Stat 850: midterm exam

2019-04-01

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 computations have 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. (12 points) Researchers have planned an experiment with 6 treatments arranged in a randomized complete block design, with 5 blocks and with 2 subsamples per experimental unit (so 30 experimental units and 60 measurements total). They anticipate a standard deviation of 11 between experimental units and a standard deviation of 8 between subsamples.
   Will their power to detect treatment differences increase if they collect data from 4 blocks and 3 subsamples per experimental unit instead (i.e. with 72 measurements total instead of 60)?

2. Consider a **balanced**, completely randomized experiment, in which birds are given a gabazine infusion at *one* of 3 doses (0.00, 0.01 or 0.02 mM), and the spiking precision is measured on several neurons per bird. Shown below is an ANOVA table where dose was considered as a factor (left). Also shown is the average values across neurons and birds, at each dose level (right).

```
Analysis of Variance Table
Response: neuron_precision
            Df  Sum Sq  Mean Sq F value   Pr(>F)
dose_factor  2 0.51768 0.258839 50.3021 2.23e-07
bird        12 0.23155 0.019296  3.7499 0.009021
Residuals   15 0.07719 0.005146
```

```
dose mean_neuron_precision
0.00              0.5019421
0.01              0.7305155
0.02              0.8123599
```

(a) (8 pts) How many birds were used *per dose* in this experiment, and how many neurons were used per bird?

(b) (12 pts) Test the null hypothesis of no dose effect.

(c) (15 pts) Test the null hypothesis of a linear dose effect.

For instructor's use:

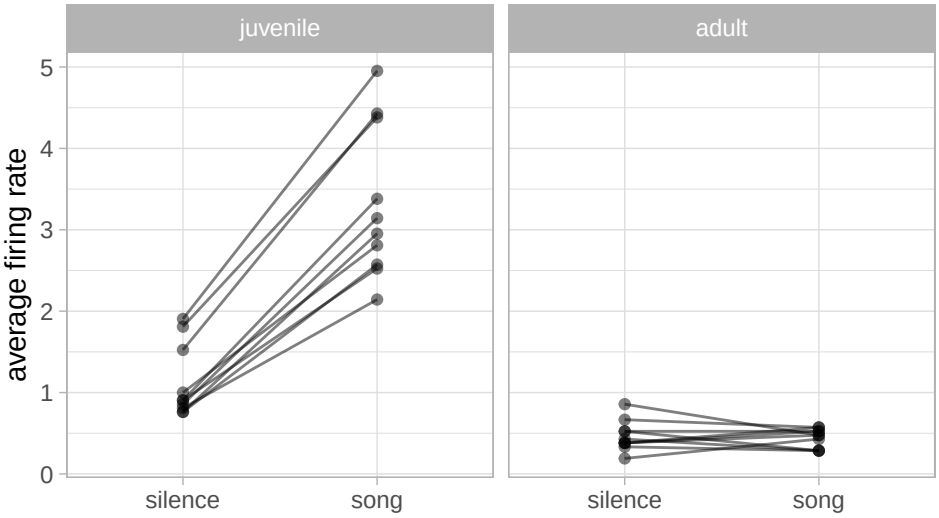
question	total points
1	12
2	35
3 a-b	16
3 c-e	37

3. Vallentin et al. (2016, Science) studied the neurological mechanism by which male zebra finches (birds) learn to sing when they are young. In one experiment, the researchers used one recording of a tutor's song, which was played 7 times to each of 10 juvenile birds and 10 adult birds. The recording contained 1000 ms of silence followed by 1000 ms of a recorded tutor song. In each bird, 3 neurons from a particular region of the brain were measured. The same 3 neurons were used for all 7 trials (i.e. each time that the recording was played to that bird). The neuron's firing rate was measured during the first silent half of the recording, and also during the second "song" half of the recording (for each neuron of each bird). The researchers wanted to know if the effect of the song on the neuron firing rate differs between juveniles (who are supposed to learn the song) and the adults (who are supposed to know how to sing already).
   (a) (8 pts) Focus on one single neuron from one single bird. Using the 14 firing rates from this neuron, what test would you recommend to detect a song effect in that particular neuron (assuming assumptions hold)? Be specific.

   (b) (8 pts) Imagine averaging the firing rate across all 3 neurons and 7 trials for each bird, separately for the two halves of the recording (during silence and during the song), for a total of 40 firing rate values (2 values for each of 20 birds). After this averaging, give the name of the experimental design, and explain concisely.

(c) (12 pts) The data averaged across trials and neurons, as explained in (b), are displayed below. Write an R command to proceed with an appropriate model fit for these averaged data.

```
ggplot(dat_ave, aes(y=firing_rate_average, x=song, group=bird)) + facet_grid(~age) + ...
```



(d) (15 pts) Now consider the full data set, and without any averaging. Build an adequate model to predict firing rate from these data. To describe your model, make a table with 3 columns, listing: the terms to be included in the model, their degrees of freedom, and whether each term should be considered as a fixed effect or as a random effect.

(e) (10 pts) Describe the appropriate F test from your model in (d), including the numerator & denominator terms and the degrees of freedom, to determine if the song effect (song versus silence) differs between juvenile and adult birds.