Exam 2

Stat 512 SP 2020

Honor Pledge: I have not given, received, or used any unauthorized assistance on this exam. By submitting this exam to be graded you agree to this statement.

Instructions:

- **Open book, open notes, calculator required.**
- If an answer is in the computer output, use it; don't calculate it by hand.
- Show your work where appropriate.
- Make explanations brief and legible.
- 24 questions (4 pts each) + score for legibility (0 – 4pts). Maximum score is 100.
- Computer input/output is provided in the posted handout.
- The exam contains a total of 7 pages.
- There is an additional **7 pages of R output**.

0. How clearly are your answers indicated? Graded on a scale of 0 – 4 pts. Consider color coded answers. Handwritten can receive full credit, but must be legible. This is not a timed exam, so please make my life easier. If handwritten, please see links posted for guidance on how to submit a scan or photos as 1 document.

Questions 1 through 9: In 1846 a party of settlers led by the Donner and Reed families left Illinois in covered wagons headed for California, traveling via a new and untested route. After many delays they were stranded in the Sierra Nevada Mountains near Lake Tahoe by an unusually early and intense snowstorm. A few members escaped or were rescued; however, by the spring many people had died. The episode is famous (or infamous) because it was discovered that some of the party members had survived by cannibalism. Grayson (1990) compiled **Age** (in years), **Sex** (male or female) and **Survival** (died or survived) records on 45 adult Donner party members and used that data to address the question of whether women are biologically better able to withstand harsh conditions than men. Two logistic regression models are considered. The R analysis is included at the end of the exam as "**Donner**". Use a=0.05.

1. A chi-square test is provided in the output. Provide the null hypothesis being tested. Give the test statistic and p-value.

H0: No association between Sex and Survival

Stat = 3.25

p-value = 0.071

2. Considering Models 1 and 2, what is considered an "event" for the purposes of these logistic regression analyses? In other words, are we modeling the probability of survival or probability of death? Just circle one answer; no need to justify.

DIED SURVIVED

3. **Consider Model 1,** What do we learn from the hoslem.test() result? Be specific about your conclusions here.

H0: Model fits
We fail to reject, and find no evidence of a lack of fit.

4. **Using Model 2**, provide an interpretation for the effect associated with Age. Be specific.

A one year increase in Age multiplies the odds of death by **1.081**, with sex held constant.

5. Using Model 2, the 95% CI for `exp(confint(Model2))` shows the interval associated with age to be (1.014, 1.176). Based on this information, can we conclude that there is a statistically significant effect of Age? Justify your response.

Yes, since the CI does NOT include 1.

6. The data were compiled to study the research hypothesis that women are better biologically adapted to harsh conditions than men. **Using Model2**, discuss whether these results support that hypothesis. Note: To get full credit you should reference (and interpret) a particular estimate(s) and test (or confidence interval).

Evidence options:

- Odds Ratio (M vs F) = 4.93; OR >1
- 95% CI (1.215, 25.25) does not include 1;
- P-value = 0.0345
- Est probability: F = 0.322 < M = 0.701

Females have lower probability of death.

7. Using the `prop.table()` output, the probability of death for Males is estimated to be 66.6%. Using **Model 2**, the probability of death for Males is estimated to be 70.1% (see the emmeans() output). Explain the difference between these two estimates. Note: I am looking for more than just that the fact that one estimate is from logistic regression.

Prop.table() gives simple proportion **ignoring Age**
Emmeans() estimate is adjusted for Age (predicted prop of both M and F is at average Age)

Questions 8 through 14 (Wood Trts): A study was done to compare wood hardness after 5 different beetle-repellant treatments (**Trt** = A, B, C, D or E). The research goal is to look for differences between treatments. A total of **8 logs** are used in the study. Each log is cut into five pieces and the treatments are randomly assigned to pieces such that each Trt is represented exactly once with each Log. The hardness of the wood (Y) is measured two-weeks after trt application. The R analysis is included at the end of the exam as "**Wood Trts**". Use $\alpha=0.05$.

8. Write a statistical model corresponding to Model1. You can use greek letters (or not), but indicate a which factor corresponds to each letter.

$Y_{ij} = \mu + \alpha_i + \beta_i + \varepsilon_{ij.}$

$Y_{ij} = \beta_0 + Trt_i + Log_i + epsilon_{ij}$ -2 missing block or extra term

i: 1,…,5
j = 1,…,8

9. Considering Model1, a colleague considers your analysis and says "Since differences between logs are not of research interest, you should drop log from the model." Justify why log should be kept in the model. I am looking for more than just a statement about significance!

Log is a blocking variable. Run as RCB, analyze as RCB. By accounting for log to log differences we can better detect differences between treatments.

10. Briefly explain the difference between Models 1 and 2. Be specific!

2pts: Model 1: Log is fixed
Model 2: Log is random;

2pts: Model 2 will generalize to the population represented by these logs. Model 1 will only apply to the 8 logs in the study.

11. Comparing Models 1 and 2, the test of Log is different. Specifically, for Model1, F= 12.569, p-value < 0.001. For Model2, Chi.sq = 24.2, p-value < 0.001. State the null hypothesis corresponding to the test of Log for each model.

H0 model 1: $\beta_1 = \beta_2 = ... = \beta_8$

H0 model 2: $\sigma^2_{Log} = 0$

–2: $\sigma^2_1 = \sigma^2_2 = \sigma^2_3$

12. If the goal of the analysis is the compare treatments, are the results affected by whether Model 1 or 2 is used? Justify your response.

No, both models give exactly the same results for comparisons of means.

13. If the goal of the analysis is to construct confidence intervals for the individual treatment means, are the results affected by whether Model 1 or 2 is used? Justify your response.

Yes. CI's for the individual treatment means are wider when Log(Block) is random (because we account for log to log variability)

14. Suppose the investigators are planning another experiment, and would like to compare 6 treatments with 12 logs (but still 5 pieces per log). Is a Balanced Incomplete Block design possible for this scenario? Justify your response.

BIBD possible: Yes No

Justification: Both conditions below are met.

t = 6; b = 12, k = 5, r = 60/6 = 10

Cond #1: bk = 12(5) = 60; tk = 6(10) = 60. Check!

Cond #2: 10(5-1)/(6-1) = 8. Interger, Check!

Questions 15 through 25 : An investigator is interested in comparing the **Yield** for 5 wheat **Varieties** (V1, V2, V3, V4, V5) and 3 **Fertilizers** (F1, F2, F3). A single field was divided into 45 uniform plots and the 15 treatment combinations (5 Varieties x 3 Fertilizers) were randomly assigned to plots with the only restriction that each treatment be assigned to three plots (n=3 plots per treatment combination). It is known that there are differences between the Varieties, so the primary interest is in comparing Fertilizer treatments.

The R analysis is included at the end of the exam as "**Wheat**". Use a=0.05.

15. Do the model assumptions (of equal variance and normally distributed residuals) appear to be satisfied? Discuss specific evidence from the output.

Equal Variance: Resids vs Fitted shows equal skatter

Normally Distributed Residuals: QQplot linear

16. In the Type 3 ANOVA table, look at the line labeled "Fertilizer" with F= 17.38 and p-value <0.0001. Explain in words what is being tested here. Note: I am looking for more than "main effect" of Fertilizer.

Testing for differences between Fertilizers averaging over varieties.

17. In the Type 3 ANOVA table, look at the line labeled "Variety*Fertilizer" with F= 1.50 and p-value = 0.1992. Explain in words what is being tested here. Note: I am looking for more than "interaction between Fertilizer and Variety".

Testing whether the effect of Fertilizer depends on the level of variety

18. Looking at the interaction plot and the Type 3 ANOVA table, a colleague expresses surprise that the Variety*Fertilizer interaction is not significant (p-value = 0.1992) given that the lines in the interaction plot are not parallel. Explain how this is possible.

Interaction plot does not give any information about variability/uncertainty, cannot be used for testing.

19. Considering the Type 3 ANOVA table, what follow-up pairwise comparisons would be of interest to compare the Fertilizer treatments? Give an appropriate emmeans statement that could be used. Justify briefly.

No evidence of interaction indicates we can focus on main effects, and average over the levels of the other factors:

Emmeans(Model, pairwise ~ Fertilizer)

20. Suppose (just for this question) that there had been a significant Variety*Fertilizer interaction. What follow-up pairwise comparisons would be of interest to compare the Fertilizer treatments. Give an appropriate emmeans statement that could be used. Justify briefly why this is either the same or different from your answer in 17.

Because there is evidence of a main effect, it could be misleading to average over the levels of the other factors:

Emmeans(Model, pairwise~ Fertilizer|Variety)

-1: Emmeans(Model, pairwise~ Fertilizer:Variety) (unnecessary comparisons leads to over correction with Tukey adjustment)

21. Suppose (just for this question) that there had been some missing data. For example, 3 (out of 15) treatment combinations had 2 observations (instead of 3) for a total of 3 missing observations for the experiment. What modifications to the R code would be required?

No change required, already using Type 3 test and emmean().

22. Would it have been acceptable to run the analysis as a one-way ANOVA with t=15 treatments? What would be a disadvantage of this approach?

Yes. One-way analysis is equivalent. However, (one of the following okay)

- it does not make use of factorial structure for ease of estimating main effects
- Difficult to test interaction
- Reduces power

23. Suppose (just for this question), that the experiment had been run in an RCB design. Specifically, 3 Fields (blocks) were used. Within each field, each of the 15 treatment combinations were randomly assigned to a single plot. Provide the one line of R code for fitting an appropriate model.

* lm(Yeild ~ Field + Variety*Fertilizer) ("Field" or "Block" acceptable)

* lmer(Yield ~Variety*Ferilizer + (1|Block))

24. Since this study has 2 factors, a split-plot design could have been used. Discuss one reason why the investigator might have chosen a split-plot design.

* Convenience or logistical constraints: Whole plot factor applied to larger areas, application to small areas is difficult

* Better accuracy for subplot factor