Assignment #4

Stat 332: Sampling and Experimental Design (Spring 2024)

Professor: Christian Boudreau   Due at 11:59 pm on Sunday Jul 28, 2024

Instructions:

- Teamwork is allowed and encouraged; however, everyone must hand in their own assignment showing that they understood what they wrote. You are not allowed to collaborate with anyone who is not a Stat 332 classmate.
- Justify all your answers.
- One solution per question; if you provide multiple solutions, you will receive 1/2 the points of the inferior solution.
- Write legibly; if the TA or I have to guess at your answers, it's unlikely to be in your favour.
- Make sure to include your `R` code and relevant output for Q#3 and Q#4.
- Crowdmark:
  - You should soon received an automated email from Crowdmark saying that you can now submit/upload your solution. If you have not received the email by noon on Wednesday, please start by checking your spam or junk mail folder. If the email is not there, email me. Note that, even if you did not get the email, you should still be able to submit/upload your assignment by simply going to Crowdmark, and signing into your account.
  - To ensure smooth and efficient grading, please make sure to upload your solution separately for each question.
  - Crowdmark allows you to resubmit anytime before the due date; **once the due date has passed, your submission is locked in and you will thus not be able to resubmit**.
  - Late submissions will be accepted, but are subject to a 5% per hour penalty*; click here for more information on Crowdmark and late submissions.
  - Technical difficulties with Crowdmark:
    - Consult Crowdmark Help
    - Watch this short video about submitting an assignment on Crowdmark
    - E-mail me
    - If tragedy strikes and you cannot upload your assignment to Crowdmark, email it to me, so I have proof of when you completed your assignment (otherwise, the above mentioned lateness penalty will apply).
- Recall that each of you is allowed a one-time extension with a reduced lateness penalty. You have 3 option: 1) a 5% reduced penalty for an additional 24 hours, 2) a 10% reduced penalty for an additional 48 hours, and 3) a 0% reduced lateness penalty if you submit your assignment within 3 hours past the deadline. To take advantage of the first two options (i.e., 5% and 10%), you have to email me (at least) 12 hours before the original due date. Please note that if you do not submit your assignment by the updated due date, you will incur both the reduced lateness penalty (i.e., 5% or 10%) plus the standard penalty of 5% per hour.

  Since this is meant to be a last minute thing, you do not need to email me in advance to take advantage of the 0% (3 hours) lateness penalty. However, since it counts towards your one-time reduced lateness penalty, you still need to email me me to inform me that you want to take advantage of this option.
- See course outline for more information on assignments, and how much each of them contributes to your overall grade.

*This is the default penalty, and is applied when there are no extenuating circumstances.

Question #1 (10 points)**:**

(a) In a two-way ANOVA, let S_{ij}^2 be the sample variance of the n_0 observations of the (i, j) treatment (i.e.; i^{th} level of factor A and j^{th} level of factor B); in other words,

$$S_{ij}^2 = \frac{1}{n_0 - 1} \sum_{k=1}^{n_0} (y_{ijk} - \bar{y}_{ij\bullet})^2 \ .$$

Prove that S_{ij}^2 is an unbiased estimator of σ^2.

(b) When the F test for interactions is not statistically significant, an alternate estimator of σ^2 is obtained by pooling SSE and SSAB. Let SSP = SSE + SSAB, and

$$\text{MSP} = \frac{\text{SSP}}{n - I - J + 1} \ .$$

be the corresponding mean squares. Show that MSP is an unbiased estimator of σ^2 when there are no interactions.

Question #2 (10 points)**:**

Suppose that you have a sample of n paired observations. For each pair, one unit is randomly assigned to treatment T1 and the other to treatment T2. One option to test if treatments T1 and T2 are the same is to use the paired t-test (see description below). However, this can also be viewed as a randomized block design where the treatment/factor A has two levels (i.e., $I = 2$). Hence, another option to test if treatments T1 and T2 are the same is to use the F test of section ~~8~~ 9.5.1 of the slides. Show that those two tests are equivalent. In other words, show that, when treatment/factor A has only two levels, $F^* = \text{MSA}/\text{MSE}_{\text{bl}}$ is the same as $(t^*)^2$ (i.e., the square of t^*).

Paired t-test

Let $(y_{11}, y_{21}), \ldots, (y_{1n}, y_{2n})$ be the n paired observations. For the j^{th} pair $(j = 1, \ldots n)$, y_{1j} is the observation for treatment T1, and y_{2j} is the observation for treatment T2. Also, let $d_j = y_{1j} - y_{2j}$ be the difference between treatments T1 and T2 for the j^{th} pair. The test statistics for the testing $H_0\colon \mu_1 = \mu_2$ (i.e., T1 and T2 are the same) is given by

$$t^* = \frac{\bar{d}}{S_d/\sqrt{n}}$$

where

$$\bar{d} = \frac{1}{n} \sum_{j=1}^{n} d_j = \frac{1}{n} \sum_{j=1}^{n} (y_{1j} - y_{2j}) = \bar{y}_{1\bullet} - \bar{y}_{2\bullet}$$

and

$$S_d^2 = \frac{1}{n-1} \sum_{j=1}^{n} (d_j - \bar{d})^2$$

It can be shown that t^* follows an t-distribution with $n - 1$ df under H_0. Hence, the paired t-test consists in rejecting H_0 (in favour of H_1) at the $(1-\alpha)$ level if $t^*_{\text{obs}} > t_{n-1;1-\alpha}$, where $t_{n-1;1-\alpha}$ is the $(1 - \alpha)$ quantile of a t-distribution with $n - 1$ df.

Question #3 (15 points)**:**

I made a small experiment to see how long it takes to boil (in minutes) a given amount of water on the four different burners of my stove (1=right–back burner, 2=right–front, 3=left–back and 4=left–front). Since salt affects the boiling point of water, I ran the experiment with 0, 2, 4 and 6 teaspoons of salt added to the water. Hence, my experiment has two factors (i.e., burner and salt, with 4 levels each). In addition, I repeated the experiment 3 times for each of the 16 burner/salt combinations; yielding a total of 48 observations. The corresponding dataset is available on the STAT 332 LEARN website; see file `water.txt`.

(a) Produce the ANOVA table.

(b) Perform the usual F tests in a two-way ANOVA. What are your conclusions on the effects of salt, burner and their interaction?

(c) Using pairwise comparisons, which burner or burners is/are the fastest at boiling water? (Make sure that you clearly justify why only this/those burner/s is/are faster than the others.)

(d) Use a contrast to test (at the 5% level) if the two right burners take the same time to boil water than the left ones.

(e) Use a contrast to test (at the 5% level) the hypothesis that there is a linear trend in the time to boil water due to the level of salt.

Question #4 (20 points)**:**

I took the follow-up course to Part 2 (Experimental Design) of STAT 332 a long, long, long, . . . , long time ago here at the University of Waterloo. We had to do a project in that course, and my classmates were all doing fancy experiments (e.g., growing things or building things). My partner and I decided that our project would consists in investigating how far paper frogs can jump (you can watch this short video if you're unsure of what I mean). Though our project might have initially looked a little silly, it consisted of a three-way Anova mode with blocking to investigate the effects of paper thickness, paper dimensions and paper colour, as well as their interactions, on the jumping distance of paper frogs.

Rest assured that this question does not involve fitting three-way Anova models, as this is beyond the scope of STAT 332. However, it uses the same dataset, which can be dowloaded from the STAT 332 LEARN website (see file `frogs.csv`). This file can be easily read into `R` by using the command `read.csv`, and it contains the following variables:

Variable	**Description**
`thickness`	Paper thickness (photocopy vs resume vs cardstock)
`dimensions`	Paper dimensions (10cm $\times$ 10cm vs 15cm $\times$ 15cm vs 20cm $\times$ 20cm)
`colour`	Paper colour (white vs blue)
`jump`	How far the frog jumped (in cm)
`person`	Who launched the paper frog (Christian/me vs Michelle/my partner)

(a) Fit a two-way Anova model where the two factors are paper thickness (variable `thickness`) and colour (variable `colour`), and answer the following: *Note: this model simply ignores paper dimensions and the person who launched the paper frog.*

- a.i) Produce the two-way Anova table, and perform the usual F tests. What are your conclusions on the effects of paper thickness, colour and their interaction?
- a.ii) What paper thickness (i.e., cardstock, photocopy or resume) jumps the furthest?
- a.iii) Obtain the contrast (and corresponding 95% CI) to test he hypothesis that there is a linear trend in paper thickness. Photocopy paper is the lightest (at 75 g/m^2), resume is in the middle (at 97 g/m^2) and cardstock is the heaviest (at 181 g/m^2).

  Note: no need for multiple comparisons adjustment, and no need to do a 5% test.
- a.iv) Check if the assumptions of the ANOVA models are satisfied. Check if there are outliers?

(b) Fit another two-way Anova model where the two factors are paper thickness (variable `thickness`) and dimensions (variable `dimensions`), and answer the following:

- b.i) Produce the two-way Anova table, and perform the usual F tests. What are your conclusions on the effects of paper thickness, dimensions and their interaction?
- b.ii) Compute all pairwise comparisons between the 9 groups (i.e., thickness/dimensions of experience combinations) and, using Tukey's method, obtain their CI's with a family level of 95%.
- b.iii) Using part b.ii), determine which one (or ones) of the 9 groups jumps the furthest?

(c) Expand on the two-way Anova model fitted in part (b) by including `person` as a blocking factor. In other words, fit a two-way Anova model with `thickness` and `dimensions` as factors (don't forget to include their interaction), and `person` as a blocking factor.

c.i) Produce the two-way Anova table, and perform the usual F tests. What are your conclusions on the effects of paper thickness, dimensions and their interaction? Are they the same as in b.i)?

c.ii) In section 9.6 ~~8.6~~ of the slides, I mentioned that estimating and comparing the blocks is usually not of interest. Though this is true, for this particular problem, I nevertheless want to know if I was better than my partner Michelle at making the paper frogs jump further. Compute a 95% CI to test this?