Stat 104: Exam 2 Solutions

True/False #1

If a population is known to be normally distributed, then it follows that the sample standard deviation must equal the population standard deviation for any size sample.

a) True

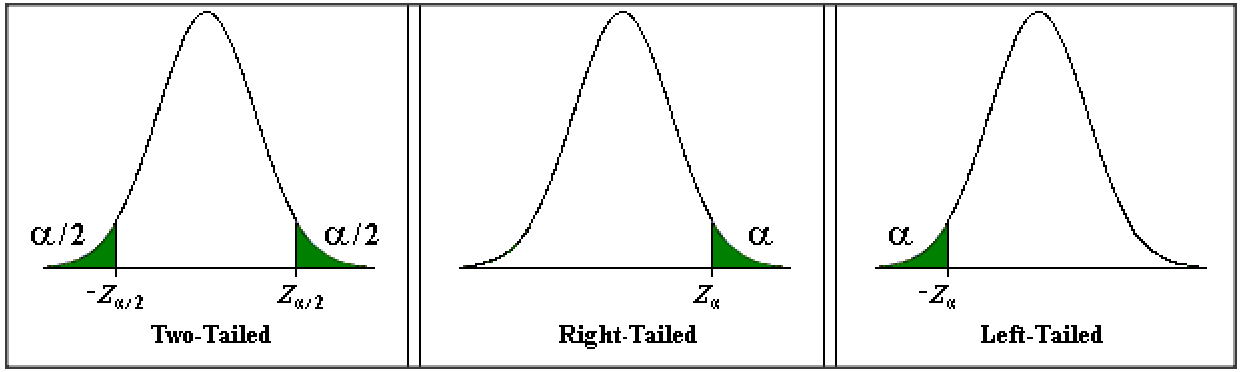
b) False

True/False #2

The smaller the p-value, the stronger the evidence against the null hypothesis

a) True

b) False



True/False #3

The larger the sample size, the more the sampling distribution of sample means resembles the shape of the population distribution.

a) True

$\boxed{\text{b) False}}$

True/ False #4

If we fail to reject the null hypothesis (Ho) at a significance level of $\alpha=0.05$, then we also must fail to reject it at a significance level of $\alpha=0.10$.

a) True

$\boxed{\text{b) False}}$

If the p-value=0.07:

- At a 5% level of significance, 0.07>0.05 so we fail to reject the null
- At a 10% level of significance 0.07<0.1 so we reject the null

True/False #5

All t distributions have a mean zero and a standard deviation of 1.

a) True

$\boxed{\text{b) False}}$

True/False #6

If a 95% confidence interval for the mean number of hours that students study per week is (18, 23), then there is a 95% chance that a randomly selected student will study between 18 and 23 hours per week.

a) True

b) False

True/False #7

Consider testing the hypothesis $H_o: \mu = 50$ $\mathrm{H}_a: \mu \neq 50$. If $n = 64$, $\bar{x} = 53.5$ and $s = 10$, then the value of the test statistic is 2.10.

a) True

$\boxed{\text{b) False}}$

$$tstat = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{53.5 - 50}{10/\sqrt{64}} = 2.8$$

True/False #8

If Ho is rejected, the probability of making a Type II error is 0.

a) True

b) False

	Fail to reject H_0	Reject H_0
If H_0 is true	Correct Decision	Type I Error
If H_0 is false	Type II Error	Correct Decision

Multiple Choice #9

According to the Law School Admission Council, in the fall of 2016, 66% of law school applicants were accepted to some law school. The training program LSATisfaction™ claims that 163 of the 240 students trained in 2016 were admitted to law school. You can safely consider these trainees to be representative of the population of law school applications. What is the value of the test statistic to test if LSATisfaction™ has demonstrated a real improvement over the national percentage?

$\boxed{\text{a) } 0.63}$

b) 0.94

c) 0.48

d) 1.27

e) 1.57

$$tstat = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{163}{240} - 0.66}{\sqrt{\frac{0.66(1-0.66)}{240}}} = 0.63$$

Multiple Choice #10

A student wishes to know if SUV drivers at his school are more likely to be male than female. He takes a simple random sample of 60 students from a large list of students that consists of all the SUV drivers at the school, and records the gender of each student in the sample. What would be the appropriate inference procedure to use?

a) Two sample t-test; testing to see if the mean number of females differs from the mean number of males driving SUVs
b) Chi-square test of independence
c) Test for a population proportion; testing to see if the proportion of SUV drivers who are males is greater than 50%
d) Test for the difference of two proportions; testing to see if the proportion of SUV drivers who are male is greater than the proportion of SUV drivers who are female
e) Paired t-test

Multiple Choice #11

Which of the following is a required condition for ANOVA?

a) The populations are not normally distributed.

b) The population variances are equal.

c) The samples are dependent.

d) All of these choices are not required conditions for ANOVA.

Multiple Choice #12

A Gallup poll of 1089 adults found 326 supported the policies of a particular political party. A 95% confidence interval for the true level of support in the entire population is:

a) (0.299, 0.300)

b) (0.272, 0.327)

c) (0.285, 0.313)

d) (0.267, 0.332)

$$95\%\ CI\colon\ \hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.30 \pm 1.96\sqrt{\frac{0.30(1-0.30)}{1089}}$$

$$= (0.272, 0.327)$$

$$\hat{p} = \frac{326}{1089} = 0.30$$

Multiple Choice #13

A financial analyst wanted to determine the mean annual return on mutual funds. A random sample of 60 returns show a mean of 12% and standard deviation of 4%. Give a 95% confidence interval for the mean annual return on all mutual funds.

a) (15%, 17%)
b) (9%, 11%)
c) (13%, 15%)
d) $\boxed{\text{(11\%, 13\%)}}$

$$95\%\ CI:\ \bar{X} \pm 1.96\frac{s}{\sqrt{n}} = 12 \pm 1.96\frac{4}{\sqrt{60}} = (11, 13)$$

Multiple Choice #14

The owner of a travel agency would like to determine whether or not the mean age of the agency's customers is over 24. If so, he plans to alter the destination of their specific cruises and tours. If he concludes the mean age is over 24 when it is not, he makes a _______ error. If he concludes the mean age is not over 24 when it is, he makes a _______ error.

a) Type II; Type II

b) Type I; Type I

c) $\boxed{\text{Type I, Type II}}$

d) Type II; Type I

	Fail to reject H_0	Reject H_0
If H_0 is true	Correct Decision	Type I Error
If H_0 is false	Type II Error	Correct Decision

Multiple Choice #15

The mean cost of renting an apartment in a city is $2000 per month with a standard deviation of $300. Suppose we take a sample of 60 apartments in the city. The probability that the sample mean is larger than $2050 is

a) 0.0985
b) 0.1783
c) 0.4013
d) 0.5987
e) 0.9015

$$z = \frac{2050 - 2000}{300/\sqrt{60}} = 1.29$$

$P(Z>1.29)=1-0.9015= 0.0985$

Multiple Choice #16

Typically about 57.5% of eligible American adults vote in presidential elections. Ted, a political science student, would like to perform a test of hypothesis to see if this percentage is different for his classmates. He is taking a lecture class and is connected to all his classmates though Canvas. He decided to randomly select 42 of his classmates for his sample and finds 35.7% of them voted in the last presidential election. Suppose a 95% confidence interval was formed for the proportion of all his classmates that voted in the last presidential election and it was found to be (0.212, 0.502). Had he performed a two-tailed hypothesis test at a 5% level of significance, would the null hypothesis have been rejected?

a) No, since 0.357 falls inside the confidence interval.

b) No, since 0.575 is larger than 0.05.

c) Yes, since 0.575 falls outside of the confidence interval.

d) Yes, since 0.05 falls outside of the confidence interval.

e) No conclusion can be made due to insufficient information.

Multiple Choice # 17

Suppose that a manufacturer is testing one of its machines to make sure that the machine is producing more that 97% good parts (Ho: p=0.97 and Ha: p>0.97). The test results in a P-value of 0.122. Unknown to the manufacturer, the machine is actually producing 99% good parts. Given that the manufacturer has chosen a significance level of 0.05, what happens as a result of the testing?

	Fail to reject H_0	Reject H_0
If H_0 is true	Correct Decision	Type I Error
If H_0 is false	Type II Error	Correct Decision

a) They correctly fail to reject the Ho.

b) They correctly reject Ho.

c) They reject Ho, making a Type I error.

d) They fail to reject Ho, making a Type I error.

e) They fail to reject Ho, making a Type II error.

Multiple Choice # 18

A recent Pew Research Center poll of teen “sexting” trends found that 20% of U.S. teens ages 12-17 have received sexually explicit photos on their cellular phones. The poll was based on a confidence level of 95%, and reported margin of error of ± 2.5%. Find the sample size used by Pew to conduct this poll.

a) 1061

b) 1911

c) 722

d) $\boxed{984}$

e) 865

$$n = \frac{p(1-p)1.96^2}{E^2} = \frac{0.2(1-0.2)1.96^2}{0.025^2} = 984$$

Multiple Choice # 19

In 2004, a random sample of 46 Coyotes in a region of northern Minnesota showed the average age to be 2.05 years with a standard deviation of 0.82 years. However, it is thought that the overall population mean age of coyotes is 1.75 years. Suppose we want to test the claim that coyotes in this region of northern Minnesota live longer than the average of 1.75 years. What is the value of the test statistic?

a) 2.21
b) 1.52
c) 0.83
d) 0.61
e) $\boxed{2.48}$

$$tstat = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{2.05 - 1.75}{\frac{0.82}{\sqrt{46}}} = 2.48$$

Multiple Choice # 20

A controversy that Americans have debated in recent years is the issue of gay marriage. In a random sample of 180 American men, 90 of them claimed to have supported gay marriage. In an independent random sample of 200 American women, 150 claim to support gay marriage. Assume we are interested in the proportion of men and women who support gay marriage. Which of the following represents a pooled estimate for the proportion of all Americans who support gay marriages under the null assumption of no difference?

a) 0.63
b) 0.45
c) 0.83
d) 0.33
e) 0.54

$$\hat{p} = \frac{90 + 150}{180 + 200} = 0.63$$

Multiple Choice # 21

Explain what the phrase "95% confident" means when we interpret a 95% CI for μ.

a) 95% of the observations in the population fall within the bounds of the calculated interval.

b) In repeated sampling, 95% of constructed intervals would contain the value of μ.

c) The probability that mean falls in the calculated interval is 0.95.

d) 95% of similarly constructed intervals would contain the value of the sampled mean.

Multiple Choice # 22

Psychologists are attempting to discern if children from urban schools perform worse on state standardized tests as compared to children from rural areas. They take two random samples of school children from two schools in a large city. They then administer a 50 point standardized test to every student. The summary statistics are shown below:

Urban		Rural
38	Mean	45
2.5	Standard Deviation	3.1
15	Sample Size	18

Assume that urban students are the 'first' group and rural students are the 'second' group. You calculate a confidence interval for the differences in population means to be (-9.2, -5.4). Which of the following is an appropriate conclusion that can be drawn from this interval?

a) The urban students score significantly higher on these tests than the rural students, on average.

b) The urban students score significantly lower on these tests than the rural students, on average.

c) There is no significant difference between the scores of the students, on average.

d) There is a significant difference between the scores of the students, but the direction of the difference cannot be determined.

e) None of the above is a true statement.

Multiple Choice # 23

An unbiased estimator of a population parameter is defined as:

a) an estimator whose expected value is equal to the parameter.

b) an estimator whose variance is equal to one.

c) an estimator whose expected value is equal to 0.

d) an estimator whose variance goes to zero as the sample size goes to infinity.

Multiple Choice # 24

A large study has determined that the diastolic blood pressure among women ages 18-74 is normally distributed with mean 70 mmHg and a standard deviation of σ=10 mmHg. Suppose that you measure the diastolic blood pressure in n=25 women. What is the point below which the mean diastolic blood pressures of these 25-women samples will be 5% of the time? That is find *x* so that $P(\bar{X} < x) = 0.05$

a) 56.08 mmHg

$\boxed{\text{b) 66.71 mmHg}}$

c) 50.40 mmHg

d) 27.55 mmHg

e) 40.21 mmHg

Reverse look-up, z=-1.645

$$-1.645 = \frac{\bar{x} - 70}{\frac{10}{\sqrt{25}}}, \bar{x} = 66.71$$

Multiple Choice # 25

A truck company wants on-time delivery for 98% of the parts they order from a metal manufacturing plant. They have been ordering from Hudson Manufacturing but will switch to a new, cheaper manufacturer (Steel-R-US) unless there is evidence that this new manufacturer cannot meet the 98% on-time goal. As a test the truck company purchases a random sample of metal parts from Steel-R-US, and then determines if these parts were delivered on-time. Which hypothesis should they test?

a) H_o: $p<0.98$ H_a: $p>0.98$

b) H_o: $p>0.98$ H_a: $p=0.98$

c) H_o: $p=0.98$ H_a: $p<0.98$

d) H_o: $p=0.98$ H_a: $p\neq0.98$

e) H_o: $p=0.98$ H_a: $p>0.98$

Multiple Choice # 26

Suppose that 58% of all gold dealers believe next year will be a good one to speculate in South African gold coins. In a simple random sample of 150 dealers, what is the probability that between 55% and 60% believe that it will be a good year to speculate?

a) 0.46
b) 0.58
c) 0.12
d) 0.31
e) 0.92

$$P(55\% < p < 60\%)$$

$$z = \frac{0.55-0.58}{\sqrt{\frac{0.58(1-0.58)}{150}}} = -0.74,\ z = \frac{0.60-0.58}{\sqrt{\frac{0.58(1-0.58)}{150}}} = 0.50$$

$$P(-0.74 < Z < 0.50) = 0.6915 - 0.2296 = 0.46$$

Multiple Choice # 27

The following output comes from an experiment comparing the levels of hemoglobin for a group of cigarette smokers of size n_1=75 and a group of non-smokers of size n_2=121. Researches want to test the claim that the mean hemoglobin level of smokers is higher that the mean level of non-smokers.

What are the null and alternative hypotheses for this test? μ_1 is the mean hemoglobin level among smokers, and μ_2 is the mean level among non-smokers.

a) Ho: $\mu_1 = \mu_2$; Ha: $\mu_1 \neq \mu_2$

b) Ho: $\mu_1 \neq \mu_2$; Ha: $\mu_1 = \mu_2$

c) Ho: $\mu_1 = \mu_2$; Ha: $\mu_1 > \mu_2$

d) Ho: $\mu_1 = \mu_2$; Ha: $\mu_1 < \mu_2$

Multiple Choice # 28

Based on the output to the previous questions, what is the decision?

```
Smokers: Number of obs =        75
Non-smokers: Number of obs =      121

------------------------------------------------------------------------------
Variable |      Mean    Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
 Smokers |       4.1    .2309401   17.7535   0.0000       3.639842    4.560158
Non-smok |       1.3    .1818182      7.15   0.0000       .9400127    1.659987
---------+--------------------------------------------------------------------
    diff |       2.8    .2939238   9.52628   0.0000       2.220304    3.379696
------------------------------------------------------------------------------

Ho: mean(smoker) - mean(non-smoker) = diff = 0

     Ha: diff < 0           Ha: diff != 0           Ha: diff > 0
       t =    9.5263          t =    9.5263          t =    9.5263
   P < t =    1.0000    P > |t| =    0.0000      P > t =    0.0000
```

a) Reject Ho; The hemoglobin levels are higher among non-smokers.

b) Do not reject Ho; The hemoglobin levels are equal.

c) Do not reject Ho; The hemoglobin levels are lower among non-smokers.

d) Reject Ho; The hemoglobin levels are higher among smokers.

Multiple Choice # 29

The mean cholesterol level of healthy 20-74 year-old males in the US is 211 mg/ml. The average cholesterol level among 16 *hypertensive* males was 232 mg/ml. Conduct a statistical test at the **10% level** about whether hypertension is associated with *higher* cholesterol levels (consult the STATA output below). That is we want to test the claim that the mean cholesterol level of hypertensive males is higher than 211.

```
Number of obs =       16
------------------------------------------------------------------------------
Variable |      Mean    Std. Err.        t      P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |       232          12   19.3333    0.0000       206.4226    257.5774
------------------------------------------------------------------------------
Degrees of freedom: 15
                            Ho: mean(x) = 211

      Ha: mean < 211          Ha: mean ~= 211          Ha: mean > 211
        t =    1.7500           t =    1.7500            t =    1.7500
    P < t =    0.9497     P > |t| =    0.1005        P > t =    0.0503
```

a) The *p* value of the test is 0.1005 and thus we reject the null hypothesis and thus conclude that hypertensives have a higher cholesterol level.

b) The *p* value of the test is 0.9496 and thus we fail to reject the null hypothesis and thus conclude that hypertensives have a lower cholesterol level.

c) The p value of the test is 0.0503 and thus we fail to reject the null hypothesis and conclude that hypertension is not associated with higher cholesterol levels

d) The p value of the test is 0.0503 and thus we reject the null hypothesis and conclude that the mean cholesterol level for hypertensives is higher than healthy 20-74 year-old US males.

Multiple Choice # 30

Suppose X, Y, and Z are independent random variables such that $X \sim N(\mu,1)$, $Y \sim N(2\mu,1)$, and $Z \sim N(10\mu,1)$. Which of the following is an unbiased estimator of μ from X, Y, and Z?

a) $(1/3)X+(1/3)Y+(1/3)Z$	$=1/3\mu+2/3\mu+10/3\mu= 4.33\mu$
b) $(1/1)X+(1/2)Y+(1/10)Z$	$=1/1\mu+2/2\mu+10/10\mu= 3\mu$
$\boxed{\text{c) } (1/3)X+(1/6)Y+(1/30)Z}$	$=1/3\mu+2/6\mu+10/30\mu= \mu$
d) $(1/2)X+(1/2)Y$	$=1/2\mu+2/2\mu = 1.5\mu$
e) $(1/3)X+(2/3)Y+(10/3)Z$	$=1/3\mu+4/3\mu+100/3\mu= 35\mu$

Multiple Choice # 31

Some have argued that throwing darts at the stock pages to decide which companies to invest in could be a successful stock-picking strategy. Suppose a researcher decides to test this theory and randomly chooses 150 companies to invest in. After one year, 81 of the companies were considered winners; that is, they outperformed other companies in the same investment class. To assess whether the dart-picking strategy resulted in a majority of winner, the researcher tested: Ho: π=0.5 versus Ha: π>0.5 and obtained a P=value of 0.1636. Write a conclusion for the researcher.

a) Because the P-value is not small, reject the null hypothesis. There is sufficient evidence to conclude that the dart-picking strategy results in a majority of winners.

b) Because the P-value is small, reject the null hypothesis. There is sufficient evidence to conclude that the dart-picking strategy resulted in a majority of winners.

c) Because the P-value is not small, do not reject the null hypothesis. There is not sufficient evidence to conclude that the dart-picking strategy resulted in a majority of winners.

d) Because the P-value is small, do not reject the null hypothesis. There is not sufficient evidence to conclude that the dart-picking strategy resulted in a majority of winners.

Multiple Choice # 32

A national organization has been working with utilities throughout the nation to find sites for large wind turbines that generate electricity. Wind speeds must average more than 13 miles per hour (mph) for a site to be acceptable. Recently, the organization conducted wind speed tests at a particular site. Based on a sample of 101 wind speed recordings (taken at random intervals), the wind speed at the site averaged 12.6 mph, with a standard deviation of 2.6 mph. To determine whether the site meets the organization's requirements, consider the test Ho: μ=13 vs. Ha: $\mu > 13$, where μ is the true mean wind speed at the site and α=0.05. What is the conclusion of the test?

a) There is sufficient evidence to conclude that the site meets the organization's requirements.

b) There is not sufficient evidence to conclude that the site meets the organization's requirements.

c) There is sufficient evidence to conclude that the average wind speed at the site exceeds 13 miles per hour.

d) None of the above.

$tstat = \frac{12.6-13}{2.6/\sqrt{101}} = -1.55 < 1.645$, so we fail to reject the null

Multiple Choice # 33

A company claims that lifetime of the lightbulb they manufacture is normally distributed with a mean of 240 hours and a standard deviation of 20 hours. An assistant at the company makes an improvement to the lightbulb and claims that the lifetime of the newer bulbs is normally distributed with a mean of 280 hours and a standard deviation of 20 hours. If we let X represent the lifetime of the lightbulb, this leads to the following null and alternative hypotheses.

- Ho: X is N (240; 20)
- Ha: X is N (280; 20)

Suppose the decision rule is that the null hypothesis will be rejected if the lifetime of the lightbulb is greater than 280 hours. Determine the probability of a Type I error.

a) 0.067
$\boxed{\text{b) } 0.023}$
c) 0.084
d) 0.075
e) 0.097

Type I error=P(reject the null | null is true)

$z = \frac{280-240}{20}$= 2, P(Z>2)= 1-0.9772= 0.023

Multiple Choice # 34

In conducting a chi-square test of association between gender and grade, what is the expected count for the number of males who earned a grade of B?

	Grade			
	A	B	C	D
Male	10	32	25	2
Female	5	41	14	12

a) (73*69)/141

b) (73*69)/140

c) (32*73)/141

d) (32*69)/110

e) (69*32)/141

Total # of B's: 32+41=73

Total # of males: 10+32+25+2 = 69

Total in table: 10+5+32+41+25+14+2+12= 141