Introduction to Statistics in Psychology: PSY 201

Greg Francis, PhD
Department of Psychological Sciences

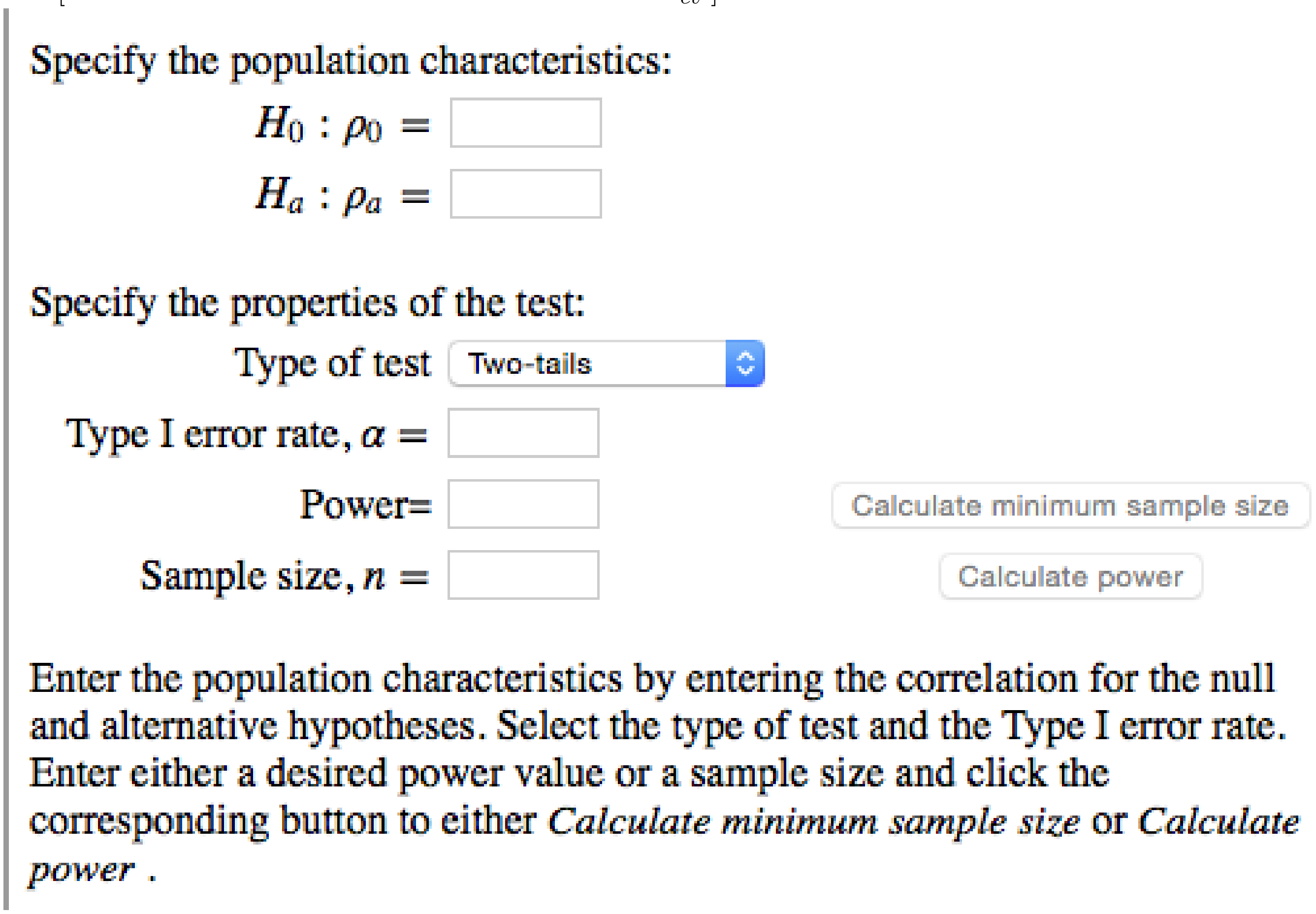
EXAM 3

Name ________________________________

Total points on the exam is 100. The exam will count as 10% of your class grade. Write your answers on the exam. Including all of your intermediate work will give you the best chance of getting partial credit (if necessary). Use the back of the page if necessary. The problem with the lowest score will be dropped.

(1) Athletes sometimes report that their percepts change in a way related to their performance. Witt and Proffitt (2005) asked $n = 47$ softball players to estimate the size of a softball just after finishing a game. They computed a correlation between the size estimate and the batting average of each player in the game, and found $r = 0.29$ ($z_r = 0.299$). Test $H_0 : \rho = 0$ against $H_a : \rho \neq 0$ with $\alpha = 0.05$. Is this correlation statistically significant? Is there a relationship between batting average and size estimate? [Hint: You cannot compute the p value, but a test statistic bigger than $z_{cv} = 1.960395$ will correspond to $p < 0.05$.]

(2) You want to repeat the Witt and Proffitt (2005) study described in question (1), but you want to find the sample size needed to have 90% power for your study. You will use the correlation found by Witt and Proffitt (2005) as your estimate of the population correlation. Fill in the details of the power calculator shown below as you would to compute the minimum sample size. Will you need a larger or smaller sample size than what was used by Witt and Proffitt (2005)? Explain your answer. [Hint: Consider how close the test statistic is to z_{cv}.]



(3) Are people more motivated to study when giving or receiving advice? In a study by Eskreis-Winkler, Fishbach & Duckworth (2018) middle-school students either received advice or gave advice about how to study. The table below summarizes how much time (minutes) the subjects then spent on line to study vocabulary words. Use Welch's test to test the null hypothesis for a difference of means. Use $\alpha = 0.05$. Does there seem to be a difference in time spent studying between the two groups? [Hint: $df = 265.7$. Although you cannot compute the p-value, the size of the test statistic should indicate what kind of decision to make.]

$$H_0 : \mu_1 - \mu_2 = 0 \qquad H_a : \mu_1 - \mu_2 \neq 0$$

	Gave advice	Received advice
n	154	164
$\overline{X}$	26.58	23.27
s	12.33	8.30

(4) Humans tend to quickly engage in prosocial behavior (e.g., open a door for someone), and they are slower to act selfishly (e.g., eat the last cookie on a tray). In a study of the prosocial behavior of chimpanzees, Rosati, DiNicola & Buckholtz (2018) looked for a similar difference in the time to act prosocially or selfishly. Across several different situations, they found that the $n = 40$ chimpanzees in their sample took $\overline{X}_1 = 0.36$ seconds to make a selfish choice and $\overline{X}_2 = 0.30$ seconds to make a prosocial choice. The standard deviation of the difference in choice times is $s_d = 0.181$. Test for a difference of dependent means. Is there a significant difference? Are chimpanzees similar to humans in regard to how they make these choices? [Hint: to get $p < 0.05$, you need a t value more extreme than ± 2.0227.]

$$H_0 : \mu_1 - \mu_2 = 0 \qquad H_a : \mu_1 - \mu_2 \neq 0$$

(5) In a study of Autism Spectrum Disorder (ASD), Oruc, Shafai & Iarocci (2018) asked people with and without ASD to identify the emotion expressed in a face. They created a range of "morphed" faces that expressed an emotion to varying degrees. Across trials, they identified a *threshold* morph-level that estimates the smallest change from neutral that can be identified as indicating the emotion. Bigger thresholds indicate that the subject needs a bigger morph from neutral before they can detect the emotion. The data was entered into the on-line calculator for a two-sample means test to compare thresholds for persons with ASD and Controls (without ASD). The image below shows the results; with four parts blanked out. Fill in the missing parts. Why does the calculator use the *Standard* test rather than Welch's test?

Test summary	
Type of test	Standard
Null hypothesis	$H_0 : \mu_1 - \mu_2 = 0$
Alternative hypothesis	$H_a : \mu_1 - \mu_2 \neq 0$
Type I error rate	$\alpha = 0.05$
Label for group 1	ASD
Sample size 1	$n_1 = 34$
Sample mean 1	$\overline{X}_1 = 8.0000$
Sample standard deviation 1	$s_1 = 8.160000$
Label for group 2	Control
Sample size 2	$n_2 = 34$
Sample mean 2	$\overline{X}_2 = 4.0000$
Sample standard deviation 2	$s_2 = 8.160000$
Pooled standard deviation	$s =$
Sample standard error	$s_{\overline{X}_1 - \overline{X}_2} =$
Test statistic	$t = 2.021130$
Degrees of freedom	$df = 66$
p value	$p = 0.047326$
Decision	
Confidence interval critical value	$t_{cv} = 1.996564$
Confidence interval	$CI_{95} =$

(6) Identify the formula for the test statistic and the shape of its sampling distribution for each of the following hypothesis testing situations. If the sampling distribution is a t distribution, also give the formula for its degrees of freedom.

a) One sample mean, $H_0 : \mu = a$.

b) One sample proportion, $H_0 : P = a$.

c) One-sample correlation, $H_0 : \rho = a$.

d) Two-sample, dependent means, $H_0 : \mu_1 - \mu_2 = 0$.

e) Two-sample, correlations, $H_0 : \rho_1 - \rho_2 = 0$.

(7) From a random sample, a scientist computes a 99% confidence interval for a proportion and gets

$$CI_{99} = [0.11, 0.34]$$

The scientist's colleague wants to run the following hypothesis test

$$H_0 : P = 0.5$$

with $\alpha = 0.05$. Without actually doing the test, you can figure out the decision of the hypothesis test. Describe the decision and explain how you know.

(8) Explain why the Fisher z transform is needed to run a hypothesis test and to build a confidence interval that involves a correlation.

(9) Kouchaki and colleagues (2018) were interested in how morality is related to a sense of humanity. Subjects were asked to write a brief essay about themselves. For the experimental group, the essay prompt was to write about a situation where they did not feel like a full human being (e.g., unable to have self-control, to make plans, or to feel emotions). In the neutral group, the essay prompt was to write about their morning routine. After writing the essay, a subject was asked to work on an anagram task (unscramble letters to form words), and they received 25 cents for each of the four anagrams that they self-reported solving. The anagram task was actually a measure of honesty. The fourth anagram was unsolvable (the letters could not form a word), but some subjects reported that they solved the anagram. The scientists compared the proportion of subjects that lied in the experimental and neutral groups to see how the essay topic influenced honesty. The table below summarizes the findings.

	Experimental group	Neutral group
n	73	77
f falsely reported solving the anagram	34	22

Test the hypotheses:

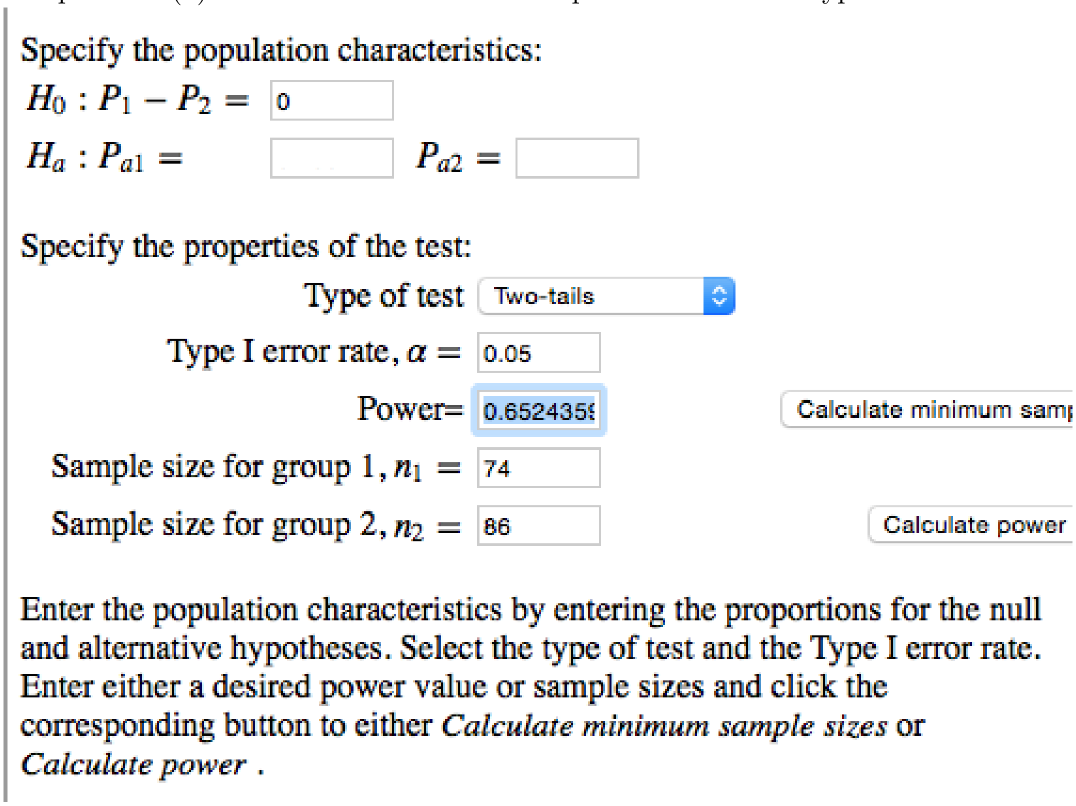
$$H_0 : P_1 - P_2 = 0 \qquad H_a : P_1 - P_2 \neq 0$$

with $\alpha = 0.05$. Does the essay prompt topic seem to influence honesty? [Hint: You cannot compute the p-value, but a test statistic more extreme than $z_{cv} = \pm 1.96$ would produce $p < 0.05$.]

(10) In a follow-up experiment, Kouchaki and colleagues (2018) ran a second experiment that used the same essay writing task but used a different measure of morality: antisocial behavior. After finishing the writing task, subjects were asked to assign two tasks: one to themselves and one to another subject. One task was a bit harder (takes 3 minutes longer) than the other task. The experiment measured how many subjects in each condition assigned the more burdensome task to another subject (antisocial behavior). The null hypothesis would be

$$H_0 : P_1 - P_2 = 0$$

with $\alpha = 0.05$. The screen shot below shows the Two Sample Independent Proportions Power Calculator with the sample sizes used in the follow-up experiment. Use the information in the table of problem (9) to fill-in the values for the specific alternative hypothesis.



What is the probability that the follow-up experiment will reject H_0 if the effect of the essay topic on antisocial behavior is similar to the effect on honesty in problem (9)? What advice would you give Kouchaki and colleagues (2018) about their experiment?

(11) Describe three ways to increase the power of an experiment that compares two independent means.