AP Statistics
STUDY GUIDE

Name Key
Date: Blk:

C **1.** The heights (in inches) of males in the United States are believed to be Normally distributed with mean μ. The average height of a random sample of 25 American adult males is found to be $\bar{x} = 69.72$ inches, and the standard deviation of the 25 heights is found to be $s = 4.15$. The standard error of $\bar{x}$ is

(a) 0.17 (b) 0.69 (c) 0.83 (d) 1.856 (e) 2.04

B **2.** An analyst, using a random sample of $n = 500$ families, obtained a 90% confidence interval for mean monthly family income for a large population: ($600, $800). If the analyst had used a 99% confidence level instead, the confidence interval would be:

(a) Narrower and would involve a larger risk of being incorrect
(b) Wider and would involve a smaller risk of being incorrect
(c) Narrower and would involve a smaller risk of being incorrect
(d) Wider and would involve a larger risk of being incorrect
(e) Wider but it cannot be determined whether the risk of being incorrect would be larger or smaller

A **3.** I collect a random sample of size n from a population and from the data collected compute a 95% confidence interval for the mean of the population. Which of the following would produce a new confidence interval with larger width (larger margin of error) based on these same data?

(a) Use a larger confidence level.
(b) Use a smaller confidence level.
(c) Use the same confidence level, but compute the interval n times. Approximately 5% of these intervals will be larger.
(d) Increase the sample size.
(e) Nothing can guarantee absolutely that you will get a larger interval. One can only say the chance of obtaining a larger interval is 0.05.

E **4.** You construct a 95% confidence interval for a mean and find it to be 1.1 ± 0.8. Which of the following is true?

(a) A test of the hypotheses H_0: $\mu = 1.2$, Ha: $\mu \neq 1.2$ would reject H_0 at the 0.05 level.
(b) A test of the hypotheses H_0: $\mu = 1.1$, Ha: $\mu \neq 1.1$ would reject H_0 at the 0.05 level.
(c) A test of the hypotheses H_0: $\mu = 0$, Ha: $\mu \neq 0$ would reject H_0 at the 0.05 level.
(d) All three tests above would reject H_0 at the 0.05 level.
(e) A test of hypothesis cannot be performed from only a confidence interval.

A **5.** Does listening to music increase the speed at which people complete routine tasks? Fifteen volunteers are asked to sort 100 red and white beads into two piles according to color, once while listening to Handel's Water Music and once in silence (the order—music or silence first—is determined for each subject by the flip of a coin). Here are the data (times are in seconds), along with summary statistics in the last two columns:

	Subject																
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	$\bar{x}$	s
Handel	14	15	17	21	13	18	21	24	15	25	22	30	13	15	16	18.6	5.04
Silence	16	18	17	20	15	17	22	27	14	22	20	32	15	13	17	19.0	5.13
Difference	2	3	0	−1	2	−1	1	3	−1	−3	−2	2	2	−2	1	0.4	1.96

We wish to test the hypothesis that the mean difference in time to sort the beads with and without music is 0. Which of the following is the appropriate test statistic?

(a) $\dfrac{19.0-18.6}{\sqrt{\dfrac{5.04^2}{15}+\dfrac{5.13^2}{15}}}$ (b) $\dfrac{19.0-18.6}{\sqrt{\dfrac{5.04^2}{14}+\dfrac{5.13^2}{14}}}$ (c) $\dfrac{19.0-18.6}{\dfrac{5.04}{\sqrt{15}}+\dfrac{5.13}{\sqrt{15}}}$

(d) $\dfrac{0.4}{\dfrac{1.96}{\sqrt{30}}}$ (e) $\dfrac{0.4}{\dfrac{1.96}{\sqrt{15}}}$

A **6.** A kennel club argues that 50% of dog owners in its area own Golden Retrievers, 40% own Shepherds of one kind or another, and 10% own a variety of other breeds. A random sample of 50 dogs from the area turns up the data in the following table:

GOLDEN RETRIEVER	SHEPHERD	OTHER
27	22	1

What is the value of the X^2 statistic for the goodness-of-fit test on these data?

a. 3.56 b. 2.12 c. 4.31 d. 3.02 e. 2.

A **7.** A researcher wants to see if birds that build larger nests lay larger eggs. She selects two random samples of nests: one of small nests and the other of large nests. She weighs one egg from each nest. The data are summarized below.

	Small nests	Large nests
sample size	60	159
sample mean (g)	37.2	35.6
sample variance	24.7	39.0

A 95% confidence interval for the difference between the average mass of eggs in small and large nests is:

(a) $(37.2-35.6)\pm 2.000\sqrt{\dfrac{24.7^2}{60}+\dfrac{39.0^2}{159}}$ (b) $(37.2-35.6)\pm 2.009\sqrt{\dfrac{24.7^2}{59}+\dfrac{39.0^2}{158}}$

(c) $(37.2-35.6)\pm\sqrt{\dfrac{24.7}{59}+\dfrac{39.0}{158}}$ (d) $(37.2-35.6)\pm\sqrt{\dfrac{24.7^2+39.0^2}{59+158}}$ (e) None of these

E **8.** Two types of tennis balls were tested to determine which one goes faster on a serve. Eight different players served one of each type of ball and the results were recorded.

Server	Raphael	Roger	Serena	Venus	Andy	Justine	LLeyton	Maria
Type A	120	125	119	110	118	82	115	105
Type B	115	122	114	114	115	91	110	106

Assuming that the speeds are approximately normally distributed, how many degrees of freedom will there be in the appropriate t-test used to determine which type of tennis ball travels faster?

(a) 6 (b) 7 (c) 16 (d) 15 (e) 14

D **9.** A survey of randomly selected college students found that 50 of the 95 freshmen and 52 of the 107 sophomores surveyed had purchased used textbooks in the past year. Construct a 98% confidence interval for the difference in the proportions of college freshmen and sophomores who purchased used textbooks.

A) (-0.098, 0.664) B) (0.362, 0.664) C) (0.388, 0.664)
D) (-0.124, 0.204) E) (0.362, 0.690)

A **10.** An SAT test preparation program advertises that its program will improve scores on the SAT test by at least 30 points. Twelve students who have not yet taken the SAT were selected for the study and were administered the test. The 12 students then went through the 3-week testprep course. The results of the testing were as follows:

Student	1	2	3	4	5	6	7	8	9	10	11	12
Before	475	500	499	477	540	608	510	425	495	502	530	487
After	495	540	495	522	555	684	535	460	522	529	560	512

Assuming that the conditions necessary to conduct the test are present, which of the following significance tests should be used to determine if the test-prep course is effective in raising score by the amount claimed?

a.) Two sample t test
b.) Chi-Square Test of Independence
c.) One sample t test
d.) t test for the slope of a regression line
e.) Two sample z test

11. A researcher wished to test the effect of the addition of extra calcium to yogurt on the "tastiness" of yogurt. Eighty-two adult volunteers were randomly divided into two groups of 41 subjects each. Group 1 tasted yogurt containing the extra calcium. Group 2 tasted yogurt from the same batch as group 1 but without the added calcium. Both groups rated the flavor on a scale of 1 to 10, with 1 being "very unpleasant" and 10 being "very pleasant." The mean rating for group 1 was $\bar{X}_1 = 6.5$ with a standard deviation of $s_1 = 1.5$. The mean rating for group 2 was $\bar{X}_2 = 7.0$ with a standard deviation of $s_2 = 2.0$. Let μ_1 and μ_2 represent the mean ratings we would observe for the entire population represented by the volunteers if all members of this population tasted, respectively, the yogurt with and without the added calcium. Researchers claim that the addition of calcium to the yogurt does not negatively affect the taste of the yogurt.

(A) Construct and interpret a 90% confidence interval for $\mu_1 - \mu_2$.

-1.15 to 0.15

(B)Use the confidence interval you constructed in (a) to comment on whether you agree with the researcher's claim. Explain your reasoning clearly.

12. Acid rain is a serious problem in Canada. In many cases, lakes become so acidified that they cannot support any significant fish life. One possible (and very costly!) solution is to try to mitigate the effects by dumping crushed limestone into the lakes. This will neutralize the acidity. The following are actual data from a study of such an intervention in a lake.
From enormous samples at other control lakes, it is reasonable to assume that under the acidic conditions the weight of individual fish of a particular age class is Normally distributed with a known mean μ of 3250 grams (g). One year after the addition of limestone, a sample of 22 fish was taken and the weight of the individual fish was obtained. Here are the sorted data (g):

1595	1605	1634	2633	2864	2924	3035	3051	3293	3344	3381
3398	3421	3446	3514	3614	3694	3739	3756	3788	3898	3952

Before the analysis began, it was noticed that several fish had abnormally low weights (below 2000 g). After further investigation it was noted that these fish had ingested pieces of plastic from litterbugs' foam cups and could not properly digest food. The study's method of analysis was to delete all values less than 2000 g. After deleting such values, $\bar{x}$ = 3407.6 g. Assume that σ = 370 grams for the weights of fish after limestone is added.

(A) Carry out an appropriate significance test at the 5% significance level to determine whether the mean weight of the fish in the lake increased after the limestone was added.

$$t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{3407.6 - 3250}{\frac{370}{\sqrt{19}}} = 1.857 \qquad \text{Pvalue} = .039$$

* Support the idea that fish weighed less!

(B)Would your conclusion in (a) have changed if the outliers had been included in the analysis? Justify your answer with appropriate statistical evidence.

$\bar{x} = 3162.68$ $\quad S_x = 717.95$

P-value = .712

You would not reject the claim that the fish's weights did not change!

13. A company hopes to improve its engines, setting a goal of no more than 3% of customers using their warranty on defective engine parts. A random survey of 1400 customers found only 30 with complaints. Create a 95% confidence interval for the true level of warranty users among all customers

$.0214 \pm .007 =$ .0144 to .0284

14. You need to find a new hair stylist and know that there are two terrific salons in your area, Hair by Charles and Curl Up & Dye. You want a really good haircut, but you do not want to pay too much for the cut. A random sample of costs for 10 different stylists was taken at each salon (each salon employs over 100 stylists).

(A) Indicate what inference procedure you would use to see if there is a significant difference in the costs for haircuts at each salon. Check the appropriate assumptions and conditions and indicate whether you could or could not proceed.

Two-sample t-test

Difference of Means

(B) A friend tells you that he has heard that Curl Up & Dye is the more expensive salon.

i. Write hypotheses for your friend's claim.

$H_0: M_1 - M_2 = 0$

$H_A: M_1 < M_2$

ii. The following are computer outputs. Which output is the correct one to use for this test? ~~Explain~~.

(*) Output A

Output A:

```
Two-sample T for Hair by Charles vs Curl Up & Dye

                    N   Mean  StDev  SE Mean
Hair by Charles    10  22.10   6.33      2.0
Curl Up & Dye      10  26.00   4.81      1.5


Difference = mu (Hair by Charles) - mu (Curl Up & Dye)
Estimate for difference:  -3.90000
95% CI for difference:  (-9.22983, 1.42983)
T-Test of difference = 0 (vs not =): T-Value = -1.55  P-Value = 0.140
DF = 16
```

Output B:

```
Paired T for Hair by Charles - Curl Up & Dye

                    N      Mean    StDev  SE Mean
Hair by Charles    10   22.1000   6.3325   2.0025
Curl Up & Dye      10   26.0000   4.8074   1.5202
Difference         10  -3.90000  7.37036  2.33071


95% CI for mean difference: (-9.17244, 1.37244)
T-Test of mean difference = 0 (vs not = 0): T-Value = -1.67  P-Value =
0.129
```

iii. Use the appropriate output to make a conclusion about the hypothesis test based on the data. Make sure to state your conclusion in context.

p-value = 0.14

* There is not enough evidence to say there is a difference in prices!

15. A state university wants to increase its retention rate of 4% for graduating students from the previous year. After implementing several new programs during the last two years, the university reevaluated its retention rate using a random sample of 352 students and found the retention rate at 5%. Test an appropriate hypothesis and state your conclusion. Be sure the appropriate assumptions and conditions are satisfied before you proceed.

$H_0: p = .04$

$H_A: p > .04$

$z = 1.06$

$P\text{value} = .143$

Support	H_0
Reject	H_A