

University of Alberta

Faculty of Science

Department of Mathematical Sciences

Stat 151 SECTION P01
FINAL Winter 2010

Instructor: Beda Dessens

HAND-IN ENTIRE BOOKLET

DO NOT DETACH PAGES

For Multiple choice CHOOSE ONLY ONE ANSWER FOR EACH QUESTION

CLEARLY INDICATE YOUR CHOICE

CHOOSE THE BEST CHOICE

No Marks Deducted for Guessing.

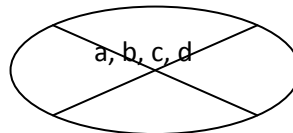
Total Points = _____
25

ANSWER PAGE**PROVIDE YOUR ANSWERS ON THIS PAGE and NEXT PAGE****DO NOT** detach **BOOKLETS****YOU MUST HAND -IN THE ENTIRE BOOKLET**

Multiple Choice QUESTION	<u>ANSWER</u> <u>In This</u> <u>COLUMN</u> ↓
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	

Use **Clearly Marked**
Capital Letters**USE ONLY**

A, B, C, D

**You Must Answer**
multiple choice
questions ON THIS
Page**DO NOT USE****lowercase**
LETTERS:**Fully Form All Letters**
CLEARLY

Question 1:

Suppose the chance that a relay closes is 80%. Assuming that whether or not one relay closes is independent of any other relay closing, what is the probability that at least one of the next 5 relays does Not close?

- a) $1 - (0.8)^5$
- b) $5(0.2)^4$
- c) $1 - 5(0.2)^4$
- d) $1 - (0.2)^4$

Question 2:

Suppose that duration in minutes for wait times in a check-out have a normal distribution with mean of 14.2 minutes with standard deviation of 2.7 minutes. Which of the following is True?

- a) Approximately 95% of wait times will take between 8.908 minutes and 9.492 minutes.
- b) The slowest 1% of wait times will take approximately 7.909 minutes or less
- c) There is approximately a 100% chance a wait time will take more than 34.703 minutes
- d) All of the above

Question 3:

Which of the following is True?

- a) Three type of Sampling are: Simple Random, Cluster, and Stratified
- b) Three types of Sampling Bias are Response Bias (Measurement), Non-Response Bias and Undercoverage (Underselection) Bias
- c) 2 variables can be correlated even though they are measured in different units.
- d) All of the above

Question 4:

You are performing a hypothesis test on a Linear Regression model to determine if a sample provides evidence if the true slope is different from 0. You have computed a test statistic of -1.826 based on a random sample of $n = 20$ data points.

Which of the following is Not True?

- a) You would reject the null hypothesis at the 5% level of significance
- b) The degrees of freedom for the test is 18
- c) The p-value of test statistic is between 5.0% and 10.0%
- d) The test will be a 2-tailed test

Question 5:

Consider the following Matched Pair data for Number of collisions at various locations, After and Before traffic cameras are installed.

Location	A	B	C	D	E
After	1	6	8	4	7
Before	20	20	8	15	5

You wish to test to determine if the number of collisions Before the installation is higher than After the installation using the data, you would

- a) Conclude the average number of collisions is Not higher Before than After when testing at the **10%** level of significance.
- b) Conclude the average number of collisions is Not higher Before than After when testing at the **2.5%** level of significance
- c) Both a) and b) are true
- d) Both a) is false and b) is false

Question 6:

Consider the following probability distribution for X= number of cars in a drive through.

X	0	1	2	3
P(x)	0.40	0.30	0.2	0.10

Which of the following is True?

- a) The expected number of cars in the checkout is 0
- b) The chance that there are at least 2 cars in the drive through is 0.70
- c) The standard deviation for the distribution is 1
- d) All of the above

Question 7

:Anova: Single Factor tested at 5%						
SUMMARY						
	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>	
A		4.000	16.000	4.000	3.333	
B		6.000	77.000	12.833	6.167	
C		5.000	35.000	7.000	2.500	
ANOVA						
	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i> <i>F critical</i>
	Between Groups	204.900	2.000	102.450	24.185	0.000 3.885
	Within Groups	50.833	12.000	4.236		
	Total	255.733	14.000			

Consider the above output regarding an ANOVA test for number of returns at stores A, B, and C

Which of the following is true in testing to determine if differences exist in the average number of returns for the stores?

- a) The result of the ANOVA test at the 5% level of significance would be that differences Do Not exist in population averaged for number of returns between the stores
- b) The alternate hypothesis is
- c) We can safely assume (using the rules of Thumb concerning sample variances) that the population variances are all equal
- d) All of the above

Question 8:

Parts manufactured by an assembly plant are either produced for the following markets: Canada or USA. The parts are either built for Standard use or for Industrial use. A random sample of 500 parts is taken. It is determined that 160 of the 500 parts are for Canada. It is determined that 240 of the 500 parts are for Standard use. It is determined that 120 parts of the 500 are from Industrial use and are for Canada. Using this information and the probability rules, we can conclude:

- a) The probability that a part is for Canada or for Standard use is 80%
- b) The events: Canada, Standard are independent.
- c)
- d) All of the above

Question 9:

A manufacturer purchases cartridges from a vendor. When a large shipment is received, a random sample of 400 cartridges is selected and each is inspected. If the sample proportion of defectives is more than 0.04, the entire shipment will be returned. Otherwise, the entire shipment will Not be returned. What is the approximate (use closest) probability that the entire shipment will Not be returned if the true proportion of defectives is 6%.

- a) 0.0009
- b) 0.0465
- c) 0.9535
- d) 0.9991

Question 10:

Five years ago, the average number of overtime hours was 8.5 hours per month. You now take a random sample of 16 months and determine a sample average of 5.1 hours per month and sample standard deviation of 4.0 hours per month. Assume that overtime hours per month are normally distributed. Which of the following is Not True when testing to determine if the sample provides evidence that the average number of overtime hours per month has decreased since 5 years ago?

- a) The critical value for the test at the 10% level of significance is -1.341
- b) The p-value of the test statistic is More than 5%
- c) At the 1% level of significance, you conclude the average number of overtime hours per month has decreased over the last 2 years
- d) The test statistic is approximately -3.4

Question 11:

Consider the following data regarding Income level (High, Moderate, Low) for Region A. You wish to test if income levels are uniformly distributed in the region. You count the number of families in each group for the region and test if income levels are uniformly distributed in the region.

	Region A
Low	50
Moderate	80
High	110

Which of the following is True?

- a) The test would be a 2-tailed test.
- b) The degree of freedom for the test is 2.
- c) The Income Levels are uniformly distributed at the 5% level of significance.
- d) All of the above

Question 12:

Which of the following is True about p-values in general?

- a) If the p-value of the test statistic is less than that level of significance, then we would reject the null hypothesis at specified level of significance
- b) The level of significance will always be more than the p-value of the test statistic in a 1-tailed test.
- c) If we did Not reject the null hypothesis at a given level of significance, then the p-value of the test statistic was less than the given level of significance
- d) The p-value of a test statistic is always double the level of significance in a 2-tailed test.

Question 13:

Consider the following sample data :

55, 52, 70, 64, 76, 78, 54, 82

Which of the following is Not True?

- a) There is range is 30
- b) 82 is Not an outlier
- c) The median is 67
- d) The mode is 76

Question 14:

Two years ago, the proportion of people interested in starting a small business was 20%. You wish to test to determine if the proportion of people interested in starting a small business has increased. You now take a random sample of 200 people and find that 70 are interested in starting a small business. Which of the following is True in testing to determine if the proportion of people wanting to start a small business has increased?

- a) At the **5%** level of significance, you conclude the proportion of people interested in starting a small business has increased over the last 2 years
- b) The p-value of the test statistic is approximately equal to 99.90%.
- c) The alternate hypothesis is: $p < 0.20$
- d) All of the above

Question 15:

In trying to determine “Does the data provide evidence two sensors (Device A and Device B) have the different rates of incorrect measurements”, you have taken a random sample of 200 measurements by Device A and 200 detections by Device B. The sample data showed Device A had 20 incorrect measurements and Device B had 26 incorrect measurements. The p-value for the test statistic would be

- a) More than 60%
- b) More than 25% but not more than 60%
- c) More than 5% but not more than 25%
- d) Less than or equal to 5%

Question 16:

You wish to estimate the average time it takes complete a task. Assuming times are normally distributed, you take a random sample of times and find (minutes) {20, 30, 30, 40, 40}. The 95% confidence interval for the mean time complete the task approximately (use closest):

- a) (21.6, 42.4) minutes
- b) (24.7, 49.3) minutes
- c) (24.4, 49.6) minutes
- d) (22.4, 41.6) minutes

Question 17:

In order to compare the mean time (min) to complete task using methods A and B, independent random samples of completion times are taken for the 2 methods. The following data is collected.

Assume

	A	B
Mean	50	70
Std dev	10	12
Sample size	5	8

Which of the following is Not True when testing if the average completion times differ for the cities at the 5% level of significance?

- a) The degrees of freedom for the test is 11
- b) The critical values for the test are 2.201 and -2.201
- c) The data Does Not show there is evidence to conclude a difference exists in the average completion times.
- d) The data Does Show there is evidence to conclude a difference exists in the average completion times

Question 18:

We wish to determine the true proportion of people who travel by airplane regularly for a certain area of work. What is the minimum sample size needed (use closest) to estimate the true proportion to within 10% of the true proportion with 95% confidence? Use a guessed value for the true proportion to be 40%.

- a) 93
- b) 65
- c) 40
- d) 26

Question 19:

Consider the following data which resulted from sample data involving Y= recovery time (days) and X= Running Race distance (km)

$$\bar{x} = 5.5$$

$$\bar{y} = 5.1$$

$$n = 10$$

Provided:

	Standard Error
y-intercept	1.920
Temperature	0.309

Which of the following is Not True?

- a) The model predicts the average time to recover increases by 5.122 for each extra km ran.
- b) The degrees of freedom when testing the model will be 8
- c) When testing if the slope differs from 0 at the 5% level of significance, you would conclude that distance is useful to predict the Recovery Time.
- d) The p-value of the test statistic in testing if the slope differs from 0 will be More than 5%

Question 20:

It is known that the wait times for a windshield repair service have an average of 35 minutes and a standard deviation of 0.8 minutes. What is the probability that the average wait time for a random sample of 16 is more than 34.6 minutes but not more than 35.4 minutes? Assume wait times are normally distributed. Choose the closest answer.

- a) Approximately 50.30%
- b) Approximately 68.30%
- c) Approximately 95.44%
- d) Approximately 100.00%

Question 21:

Which of the following is Not True for the ANOVA test when testing if differences exist in the averages of several populations?

- a) The smaller the p-value of the test statistic, the more evidence to conclude differences exist in the averages
- b) The test is a 1-tailed test
- c) A Large Test statistic implies the variance Between the groups is More than the variance Within the Groups
- d) The Null hypothesis states that differences exist in the population averages

Question 22:

The result of a 2 tailed hypothesis proportion test had a test statistic of $Z=2.24$. The associated p-value would be

- a) 0.9875
- b) 0.0125
- c) 0.0250
- d) 0.9750

Question 23:

The 95% confidence interval for the average thickness in coating of a wire (27mm , 41mm). Which of the following is True?

- a) The 90% confidence interval would be wider than the 95% confidence interval
- b) The Margin of Error for the average thickness is 14 mm
- c) The sample average thickness was 34 mm.
- d) All of the above

Question 24:

Which of the following is True for Simple Linear regression?

- a) The theoretical mean of the residuals is 1.
- b) The Simple Linear Regression line has the Highest sum of the square of the residuals of all lines that pass through
- c) Residuals must Not Have a constant variance for the model to be valid
- d) The assumption (condition) of Linearity must be met for the model to be valid.

Question 25:

You are comparing the mean amounts spent (per month) for cellular telephone service with 2 cellular telephone service providers. The 98% confidence interval for the difference between the 2 cellular telephone service providers in mean amounts spent for cellular telephone service (per month) (\$) is (\$3.53, \$12.42). From this information, which of the following could we conclude?

- a) The hypothesis test at the 2% level of significance would result in showing the mean amounts spent (per month) for cellular telephone service are different for the 2 cellular telephone service providers.
- b) The hypothesis test at the 5% level of significance would result in showing that the mean amounts spent (per month) for cellular telephone service are different for the 2 cellular telephone service providers.
- c) Both a) and b) are True
- d) Both a) and b) are Not True