KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DEPARTMENT OF MATHEMATICAL SCIENCES

DHAHRAN, SAUDI ARABIA

STAT 212: BUSINESS STATISTICS II

Term 211

Second Major Exam

Tuesday November 9, 2021

Section #	Class time
1	10:00 AM
2	11:00 AM
3	9:00 AM

Name: ID#:   Section: Serial:

The correct answer is **a**

A research center provides information on wholesale and retail price of cars. The following age and price data for 10 randomly selected sports cars (from a certain brand) between 1 and 6 years old. Here x denotes the Age (in years) and y denotes the Price (in hundreds of dollars)

Age (x)	6	6	6	2	2	5	4	5	1	3
Price (y)	175	165	180	310	269	200	240	213	310	210

You have calculated some of the necessary summary information to carry out the analyses as follows:

$$\sum x = 40 \ , \ \sum x^2 = 192 \ , \ \sum y = 2272 \ , \ \sum y^2 = 541880 \ , \ \sum xy = 8243$$

Use the above information to answer the following questions

1. The correlation coefficient of x and y is
    a. -0.932
    b. 0.932
    c. 0.329
    d. -0.623
    e. 0.623

2. The correlation coefficient of x and y can be interpreted as
    a. The car Age and the car Price are <u>inversely</u> <u>strongly</u> correlated.
    b. The car Age and the car Price are <u>inversely</u> <u>weakly</u> correlated.
    c. The car Age and the car Price are <u>directly</u> <u>strongly</u> correlated.
    d. The car Age and the car Price are <u>directly</u> <u>weakly</u> correlated.
    e. The car Age and the car Price are <u>inversely</u> <u>moderately</u> correlated.

3. If you want to test, at 1% level of significance, that the data provide sufficient evidence to conclude that the Age and the Price of the cars are negatively linearly correlated, then the test statistic is
    a. -7.279
    b. 7.279
    c. -2.799
    d. 2.799
    e. 2.979

4. If you want to test, at 1% level of significance, that the data provide sufficient evidence to conclude that the Age and the Price of the cars are negatively linearly correlated, then the critical value is
    a. 2.897
    b. 3.355
    c. 2.306
    d. 2.325
    e. 2.575

5. If you want to test, at 1% level of significance, that the data provide sufficient evidence to conclude that the Age and the Price of the cars are negatively linearly correlated, then the test conclusion is
    a. There is sufficient evidence that Age and Price are negatively linearly correlated.
    b. There is NO sufficient evidence that Age and Price are negatively linearly correlated.
    c. There is sufficient evidence that Age and Price are NOT negatively linearly correlated.

    d.   There is NO sufficient evidence that Age and Price are NOT negatively linearly correlated.

    e.   There is sufficient evidence that Age and Price are directly linearly correlated.

The managers of a manufacturing company believe that they can use production rate (x) to predict the number of defects (y) produced per hour. The following data were collected for 10 randomly selected hours.

Defects (y)	40	60	20	40	60	50	60	40	20	80
Production Rate Per Hour (x)	800	900	700	750	800	800	900	600	590	820

Also, the following summary statistics are obtained by the

$\sum x = 7660$, $\sum y = 470$, $S_{xy} = 13380$, $S_{xx} = 105440$, and $SSE = 1512.12064$

Use the above information to answer the following questions

6. The estimated <u>slope</u>, of the predicted regression equation, is

    a. 0.1269

    b. -0.1269

    c. 50.203

    d. -0.1926

    e. -50.203

7. The estimated <u>intercept</u>, of the predicted regression equation, is

    a. -50.203

    b. 0.1269

    c. -0.1269

    d. 50.203

    e. -0.1926

8. The amount of <u>error</u>, in estimating the number of Defects if 600 units is produced per hour, is

    a. 14.063

    b. 13.748

    c. 52.893

    d. 25.935

    e. 3.7079

9. If the number of units Produced, per hour, increases by one, then, on average, the number of Defects would change by _____ defective units per hour

    a. 0.1269

    b. -0.1269

    c. 50.203

    d. -0.1926

    e. -50.203

10. _____ of the variation in the number of Defects produced is explained by the variation in the Production rate per hour
   a. 52.893
   b. 14.065
   c. 13.748
   d. 25.935
   e. 3.7079

11. A 98% confidence interval for the <u>mean</u> number of Defects produced when 600 units are Produced per hour is
   a. (1.998 , 49.872)
   b. (-20.527 , 72.397)
   c. (0.0043 , 0.2495)
   d. (-144.983 , 44.577)
   e. (5.309 , 81.001)

The following is the ANOVA summary table for a multiple regression model to regress Y on two independent variables $X1$ & $X2$:

```
Analysis of Variance
Source          DF      Adj SS      Adj MS      F-Value     P-Value
Regression       2          60
Error           18         120
Total           20         180
```

Using the above Minitab output and given that $SSR_{(x1)} = 45$ and $SSR_{(x2)} = 25$, answer the following questions

12. The <u>standard error of the estimate</u> is
   a. 2.582
   b. 5.477
   c. 3.000
   d. 0.707
   e. 6.667

13. The *percentage of variation in Y that is explained by the variation in both $X1$ & $X2$ adjusted to a sample of size 21 and two predictors,* is
   a. 25.93 %
   b. 33.33 %
   c. 40.00 %
   d. 70.00 %
   e. 62.96 %

14. To test the overall significance of the model at the 2.5% significance level, the **test statistic** is
   a. 4.5
   b. 2.25

    c. 3.333
    d. 4.56
    e. 5.25

15. To test the overall significance of the model at the 2.5% significance level, the **critical value** is
    a. 4.56
    b. 4.5
    c. 5.25
    d. 5.98
    e. 3.55

16. To test that the predictor $X1$ makes a significant <u>contribution</u> to the model at the 2.5% significance level, the **test statistic** is
    a. 5.25
    b. 4.5
    c. 2.25
    d. 5.98
    e. 3.55

17. To test that the predictor X_1 makes a significant <u>contribution</u> to the model at the 2.5% significance level, the **critical value** is
    a. 5.98
    b. 4.5
    c. 5.25
    d. 3.333
    e. 3.55

18. The <u>partial coefficient of determination</u> for X_1 given that X_2 is already in the model, is
    a. 22.581%
    b. 11.111%
    c. 33.333%
    d. 66.667%
    e. 58.221%

A study is conducted to determine the effects of company size and the presence or absence of a safety program on the number of hours lost due to work-related accidents. <u>**Forty**</u> companies are selected for the study. The variables are as follows:
 Y = lost work hours
 $X1$ = number of employees
 $X2$ = 1 if a safety program is used, 0 if no safety program is used

The following regression Minitab output was obtained:

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant		9.90	3.17	0.003
X1	0.014		10.15	0.000
X2	-54.21	7.24	-7.48	0.000

If you want to estimate the lost work hours (Y) using $X1$ & $X2$, then using the above information, answer the following questions

19. The <u>intercept</u> of the estimated regression equation is
   a. 31.4
   b. 0.014
   c. 0.001
   d. 9.90
   e. -54.21

20. The <u>standard error</u> of b_1 is
   a. 0.0014
   b. 0.014
   c. 10.15
   d. 3.17
   e. 7.24

21. The estimated regression equation is
   a. $\hat{Y} = 31.4 + 0.014X_1 - 54.21X_2$
   b. $\hat{Y} = 3.17 + 10.15X_1 - 7.48X_2$
   c. $\hat{Y} = 31.4 + 0.001X_1 - 7.48X_2$
   d. $\hat{Y} = 0.014 + 0.001X_1 + 10.15X_2$
   e. $\hat{Y} = -54.21 + 7.24X_1 - 7.48X_2$

22. The predicted lost work hours for a company with 7000 employees which is NOT using a safety program is
   a. 129.4
   b. 75.19
   c. 192.4
   d. 57.91
   e. 91.75

23. Is it reasonable to drop the dummy variable from model?
   a. No, because X2 is significantly related to Y since the p-value < 0.05.
   b. Yes, because X2 is significantly related to Y since the p-value < 0.05.
   c. No, because X2 is not significantly related to Y since the p-value < 0.05.
   d. Yes, because X2 is not significantly correlated to X1 since the p-value < 0.05.
   e. No, because X2 is significantly correlated to X1 since the p-value < 0.05.

24. What is a 95% confidence interval estimate of the population slope for the relationship between lost work hours and the use of a safety program?
    a. [-68.879 , -39.540]
    b. [-66.425 , -41.995]
    c. [-71.813 , -36.607]
    d. [-61.025 , -50.375]
    e. [-80.004 , -29.996]

25. What is the interpretation of the coefficient -54.21?
    a. When the safety program is used, the average number of lost work hours is 54.21 hours less than the number of lost work hours when the safety program is not used, given that the number of employees is kept constant.
    b. When the safety program increases by one time, the average number of lost work hours decreases by 54.21 hours, given that the number of employees is kept constant.
    c. When the safety program is not used, the average number of lost work hours is 54.21 hours less than the number of lost work hours when the safety program is not used, given that the number of employees is kept constant.
    d. When the safety program decreases by one time, the average number of lost work hours decreases by 54.21 hours, given that the number of employees is kept constant.
    e. When the safety program is used, the average number of lost work hours is 54.21 hours less than the number of lost work hours when the safety program is not used.