# STA 3032 - Exam 2

Code: B

Last Name: First Name:

1. As $\alpha$ decreases (all other variables constant), the width of the confidence interval for $\mu$ becomes
   (a) Smaller.
   (b) Larger.

2. For a particular sample the 95% confidence interval for the population mean is (23, 37). Which of the following statements is true?
   (a) 95% of the population lines in the interval (23, 37).
   (b) 95% of the samples from this population have means in the interval (23, 37).
   (c) This interval was calculated using a procedure that gives correct intervals 95% of the time.
   (d) The probability the population mean lies in the interval (23,37) is 0.95.

3. When performing inference on population means we require that the data are normal or that the sample size is large (to use the Central Limit Theorem) so that the sample means have a sampling distribution that is at least approximately normal. The same conditions hold for performing inference on population variance(s).
   (a) False
   (b) True

4. In an experiment two independent groups were asked to perform a task. In group A there were 15 subjects and in group B there were 13 subjects There were 15 successes in A and 0 successes in B. The appropriate **Agresti-Coull** 95% C.I. for $p_A - p_B$ is (using $z_{0.975} = 1.96$)
   (a) (0.7058, 1.0432)
   (b) (0.0000, 0.7888)
   (c) (0.0000, 1.0432)
   (d) (0.6500, 1.0423)
   (e) (0.7058, 1.0000)

5. Two random samples $A$ and $B$ of size $n_A = 10, n_B = 8$ were obtained with sample variances
$$s_A^2 = 2.87,\ s_B^2 = 1.91$$
   What is lower limit of the two-sided 95% C.I. for $\sigma_B^2/\sigma_A^2$
   (a) $1.5026(1/F_{0.95,(10,8)})$
   (b) $0.6655(1/F_{0.975,(9,7)})$
   (c) $1.5026(1/F_{0.975,(7,9)})$
   (d) $0.6655(1/F_{0.025,(9,7)})$
   (e) $0.6655(1/F_{0.975,(7,9)})$

6. For the Wilcoxon rank-sum and signed rank tests, the data needs to be ranked.

| Data | 16 | 18 | 33 | 26 | 18 | 21 | 31 | 18 | 28 |
|---|---|---|---|---|---|---|---|---|---|
| Rank | | | | | | | | | |

What is the rank for the value of 21?

(a) 7
(b) 4
(c) 4.5
(d) 6
(e) 5

7. A fertilizer plant has a goal of the 40th percentile of its daily total fertilizer that does not meet expectations to be less than 50 kilogramms. Data was collected over a period of 23 days. The test statistic $B$ (as defined in class) is the number of observations greater than the said percentile's value, i.e. 50 kg. What is the p-value for testing the plant's goal?

(a) $P\left(\text{Bin}(23, 0.4) \geq B\right)$
(b) $P\left(\text{Bin}(23, 0.4) \leq B\right)$
(c) $P\left(\text{Bin}(50, 0.4) > B\right)$
(d) $P\left(\text{Bin}(23, 0.6) \leq B\right)$
(e) $P\left(\text{Bin}(23, 0.6) < B\right)$

8. A 95% confidence interval for the population variance yields $(8.175, 20.588)$, while a 90% confidence interval yields $(8.720, 18.903)$. For a hypotheses of

$$\text{H}_0 : \sigma^2 = 8.5 \quad \text{vs} \quad \text{H}_a : \sigma^2 \neq 8.5$$

the p-value would be

(a) $< 0.025$
(b) $0.025 <$ p-value $< 0.05$
(c) $0.05 <$ p-value $< 0.10$
(d) $> 0.10$

9. We saw that Wilcoxon signed rank tests was used for one sample, and for two dependent samples. The sign-test for inference on population percentiles (location statistics) can theoretically be used in which of the following settings?

   [1 ] One population

   [2 ] Two independent populations

   [3 ] Two dependent populations

   (a) Answer Only (1)
   (b) Answers (1) and (2)
   (c) Answers (1) and (3)
   (d) Answers (2) and (3)
   (e) Answers (1) and (2) and (3)

10. The claim "the standard deviation of a population is greater than 12.6" is tested using

    (a) Chi-squared statistic.
    (b) q-statistic.
    (c) F-statistic.
    (d) t-statistic.
    (e) z-statistic.

11. A sample of 763 customers with non-zero balances yielded a sample mean of 115 (in hundreds of dollars). Assume the population standard deviation is 25. To create a 95% confidence interval for the *total* (in hundreds of dollars) of the non-zero balances owed by these customers we need to know the distribution of the **sum**. It's distribution will be

    (a) $t_{762}$
    (b) $\chi^2_{762}$
    (c) $N(87745, 25^2)$
    (d) $N(87745, 476875)$
    (e) $t_1$

12. A bored statistician believes the average length of commercial breaks on cable channels (ESPN, TNT, etc.) is longer than the average length of commercial breaks on the networks (ABC, CBS, NBC, FOX). She times a random sample of commercials from each of the two types of broadcasts.

| Source | $n$ | Mean(min.) | Std. Dev.(min.) |
|---|---|---|---|
| Cable (C) | 115 | 2.36 | 0.16 |
| Network (N) | 112 | 2.18 | 0.13 |

Which are the appropriate hypotheses?

(a) $\text{H}_0 : \mu_C - \mu_N \leq 0 \quad \text{vs} \quad \text{H}_a : \mu_C - \mu_N > 0$

(b) $\text{H}_0 : \mu_C - \mu_N \geq 0 \quad \text{vs} \quad \text{H}_a : \mu_C - \mu_N < 0$

(c) $\text{H}_0 : \mu_C - \mu_N = 0 \quad \text{vs} \quad \text{H}_a : \mu_C - \mu_N \neq 0$

(d) $\text{H}_0 : \mu_C - \mu_N > 0 \quad \text{vs} \quad \text{H}_a : \mu_C - \mu_N \leq 0$

13. For question 12, which distribution is most (technically) appropriate to be used to find the p-value?

(a) $Z \sim N(0, 1)$

(b) $t_{225}$

(c) $t_{218.0379}$

(d) Uniform(0,1)

14. (Extra credit, half value of regular question). For the following hypotheses

$$\text{H}_0 : \mu \leq \mu_0 \quad \text{vs} \quad \text{H}_a : \mu > \mu_0$$

performed at the $\alpha$ significance level, the corresponding confidence interval that would included all the $\mu_0$ values for which one would fail to reject the null is

(a) $100(1-\alpha)\%$ two-sided confidence interval

(b) $100(1-\alpha)\%$ one-sided confidence interval with only upper limit, i.e. $(-\infty, U)$

(c) $100(1-\alpha)\%$ one-sided confidence interval with only lower limit, i.e. $(L, \infty)$

**END**