WINTER 2019/20 TERM 2 EPSE 592: ASSIGNMENT 1
Due: Week 4 in class: Jan. 30th

- Please make sure you write your answers to these questions in your own words. Even if you work with a group to formulate your responses, do not just copy someone else's sentences/words.
- There is no need to record more than 3 decimal places for any of these problems.

Question 1: **(Randomization, control, and confounders)** Suppose you are part of a team that has designed a study to assess the potential effects of social media exposure on anxiety levels in people aged 18-34. Your team will submit repeated calls for participants across the UBC campus until you have enlisted 9 people who report high levels of social media use, 9 with medium levels of use, and 9 with low levels of use. (Suppose the definitions of "high/medium/low levels of use" are well-defined and uncontroversial.) Each subject will be compensated financially for agreeing to participate in the study.

Once these 27 participants are secured, they will each complete a 10-minute questionnaire to assess their baseline level of anxiety. Afterwards, they will receive a random assignment to one of three treatment groups: low, medium, or high social media use. Each subject will then have to adjust their social media use accordingly for the next 24 hours, afterwhich time they will return to your lab and complete another 10-minute questionnaire assessing their post-treatment anxiety levels.

(a) Does this study design use random sampling? Is the treatment of interest (social media use) randomly assigned to experimental units? Could you foresee any potential problems with implementing the proposed randomizations?

(b) The study design recognizes that baseline social media usage could be a significant confounding variable; i.e. if higher social media usage is correlated with higher anxiety levels, then individuals who regularly spend lots of time on social media will likely already have higher levels of anxiety when entering the study. Can you think of any other possible confounding variables that might affect our results given how our team has decided to sample experimental units or assign experimental treatments?

(c) Does this study have a control group? If so, who is in the control group? If not, why not?

Question 2: **(P-values, priors, posterior believability, and the truth)** In this sequence of problems, you will get some practice thinking about the different kinds of conditional probabilities that scientists must consider in their analyses. You will also explore the relationship between p-values (the most common quantitative tool for inference) and evidence for the likelihood of a hypothesis of interest (what we actually care about when we do science). While this discussion could get very complicated very quickly, we will explore the issues with a simple, though entirely representative, example.

Suppose I claim to have extrasensory perception (ESP); i.e. the supernatural ability to see things without actually witnessing them for myself. I challenge you to test my claim and so you devise a simple experiment: you will randomly draw one playing card from a well-shuffled standard deck of 52 cards, making sure I cannot possibly see the card, and then ask me what card you drew. This experiment has two possible outcomes: (1) I guess your card correctly, or (2) I guess incorrectly. The null hypothesis (i.e. the hypothesis of no effect) is that I do *not* have ESP, and thus, my guesses are random.

(a) Let Y denote the event that I guess your card correctly, and let N denote the event that I guess incorrectly. Assuming the null hypothesis, what is the probability that I guess your card correctly: $\Pr(Y \mid H_0)$? Given the null, what is the probability that I guess your card incorrectly: $\Pr(N \mid H_0)$?

(b) Suppose we run the experiment once and I guess your card correctly. The p-value for this observed outcome is what you calculated above as $\Pr(Y \mid H_0)$; i.e. the probability of observing an experimental outcome as or more extreme than the one we actually did observe, given the null hypothesis. Since this experiment has only two possible outcomes, and we observed the "more extreme" one, the p-value is simply $\Pr(Y \mid H_0)$. Using the typical rule of thumb that *a p-value of less than 0.05 is "significant" (i.e. decent evidence of an effect)*, is the p-value for this particular experimental outcome significant? If so, would this experimental outcome convince you that I probably have ESP?

(c) Recall that $\Pr(H_0 \mid Y) \neq \Pr(Y \mid H_0)$. The latter probability is our p-value, while the former probability is what we actually need to evaluate if we want to make an informed *decision* about the believability of H_0 given the observed evidence; this former probability is called a *posterior probability*. [Go back and review the Sally Clark case: $\Pr(Innocence \mid Evidence)$ is exactly a posterior probability.] There is a famous mathematical result (Bayes' Theorem) that allows us to relate our p-value to the posterior probability of interest:

$$\Pr(H_0 \mid Y) = \frac{\Pr(Y \mid H_0)\Pr(H_0)}{\Pr(Y \mid H_0)\Pr(H_0) + \Pr(Y \mid H_A)\Pr(H_A)},$$

where H_A denotes the *alternative hypothesis* to the null. For our example, H_A would be the hypothesis that I *do* have ESP. The quantities $\Pr(H_0)$ and $\Pr(H_A)$ are called *priors*; they represent how likely each hypothesis is *before* we conduct our experiment to test them.

Given everything that we know about general physics, chemistry, and biology - and given all the failed attempts at proving the existence of ESP in other people in the past - it is quite reasonable to assign a very large value to $\Pr(H_0)$. Let's say, conservatively, that $\Pr(H_0) = 0.999$. What then must be the value of $\Pr(H_A)$? Moreover, what is $\Pr(Y \mid H_A)$; i.e. what is the probability that I identify your card correctly given that I actually do have ESP?

(d) Now we can put all the pieces together to calculate the posterior probability we are really interested in. Use Bayes' Theorem (equation above) and your results from parts (a) and (c) to calculate $\Pr(H_0 \mid Y)$. Is the fact that I guessed your card correctly during the experiment convincing evidence that I have ESP?

Question 3: (The effect of experimental replication and the importance of critically assessing your quantitative inferences) Suppose that I challenge you to test my ESP claim again. We repeat the same card experiment as before, and this time I again guess the correct card. Note that H_0, H_A, and the p-value $\Pr(Y_2 \mid H_0)$ are all the same as in the first experimental run. Here, we use the subscript to denote which experiment we are talking about: the first or the second. The probability of the event Y_i under the null is unchanged by the value of i, however.

The first experiment resulted in me correctly identifying your card: the event Y_1. We would like to combine that knowledge with the information we have gained from the second experiment to make a more informed assessment about the believability of H_0.

(a) Before we ran our first experiment, we reasoned that the *prior* believability of our null hypothesis was about $\Pr(H_0) = 0.999$. In part (d) of the previous problem, you "updated" this probability given the outcome of the first experiment: $\Pr(H_0 \mid Y_1)$. So if we now want to update our belief in H_0 given the outcome of the *second* experiment, we can take $\Pr(H_0 \mid Y_1)$

as our new *prior*, since it encodes all the information about the first experiment *and* our prior assessment of the hypothesis before we ran any experiments. Applying Bayes' Theorem again, we can compute the posterior probability as

$$\Pr(H_0 \mid Y_2) = \frac{\Pr(Y_2 \mid H_0)\Pr(H_0 \mid Y_1)}{\Pr(Y_2 \mid H_0)\Pr(H_0 \mid Y_1) + \Pr(Y_2 \mid H_A)\Pr(H_A \mid Y_1)}.$$

Write down the prior believability of the null hypothesis, $\Pr(H_0 \mid Y_1)$, and use this value to find $\Pr(H_A \mid Y_1)$. Write down the p-value of the outcome of the second experiment, $\Pr(Y_2 \mid H_0)$. Finally, write down $\Pr(Y_2 \mid H_A)$.

(b) Putting the pieces together again, calculate $\Pr(H_0 \mid Y_2)$. How does this value compare to $\Pr(H_0 \mid Y_1)$ that you found in part (d) of the previous question? Is the fact that I guessed your card correctly in *both* experiments convincing evidence that I have ESP?

(c) You may still be skeptical, so suppose I challenge you one last time to test my ESP claim. We repeat the same experiment a third time, and yet again I guess the correct card. Repeat the updating procedure we performed above to calculate the new posterior probability $\Pr(H_0 \mid Y_3)$.

(d) How convinced are you that I have ESP given the outcome of the three experiments and the quantitative evidence in part (c)? *Think critically, not like you are following a recipe* (quantitative inference is *not* a recipe to be followed). If you are still skeptical about my claim of ESP, are there any other more reasonable explanations for my ability to guess your card correctly three times in a row?

Question 4: (T-tests: heteroskedasticity means loss of power) In this exercise, you will perform some t-tests and investigate how inequality of variances across test groups can affect your results.

Suppose you are trying to compare average test scores between two sections of the same class. The data are presented in the following table (tests were scored out of 50 points).

Section I	22	23	19	24	20	20	26	19	21
Section II	12	22	29	42	47	33	24	19	38

(a) Calculate the sample mean and sample standard deviation for the Section I scores. Do the same for the Section II scores. Informally, does there appear to be a difference in the sample means? Does there appear to be a difference in the sample standard deviations?

(b) Apply a formal independent samples t-test to test the hypothesis that the mean scores of the two sections are equal. Report your p-value and interpret.

(c) Apply a formal test to check the equality of variances assumption for your above t-test. In Jamovi, you can produce a simple F-test (called Levene's test here) by clicking in the appropriate box in the "Assumptions Check" options. Report the p-value of this test and interpret.

(d) Now suppose you have data on a third section of the same class, given by the data below.

Section III	25	26	29	34	32	33	27	29	31

Calculate the sample mean and sample standard deviation for the test scores from this section. How do these compare to the same statistics for Sections I and II?

(e) Apply an independent samples t-test to test if the mean scores from sections I and III are equal. Report the p-value and interpret. Again, check the equality of variances assumption formally.

(f) What kind of effect(s) did hetero/homoskedasticity have on your analysis and interpretations?