

Your name: _____

Your lab time (Circle one): 8:00 10:00 12:00

This is a 50-minute open-book, open-notes test. Show work where appropriate.

(3 points each) Write the letter from the list below that best matches the description. (Letters may be used more than once.)

1. _____ The t-ratio when the parameter value is replaced by its hypothesized value.
2. _____ The square root of the population variance.
3. _____ The set of all possible parameter values that lead to two-sided p-values greater than or equal to .05
4. _____ A term to describe a variable whose possible presence prevents causal statements in observational studies.
5. _____ The probability distribution of outcomes of a statistic from all possible random samples.
6. _____ A property of a statistical tool that implies the tool still works well even if the assumptions on which it is based are not exactly met.
7. _____ The probability that randomization or random sampling leads to a test statistic as extreme in evidence against the null hypothesis as the observed test statistic.
8. _____ A test for comparing the means of two normal populations with unequal variances.
9. _____ A distribution-free alternative to the two-sample t-test.
10. _____ A statistical measure to express the degree of credibility in a null hypothesis.

- | | | |
|----------------------------------|-------------------------------|----------------------|
| A. 95% confidence interval | H. randomization distribution | O. signed-rank test |
| B. confounding variable | I. rank-sum test | P. t distribution |
| C. p-value | J. resistance | Q. t-ratio |
| D. permutation distribution | K. robustness | R. t-statistic |
| E. population distribution | L. sample standard deviation | S. two-sample t-test |
| F. population standard deviation | M. sampling distribution | T. Welch's t-test |
| G. population variance | N. standard error | U. z-statistic |

11. (8 points) A sociologist identified 15 days in which there was a newspaper article about a suicide in New York City, and 15 other days in which there was no article about suicide. For each of these she determined, from public health records, the number of suicides in New York City in the following week. She wished to see whether the mean number of suicides is greater in weeks following a newspaper article about suicide than in weeks that don't follow a publicized suicide.

Is the data structure paired or two independent samples? (Circle one) Paired Two-sample

Can a statistical statement of causation be made? (Circle one:) Yes No

12. (8 points) Researchers found a set of 30 twins in which one twin smoked marijuana and the other didn't. After 5 years, the researcher asked all 60 participants to report the number of times they had taken harder drugs. He wished to study whether smoking marijuana caused a person to be more susceptible to harder drugs.

Is the data structure paired or two independent samples? (Circle one) Paired Two-sample

Can a statistical statement of causation be made? (Circle one:) Yes No

13. (8 points) Ten tidal pools of the same size were found in a certain coastal region. It was randomly determined which 5 would receive a treatment (removal of limpets, a type of seaweed grazer) and which 5 would serve as controls. The amount of seaweed covering the floor of the tide pools was measured at the end of the study period. It was desired to see whether the amount of seaweed was affected by limpet removal.

Is the data structure paired or two independent samples? (Circle one) Paired Two-sample

Can a statistical statement of causation be made? (Circle one:) Yes No

14. (8 points) Two fifth-grade classes in each of ten schools agreed to participate in an experiment. In each school it was randomly decided which fifth grade class would receive a new method of teaching mathematics and which would use the standard teaching method. After the teaching of the material, students took a common test. The average test scores in the 20 classes were used to investigate whether the new method tended to lead to higher scores.

Is the data structure paired or two independent samples? (Circle one) Paired Two-sample

Can a statistical statement of causation be made? (Circle one:) Yes No

15. (6 points) Calculate the rank-sum statistic from the following two groups of numbers. Show work. Circle your answer.

group 1: 1, 4, 9

group 2: 1, 2, 3

16. (8 points) Suppose that from samples of size $n_1 = 12$ and $n_2 = 17$, it is found that $(\bar{Y}_2 - \bar{Y}_1) = 5.7$ and $SE(\bar{Y}_2 - \bar{Y}_1) = 2.1$. What is a t-statistic for the hypothesis that $\mu_2 - \mu_1 = 3$?

17. (5 points) What are the degrees of freedom associated with the t-statistic in problem 16?

18. (6 points) In a sex discrimination problem, the following averages were obtained for comparing log salaries (natural log):

Males: average log salary = 10.90

Females: average log salary = 10.71

Fill in the blank: It is estimated that the median salary for males is _____ percent more than the median salary for females.

19. (8 points) The following are summary statistics for random samples from 3 populations, which are thought to have the same variance:

Sample 1: $n_1 = 20$, $\bar{Y}_1 = 19.5$, $s_1 = 4.5$

Sample 2: $n_2 = 18$, $\bar{Y}_2 = 24.2$, $s_2 = 5.4$

Sample 3: $n_3 = 20$, $\bar{Y}_3 = 17.5$, $s_3 = 5.0$

What is the pooled estimate of the common standard deviation?

20. (5 points) If it is desired to construct a confidence interval for $\mu_3 - \mu_2$ from the data in problem 19, the t-multiplier should be found from a t-distribution with how many degrees of freedom?