

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICS

STAT 212 BUSINESS STATISTICS II
Semester 222, Major Exam II
Monday March 28, 2023

Time allowed 90 minutes.

Name: _____ ID #: _____

Section #: _____ Serial #: _____

Important Notes:

- Formula sheet will be provided to you in exam. You are not allowed to bring, with you, formula sheet or any other printed/written paper.
- Make sure you have 12 pages of exam paper (including this title page) and 20 questions.
- If exact answer is not in the given options then go for the nearest option.
- Students are not allowed to enter the exam hall without either KFUPM ID or Saudi ID/ Iqama ID.
- Students must take the exam in the place assigned to them.
- Students are not allowed to carry mobiles, smart watches, or electronic devices to the exam halls/rooms.
- Violations of these rules will result in a penalty decided by the chairman of Math Department.

A study indicated where different age groups primarily get their news, the results are shown in the following table:

Source	Age Groups						Total
	< 36		36 - 50		50+		
	f_o	f_e	f_o	f_e	f_o	f_e	
TV	46	103.89	98	99.92	131		275
Newspapers	19	45.33	39		62	31.07	120
Internet	275	190.78	190	183.48	40	130.74	505
Total	340		327		233		900

From the data above answer the following 3 questions:

- At the 0.05 level of significance, we want to test whether there is evidence of a significant relationship between the age group and where people primarily get their news. What is the expected frequency of people of between the **ages 36 – 50** and get their news from **newspapers**?
 - 43.6
 - 71.19
 - 45.33
 - 31.07
 - 39

- At the 0.05 level of significance, we want to test whether there is evidence of a significant relationship between the age group and where people primarily get their news. What is value of the **test statistic**?
 - 229.6
 - 84.8
 - 103.3
 - 292.6
 - 25.9

3. At the 0.05 level of significance, we want to test whether there is evidence of a significant relationship between the age group and where people primarily get their news. The degrees of freedom of the test is equal to:
- A. 4
 - B. 1
 - C. 9
 - D. 898
 - E. 16

Two candidates for governor participated in a televised debate. A political pollster recorded the preferences of 500 registered voters in a random sample prior to and after the debate:

Preference Prior the debate	Preference After the debate		
	Candidate A	Candidate B	Total
Candidate A	269	21	290
Candidate B	36	174	210
Total	305	195	500

Using the data above, answer the following 2 questions:

4. At the 0.01 level of significance, if a researcher wanted to test whether there is evidence of a difference in the proportion of voters who favoured Candidate A prior to and after the debate. The proportion of people who preferred Candidate A **prior to** the debate is:
- A. 0.58
 - B. 0.61
 - C. 0.39
 - D. 0.42
 - E. -1.9868

5. At the 0.01 level of significance, if a researcher wanted to test whether there is evidence of a difference in the proportion of voters who favoured Candidate A prior to and after the debate. **The test statistic** is:
- A. -1.9868
 - B. 7.5498
 - C. 0.2632
 - D. -3.8730
 - E. -7.5498

Has the ease of removing your name from an e-mail list changed? A study of 100 large online retailers revealed the following:

YEAR	NEED THREE OR MORE CLICKS TO BE REMOVED	
	Yes	No
2009	39	61
2008	7	93

Source: Data extracted from “More Clicks to Escape an Email List,”
The New York Times, March 29, 2010, p. B2.

Based on the above information, answer the next **two** questions.

6. If we want to test whether the effort it takes to be removed from an e-mail list has changed, then the **alternative hypothesis** is
- A. $H_1: \pi_{2008} \neq \pi_{2009}$
 - B. $H_1: \pi_{2008} > \pi_{2009}$
 - C. $H_1: \pi_{2008} < \pi_{2009}$
 - D. $H_1: \mu_{2008} = \mu_{2009}$
 - E. $H_1: \mu_{2008} \neq \mu_{2009}$

7. If we want to test whether the effort it takes to be removed from an e-mail list has changed, then the **critical value** of the test at 0.05 is
- A. 3.841
 - B. 5.991
 - C. 5.024
 - D. 7.378
 - E. 2.706

Undergraduate students at Miami University in Oxford, Ohio, were surveyed in order to evaluate the effect of price on purchasing a pizza from Pizza Hut. Students were told to suppose that they were planning to have a large two-topping pizza delivered to their residence that evening. The students had to decide between ordering from Pizza Hut at a reduced price of \$8.49 (the regular price for a large two-topping pizza from the Oxford Pizza Hut at this time was \$11.49) and ordering a pizza from a different pizzeria. A survey evaluated purchase decisions at other prices. These results are summarized in the following contingency table:

	PRICE			
PIZZERIA	\$8.49	\$11.49	\$14.49	Total
Pizza Hut	10	5	2	17
f_e			5.36	
Other	25	23	27	75
f_e	28.53	22.83	23.64	
Total	35	28	29	92

Using the above table answer the following three questions:

8. What is the expected number of orders Pizza Hut and the price is \$8.49?
- A. 6.47
 - B. 5.17
 - C. 1.93
 - D. 10
 - E. 17

9. What is the expected number of orders Pizza Hut and the price is \$11.49?
- A. 5.17
 - B. 6.46
 - C. 1.93
 - D. 10
 - E. 17
10. A test was conducted to determine if a relationship exists between price and pizzeria selection, which of the following is true **at 5% level of significance**:
- A. Do not Reject H_0 Since the Test statistic < 5.991 (the Critical value at 5% **level of significance**) and we conclude that there is not enough evidence to conclude that there is a relationship between price and pizzeria selection.
 - B. Reject H_0 Since the Test statistic > 5.991 (the Critical value at 5% **level of significance**) and we conclude that there is enough evidence to conclude that there is a relationship between price and pizzeria selection.
 - C. Reject H_0 Since the Test statistic > 3.841 (the Critical value at 5% **level of significance**) and we conclude that there is not enough evidence to conclude that there is a relationship between price and pizzeria selection.
 - D. Reject H_0 Since the Test statistic > 3.841 (the Critical value at 5% **level of significance**) and we conclude that there is enough evidence to conclude that there is a relationship between price and pizzeria selection.
 - E. Do not Reject H_0 Since the Test statistic > 5.991 (the Critical value at 5% **level of significance**) and we conclude that there is not enough evidence to conclude that there is a relationship between price and pizzeria selection.

A company that holds the DVD distribution rights to movies previously released only in theaters wants to estimate sales revenue of DVDs based on box office success. The box office gross (in \$ millions) for each of **22** movies in the year that they were released and the DVD revenue (in \$ millions). Here y denotes the DVD revenue (in \$ millions) and x denotes the box office gross (in \$ millions). The summary of this data is given below:

$$\sum y = 900.40, \quad \sum x = 1745.21, \quad \sum xy = 86195.43, \quad \sum y^2 = 47250.47, \quad \sum x^2 = 176913.38, \quad \text{and} \\ SSR = 5669.7847$$

Using the data above answer the following 4 questions:

11. The regression equation is given by:

- A. $\hat{y} = 10.4734 + 0.3839 x$
- B. $\hat{y} = 0.3839 + 10.4734 x$
- C. $\hat{y} = 12.4734 - 0.3839 x$
- D. $\hat{y} = 0.3839 - 10.4734 x$
- E. $\hat{y} = 12.4734 + 0.6639 x$

12. The estimated correlation, of x and y , is:

- A. 0.7384
- B. 0.9998
- C. -0.7384
- D. -0.2231
- E. 0.0002

13. The standard error of the estimate is:

- A. 15.3782
- B. 22.8030
- C. 11.3022
- D. 8.7789
- E. 12.6008

14. The percentage of the variation in the DVD revenue that is explained by the variation in the box office gross is:

- A. 54.52%
- B. 45.48%
- C. 15.38%
- D. 79.88%
- E. 95.62%

In the Sunflowers Apparel chain of stores, the business objective of the director of planning is to forecast annual **Sales (Y)** (in millions \$) for all new stores, based on store **Size (X)** (in 1000's square feet). To examine the relationship between the store size in square feet and its annual sales, data were collected from a sample of 14 stores. The data summaries are shown below;

$$\bar{X} = 2.9214, \bar{Y} = 5.8429, S_{xx} = 37.9236, S_{yy} = 116.9543, b_1 = 1.6699$$

Using the above information, answer the next **two** questions.

15. The average change in the Sales, when the Size of the store increases by 1000 square feet, is

- A. 1.6699 million \$
- B. 0.9645 million \$
- C. 0 million \$
- D. 0.4596 million \$
- E. 1.9966 million \$

16. A 95% Confidence interval for the slope is

- A. [1.3280, 2.0118]
- B. [-1.3280, 2.0188]
- C. [2.1936, 3.0675]
- D. [-2.0118, 1.3290]
- E. [-3.0675, 2.1936]

A highway employee performed a regression analysis of the relationship between the number of construction work-zone fatalities and the number of unemployed people in a state. The Minitab output of the regression analysis is shown below:

The regression equation
Fatalities = 12.7 + 0.000114 (Unemp)

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1		10354	15.4614	0.001
Error		12054			
Total		22408			

Coefficients					
Term	Coef	SE Coef	T-Value	P-Value	
Constant	12.726	8.115	1.57	0.134	
Unemp	0.00011386	0.000002896	3.93	0.001	

Using the above information, answer the next **two** questions.

17. How many states were in the sample?

- A. 20
- B. 19
- C. 18
- D. 21
- E. 22

18. What is the standard error of the estimate?

- A. 25.878
- B. 669.667
- C. 5.087
- D. 8.115
- E. 12.726

You want to develop a model to predict the assessed value of houses, based on heating area. A sample of 15 single-family houses in a city is selected.

The assessed value (in thousands of dollars) and the heating area of the houses (in thousands of square feet) are recorded.

Use the results below to answer the following two questions:

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.81199569
R Square	0.65933699
Adjusted R Square	0.63313215
Standard Error	2.91892772
Observations	15

ANOVA

	df	SS	MS	F	Significance F
Regression	1	214.374192	214.374192	25.1608796	0.00023616
Residual	13	110.761808	8.52013905		
Total	14	325.136			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	151.91534	5.56252089			139.898244	163.932436
Heating Area	16.6333695				9.46954077	23.7971982

19. Interpret the meaning of the slope:

- A. For each additional thousand square foot increase in the heating area of a house, the estimated mean assessed value increases by \$16633.4.
- B. For each additional thousand square foot increase in the heating area of a house, the estimated mean assessed value increases by \$16.6334.
- C. For each additional square foot increase in the heating area of a house, the estimated mean assessed value increases by \$16.6334.
- D. For each additional square foot increase in the heating area of a house, the estimated mean assessed value increases by \$16633.4.
- E. For each additional thousand square foot increase in the heating area of a house, the assessed value increases by \$1663.3.

20. Predict the assessed value for a house whose heating area is 1,750 square feet.

- A. \$181023.74
- B. \$29260.31
- C. \$442.99
- D. \$3062.76
- E. \$181.02