BIOS 830: Homework 3

Due on March 14, 2017

February 27, 2017

<u>**Instructions**</u>: Students are encouraged to work together on this problem set. However, each student is expected to *independently* write up the assigned problems. Please provide the code for Questions 3 in a clean and readable document; points will be deducted for "messy code". **Assignments are to be turned in at the beginning of lecture on the due date above. Any assignments not turned in at this time will be considered late.**

Question 1:
Consider a balanced experiment with two crossed treatment factors, A and B. The two-way complete model is given by:

$$Y_{ijt} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijt},$$

where, $i = 1, 2, \ldots, a$, $j = 1, 2, \ldots, b$, $t = 1, 2, \ldots, n$ and $\epsilon_{ijt} \sim N(0, \sigma^2)$. When observations are extremely time-consuming or expensive to collect, an experiment may be designed to have $n = 1$ observations on each treatment combination.

(a) Assuming the two-way complete model, explain why such a scenario would be problematic. (Hint, what are the degrees of freedom for the error?)

(b) For the two-way complete model with $n = 1$ observations per treatment combination, the sums of squares for testing the null hypothesis that a particular interaction contrast is negligible, say $\sum_i \sum_j d_{ij}(\alpha\beta)_{ij} = 0$ (with $\sum_i d_{ij} = \sum_j d_{ij} = 0$), against the alternative that the contrast is non-negligible is:

$$ssc = \frac{\left(\sum_i \sum_j d_{ij} y_{ij}\right)^2}{\sum_i \sum_j d_{ij}^2}$$

Suppose it is known in advance that e orthogonal contrasts are likely to be negligible. The sums of squares for these e contrasts can be pooled together to obtain an estimate of the error variance based on e degrees of freedom:

$$ssE = \sum_{h=1}^{e} ssc_h \text{ and } msE = ssE/e$$

Prove that MSE is an unbiased estimator for σ^2. (Hint, recall the set of restrictions for the two-way complete model, i.e., $\sum_i \alpha_i = \sum_j \beta_i = \sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0$)

(c) The sums of squares for the remaining interaction contrasts for the remaining contrasts can be added together to obtain an interaction sum of squares:

$$ssAB_m = \sum_{k=e+1}^{(a-1)(b-1)} ssc_h$$

What is the decision rule for testing the hypothesis H_0^{AB}: (the interaction AB is negligible) against the alternative hypothesis that the interaction is not negligible?

Question 2:
In lecture, a formula was presented for computing sample size for the two-way complete model based on a pre-specified power for a hypothesis test:

$$n = 2a\sigma^2\phi^2/(b\Delta_A^2)$$

where Δ_A is the smallest difference in the α_i's (or $\alpha_i^\star$'s), a is the number of levels of factor A, b is the number of levels of factor B, and σ^2 is the error variance. The goal of this question is to derive the above equation. Assume equal sample sizes across the ab treatment combinations (e.g., $n_{ij} = n$ for all i, j) and a total sample size given by $N = abn$.

(a) Show that $E(MSA) = \sigma^2 + Q(\alpha_i)$, where $Q(\alpha_i) = \frac{bn}{a-1}\sum_i(\alpha_i - \bar{\alpha}_.)^2$ and $\bar{\alpha}_. = \sum_i \alpha_i/a$.

(b) What is $E(MSA)$ equal to when all of the main-effect parameters for factor A are all equal?

(c) The non-centrality parameter λ^2 is defined as $\lambda^2 = (a-1)Q(\alpha_i)/\sigma^2$. Using this, provide a step-by-step justification for how one arrives at the sample size formula given above. (Hint: follow the logic given on pages 51-52 of Dean & Voss).

Question 3:
Suppose that a group of researchers are interested in the effect of two different treatment factors (i.e., Diet (2 levels) and Exercise intensity (3 levels)) on weight loss over the course of a six-week study period. Broadly speaking, these questions can be addressed using one of two possible strategies: (1) two separate CRDs, each focusing on a different treatment factor or (2) a two-factor crossed study where subjects are randomly assigned to one of 6 possible treatment combinations. Aside from the obvious advantage that the later allows for the examination of interaction effects between the two treatment factors, it was mentioned in lecture that two-factor crossed studies are generally more powerful than running two separate studies.

(a) Assuming that the researchers have enough money to enroll a total of 100 subjects, conduct a simulation study to empirically evaluate the claim that "two-factor crossed studies are generally more powerful than running two separate studies". Be sure to state all of your assumptions and thoroughly describe what conclusions you reach.

In addition to the above questions, questions 8, 12, and 19 from Dean & Voss Chapter 6 and question 8 from Dean & Voss Chapter 7.