Introduction to Statistics in Psychology: PSY 201

Greg Francis, PhD
Department of Psychological Sciences

Practice **EXAM** ~~**1**~~ 2

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 your lowest score will be dropped from your grade.

(1) In 1997, researchers used hydrophones to listen to underwater volcanic activity in the south Pacific. They recorded an extremely loud, ultra-low frequency sound that is now called "the Bloop." Although nearly all sounds picked up on the hydrophones are likely to originate from geophysical activity, ice, animals, or human-related activity (e.g., ships, airguns), the Bloop's origin was initially unknown. Since then, a few more Bloop-like sounds have been recorded, and many people now think the Bloop is produced by an ice quake.

a) In this example, when you are looking for a Bloop sound, what is the noise? Justify your answer using the definition of "noise."

b) In this example, when you are looking for a Bloop sound, what ~~is~~ the signal? Justify your answer using the definition of "signal."

c) Using data from the hydrophones, we compute some sound measure for each sound. The mean of the distribution for non-Bloop-like sounds is 55. The mean of the distribution for Bloop-like sounds is 72. The standard deviation of both distributions is 26. Compute d'.

d) What does the d' value tell you about distinguishing Bloop and non-Bloop sounds? Explain this concept with reference to the example. (Hint: It may help to draw a plot of the distributions from part (c).)

(2) A population has a mean of $\mu = 25$ and a standard deviation of $\sigma = 4$. For samples of size $n = 10$

a) Compute the mean of the sampling distribution of the mean.

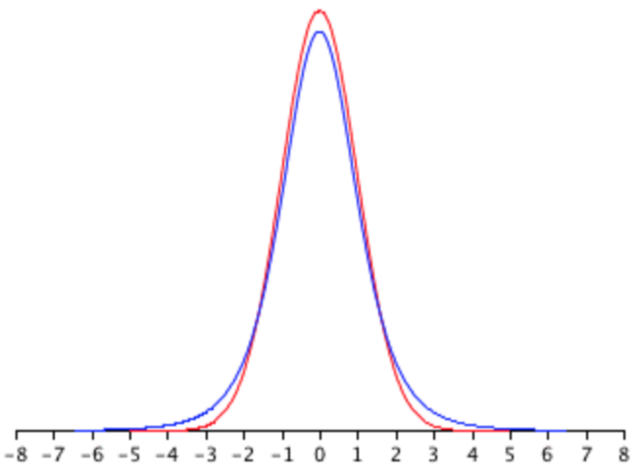
b) Compute the standard error of the mean.

c) Compute the variance of the sampling distribution of the mean.

d) What does the Central Limit Theorem imply about this sampling distribution?

e) Will increasing the sample size increase the variability of the sample means? Justify your answer.

(3) Mark answers on the image below



a) Label the distribution that is the t distribution for a sample size of $n = 5$ with the letter "A."

b) Label the distribution that is the t distribution for a sample size of $n = 50$ with the letter "B."

c) What is the number of degrees of freedom for t in part (a)? What is the number of degrees of freedom for t in part (b)?

d) Which distribution is more similar to a normal distribution? Report the corresponding letter. Explain how you know the answer.

(4) Across the population of college students, the time needed to respond to a clear visual stimulus is around 300 milliseconds. Experimenters want to test whether students in a class differ from the population by using data from the STATLAB *Speeded Response Task* experiment.

a) In this study, what is the null hypothesis? Justify your answer using the definition of "null hypothesis."

b) In this study, what is the alternative hypothesis? Justify your answer using the definition of "alternative hypothesis."

c) Suppose that the experimenters made a Type I error. Given the definition of a Type I error, did they reject or not reject H_0? Is H_0 true or false?

d) Suppose that the experimenters made a Type II error. Given the definition of a Type II error, did they reject or not reject H_0? Is H_0 true or false?

e) Which kind of error is controlled by hypothesis testing? Justify your answer. (Hint: What does the α level indicate?)

(5) In hypothesis testing, Step 1 consists of stating the null hypothesis, alternative hypothesis, and criterion. Suppose $H_0 : \mu = 51$, $H_a : \mu < 51$, and $\alpha = 0.05$. Further, you are given $s = 8.24$, $n = 105$, $\overline{X} = 50$.

a) Step 2 is to compute the test statistic. Perform this step.

b) Step 3 is to compute the p-value. Using the online calculator, we find $p = 0.108228$. Step 4 is to make a decision. Given this p-value, what decision do you make? Justify your decision.

(6) In the STATLAB *Sense of Humor* experiment, you rated jokes on a scale from 1-10, where 0 is the minimum ("Not funny at all."). The number of participants from our class was $n = 33$. The average rating for jokes was $\overline{X} = 4.939$ with $s = 1.269$. Use these ratings as a random sample to test $H_0 = 5$ against $H_a \neq 5$, with $\alpha = 0.05$.

What decision do you make? What should we conclude about our class's sense of humor? [Hint: Although you cannot compute the p-value, you can do every other part of the hypothesis test, including the decision.]

(7) Suppose you plan to conduct a study to see whether the sudden appearance of electric scooters in West Lafayette lessens average commute time. You know that the mean commute time prior to the scooters is $\mu = 14$ minutes. To find the power of a study in which 20 people are surveyed, you enter the relevant data into the One Sample Mean Power Calculator:

Specify the population characteristics:

$H_0 : \mu_0 =$ 14

$H_a : \mu_a =$ 10

$\sigma =$ 7

Or enter a standardized effect size

$\frac{\mu_a - \mu_0}{\sigma} = \delta =$ -0.571428

Specify the properties of the test:

Type of test: Negative one-tail

Type I error rate, $\alpha =$ 0.05

Power= 0.793330(— Calculate minimum sample size

Sample size, $n =$ 20 — Calculate power

a) How do you determine what value to enter for H_a?

b) Why is the "Type of test" a *Negative one-tail*?

c) What does the Power value mean?

d) What is an advantage of running a study with large power?

e) What is a disadvantage of running a study with large power?

(8) For a random sample of 5 ladybugs, the number of spots per ladybug is: 9, 9, 2, 4, 24. The standard deviation is 8.6197.

a) Compute the point estimate of the population mean of number of spots per ladybug.

b) Compute the 99% confidence interval. (Hint: $t_{cv} = 4.604$)

c) What is a benefit of using a confidence interval instead of a point estimate of a population mean?

(9) The lifespan of a worker bee is influenced by season. The lifespan of a summer worker bee ranges from 15 to 38 days, while a winter worker bee can live for nearly a year. Suppose we have a random sample of 16 summer bees. Their mean lifespan is 23.275 days. Standard deviation of the sample is 5.8295.

a) Compute the 95% confidence interval for this sample. (Hint: $t_{cv} = 2.13145$)

b) Suppose instead that the confidence interval instead ranges from 20.82 to 25.93. Compute the sample standard error. [Hint: use algebra.]

c) Which confidence interval is wider: the interval computed in part (a) or the interval given in part (b)?

d) Can we way that the 95% confidence interval computed in part (a) contains μ with a probability of 0.95? Justify your answer.

(10) Explain what a sampling distribution is.

(11) Suppose you flip a coin 10 times. The coin is fair. So, the probability that any flip will result in heads is 0.5.

a) For 10 coin flips, sketch a graph that plots probability against the number of heads. What is this distribution called? (Hint: It is not necessary to compute the probability of getting 0 heads, 1 head, etc.)

b) What is the probability of getting more than 5 heads?

c) Are coin flips independent or dependent? Explain what that means.