

**Sample Exam 4 - STAT 303 Session 201**  
**Summer 2009**

**Name:**

**UIN:**

**Signature:**

1. Do not open this test until told to do so.
2. Turn in your exam with your answers circled when you are done with the exam. You should not take the exam with you.
3. This is a closed book examination. You may use one both-sided sheet of formulas that you have brought with you. You should have no other printed or written material with you on the exam.
4. You have 60 minutes to work on this exam. There are 16 multiple choice questions and 4 work out questions, each worth 5 points.
5. You may use a calculator but not a phone during the exam.
6. If you are unsure of what a question is asking for, do not hesitate to ask the instructor or course assistant for clarification.
7. Do not sit directly next to another student.
8. Good Luck!!!

1. A researcher is interested in the tensile strength of a synthetic fiber used to make cloth for men's shirts. It is suspected that the strength is affected by the percentage of cotton in the fiber. Five levels of cotton percentage are considered, and five observations are taken at each level. The 25 experiments are run in random order. The results are as follows.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	475.76	118.940	14.76	<.0001
Error	20	161.20	8.060		
Total	24	636.96			

Which is an appropriate conclusion? (TWO ANSWER)

- (a) **P-value in the ANOVA is quite small. Therefore, we reject the null and conclude that the tensile strength is affected by the percentage of cotton in the fiber.**
  - (b) P-value in the ANOVA is quite large. Therefore, we fail to reject the null and conclude that the tensile strength is not affected by the percentage of cotton in the fiber.
  - (c) P-value in the ANOVA is quite small. Therefore, we fail to reject the null and conclude that the tensile strength is not affected by the percentage of cotton in the fiber.
  - (d) P-value in the ANOVA is quite large. Therefore, we reject the null and conclude that the tensile strength is affected by the percentage of cotton in the fiber.
  - (e) **Checking assumption for ANOVA is the first step. Before analyzing ANOVA, we need to check the assumption.**
2. A civil engineer wishes to compare the strength properties of three different types of beams. Type A is made of steel, while types B and C are made of two different and more expensive alloys. The engineer measures the strength of a beam by applying 3000 pounds at the center, and then measuring the deflection. The data follows normal distribution but each group does not have equal variance. Test result are given as follows:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	627.083	313.541	12.45	0.0003
Error	21	528.750	25.178		
Total	23	1155.833			

Which is an appropriate conclusion?

- (a) P-value in the ANOVA is quite small. Therefore, we reject the null and conclude that there exists a difference in strength properties for three types of beams.
- (b) P-value in the ANOVA is quite large. Therefore, we do not reject the null and conclude that there is not a difference in strength properties for three types of beams.

- (c) **Since an equal variance assumption is violated, we have to be careful in using ANOVA. It is recommendable to try another approach such as transformation of data.**
- (d) Since a normality assumption is violated, we have to be careful in using ANOVA. It is recommendable to try another approach such as transformation of data.
3. An experiment was conducted to study the relationship between baking temperature  $x$  (in units of 10 Fahrenheit) and yield  $y$  (as a percentage) of a popular cake mix. Regression estimation results are as follows,

Parameter Estimates

Parameter	Estimate	Std. Err.	DF	T-Stat	P-Value
Intercept	0.988	0.858	12	1.151	0.2719
Slope	1.994	0.045	12	44.097	<0.0001

Analysis of Variance Table

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	2306.416	2306.416	1944.626	< 0.0001
Error	12	14.233	1.186		
Total	13	2320.649			

Predicted values

X value	Pred. Y	s.e.(Pred. y)	95% C.I.	95% P.I.
27	54.844	0.505	(53.74, 55.94)	(52.22, 57.45)

Which is incorrect?

- (a) P-value for testing the alternative that an slope is near zero is so large that we reject the null.
- (b) 95% Confidence Interval for expected value at  $X = 27$  is (53.74, 55.94).
- (c) 95% Prediction Interval for observation  $Y$  at  $X = 27$  is (52.22, 57.45).
- (d) **Since P-value in ANOVA table for regression is so small that we can conclude that intercept is not zero.**
4. Which is an appropriate conclusion of above question?
- (a) P-value for testing slope is quite large so that we do not reject the null. Hence we conclude that there is not a linear relationship between temperature and yield.
- (b) P-value for testing intercept is quite large so that we do not reject the null. Hence we conclude that there is not a linear relationship between temperature and yield.

(c) **P-value for testing slope is quite small so that we reject the null. Hence we conclude that there is a linear relationship between temperature and yield.**

(d) P-value is for testing slope is quite small so that we reject the null. Hence we conclude that there is not a linear relationship between temperature and yield.

5. The data below show Olympic triple jump distances for men in meters for the years 1896 to 1992 (there were no Olympic games in 1916, 1940, 1944).

year	Distance	year	Distance	year	Distance
1896	13.71	1932	15.72	1968	17.39
1900	14.47	1936	16.00	1972	17.35
1908	14.92	1948	15.40	1976	17.29
1912	14.64	1952	16.22	1980	17.35
1920	14.50	1956	16.35	1984	17.25
1924	15.53	1960	16.81	1988	17.61
1928	15.21	1964	16.85	1992	18.17

Using the data, an a researcher tried to fit a simple linear regression between year and jump distance (response variable) as follows;

Parameter	Estimate	Std. Err.	DF	T-Stat	P-Value
Intercept	-62.312	4.664	19	-13.36	<0.0001
Slope	0.040	0.002	19	16.82	<0.0001

What is your conclusion on the research hypothesis that there is not a linear relationship between them?

(a) P-value is quite large so that we do not reject the null. Hence we conclude that there is not a linear relationship between year and jump distance.

(b) P-value is quite large so that we reject the null. Hence we conclude that there is a linear relationship between year and jump distance.

(c) **P-value is quite small so that we reject the null. Hence we conclude that there is a linear relationship between year and jump distance.**

(d) P-value is quite small so that we reject the null. Hence we conclude that there is not a linear relationship between year and jump distance.

(e)

6. A traffic engineering study on traffic delay was conducted at intersection with signals on urban streets. Three types of traffic signals were utilized in the study: (1) pretimed, (2) semi-actuated, and (3) fully

actuated. Five intersections were used for each type of signal. The measure of traffic delay used in the study was the average stopped time per vehicle at each of the intersections (seconds/vehicle). To see if there is any difference in traffic delay for three types of traffic signals, a researcher obtained ANOVA table and tests associated with it. The data is normally distributed and the variance of each group is similar.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1131.327	565.664	45.74	< 0.0001
Error	11	136.028	12.366		
Total	13	1267.355			

What is a null and an alternative hypothesis of ANOVA table?

- (a)  $H_0: \mu_1 = \mu_2 = \mu_3, H_1: \mu_1 \neq \mu_2 \neq \mu_3$
- (b)  $H_0: \mu_1 = \mu_2 = \mu_3, H_1: \mu_1 \neq \mu_2$  **or**  $\mu_1 \neq \mu_3$  **or**  $\mu_2 \neq \mu_3$
- (c)  $H_0: \mu_1 = \mu_2 = \mu_3, H_1: \mu_1 \neq \mu_2$  and  $\mu_1 \neq \mu_3$  and  $\mu_2 \neq \mu_3$
- (d)  $H_0: \mu_1 \neq \mu_2 \neq \mu_3, H_1: \mu_1 = \mu_2 = \mu_3$

7. What is your conclusion of above question?

- (a) **P-value in the ANOVA is quite small. Therefore, we reject the null and conclude that there exists a difference in traffic delay for three types of signal.**
- (b) P-value in the ANOVA is quite large. Therefore, we do not reject the null and conclude that there is not a difference in traffic delay for three types of signal.
- (c) Since an equal variance assumption is violated, we have to be careful in using ANOVA. It is recommendable to try another approach such as transformation of data.
- (d) Since a normality assumption is violated, we have to be careful in using ANOVA. It is recommendable to try another approach such as transformation of data.