Solutions

Experimental design (MSA250/TMS031) August 22, 14:00 - 18:00
Tools: A Chalmers accepted pocket calculator with emptied memory. At the examination, sheets with statistical distributions and tables will be handed out.
Maximum number of points: 30p
Limits GU: G (15p) and VG (22p)
Limits Chalmers: 3:a (15p), 4:a (20p) and 5:a (25p)
Give explanations to the notation you use and motivation to your conclusions.

1. (5 p) We want to compare two treatments, i.e. two means (or medians) of some response variable. The aim is to find out whether the differences that are found are likely to be genuine or due to chance.

   a) Mention three different methods to make the comparison.

   b) Give the assumptions and the form of data needed for the three methods you mention in a). Discuss which of the assumptions are important and which are not as important.

   Short answer:

a) To compare means/medians of two independent samples, you can use T-test, a non-parametric test (Willcoxon), randomization or base the test on a reference distribution based on historical data. (You could have chosen to describe pairwice data as well.)

b) T-test (observations independent, each from a normal distribution, equal variances); Non-parametric test (observations independent, continuous distribution, equal variances); Randomization (equal variances); Reference distribution (historical data).

2. (3 p) Give an example of a split-plot design? When should it be used?

   Short answer: See the book.

3. (5 p) A filtration time (y) of a newly constructed plant is studied. The three factors that are believed to affect y are **A** water supply (town reservoir/well), **B** raw material (on site/other), and **C** temperature (low/high). The design and the results of the experiment (two independent replications of each run) are as follows:

Run number	**A**	**B**	**C**	y_1	y_2	$\bar{y}$
1	-	-	-	65.1	71.7	68.4
2	+	-	-	79.4	76.0	77.7
3	-	+	-	72.3	60.5	66.4
4	+	+	-	78.9	83.1	81.0
5	-	-	+	77.8	79.4	78.6
6	+	-	+	48.1	34.3	41.2
7	-	+	+	76.2	61.2	68.7
8	+	+	+	51.2	26.2	38.7

The experiments were run in random order. Which effects are active? Do you see something suspicious among the measured values above?

Short answer:

The estimated effects are $L_A = -10.875$, $L_B = -2.775$, $L_C = -16.575$, $L_{AB} = 3.175$, $L_{AC} = -22.825$, $L_{BC} = -3.425$, and $L_{ABC} = 0.525$. Furthermore, $Var(effect) = (\frac{1}{8} + \frac{1}{8})s^2 = (\frac{1}{8} + \frac{1}{8}) \sum s_i^2/8 = 19.61$ and $sd(effect) = 4.43$. This gives that A, C and AC are active. The two measurements for run 8, 51.2 and 26.2, are very different, and (at least) one of them may be incorrect.

4. (4 p) Suppose you wish to estimate the difference between two mean acidity (pH) for rainfalls at two different locations, one in a relatively unpolluted area along the ocean and the other in an area subject to heavy air pollution. If you wish your estimate to be correct to the nearest 0.1 pH with probability 90%, approximately how many rainfalls (pH values) would have to be included in each sample? Assume that the variance of the pH measurements is 0.16 at both locations and that the samples are of equal size.

   Short answer: Let X_1 be the pH in the unpolluted area and X_2 the pH in the polluted area. Also, let $X_1 \sim N(\mu_1, \sigma^2)$ and $X_2 \sim N(\mu_2, \sigma^2)$. To find the common sample size, one should compute the 90% confidence interval for $\mu_1 - \mu_2$ which has the length 0.2, i.e. $\bar{X}_1 - \bar{X}_2 \pm z_{0.05}\sqrt{2}\sigma/\sqrt{n}$ ($\text{Var}(\bar{X}_1 - \bar{X}_2) = 2\sigma^2/n$). Now $z_{0.05}\sqrt{2}\sigma/\sqrt{n} = 1.645 \cdot \sqrt{2} \cdot 0.4/\sqrt{n} = 0.1$ which gives $n = 87$ (nearest integer that is larger than the computed n).

5. (6 p)

   a) Write an eight-run two-level design of resolution III in seven factors and give its generators.
   b) How may members of this family of designs be combined to isolate (free from other effects) the main effects? Give such a combined design, its defining relations and resolution. State also the assumptions you have made.
   c) What is the advantage of running a set of experiments first and then running another set of experiments, instead of doing all the runs at once?

   Short answer:

   a) Generators are D=AB, E=AC, F=BC, G=ABC
   b) We can make another set of eight runs by changing all the signs. This would free the main effects. The defining relation of the combined design is I=ABCG=BCDE=ACDF=ABEF and the resolution is IV.
   c) The first set of run may indicate which effects are active so that it becomes easier to interpret the higher order effects after the second set of runs.

6. (7 p) Acrophobia is a fear of heights. It can be treated in a number of different ways. Here, we compare three techniques with 15 volunteers (subjects) with a history of severe acrophobia. It was noticed, however, that some of the subjects were more afraid of heights than others and that this heterogeneity might affect the therapy comparison. The experiment started with each subject being given the Height Avoidance Test (HAT). Based on the results, the subjects were divided into five groups (A, B, C, D and E) each of size 3. The subjects in group A had the lowest scores (the most severe acrophobia), those in group B the second lowest and so on. Each of the three

therapies was then assigned at random to one of the three subjects in each group. When the counseling sessions were over, the subjects retook the HAT. The changes of the scores are

Group	Therapy 1	Therapy 2	Therapy 3
A	8	2	-2
B	11	1	0
C	9	12	6
D	16	11	2
E	24	19	11

and the corresponding ANOVA table is

Source	df	SS	MS
Groups	?	438.0	?
Therapies	?	260.9	?
Residual	?	68.4	?
Total	?	767.3	

a) What is the design? Why has this particular design been chosen?

b) Fill in the ?'s in the ANOVA table and explain the numbers of degrees of freedom.

c) Is there evidence of the therapies being different? Why?

d) Can you say something about the difference between the groups based on the ANOVA table?

Good luck!

Short answer:

a) Random block design. This has been chosen in order to separate the treatment effect from the group (block) effect.

b) The ANOVA table becomes

Source	df	SS	MS
Groups	4	438.0	109.5
Therapies	2	260.9	130.4
Residual	8	68.4	8.6
Total	14	767.3	

c) Yes, since the $F = 15.2$ (MS(therapies)/MS(residuals)) is significant at the 5% level.

d) Yes, by looking at the F-ratio MS(groups)/MS(residuals), which is also significant.