Stat 850: final exam

2019-05-08

Name:

- The exam is open book: you may use textbooks, notebooks, and a calculator.
- Laptops, tablets and phones must be turned off and put away.
- To get full credit, you must show your work. Partial credit will be awarded. Simply writing down the formulas for various quantities will not get you any partial credit.
- Some **partial** computer output has been provided on some questions. You may find some but **not necessarily all** of these computations useful. You may assume that these computations are correct.
- Do not dwell too long on any one question. Answer as many questions as you can.
- The parts within a problem are not necessarily sequential.
- For each hypothesis test that you perform, **state the corresponding p-value**, or give a range for the p-value, and interpret the p-value in words. Do not say "statistically significant" please.

---

1. In a certain species of damselflies, males are red. Most females are green, although some take on a red color. Researchers wanted to test various hypotheses that might explain color variation. They collected 5 females from each of 7 lakes, located in the same mountain area but at different elevations. Lake elevation was measured in meters (m). For each damselfly, they measured its wingspan (in mm), redness (from 0 = full green to 100 = pure red) and brightness (from 0 = most dark to 100 = most bright).

```
damselfly[c(1,2, 34,35),]
```

```
   lake elevation wingspan redness brightness
1    L1       510     35.8     0.6       22.2
2    L1       510     39.9    49.9       82.0
34   L7      1020     33.1    73.9       45.8
35   L7      1020     34.0    95.4       74.0
```

```
summary(lm(redness ~ brightness, damselfly))
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    8.167     12.649    0.65   0.5229
brightness     0.776      0.239    3.25   0.0027
```

```
summary(lm(redness ~ wingspan, damselfly))
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   190.03      67.25    2.83   0.0079
wingspan       -4.02       1.87   -2.14   0.0395
```

```
anova(lm(redness ~ elevation + lake, damselfly))
```

```
          Df Sum Sq Mean Sq F value  Pr(>F)
elevation  1  19736   19736   51.86 7.7e-08
lake       5   3478     696    1.83    0.14
Residuals 28  10655     381
```

```
fit = lmer(redness ~ brightness+wingspan+elevation+(1|lake), damselfly)
ranova(fit)
```

```
ANOVA-like table for random-effects: Single term deletions
Model: redness ~ brightness + wingspan + elevation + (1 | lake)

           npar logLik AIC   LRT Df Pr(>Chisq)
<none>        6   -147 306
(1 | lake)    5   -147 304 0.162  1       0.69
```

```
summary(fit)
```

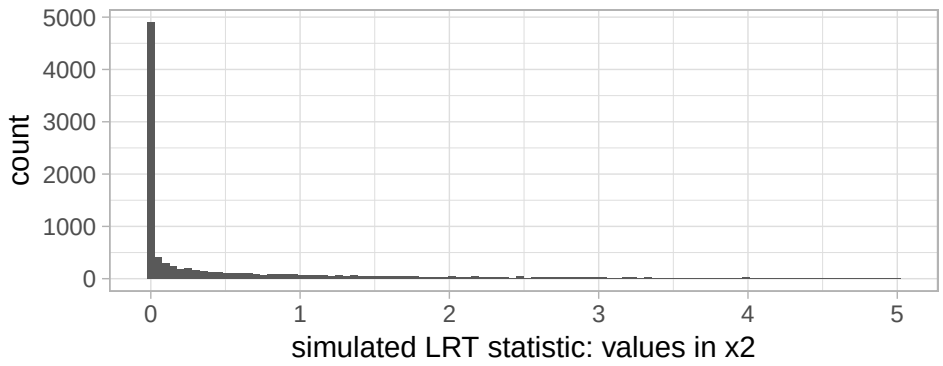
```
Random effects:
 Groups   Name        Variance Std.Dev.
 lake     (Intercept)  16.7     4.08
 Residual             259.0    16.09
Number of obs: 35, groups:  lake, 7

Fixed effects:
             Estimate Std. Error       df t value Pr(>|t|)
(Intercept) -133.9274    57.1636  30.7398   -2.34  0.02579
brightness     0.6036     0.1494  22.8529    4.04  0.00052
wingspan       1.6597     1.3497  30.9946    1.23  0.22809
elevation      0.1258     0.0198   9.4703    6.36  0.00011
```

```
ranova(fit)$LRT[2]
```

```
[1] 0.16229
```

```
set.seed(5432)
oneX2 = function(){
  y = simulate(fit)$sim_1
  m = lmer(y~brightness+elevation+wingspan+(1|lake), data=damselfly)
  ranova(m)$LRT[2]
}
x2 = replicate(10000, oneX2())
```




```
quantile(x2, probs=c(.05,.40,.45,.50,.55,.60,.95))
```

```
       5%       40%       45%       50%       55%       60%       95%
-3.41e-13 -1.71e-13  1.98e-03  3.49e-02  1.04e-01  2.22e-01  3.94e+00
```

(a) (5 points) Using the model named `fit` (defined on page 2), predict the expected redness of a new female damselfly D_1 of brightness 20, wingspan 35mm, sampled from a new lake at elevation 700m.

(b) (7 points) Consider another damselfly D_2 of same wingspan and brightness as D_1 in (a). Assuming model `fit` again, what is the estimated covariance between the redness of D_2 and D_1

(i) if D_1 and D_2 are sampled from different lakes?

(ii) if D_1 and D_2 are sampled from the same lake?

(c) (4 points) Consider model `fit` again. Explain why the approximate degree of freedom for `brightness` is quite higher than the approximate degree of freedom for `elevation`.

(d) (8 points) Consider the null hypothesis of no differences in average damselfly redness between lakes, beyond differences explained by lake elevation. Test this null hypothesis using the most appropriate method, given the available output. Be explicit about which method is used.

(e) (8 points) What is the estimated effect of `wingspan` on `redness`? Test the null hypothesis that this effect is zero.

2. Answer the following questions for each of the following experiments.

   (a) Identify the experimental / observational units, and, if appropriate, indicate the name of the design.
   (b) Give a skeleton table with the sources of variation and the degrees of freedom only.
   (c) Indicate if the data type suggests that a transformation should be investigated a priori, or if the analysis should be using a non-normal distribution.

   Experiment 1. (12 points) A researcher investigated the effect of the common pesticide carbaryl on different frog species. Also, the researcher wanted to determine whether carbaryl effect was exacerbated in the presence of predatory stress, a common environmental condition. The biologist sampled tadpoles from several ponds and wetlands, and from the following 3 species: Wood frog, Leopard frog, and American toad. Newly hatched tadpoles were placed in 10-liter tubs, containing one of 6 concentration of carbaryl: 0.0, 0.03, 0.3, 1.6, 3.2 or 6.5 mg/l. For half of the tubs, a red-spotted newt (a native predator) was also present in the tub. The newt was caged and could cause no direct harm, but it emitted visual and chemical cues that are known to affect tadpoles. For each of the 36 combinations (6 carbaryl concentrations $\times$ 2 predator treatments x 3 species), there were 4 replicate tubs, with a total of 144 tubs. The tubs were randomly assigned to treatments, although making sure that exactly 4 tubs were assigned to each treatment. In each tub, 10 tadpoles of the assigned species were randomly chosen to be placed in that tub at day 0. At the end of 16 days, the number of surviving tadpoles was recorded in each tub.

Experiment 2. (12 points) Smith et al. (2007) wanted to test evolutionary theories on why certain people do not trust leaders. In a lab experiment, participants played the "ultimatum game". 60 participants were enrolled, and randomly assigned a "leader" treatment: 30 randomly chosen participants were told that the leader earned power (legitimate leader). The other 30 participants were told that the leader craved power (selfish leader). Based on a questionnaire, 25 of participants were classified as generally distrustful, and 35 of them were classified as generally trusting. Each participant played the ultimatum game, individually. The response of interest was whether the participant accepted or rejected the offer from the leader. (This offer was in fact the same for all participants: \$3). There are two predictors of interest: the leader type (earned vs craved power) and the participant trust type (generally distrustful vs trusting). The theory under investigation had to do with their interaction effect.

3. A study investigated the effect of co-infection with hookworm and roundworm on the human immune system. Adult participants were recruited from a particular region of Brazil. Of those eligible, a total of 80 participants were randomly selected for immunological analysis. More specifically, 20 participants were sampled from each of 4 infection categories: uninfected, infected with hookworm only, infected with roundworm only, and infected with both. For each infection category, 10 participants were randomly sampled from each of 2 age groups (20-34 and 35-49 years old). From each participant, 2 blood samples were collected and analyzed to measure the quantity of 8 analytes (immunological responses by white blook cells, in pg/ml): IL2, IL4, IL5, IL10, IL13, CXCL10, TNFα, and IFNγ. The values were log-transformed before analysis, to better meet the model assumptions. Partial output is shown below.

```
worm[c(1,2, 159,160),1:4]
```

```
    subject   age infection logIL2
1         1 20-34      none  -0.25
2         1 20-34      none   0.00
159      80 35-49      both   2.53
160      80 35-49      both   2.64
```

```
worm_average = worm %>% group_by(subject) %>%
  summarize(age=age[1], infection=infection[1], logIL2 = mean(logIL2))
as.data.frame(worm_average[c(1, 80),])
```

```
  subject   age infection logIL2
1       1 20-34      none -0.125
2      80 35-49      both  2.585
```

```
model_average = lm(logIL2 ~ infection * age, data=worm_average)
```

```
summary(model_average) # partial output only
```

```
                            Estimate
(Intercept)                  0.202
infectionhookworm            2.361
infectionroundworm           2.364
infectionboth                2.473
age35-49                    -0.081
infectionhookworm:age35-49  -0.017
infectionroundworm:age35-49  0.068
infectionboth:age35-49      -0.074

Residual standard error: 0.1919
```

(a) (12 points) Provide an estimate and a 95% confidence interval, for the logIL2 average response among 20-34 year-old adults infected with both hookworms and roundworms.

(b) (5 points) Using output not shown here, the investigators found very strong evidence for a two-way interaction effect of hookworm infection and roundworm infection on the log IL4 response, with a p-value of 0.00015. In fact, they are going to test for a hookworm-by-roundworm interaction effect for all 8 responses (log of IL2, IL4, IL5, IL10, IL13, CXCL10, TNFα, and IFNγ). Revise their conclusion and report for the IL4 response, in light of this plan for multiple tests.

(c) (10 points) The investigators re-analyzed the IL2 data without averaging the 2 samples obtained from each subject, as shown below. Determine the σ_{subject} value erased from this partial output.

```
model_randomsubject = lmer(logIL2 ~ infection * age + (1|subject), data=worm)
VarCorr(model_randomsubject) # output edited
```

```
 Groups   Name        Std.Dev.
 subject  (Intercept)
 Residual             0.21440
```

(d) (5 points) Why would the researchers want to run the model in (c), after answering their primary question using the averaged data as used in (a)?

4. Measurements were made on plates containing bacterial colonies. There were 3 fungal treatments ($i = 1, 2, 3$). A total of 6 plates were randomized, such that 2 plates ($j = 1, 2$) were allocated to each fungal treatment. Each plate was measured at 3 time points ($k = 1, 2, 3$). The researchers used a mixed linear model of the following standard form to analyze these data, where time was considered numerical ($t_1 = 0$, $t_2 = 1$, $t_3 = 2$): $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$.

```
model_bacteria = lmer(y ~ treatment*time + (1|treatment:plate), bacteria)
summary(model_bacteria) # partial output
```

```
Number of obs: 18, groups:  treatment:plate, 6

Fixed effects:
                 Estimate
(Intercept)        45.656
treatmentF2       -10.932
treatmentF3        -2.936
time                7.097
treatmentF2:time    4.810
treatmentF3:time   -1.597
```

```
head(bacteria)
```

```
  treatment plate time     y
1        F1     1    0 46.00
2        F1     1    1 57.94
3        F1     1    2 61.33
4        F2     1    0 30.49
5        F2     1    1 56.79
6        F2     1    2 60.33
```

(a) (6 points) Determine the first 6 rows of the design matrix $\mathbf{X}$.

(b) (6 points) Determine the first 6 rows of the design matrix $\mathbf{Z}$.

For instructor's use:

question	total points
1	32
2	24
3	32
4	12