Homework 11, Handed out: 22 November 2009

     on campus   Friday, 4 Dec 2009, in lecture (11 am)
                      or by e-mailto Chuanlong, dclong@iastate.edu, no later than noon.
     off campus Monday, 7 Dec 2009, by 4 pm to Nicole Rembert, email: rembeall@iastate.edu or
                      FAX: 515-294-4040 (please include cover page with Stat 500 / Nicole Rembert).

1. **power of ANOVA tests** — This question reinforces a lecture point about the design of a factorial ANOVA experiment. Consider a 2 x 3 factorial, for example the data from the children's memory experiment from HW 10. That study had 6 treatments, all combinations of 2 levels of reinforcement and 3 isolation times. Consider the following four estimates:

   1. None - Verbal reinforcement, averaged over the three isolation times

   2. None - Verbal, estimated only for 20 min isolation. i.e.
      the simple effect: none, 20min - verbal, 20min

   3. 20min - 40min, averaged over the two reinforcement levels

   4. Interaction between 20 and 40 min and reinforcement i.e. the difference in simple effects:
      (none, 20min - verbal, 20min) - (none, 40min - verbal, 40min)

   (a) Compute the s.e. for each contrast. Replace the contrast coefficients with the appropriate values, but you may leave the sample size as n and the error variance as σ^2. Which contrast has the smallest standard error?
      Hint: Consider the design as a 1-way ANOVA and write out the appropriate coefficients for each of these contrasts. Then, remember the formula for the s.e. of a contrast in a 1-way ANOVA.
      Or: Work from the s.e.'s of the cell and marginal means.
      Caution: Make sure to use coefficients that estimate the contrasts specified above. You can reduce the s.e. by changing the coefficients, but then the contrasts aren't comparable.

   (b) Each of the four estimates leads to a t-test (e.g. estimate 2 can be used to test H0: None, 20 min. - Verbal, 20 min. = 0). Imagine that the 'interesting difference', δ, is the same value for all four tests. Which test is the most powerful? Which is the least?
      You don't need to calculate power or sample size to answer this question.

   (c) The investigators now need to determine a sample size. They decide to base this on estimate # 4 (the interaction contrast). Based on previous experience, they decide to use $\sigma^2 = 16$, $\alpha = 0.05$, and power = 0.80. Determine the sample size needed to detect a difference of 5 in the interaction contrast. You only need to iterate once.

2. **Three-factor factorial:** This problem uses made up data loosely based on the design of a recent consulting project. Freeway roadways are slabs of concrete joined together by metal connecting rods. The connecting rods are placed under a lot of stress by the pressure of heavy traffic. When the connecting rods loosen, stretch or break, the join between the concrete segments widens. Eventually drivers notice the rough road. This study was designed to evaluate 3 different designs of connecting rods.

   The experimental units are 1 mile stretches of freeway in the process of being rebuilt. The experimenters had permission to use a nine-mile stretch of freeway between two exits. They randomly assigned the three connecting rods to 1 mile segments, with each connecting rod type being used in 3 segments. The road contractors then built the freeway using the assigned connecting rods in the appropriate 1mile road segments. After 5 years, the average crack width is measured in 6 locations per freeway segment. The 6 locations are the outside, center, and inside wheel track in both the slow and the fast lane. Hence, the design is a 3 way factorial (3 x

3 x 2) treatment design. ROD has 3 levels (1, 2, and 3)), LOCATION has 3 levels (outside=1, center=2 and inside=3) and LANE has 2 levels (fast lane=2 and slow lane=1). For now, treat the design as if it were a completely randomized 3 way factorial. The data in freeway.txt contain the untransformed crack width for each road segment and treatment combination.

The investigators want to compare the performance of the three connecting rods. Can they recommend a single best type of rod?

(a) Analyze the crack width data as a 3 way factorial. Are the differences between rod types the same in all locations and lanes?

(b) Because of their prior knowledge of the nature of the response (crack width) and inspection of the data, the investigators decide to log transform the crack width. Reanalyze the data after log transforming the response. Are the log-scale differences between rod types the same in all locations and lanes?

(c) Summarize the differences between rod types in the simplest way possible. The investigators are interested in estimates of differences (or log differences), not tests.

3. **One-factor random effects:** In a pig breeding study two offspring from each of ten litters were measured for average daily weight gain (ADWG). The individual pig measurements can be thought of as having two pieces,
pig gain = common mean + litter effect + individual pig effect.

(a) Complete the d.f. in the analysis of variance table.

Source	d.f.	SS	E MS
litters	?	0.6705	$\sigma^2 + 2\sigma_\mu$
pigs within litters	?	0.3834	σ^2
total	?	1.0539	

(b) The problem gives an informal model statement (pig gain = common mean + litter effect + indiv effect). Write this as a formal model statement, specifying the assumptions made about each term.

(c) Obtain point estimates of the individual variance component (the error variance) and the litter variance component.
HINT: I give you the two expected MS; those are what you need to find method of moments (ANOVA) estimates of the variance components.

(d) Suppose I intend to select a pig at random from the population and measure its average daily weight gain. According to the model what is the variance associate with this quantity? What fraction of this variance is due to genetic factors (i.e., litter)?

(e) Suppose I intend to design an experiment where the treatments are assigned to mothers (i.e. litters), but the response is observed on pigs. Many teratology (effects of drugs or other chemicals while critters are in utero) studies are designed this way. The response of interest is a treatment mean, averaged over r litters and n pigs. The investigators can use one of three designs:
    a) r=2 litters, 4 pigs per litter
    b) r=4 litters, 2 pigs per litter
    c) r=8 litters, 1 pig per litter
Assume that the variance components estimated in part c are appropriate for this new study. Calculate the s.e. of the treatment mean for each of the three designs. Which design gives them the most precise estimate of the treatment mean?

(f) The data, as originally reported, are in pigs.txt on the class web site. Use the data to estimate the litter and individiual variance components. Report both the ANOVA estimates and REML estimates.

Bonus point: explain the ANOVA and REML estimated variance component for **individuals** are not the same.

(g) Identify a reasonable explanation why the estimated litter variance component is negative (in the ANOVA estimate).