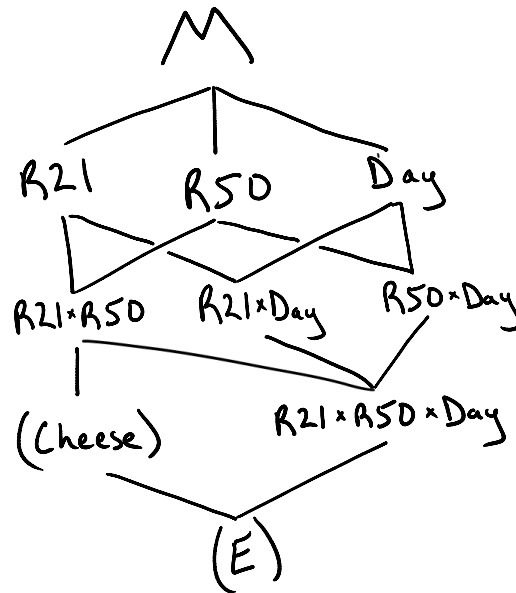


Assignment Due (by 11:59 P.M.): Mon 30 Nov

Directions: (Same as previously, with 10 points for neatness and format, including SAS code in appendix.) This assignment will be graded out of 100 points. Note that some SAS code to get you started is provided on the class website (the only changes necessary in PROC MIXED portions are putting appropriate model terms in the “...” spots). Use Type 3 (not REML) estimation here, and don’t worry about checking model assumptions.

Exercises:

1. Based on textbook Problem 11.2 (p. 277-278). For this homework, don’t account for the “pairing” that was done; i.e., treat this as a 4×5 factorial design with 2 replicates at each factor level combination. Include the interaction term in the model.
 - (a) (4 points) Explain clearly why both `lab` and `sample` are random effects.
 - (b) Consider H_0 : “no variation among labs”.
 - i. (1 point) What is the P-value for this test? (Check your work: $F=2.64$ is the test statistic.)
 - ii. (2 points) What is the conclusion (in the context of this problem; use significance threshold 0.05)?
 - (c) (3 points) With both `lab` and `sample` treated as fixed effects, consider H_0 : “no mean difference among labs”. Report (i)-(ii) as in (b). (Check: $F=16.43$ is the test statistic.)
2. Based on textbook Problem 12.7 (pp. 312-313). Use the log of free amino acids as the response. Based on the description in the textbook, the Hasse Diagram on the following page is appropriate. (Note that this diagram is only missing the degrees of freedom subscripts. Also, in the accompanying SAS code, variables `R21` and `R50` equal 1 when those bacteria strains are absent, and equal 2 when those bacteria strains are present. The textbook’s “Control” treatment corresponds to when both bacteria strains are absent (i.e., when `R21` and `R50` both equal 1).)
 - (a) (10 points) Briefly summarize which factors are crossed, which are nested, and why. (For the “why” part, refer to the description in the textbook.)
 - (b) (6 points) In terms of the mean squares (MS symbolically, not numerically), what are the test statistics for testing the effects of (i) `R21`, (ii) `R21` \times `R50`, and (iii) `R21` \times `R50` \times `Day` ?
 - (c) (6 points) What is the sampling distribution for each of the test statistics in b.i, b.ii, and b.iii above?
 - (d) Using PROC MIXED (and the design summarized in the Hasse Diagram), analyze these data. Report the P-values, conclusions, and (if appropriate) post-hoc means comparisons for the following questions:
 - i. (6 points) Is there a bacterial treatment interaction (`R21` \times `R50`)? (Check your work: P-value here is 0.6412.)
 - ii. (6 points) Is there a bacterial treatment effect due to `R21`?
 - iii. (6 points) Is there a bacterial treatment effect due to `R50`?



3. Based on textbook Problem 12.8 (p. 313). (Here, all subjects came in on one day and did both treatments (amounts taken – 10ml and 20ml), and then all the subjects returned on another day and did both treatments again.)
- (16 points) Based on the description of the experiment and the SAS code given (plus its output), draw a Hasse Diagram for this design, using the same variable (or factor) names as in the SAS code. (You may either do it on paper and scan it in, or use a line diagram in a word processing document. Be sure to include all necessary lines, indications of random effects, and subscripts to denote corresponding degrees of freedom.)
 - (3 points) From this Hasse diagram, how do you know that the test for the effect of Taken (the treatment of interest here) will involve an approximate F-test?
 - (4 points) (i) Using the PROC MIXED output, what is the test statistic for the effect of Taken, in terms of the mean squares (MS symbolically, not numerically)? (ii) Also report its sampling distribution. (iii) Finally, report what could happen (with the MS's) to make this test statistic negative (and therefore problematic).
 - (13 points) In the PROC MIXED output, the P-value for testing the effect of Taken is .1217, based on the potentially problematic test statistic in part (c) above. Using the EMS and MS values from the PROC MIXED output, construct a more stable test statistic using Cochran's approach, with no negative coefficients on MS's (recall p. 4 of Handout #15a). Report (i) the revised test statistic for the effect of Taken (in terms of the MS's symbolically), (ii) the revised sampling distribution [as a check, the numerator d.f. are approx. 1.04314], (iii) the P-value from this revised test, and (iv) your conclusion in context. [Note that part (iii) will require a SAS DATA step and a PROC PRINT statement.]
 - (4 points) What [omitted] interaction term does the error term correspond to in this model?

Appendix: (10 points) Include SAS code used for this assignment.