PSYC 2021 – Statistical methods I Name _____
November 19, 2019

Student number _____

Test 2 - answers

Aids permitted
(a) a simple calculator with arithmetic and square root (no scientific calculators)
(b) one-sided 8.5" x 11" page of formulas and notes

Time permitted: 90 minutes

A. Multiple choice questions

Answer all questions in this section. Each question is worth two marks.

1. What is the standard error of the mean?
(a) the difference between a sample mean and the population mean
(b) the standard deviation of the population that a sample is drawn from
(c) the mean of the sample mean over many samples
(d) the standard deviation of the population mean over many samples
(e) the standard deviation of the sample mean over many samples

2. We take a sample of 100 data points from a normally distributed population. The sample has mean 100 and standard deviation 100. The standard error of the mean is:
(a) 1.0
(b) 3.16
(c) 10.0
(d) 31.6
(e) 100.0

3. In which case is the sample mean most likely <u>not</u> to be normally distributed?
(a) population normally distributed, sample size $n < 30$
(b) population normally distributed, sample size $n \geq 30$
(c) population not normally distributed, sample size $n < 30$
(d) population not normally distributed, sample size $n \geq 30$
(e) none of the above; the sample mean is always normally distributed

4. The null hypothesis is:
(a) the hypothesis that the mean is zero
(b) the hypothesis that the mean is too small to measure
(c) the hypothesis that the sample has too few data points
(d) the hypothesis that we test in a statistical significance test
(e) the hypothesis that two populations are significantly different

5. The alpha level (α) is the probability of:
(a) rejecting the null hypothesis when it is true
(b) rejecting the null hypothesis when it is false
(c) accepting the null hypothesis when it is true
(d) accepting the null hypothesis when it is false
(e) the null hypothesis being true

6. When is a t distribution most similar to a normal distribution?
(a) when we do a one-tailed t test
(b) when we do a two-tailed t test
(c) when $\alpha = 0.05$
(d) when the number of degrees of freedom is large
(e) when the number of degrees of freedom is small

7. What is a type I error?
(a) failing to check whether the population is normally distributed
(b) accepting the null hypothesis when it is false
(c) rejecting the null hypothesis when it is true
(d) using the z score when the t value is appropriate
(e) using the t value when the z score is appropriate

8. A statistically significant result is one that
(a) cannot occur if the null hypothesis is true
(b) is unlikely to have occurred if the null hypothesis is true
(c) conclusively proves that the null hypothesis is true
(d) conclusively proves that the null hypothesis is false
(e) always occurs if the null hypothesis is false

9. The effect size (i.e., Cohen's d) is
(a) the difference between two means
(b) the difference between two means, measured as a multiple of the population standard deviation
(c) the difference between two means, measured as a multiple of the standard error
(d) the difference between two standard errors
(e) the α level we use in a significance test

10. A confidence interval is
(a) the length of time during which a measurement remains valid
(b) the difference between the first quartile and the median
(c) the range that we are fairly certain contains the true value of a parameter
(d) the number of data points required to estimate a parameter
(e) a number that reports how confident a subject is of their response in an experiment

11. One limitation of statistical significance testing is that
(a) it does not take into account random variations from sample to sample
(b) it assumes that we know the population standard deviation
(c) critical t values are usually difficult to determine
(d) it is strongly affected by sample size
(e) all of the above

12. Another limitation of significance testing is that
(a) it is unaffected by sample size
(b) t values are difficult to calculate
(c) the sample size is usually unknown
(d) it assumes that we know the population mean
(e) statistically significant differences are not always practically important differences

Multiple choice answers

1 e	2 c	3 c	4 d	5 a	6 d
7 c	8 b	9 b	10 c	11 d	12 e

Name _____

Student number _____

B. Written answer questions

Answer any two of the following three questions. If you answer all three, we will only mark the first two. Each question is worth ten marks. Show your work in a way we can understand.

1. We measure the reaction time of four chimpanzees at recognizing faces of family members. The reaction times are 0.45 s, 0.48 s, 0.42 s, and 0.44 s. Previous studies show that the mean reaction time for humans in this task is 0.50 s.

(a) What is the mean reaction time? (one mark)

0.448

(b) What is the standard deviation of the reaction times? (two marks)

0.025

This problem continues on the back of the page.

(c) What is the standard error of the mean that you calculated in part (a)? (two marks)

0.025 / sqrt(4) = 0.0125

(d) Is the mean reaction time for chimpanzees significantly different from the mean reaction time for humans? Use a one-sample, two-tailed t test with $\alpha = 0.05$. (five marks)

t = (0.448 - 0.5) / 0.0125 = -4.16

df = 3

$t_{crit} = 3.182$

Yes, $| t | > t_{crit}$. There is a significant difference.

Name _____

Student number _____

2. You wish to evaluate a new training course for radiologists. For radiologists who have not taken the course, the score on a standardized test has a population mean of 82.1 and a population standard deviation of 5.4. Ten radiologists take the course, and after the course their scores on the standardized test have a sample mean of 85.5 and a sample standard deviation of 5.1.

(a) What effect size does the training course have on the standardized test score, measured using Cohen's d? (four marks)

$d = (85.5 - 82.1) / 5.4 = 0.63$

(b) Is this a small, medium, or large effect size? (one mark)

This is a medium effect size.

(c) What is the 95% confidence interval for the mean score of radiologists who have taken the test? (five marks)

$s_{xbar} = 5.1 / \text{sqrt}(10) = 1.61$

$df = 9$

$t_{crit} = 2.262$

$85.5 \pm 2.262 \times 1.61 = (81.86, 89.14)$

3. Climate scientists measure the temperature at the South Pole once per day, for 365 days, in the year 2019. The mean of these 365 measurements is -47.1 ºC and the standard deviation is 10.2 ºC. You wish to use a t test to see whether this mean temperature is significantly higher than the historical average of -49.5 ºC, at the α = 0.01 level.

(a) Is this a one- or two-tailed t test? (two marks)

one-tailed

(b) What is the standard error of the mean temperature in 2019? (two marks)

10.2 / sqrt(365) = 0.53

(c) Do the t test described above: is the mean significantly greater than -49.5 ºC at the α = 0.01 level? (six marks)

t = (-47.1 - (-49.5)) / 0.53 = 2.4 / 0.53 = 4.53

df = 364, so we use the line in Appendix B for df = Inf

t_{crit} = 2.326

Yes, t > t_{crit}. The mean temperature is significantly higher than the historical average.