Statistics for Management II-STAT 362-Final Review

Multiple Choice
Identify the letter of the choice that best completes the statement or answers the question.

1. The ability of an interval estimate to contain the value of the population parameter is described by the
   a. confidence level
   b. degrees of freedom
   c. precise value of the population mean μ
   d. degrees of freedom minus 1

Exhibit 8-2
A random sample of 121 automobiles traveling on an interstate showed an average speed of 65 mph. From past information, it is known that the standard deviation of the population is 22 mph.

2. Refer to Exhibit 8-2. If the confidence coefficient is reduced to 0.9, the standard error of the mean
   a. will increase
   b. will decrease
   c. remains unchanged
   d. becomes negative

3. When constructing a confidence interval for the population mean and the standard deviation of the sample is used, the degrees of freedom for the t distribution equals
   a. n-1
   b. n
   c. 29
   d. 30

4. The p-value
   a. is the same as the Z statistic
   b. measures the number of standard deviations from the mean
   c. is a distance
   d. is a probability

5. For a two tail test, the p-value is the probability of obtaining a value for the test statistic as
   a. likely as that provided by the sample
   b. unlikely as that provided by the sample
   c. likely as that provided by the population
   d. unlikely as that provided by the population

6. The probability of making a Type II error is denoted by
   a. α
   b. β
   c. 1 - α
   d. 1 - β
7. In order to test the following hypotheses at an $\alpha$ level of significance

\[ H_0: \mu \leq 100 \]
\[ H_a: \mu > 100 \]

the null hypothesis will be rejected if the test statistic $Z$ is

a. $\geq Z_\alpha$

b. $\leq Z_\alpha$

c. $< -Z_\alpha$

d. $< 100$

8. Refer to Exhibit 10-4. The degrees of freedom for the t-distribution are

a. 22

b. 21

c. 20

d. 19

9. Refer to Exhibit 10-5. If the null hypothesis is tested at the 5% level, the null hypothesis

a. should be rejected

b. should not be rejected

c. should be revised

d. None of these alternatives is correct.

10. The value of $F_{0.05}$ with 8 numerator and 19 denominator degrees of freedom is

a. 2.48

b. 2.58

c. 3.63

d. 2.96
11. The symbol used for the variance of the population is
   a. $\sigma$
   b. $\sigma^2$
   c. S
   d. $S^2$

12. The symbol used for the variance of the sample is
   a. $\sigma$
   b. $\sigma^2$
   c. S
   d. $S^2$

13. For an F distribution, the number of degrees of freedom for the numerator
   a. must be larger than the number of degrees for the denominator
   b. must be smaller than the number of degrees of freedom for the denominator
   c. must be equal to the number of degrees of freedom for the denominator
   d. can be larger, smaller, or equal to the number of degrees of freedom for the denominator

14. The sampling distribution of the quantity $(n-1)s^2/\sigma^2$ is the
   a. chi-square distribution
   b. normal distribution
   c. F distribution
   d. t distribution

Exhibit 11-5

$n = 14 \quad s = 20 \quad H_0: \sigma^2 \leq 500 \quad H_a: \sigma^2 > 500$

15. Refer to Exhibit 11-5. The test statistic for this problem equals
   a. .63
   b. 10.4
   c. 1.04
   d. 0.52

16. Refer to Exhibit 11-5. The null hypothesis is to be tested at the 5% level of significance. The critical value(s) from the table is(are)
   a. 22.362
   b. 23.685
   c. 5.009 and 24.736
   d. 5.629 and 26.119
Exhibit 13-3
To test whether or not there is a difference between treatments A, B, and C, a sample of 12 observations has been randomly assigned to the 3 treatments. You are given the results below. Assume that $\alpha = 0.05$.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20 30 25 33</td>
</tr>
<tr>
<td>B</td>
<td>22 26 20 28</td>
</tr>
<tr>
<td>C</td>
<td>40 30 28 22</td>
</tr>
</tbody>
</table>

17. Refer to Exhibit 13-3. The null hypothesis
   a. should be rejected
   b. should not be rejected
   c. should be revised
   d. None of these alternatives is correct.

18. In a simple regression analysis (where Y is a dependent and X an independent variable), if the Y intercept is positive, then
   a. there is a positive correlation between X and Y
   b. if X is increased, Y must also increase
   c. if Y is increased, X must also increase
   d. None of these alternatives is correct.

19. In regression analysis, the variable that is being predicted is the
   a. dependent variable
   b. independent variable
   c. intervening variable
   d. is usually x

20. In simple linear regression analysis, which of the following is not true?
   a. The F test and the t test yield the same results.
   b. The F test and the t test may or may not yield the same results.
   c. The relationship between X and Y is represented by means of a straight line.
   d. The value of $F = t^2$.

Exhibit 14-4
Regression analysis was applied between sales data (in $1,000s) and advertising data (in $100s) and the following information was obtained.

$\hat{Y} = 12 + 1.8x$

n = 17
SSR = 225
SSE = 75
$S_{b1} = 0.2683$

21. Refer to Exhibit 14-4. Based on the above estimated regression equation, if advertising is $3,000, then the point estimate for sales (in dollars) is
   a. $66,000
   b. $5,412
   c. $66
   d. $17,400
Problem

22. Below you are given ages that were obtained by taking a random sample of 9 undergraduate students. Assume the population has a normal distribution.

19 22 23 19 21 22 19 23 21

a. What is the point estimate of $\mu$?
b. Construct a 99% confidence interval for the average age of undergraduate students.
c. Construct a 98% confidence interval for the average age of undergraduate students.
d. Discuss why the 98% and 99% confidence intervals are different.

23. In order to determine the average weight of carry-on luggage by passengers in airplanes, a sample of 16 pieces of carry-on luggage was weighed. The average weight was 20 pounds. Assume that we know the standard deviation of the population to be 8 pounds.

a. Determine a 97% confidence interval estimate for the mean weight of the carry-on luggage.
b. Determine a 95% confidence interval estimate for the mean weight of the carry-on luggage.

24. Consider the following hypothesis test:

$$H_0: \mu \geq 14$$
$$H_a: \mu < 14$$

A sample of 64 provides a sample mean of 13 and a sample standard deviation of 4.

a. Determine the standard error of the mean.
b. Compute the value of the test statistic.
c. Determine the $p$-value; and at 95% confidence, test the above hypotheses.

25. You are given the following information obtained from a random sample of 5 observations. Assume the population has a normal distribution.

20 18 17 22 18

You want to determine whether or not the mean of the population from which this sample was taken is significantly less than 21.

a. State the null and the alternative hypotheses.
b. Compute the standard error of the mean.
c. Determine the test statistic.
d. Determine the $p$-value and at 90% confidence, test whether or not the mean of the population is significantly less than 21.
26. Consider the following hypothesis test:

\[ \mu_1 - \mu_2 \leq 0 \]
\[ \mu_1 - \mu_2 > 0 \]

The following results are for two independent samples taken from two populations.

<table>
<thead>
<tr>
<th></th>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>35</td>
<td>37</td>
</tr>
<tr>
<td>Sample Mean</td>
<td>43</td>
<td>37</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>140</td>
<td>170</td>
</tr>
</tbody>
</table>

a. Determine the degrees of freedom for the t distribution.
b. Compute the test statistic.
c. What is your decision based on \( \alpha = 0.05 \)?

27. The daily production rates for a sample of factory workers before and after a training program are shown below. Let \( d = \text{After} - \text{Before} \).

<table>
<thead>
<tr>
<th>Worker</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>9</td>
</tr>
</tbody>
</table>

We want to determine if the training program was effective.

a. Give the hypotheses for this problem.
b. Compute the test statistic.
c. At 95% confidence, test the hypotheses. That is, did the training program actually increase the production rates?

28. We are interested in determining whether or not the variances of the starting salaries of accounting majors is significantly different from management majors. The following information was gathered from two samples.

<table>
<thead>
<tr>
<th></th>
<th>Accounting</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>Average Monthly Income</td>
<td>$3,600</td>
<td>$3,500</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>$900</td>
<td>$400</td>
</tr>
</tbody>
</table>

At 90% confidence, test to determine whether or not the variances are equal.
29. You are given the following data regarding a sample.

\[
\begin{array}{c}
X \\
16 \\
12 \\
21 \\
10 \\
13 \\
18 \\
\end{array}
\]

a. Compute the mean and the variance.
b. Using the critical value approach, test to determine if the variance of the population from which this sample is taken is significantly more than 12. Let \( \alpha = 0.05 \).

30. A major automobile manufacturer claimed that the frequencies of repairs on all five models of its cars are the same. A sample of 200 repair services showed the following frequencies on the various makes of cars.

<table>
<thead>
<tr>
<th>Model of Car</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>32</td>
</tr>
<tr>
<td>B</td>
<td>45</td>
</tr>
<tr>
<td>C</td>
<td>43</td>
</tr>
<tr>
<td>D</td>
<td>34</td>
</tr>
<tr>
<td>E</td>
<td>46</td>
</tr>
</tbody>
</table>

At \( \alpha = 0.05 \), test the manufacturer's claim.

31. The following table shows the results of recent study regarding gender of individuals and their selected field of study.

<table>
<thead>
<tr>
<th>Field of study</th>
<th>Male</th>
<th>Female</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>80</td>
<td>40</td>
<td>120</td>
</tr>
<tr>
<td>Business</td>
<td>60</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>Engineering</td>
<td>160</td>
<td>40</td>
<td>200</td>
</tr>
<tr>
<td>TOTAL</td>
<td>300</td>
<td>100</td>
<td>400</td>
</tr>
</tbody>
</table>

We want to determine if the selected field of study is independent of gender.

a. Compute the test statistic.
b. Using the \( p \)-value approach at 90% confidence, test to see if the field of study is independent of gender.
c. Using the critical method approach at 90% confidence, test for the independence of major and gender.
32. The marketing department of a company has designed three different boxes for its product. It wants to determine which box will produce the largest amount of sales. Each box will be test marketed in five different stores for a period of a month. Below you are given the information on sales.

<table>
<thead>
<tr>
<th></th>
<th>Store 1</th>
<th>Store 2</th>
<th>Store 3</th>
<th>Store 4</th>
<th>Store 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box 1</td>
<td>210</td>
<td>230</td>
<td>190</td>
<td>180</td>
<td>190</td>
</tr>
<tr>
<td>Box 2</td>
<td>195</td>
<td>170</td>
<td>200</td>
<td>190</td>
<td>193</td>
</tr>
<tr>
<td>Box 3</td>
<td>295</td>
<td>275</td>
<td>290</td>
<td>275</td>
<td>265</td>
</tr>
</tbody>
</table>

a. State the null and alternative hypotheses.
b. Construct an ANOVA table.
c. What conclusion do you draw?
d. Use Fisher's LSD procedure and determine which mean (if any) is different from the others. Let \( \alpha = 0.01 \).
1. A
2. C
3. A
4. D
5. B
6. B
7. A
8. C \[ n_1 + n_2 - 2 = 10 + 12 - 2 = 20 \]
9. B
10. A
11. B
12. D
13. D
14. A
15. B
16. A
17. B
18. D
19. A
20. B
21. A
### Matched Samples

<table>
<thead>
<tr>
<th>Method 1</th>
<th>Method 2</th>
<th>d</th>
<th>( (d-d̄)^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>5</td>
<td>+2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>-4</td>
<td>-3</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>-2</td>
<td>-1</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>-1</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
\overline{d} = \frac{\sum d}{n} = \frac{-5}{5} = -1
\]

- \( H_0: \mu_d = 0 \) (No difference)
- \( H_a: \mu_d \neq 0 \) (Significant difference exists)

\[
S_d = \sqrt{\frac{\sum (d-d̄)^2}{n-1}} = \sqrt{\frac{20}{4}} = \sqrt{5} = 2.236
\]

\[
t = \frac{\overline{d} - \mu_d}{S_d / \sqrt{n}} = \frac{-1 - 0}{2.236 / \sqrt{4}} = -1.000
\]

\[
\alpha = 0.05 \quad \alpha / 2 = 0.025
\]

- \( d.f. = n - 1 = 4 \)
- Critical \( t = \pm 2.776 \)

\[
t < t_{\alpha / 2} \text{, do not reject } H_0.
\]
15. \( n = 14 \quad S_x = 20 \quad S_{x}^{2} = 400 \)

\( H_0: \sigma^2 \leq 500 \)

\( H_a: \sigma^2 > 500 \)

Use Chi-Square Test

\[ \chi^2 = \frac{(n-1)S_x^2}{\sigma_x^2} = \frac{(14-1)(400)}{500} = 10.4 \]

16. \( \alpha = 0.05 \quad d.f. = n - 1 = 13 \)

Since \( \chi^2 < \chi^2_{critical} \), do not reject \( H_0 \).
**ANOVA - Test of 3 Different Treatment Means**

17. \( H_0: \mu_A = \mu_B = \mu_C \)  (No differences in means)

\( H_A: \mu_A \neq \mu_B \neq \mu_C \)  (A significant difference exists between the groups)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments (between groups)</td>
<td>72</td>
<td>( k-1=2 )</td>
<td>36</td>
<td>1.059</td>
</tr>
<tr>
<td>Error (within groups)</td>
<td>306</td>
<td>( N_t-k=12-3=9 )</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( N_t-1=11 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( \alpha = 0.05 \)

\( F = 1.059 \)

\( F_{0.05} = 3.98 \)  (2/11 d.f.)

\( F < F_{critical} \), Do not reject \( H_0 \).

No significant differences found between the three groups.
21. **Regression Model**

\[ x = \$ \text{ in advertising spent (100s)} \]
\[ y = \text{predicted sales dollars (1000s)} \]

\[ \hat{y} = 12 + 1.8x \]
\[ \hat{y}(30) = 12 + 1.8(30) = \$66 \text{ (thousand)} \]

\$66,000 predicted in sales dollars.

22. 19 22 23 19 21 22 19 23 21

A. **Point estimate of \( \mu_x \) is**

\[ \bar{x} = \frac{\sum x}{n} = \frac{189}{9} \]
\[ \bar{x} = 21 \]
\[ s_x = 1.658 \]

B. 99% C.I. \( \sigma_x \) is unknown d.f. = n-1 = 8

Critical \( t = 3.355 \) \( \alpha = 0.01 \)

\[ \bar{x} \pm t \frac{s_x}{\sqrt{n}} \]
\[ 21 \pm 3.355 \left( \frac{1.658}{\sqrt{9}} \right) \]
\[ 21 \pm 1.854 \]
\[ 19.15 \quad 22.85 \]

99% C.I.

C. 98% C.I. \( t = 2.896 \) 8 d.f. \( \alpha = 0.02 \)

\[ \bar{x} \pm t \left( \frac{s_x}{\sqrt{n}} \right) \]
\[ 21 \pm 2.896 \left( \frac{1.658}{\sqrt{9}} \right) \]
\[ 21 \pm 1.601 \]
\[ 19.399 \quad 22.601 \]
23. n=16, \ \bar{X} = 20, \ \sigma_x = 8

A. 97th C.I., \ \alpha = 0.03, \ \alpha/2 = 0.0125
\[ z = 2.17 \]
\[
\bar{X} \pm z \sigma_x / \sqrt{n}
\]
\[
20 \pm (2.17)(\frac{8}{\sqrt{16}})
\]
20 \pm 4.34
15.66 \hspace{1cm} 24.34

B. 95th C.I., \ \alpha = 0.05, \ \alpha/2 = 0.025
\[ z = 1.96 \]
\[
\bar{X} \pm z (\sigma_x / \sqrt{n})
\]
\[
20 \pm (1.96)(\frac{8}{\sqrt{16}})
\]
20 \pm 3.92
16.08 \hspace{1cm} 23.92
24. $H_0: \mu_x \geq 14 \quad n=64$

$H_A: \mu_x < 14 \quad \bar{x}=13 \quad s_x=4$

A. $s_x = \frac{4}{\sqrt{64}} = 0.5 \quad (\text{STANDARD ERROR})$

B. $\sigma_x$ is unknown, use $t$

$$t = \frac{\bar{x} - \mu_x}{s_x/\sqrt{n}} = \frac{13-14}{0.5} = -2.00$$

C. $\alpha=0.05$

At $t=-2.00$ $p$-value is between 0.01 and 0.025, so reject $H_0$.

At $60 \text{ d.f.}$ $t=-1.671$ (critical value)

At $60 \text{ d.f.}$ $t=-2.00$ which has probability of at most 0.025, which is less than $\alpha=0.05$.
25. \( n = 5 \)  NORMAL DISTRIBUTION ASSUMED

\[ 20 \quad 18 \quad 17 \quad 22 \quad 18 \]

\[ \bar{x} = \frac{\sum x}{n} = \frac{95}{5} = 19 \]

\[ s_x = 2 \]

A. \( H_0 : \mu_x \geq 21 \)
   \( H_A : \mu_x < 21 \)

B. \[ \frac{s_x}{\sqrt{n}} = \frac{2}{\sqrt{5}} = 0.89443 \]

C. \[ t = \frac{x - \mu_x}{s_x/\sqrt{n}} = \frac{19 - 21}{2/\sqrt{5}} = \frac{-2}{2/\sqrt{5}} = -2.2361 \]

D. 

\[ \alpha = 0.10 \]

\[ d.f. = n - 1 = 4 \]

\[ -1.533 \quad 0 \quad 1 \]

\[ t \]

CRITICAL

AT 4 d.f. -2.2361 IS BETWEEN 0.05 AND 0.025 PROBABILITIES, SO REJECT \( H_0 \).
THERE IS SIGNIFICANT EVIDENCE THE POPULATION MEAN IS LESS THAN 21.
INDEPENDENT POPULATION MEANS TEST

26. (EQUA\ VARIANCES NOT ASSUMED)

H₀: μ₁ - μ₂ ≤ 0
Hₐ: μ₁ - μ₂ > 0

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>35</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Mean</td>
<td>43</td>
<td>37</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>140</td>
<td>170</td>
</tr>
</tbody>
</table>

A. d.f. = \[ \left( \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{} \right)^2 \left( \frac{1}{n_1 - 1} \right) \left( \frac{1}{n_2 - 1} \right) \]

\[ \begin{align*}
&\text{d.f.} = \left( \frac{140}{35} + \frac{170}{37} \right)^2 \\
&\phantom{=} \left( \frac{1}{35 - 1} \right) \left( \frac{1}{37 - 1} \right) \\
&\phantom{=} \left( \frac{140}{35} \right)^2 + \left( \frac{170}{37} \right)^2 \\
&\phantom{=} \left( \frac{0.470588}{35} \right) + \left( \frac{0.580397}{37} \right) \\
&\phantom{=} 73.86705625 = 69.8846 \approx 69 \text{ d.f.}
\end{align*} \]

\[ \frac{1,056,985,458}{1} \]

B. \[ t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(43 - 37) - 0}{\sqrt{\frac{140}{35} + \frac{170}{37}}} \]

\[ t = 2.0466 \]

C. P-VALUE IS 0.0222 WHICH IS LESS THAN α = 0.05. REJECT H₀.
\[ d = \text{AFTER} - \text{BEFORE} \]

<table>
<thead>
<tr>
<th>( \text{BEFORE} )</th>
<th>( \text{AFTER} )</th>
<th>( d )</th>
<th>( d - \bar{d} )</th>
<th>((d - \bar{d})^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>9</td>
<td>3</td>
<td>-0.8</td>
<td>0.64</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>2</td>
<td>-0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>-1</td>
<td>-1.2</td>
<td>1.44</td>
</tr>
<tr>
<td>8</td>
<td>11</td>
<td>3</td>
<td>-0.8</td>
<td>0.64</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>-2</td>
<td>-0.2</td>
<td>0.04</td>
</tr>
</tbody>
</table>

\[ \bar{d} = \frac{\sum d}{n} = \frac{-11}{5} = -2.2 \]

\[ s_d = \sqrt{\frac{\sum(d - \bar{d})^2}{n-1}} = \sqrt{\frac{2.8}{4}} = 0.836666 \]

\( H_0: \mu_{\text{AFTER}} \leq \mu_{\text{BEFORE}} \) \text{ (NO IMPROVEMENT)}

\( H_a: \mu_{\text{AFTER}} > \mu_{\text{BEFORE}} \) \text{ (AFTER SCORE IS HIGHER)}

or

\( H_0: \mu_d \leq 0 \)

\( H_a: \mu_d > 0 \)

\[ B. \ t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{2.2 - 0}{0.836666 / \sqrt{5}} = \frac{2.2}{0.37417} = 5.88 \]

\( C. \ \alpha = 0.05 \)

\[ t > t_{\text{critical}}, \ \text{REJECT} \ H_0. \ \text{P-VALUE IS LESS THAN} \ \alpha = 0.05 \]
28. \( H_0: \sigma_1^2 = \sigma_2^2 \) (VARIANCES ARE EQUAL)
   \( H_A: \sigma_1^2 \neq \sigma_2^2 \) (VARIANCES ARE NOT EQUAL)

<table>
<thead>
<tr>
<th></th>
<th>ACCT</th>
<th>MGMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>Average Income</td>
<td>3600</td>
<td>3500</td>
</tr>
<tr>
<td>Standard Variance</td>
<td>900</td>
<td>400</td>
</tr>
</tbody>
</table>

\( \alpha = 0.10 \)

Use F-test of Two Population Variances

\[
F = \frac{s_1^2}{s_2^2} = \frac{900}{400} = 2.25
\]

\( F_{critical} (\alpha/2 = 0.05) \)

Two-tail test

\( F_{critical} = 2.23 \)

Our sample is greater than our critical value, reject \( H_0 \).
29. \[
\begin{array}{ccc}
X & X - \bar{X} & (X - \bar{X})^2 \\
16 & 1 & 1 \\
12 & -3 & 9 \\
21 & 6 & 36 \\
10 & -5 & 25 \\
13 & -2 & 4 \\
18 & 3 & 9 \\
\hline
90 & & 84
\end{array}
\]

A. \[
\bar{X} = \frac{\sum X}{n} = \frac{90}{6} = 15
\]
\[
S_X = \sqrt{\frac{8 \sum (X - \bar{X})^2}{n-1}} = \sqrt{\frac{84}{5}} = 4.09878
\]
\[
\text{VARIANCE} = (S_X)^2 = (4.09878)^2 = 16.8
\]

B. \(H_0: \sigma^2 \leq 12\) \(\text{(NOT SIGNIFICANTLY MORE THAN 12)}\)
\(H_a: \sigma^2 > 12\) \(\text{(SIGNIFICANTLY MORE THAN 12)}\)
\(\alpha = 0.05\)

Use \(\chi^2\) test
\[
\chi^2 = \frac{(n-1)S_X^2}{\sigma_0^2} = \frac{(6-1)(16.8)}{12} = 7
\]
\(x = 0.05\) d.f. = 5
\[
\chi^2_{\text{critical}} = 11.070
\]

Sample \(\chi^2\) is NOT GREATER THAN CRITICAL \(\chi^2\) of 11.070. DO NOT REJECT \(H_0\).
### $\chi^2$-Goodness of Fit Test

<table>
<thead>
<tr>
<th>Model</th>
<th>$f_o$</th>
<th>Expected</th>
<th>$f_o - f_e$</th>
<th>$(f_o - f_e)^2$</th>
<th>$\frac{(f_o - f_e)^2}{f_e}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>32</td>
<td>40</td>
<td>-8</td>
<td>64</td>
<td>1.6</td>
</tr>
<tr>
<td>B</td>
<td>45</td>
<td>40</td>
<td>5</td>
<td>25</td>
<td>0.625</td>
</tr>
<tr>
<td>C</td>
<td>43</td>
<td>40</td>
<td>3</td>
<td>9</td>
<td>0.225</td>
</tr>
<tr>
<td>D</td>
<td>34</td>
<td>40</td>
<td>-6</td>
<td>36</td>
<td>0.90</td>
</tr>
<tr>
<td>E</td>
<td>46</td>
<td>40</td>
<td>6</td>
<td>36</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>200</strong></td>
<td><strong>200</strong></td>
<td></td>
<td></td>
<td><strong>4.25</strong></td>
</tr>
</tbody>
</table>

**Expected Frequency of Five Models is** $\frac{200}{5} = 40$

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 4.25$$

$\alpha = 0.05$

- **$H_0$: No difference between observed and expected values**
- **$H_a$: There is a difference**

d.f. = c - 1 = 5 - 1 = 4

Critical $\chi^2 = 9.488$. Since our sample $\chi^2$ is less, there is no evidence that observed data is different from expected values.
**χ²-Test of Independence**

31. $H_0$: Field of Study and Gender are independent  
   $H_a$: They are NOT independent

<table>
<thead>
<tr>
<th>Field of Study</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>50</td>
<td>40</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>(90)</td>
<td>(30)</td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td>60</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>(100)</td>
<td>(20)</td>
<td></td>
</tr>
<tr>
<td>Engineer</td>
<td>160</td>
<td>40</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>(150)</td>
<td>(50)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>300</td>
<td>100</td>
<td>400</td>
</tr>
</tbody>
</table>

Expected values are in parentheses

A, B.  
\[
\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} 
\]

\[
\chi^2 = \frac{(50-90)^2}{90} + \frac{(40-30)^2}{30} + \frac{(60-60)^2}{60} + \frac{(20-20)^2}{20} + \frac{(160-150)^2}{150} + \frac{(40-50)^2}{50} = 7.111 
\]

$$\chi^2 = 7.111$$

\[
d.f. = (r-1)(c-1) = (3-1)(2-1) = 2 \text{ d.f.} 
\]

C.  
Reject $H_0$ if $\chi^2 > 4.605$

Since our $\chi^2(7.111)$ is greater than the critical value (4.605) and since p-value of 7.111 is between 0.023 and 0.01, reject $H_0$.  

Field of Study and Gender are NOT independent of each other.
### ANOVA

<table>
<thead>
<tr>
<th>Store</th>
<th>Box 1</th>
<th>Box 2</th>
<th>Box 3</th>
<th>Store Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>210</td>
<td>195</td>
<td>295</td>
<td>( \bar{X}_1 = 233.33 )</td>
</tr>
<tr>
<td>2</td>
<td>230</td>
<td>170</td>
<td>275</td>
<td>( \bar{X}_2 = 225.0 )</td>
</tr>
<tr>
<td>3</td>
<td>190</td>
<td>200</td>
<td>290</td>
<td>( \bar{X}_3 = 226.67 )</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>190</td>
<td>275</td>
<td>( \bar{X}_4 = 215 )</td>
</tr>
<tr>
<td>5</td>
<td>190</td>
<td>193</td>
<td>265</td>
<td>( \bar{X}_5 = 216 )</td>
</tr>
</tbody>
</table>

**Box 1**

- \( \bar{X}_1 = 200 \)
- \( S_{X_1} = 20 \)
- \( S_{X_1}^2 = 400 \)
- \( n_1 = 5 \)

**Box 2**

- \( \bar{X}_2 = 189.6 \)
- \( S_{X_2} = 11.546 \)
- \( S_{X_2}^2 = 133.3 \)
- \( n_2 = 5 \)

**Box 3**

- \( \bar{X}_3 = 280 \)
- \( S_{X_3} = 12.2474 \)
- \( S_{X_3}^2 = 149.9999 \)
- \( n_3 = 5 \)

**Grand Mean**

\[ \bar{X} = 223.2 \]

A. \( H_0: \mu_1 = \mu_2 = \mu_3 \) (No difference between group means)

**HA**: At least one mean is different with blocking (5 stores)

B. \( SSTR = 5(200 - 223.2)^2 + 5(189.6 - 223.2)^2 + 5(280 - 223.2)^2 = 24467.2 \)

With \( k = 3 \) box types

\( SSBCL = 3(233.33 - 223.2)^2 + 3(225 - 223.2)^2 + 3(226.67 - 223.2)^2 + 3(215 - 223.2)^2 + 3(216 - 223.2)^2 = 711.07 \)

\( SST = (210 - 223.2)^2 + (195 - 223.2)^2 + (295 - 223.2)^2 + (230 - 223.2)^2 + (170 - 223.2)^2 + (275 - 223.2)^2 + (190 - 223.2)^2 + (200 - 223.2)^2 + (290 - 223.2)^2 + (180 - 223.2)^2 + (190 - 223.2)^2 + (275 - 223.2)^2 + (190 - 223.2)^2 + (193 - 223.2)^2 + (265 - 223.2)^2 \)

\( SST = 27400.41 \)
\[ SSE = SST - SSTR - SSBL \]
\[ = 27400.41 - 24667.20 - 711.07 \]
\[ = 2022.14 \]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>24667.20</td>
<td>k-1=2</td>
<td>12333.6</td>
<td>48.794</td>
</tr>
<tr>
<td>Blocks</td>
<td>711.07</td>
<td>b-1=4</td>
<td>177.77</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>2022.14</td>
<td>(k-1)(b-1)=8</td>
<td>252.77</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27400.41</td>
<td>n_t-1=14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ F = \frac{MSTR}{MS_E} = \frac{12333.6}{252.77} = 48.794 \]

C.

2 d.f. / 8 d.f. Critical F = 8.65
\[ \alpha = 0.01 \]

Reject \( H_0 \). Our sample \( F \) is significant at least at the 0.01 level. At least one mean is different from the others.

D.

\[ LSD = 2.681 \sqrt{\frac{252.77}{5}} = 26.958 \]

Box 3 mean is different
\[ \bar{x}_3 - \bar{x}_2 = 90.4 \sqrt{\frac{1}{5} + \frac{1}{5}} = 26.958 \]

Greater than \[ \bar{x}_3 - \bar{x}_1 = 80 \sqrt{\frac{1}{5} + \frac{1}{5}} = 26.958 \]