KING FAHD UNIVERSITY OF PETROLEUM & MINERALS, DHAHRAN, SAUDI ARABIA

DEPARTMENT OF MATHEMATICS

STAT 319: Probability & Statistics for Engineers & Scientists

Term 212, Final Exam, Monday May 23, 2022, 7.00 PM – 9.10 PM

Name: _____ID #: _____

Instructions:

1. Formula sheet is attached at the end of this exam. You are not allowed to bring with you, formula sheet or any other printed/written paper.

2. Mobiles are not allowed in exam. If you have your mobile with you, turn it off and keep it under your seat so that it is visible to proctor. Your mobile(s) should not be in your pocket during the exam.

3. The answers are rounded. If the exact answer is not there in any of the 5 choices, then pick the one that you think is closest to correct answer.

4. Make sure you have 35 questions

1. Suppose the probability that the component A of a machine will work for 10 more years is 0.8 and the probability that the component B of the machine will work for 10 more years is 0.85. Assume that components A and B are independent events. The probability that exactly one of these two components works 10 more years is
    a. 0.8
    b. 0.85
    c. 0.97
    d. 0.29
    e. 0.45

2. Printed circuit cards are placed in a functional test after being populated with semiconductor chips. A lot contains 50 cards, and three are selected without replacement for functional testing. If 10 cards are defective, what is the probability that at least three cards in the sample are defective?

a. 0.9939
b. 0.0061
c. 0.7851
d. 0.1141
e. 0.6793

3. Suppose that the number of customers who enter a bank in an hour is a Poisson random variable, and suppose that P(X = 0) = 0.2. Determine the mean and variance of X.
   a. 1.61
   b. 1.39
   c. 0
   d. $1 - \ln(0.2)$
   e. 2.43

4. If Experience has shown that the width, in mm, of a certain type of plastic connector has the following probability distribution

$$f(y) = \begin{cases} 50y & ,0.48 < y < 0.52 \\ 0 & ,otherwise \end{cases}$$

What is the probability that the probability that the width of the connector is 0.50?

a. 0
b. 0.51
c. 0.49
d. 0.50
e. 1

5. Suppose the waiting time for a bus is a random variable X having the probability density function

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Find the median of the waiting time.

a. 0.53240
b. 0.34657
c. 24329
d. 0.0234
e. 0.9766

6. The thickness of a flange on an aircraft component is uniformly distributed between 950 and 1050 micrometers (μm). If a sample of 40 flanges is randomly selected, the 90th percentile of the mean thickness is
    a. 1037.00 μm
    b. 994.15 μm
    c. 963.00 μm
    d. 863.00 μm
    e. 1005.90 μm

7. A very large batch of components has arrived at a distributor. The actual proportion of defectives is 0.05. Suppose that the distributor decides to randomly select 10 components and to accept the batch only if the number of defective components in the sample is at most 2. What is the probability that the batch will be accepted?

a. 0.9885
b. 0.9139
c. 0.0115
d. 0.0861
e. 0.0746

8. Current estimates suggest that 75% of the home-based computers in a foreign country have access to on-line services. Suppose 20 people with home-based computers were randomly and independently sampled. Using normal approximation find the probability that fewer than 10 of those sampled currently have access to on-line services.

a. 0.0049
b. 0.0102
c. 0.0023
d. 0.9979
e. 0.9951

9. At a computer manufacturing company, the actual size of a computer chip has a mean of 3.26 mm and a standard deviation of 1.2 mm. A random sample of 100 computer chips is taken. Find the approximate probability that the mean size of the 100 chips is no more than 3.0 mm?
    a. Approximately 0
    b. 0.9849
    c. Approximately 1
    d. 0.1645
    e. 0.0150

10. The following frequency distribution shows the lifetimes of batteries (in hours).

Lifetime class	*f*		
[000, 100)	11		
[100, 200)	16		
[200, 300)	9		
[300, 400)	6		
[400, 500)	3		

The approximate mean lifetime and the approximate mode are respectively

a. 242.22 & 200 hours.
b. 142.22 & 100 hours.
c. 192.22 & 200 hours.
d. 132.22 & 250 hours.
e. 192.22 & 150 hours.

11. Box plots have been used successfully to describe
    a. center and spread of a data set
    b. the extent and nature of any departure from symmetry
    c. identification of "outliers"
    d. All of the choices.
    e. none of the choices
12. A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed with variance $1000(psi)^2$. A random sample of 10 specimens has a mean compressive strength of 3250 psi. With what degree of confidence could we say that the mean compressive strength between 3235 and 3265?

    a. 90%
    b. 87%
    c. 95%
    d. 85%
    e. 99%

13. An experimenter claims that the viscosity of a liquid detergent is supposed to average 800. A random sample of size 16 batches of detergent is collected, and the average viscosity is 812, with a standard deviation of 25. The smallest level of significance at which the claim is rejected is:

    a. Between 0.025 and 0.05
    b. Between 0.035 and 0.05
    c. Between 0.05 and 0.100
    d. Between 0.07 and 0.1
    e. Between 0.035 and 0.1

14. A chemical engineer is investigating the effect of the reaction temperature x on percent conversion of a chemical process y. A random sample of size 10 chemical processes is selected, and the following information are recorded

$\sum x = 1900, \sum y = 533,\ \sum(x-\bar{x})^2 = 33000,\ \sum(y-\bar{y})^2 = 3760,\ \sum(x-\bar{x})(y-\bar{y}) = 10530$

If the reaction temperature is 240 ^{0}C, what is the predicted percent conversion of the chemical process?

a. 31.91%
b. 57.96%
c. 24.55%
d. 38.21%
e. 69.26%

15. 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. For a sample of size 20, he believes that the appropriate model is

$$\hat{y} = 12.726 + 0.00011386\,X$$

Some additional results are given: $s_e(\beta_1) = 0.00002896$, $SSR = 10354$, $SSE = 12054$

To test $H_0: \beta_1 = 0.00008$ $vs.$ $H_1: \beta_1 \neq 0.00008$ at 1% level of significant, the test statistic equal to

a. 1.169
b. 3.93
c. -1.169
d. -3.93
e. You need more information

16. The sample correlation coefficient between X and Y is 0.375. It has been found out that the p – value is 0.256 when testing $H_0: \rho = 0 \ \ vs. \ H_1: \rho \neq 0$. To test $H_0: \rho \leq 0 \ \ vs. \ H_1: \rho > 0$ at significance level of 10%, the p – value is?

    a. 0.256.
    b. 0.872.
    c. 0.128.
    d. 0.375.
    e. We can't find the p – value depend on the information that we have.

17. The least squares method minimizes which of the following?
    a. SSR
    b. SSE
    c. SST
    d. All of the above
    e. None of the above

A microeconomist wants to determine how corporate sales are influenced by capital and wage spending by companies. She proceeds to randomly select 26 large corporations and record information in millions of dollars. The Microsoft Excel output below shows results of this multiple regression.

SUMMARY OUTPUT

Regression Statistics

Multiple R	0.830
R Square	0.689
Adjusted R Square	0.662
Standard Error	17501.643
Observations	26

ANOVA

	df	SS	MS	*F*	*Signif F*
Regression	2	15579777040	7789888520	25.432	0.0001
Residual	23	7045072780	306307512		
Total	25	22624849820			

	Coeff	StdError	*t Stat*	*P-value*
Intercept	15800.0000	6038.2999	2.617	0.0154
Capital	0.1245	0.2045	0.609	0.5485
Wages	7.0762	1.4729	4.804	0.0001

Using the above output answer the following two questions:

18. What is the p-value for testing whether Wages have a negative impact on corporate sales?
    a. 0.05
    b. 0.0001
    c. 0.00005
    d. 0.00154
    e. 0.99995

19. One company in the sample had sales of $20 billion (Sales = 20,000 million). This company spent $300 million on capital and $700 million on wages. What is the residual (in millions of dollars) for this data point?
    a. –790.69
    b. 874.55
    c. 622.87
    d. –983.56
    e. -561.43

20. A computer software developer would like to use the number of downloads for the trial version of his new shareware to predict the amount of revenue (in thousands of dollars) he can make on the full version of the new shareware. From a data set of 15 different sharewares that he has developed, the regression sum of squares SSR=234.423. The computed value of standard error of the model estimate is 2.16451. What is the value of the coefficient of determination?

a. 0.8564
b. 0.6785
c. 0.5341
d. 0.7937
e. 0.9731

21. A chemist is trying to model wear volume Y as a function of oil viscosity X, Use the following summary to answer the questions below

$n = 10$, $\sum x = 1307$, $\sum y = 216$, $S_{xx} = 23494.1$, $S_{yy} = 1132.4$ and $S_{xy} = -3593.2$

The Estimated strength of the linear relationship between wear volume and oil viscosity, is

a. – 0.4852
b. 0.8346
c. – 0.6966
d. – 0.8346
e. 0.1654

22. A consulting group was hired by the Human Resources Department at General Mills, Inc. to survey company employees regarding their degree of satisfaction with their quality of life. A special index, called the index of satisfaction, was used to measure satisfaction. Six factors were studied, namely, age at the time of first marriage (x1), annual income (x2), number of children living (x3), value of all assets (x4), status of health in the form of an index (x5), and the average number of social activities per week—such as bowling and dancing (x6). Suppose the multiple regression equation is:

$$\hat{y} = 16.24 + 0.017x1 + 0.0028x2 + 42x3 + 0.0012x4 + 0.19x5 + 26.8x6$$

What is the estimated index of satisfaction for a person who first married at 18, has an annual income of $26,500, has three children living, has assets of $156,000, has an index of health status of 141, and has 2.5 social activities a week on the average?

a. 579.736
b. 397.736
c. 607.412
d. 497.736
e. 207.412

23. In a multiple regression model, which of the following is correct regarding the value of the adjusted r^2?

a. It can be negative.
b. It has to be positive.
c. It has to be larger than the coefficient of multiple determination.
d. It can be larger than one.
e. It is always One half of r^2

A mechanical engineer is interested in studying the relationship between car mileage MPG (y) and five predictors, namely height (x_1), horse power HP (x_2), weight (x_3), max load (x_4), and turning circle (x_5). A regression model was fitted and the results are shown in the output below:

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	49.25	4.92	10.01	
Height X1	-0.0266	0.0810	-0.33	
HP X2	-0.00056	0.00700	-0.08	
Weight X3	-0.00370	0.00108	-3.41	
Max Load X4	0.00269	0.00201	1.34	
Turning Circle X5	-0.394	0.116	-3.38	

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	286.85	57.371	27.20	0.000
Error	24	50.61	2.109		
Total	29	337.47			

Answer the following two questions:

24. To test the significance of the Weight (x_3) to the MPG (y), the p-value of the test is

a. $0.02 < \text{p-value} < 0.01$
b. $0.005 < \text{p-value} < 0.01$
c. $0.001 < \text{p-value} < 0.002$
d. $0.002 < \text{p-value} < 0.005$
e. $0.998 < \text{p-value} < 0.999$

25. The percentage of variation in the MPG (y) that can be explained by the variation in all of the predictors taking into account the given number of predictors and the given sample size, is equal to

a. 81.88
b. 85.03
c. 88.18
d. 92.37
e. 95.08

The manager of a car plant wishes to investigate how plant's electricity usage depends upon the amount plant's production. He had the following data

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Production ($million)	4.51	3.58	4.31	5.06	5.64	4.99	5.29	5.83	4.7	5.61	4.9	4.2
Electricity Usage (million KWh)	2.48	2.26	2.47	2.77	2.99	3.05	3.18	3.46	3.03	3.26	2.67	2.53

A simple linear regression was fit using MINITAB and the results are:

Term	Coef	SE Coef	T-Value	P-Value
Constant	0.409	0.386	1.06	0.314
Production	0.4988	0.0784	6.37	0.000

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	1.2124	1.21238	40.53	0.000
Error	10	0.2991	0.02991		
Total	11	1.5115			

26. Find a 99.5% confidence interval for the mean electricity usage when the production is worth 5 $million:
    a. (2.790, 3.016)
    b. (2.700, 3.020)
    c. (2.612, 3.520)
    d. (2.895, 2.997)
    e. (2.721, 3.085)

Twelve specimens of cold-reduced sheet steel, having different copper contents (x_1) and annealing temperatures (x_2), are measured for hardness (y) with the following results:

y	78.9	65.1	55.2	56.4	80.9	69.7	57.4	55.4	85.3	71.8	60.7	58.9
x_1	0.02	0.02	0.02	0.02	0.1	0.1	0.1	0.1	0.18	0.18	0.18	0.18
x_2	1000	1100	1200	1300	1000	1100	1200	1300	1000	1100	1200	1300

A simple linear regression was fit using MINITAB and the results are:

Term	Coef	SE Coef	T-Value	P-Value
Constant	161.3	11.4	14.11	0.000
x1	33.0	16.8	1.97	0.081
x2	-0.08550	0.00979	-8.74	0.000

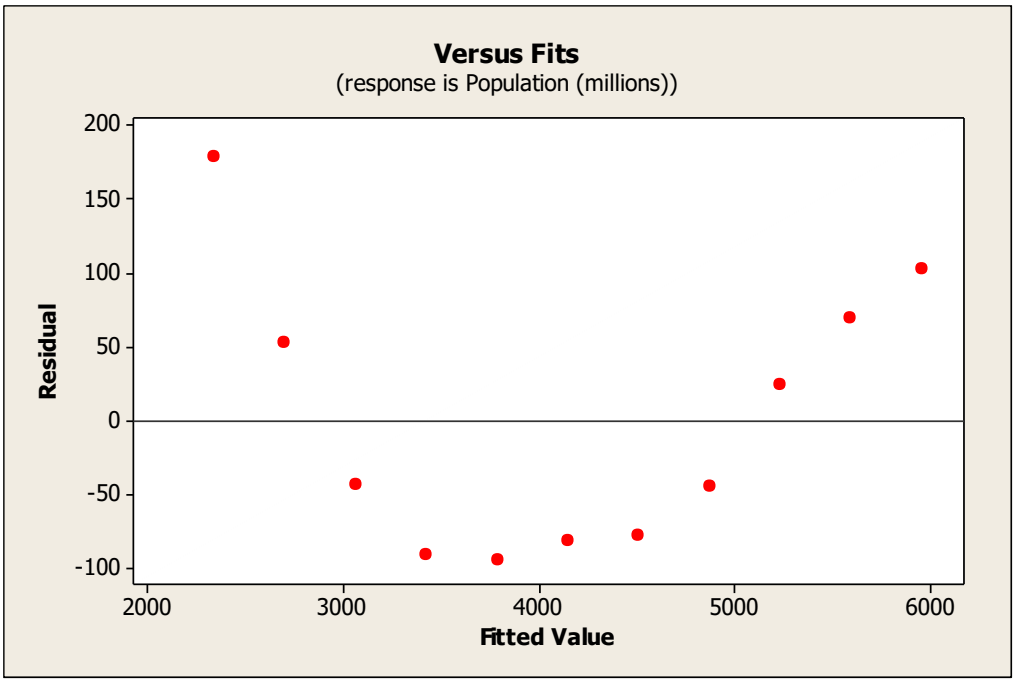
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	1152.2	576.09	40.09	0.000
Error	9	129.3	14.37		
Total	11	1281.5			

27. The percentage of variation explained by the model taking into account number of independent variables in the model and the sample size is
    a. 89.91%
    b. 44.51%
    c. 94.33%
    d. 86.73%
    e. 87.67%
28. A regression model was used to analyze the data from a study investigating the relationship between compressive strength (x) and intrinsic permeability (y) of various concrete mixes and cures. Summary quantities are:

$$\sum_{i=1}^{14} y_i = 572, \qquad \sum_{i=1}^{14} x_i = 43, \qquad S_{xx} = 25.348, \qquad S_{yy} = 159.714,$$
$$S_{xy} = -59.057$$

An estimate of the variance of the error is
    a. 1.3578
    b. 1.8433
    c. 4.7035
    d. 22.123
    e. 3.1436



29. What does the plot of residuals suggest?
    a. An outlier is present in the data set.
    b. A high leverage point is present in the data set.
    c. The data are not normal.
    d. The linearity condition is not satisfied.
    e. The equal spread condition is not satisfied.

After completing sales training for a large company, it is expected that a salesperson will generate a sale on more than 15 percent of the calls he or she makes. To make sure that the sales training process is working, a random sample of 400 sales calls made by sales representatives who have completed the training have been selected and the null hypothesis is to be tested at 0.05 alpha level. Suppose that a sale is made on 76 of the calls.

Use the information above to answer the following 3 questions.

30. The appropriate null and alternative hypotheses are

a. $H_0 : p = 0.15$ and $H_A : p > 0.15$.

b. $H_0 : p > 0.15$ and $H_A : p = 0.15$.

c. $H_0 : p = 0.15$ and $H_A : p < 0.15$.

d. $H_0 : p < 0.15$ and $H_A : p = 0.15$.

e. $H_0 : p = 0.15$ and $H_A : p \neq 0.15$.

31. The correct value of the test statistic is

a. 1.96.

b. -2.24.

c. 2.24.

d. 3.34.

e. -1.645.

32. At $\alpha = 0.05$:

a. We can conclude that the percentage of sales after training is significantly higher than 15%.

b. We can conclude that the percentage of sales after training is significantly lower than 15%.

c. We can conclude that the percentage of sales after training is equal to 15%.

d. We can conclude that the percentage of sales after training is not significantly different from 15%.

e. We can conclude that the percentage of sales after training is not significantly higher than 15%.

33. An appliance manufacturer claims to have developed a compact microwave oven that consumes a mean of no more than 250W. From the previous studies, it is believed that power consumption for microwave ovens is normally distributed. A consumer group has decided to try to discover if the claim appears true. They take a sample of 20 microwave ovens and find that they consume a mean of 247W with a standard deviation of 15W. At 5% level of significance, the p-value associated with the test statistic is between what two values?

a. $0.1<p\text{-value}<0.25$
b. $0.01<p\text{-value}<0.025$
c. $0.25<p\text{-value}<0.75$
d. $0.20<p\text{-value}<0.95$
e. $0.75<p\text{-value}<0.90$

34. The owner of a local nightclub has recently surveyed a random sample of n = 250 customers of the club. She would now like to determine whether or not the mean age of her customers is greater than 30. If so, she plans to alter the entertainment to appeal to an older crowd. If not, no entertainment changes will be made. Suppose she found that the sample mean was 30.45 years and the sample standard deviation was 5 years. What is the p-value associated with the test statistic?

a. 0.0778
b. 0.3577
c. 0.1423
d. 0.6423
e. 0.02

35. The heat evolved in calories per gram of a cement mixture is approximately normally distributed. The mean is thought to be 100, and the standard deviation is 2. You wish to test $H_0: \mu = 100 \quad vs. H_0: \mu \neq 100$ with a sample of n = 9 specimens. If the critical region is defined as $\bar{x} > 101.5 \ \ or \ \ \bar{x} < 98.5$, the probability of accepting the null hypothesis if the true mean heat evolved is 103 is

a. 0.98777
b. 0.05
c. 0.025
d. 0.01222
e. We need more information