely independent.

4 MARKS

(b) (i) Yes: this comes from Analysis 2, where the two-tailed test rejects the null hypothesis of equal medians at $\alpha = 0.05$, and the sum of positive ranks is greater than the sum of positive ranks, so the left-tailed test also rejects the null at $\alpha = 0.05$.

5 MARKS

(ii) Yes: this comes from Analysis 5, where the two-tailed p-value is 0.016.

5 MARKS

(iii) No: this comes from Analysis 3, where the two-tailed p-value is 0.239.

5 MARKS

(c) This would be a Completely Randomized Design, where the null hypothesis could be tested using one-way ANOVA (parametric) or the Kruskal-Wallis test (non-parametric).

6 MARKS