KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICAL SCIENCES
DHAHRAN, SAUDI ARABIA

STAT 212: BUSINESS STATISTICS II

Semester 043
Final Exam
Wednesday August 24, 2005
12:30 pm – 2:30 pm

Please **circle** your instructor's name: Marwan Al-Momani
Raid Anabosi

Name: ID#: Section: **Serial:**

Question No	Full Marks	Marks Obtained
1	7	
2	10	
3	22	
4	7	
5	7	
6	7	
7	20	
Total	**80**	

Q1. The following table presents data on the number of moving traffic offenses during the past five years for various age groups

Age	Number of Traffic Offenses			Total
	0	1 to 2	3 or more	
16 - 25	6	22	32	60
26 - 50	12	17	10	39
51 - 75	4	16	11	31
Total	22	55	53	130

Based on these data, can we conclude that the number of traffic offenses is independent of the age? Use $\alpha = .05$

The hypothesis are	H_0: H_A:
The test statistic Value	
The critical Value:	
Decision Rule	
Conclusion	

Q2. A marketing research study performed by the marketing division of a certain company surveyed the income levels and expenditures of recreation for a sample of 20 people. Measurements recorded the expenditures on recreation during the previous year, Y (In 100,000 S.R), and the total family income, X (In 100,000 S.R).

X	21.3	30.2	31.5	45.9	34.6	17.8	53.6	17.4	26.8	15.7
Y	1.425	1.675	1.356	4.53	3.2	1.06	4.09	1.2	1.8	0.7
X	17.6	16.89	28	14.3	9.8	24.7	20.5	31.7	47.8	8.4
Y	0.9	1	2.45	0.65	0.3	1.5	0.89	2.3	3.1	0.1

$\sum x = 514.49, \sum x^2 = 16189.432, \sum x\,y = 1151.2905$

Given that: $\sum y = 34.226, \ \sum y^2 = 86.734686$

a. $r =$	
Interpretation of $r =$	
b. Calculate the least square estimates b_0 *and* b_1	$b_1 =$ $b_0 =$
c. Is it possible to calculate R^2? If yes find it and interpret its value	$R^2 =$ Interpretation of R^2:
d. Compute a 95% confidence interval for the <u>average value of Y</u> given $x_p = 28$	

Q3. The following Minitab output is the result of a multiple regression analysis in which we are interested in explaining the variation in retail price (**Y**) of personal computers based on four independent variables, monitor included (1=Yes, 0=No) (**X1**), CPU *Speed* in Mhz (**X2**), *RAM* in MB's (**X3**), and *Hard drive* capacity in GB's (**X4**).

Regression Analysis: Y versus X1; X2; X3; X4; X2X4

```
The regression equation is
Y = 1404 + 49 X1 - 3.37 X2 + 4.72 X3 - 105 X4 + 0.644 X2X4

Predictor        Coef     SE Coef          T        P       VIF
Constant         1404        1765       0.80    0.433
X1               48.7       240.5       0.20    0.841       1.0
X2             -3.372       4.689      -0.72    0.478       8.3
X3              4.721       3.005       1.57    0.127       2.2
X4             -104.9       304.6      -0.34    0.733     133.3
X2X4           0.6442      0.6967       0.92    0.363     176.2

S = 697.0       R-Sq = 70.5%     R-Sq(adj) = 65.5%


Analysis of Variance

Source            DF          SS          MS         F        P
Regression         5    34753583     6950717     14.31    0.000
Residual Error    30    14573666      485789
Total             35    49327250

Source       DF      Seq SS
X1            1      252592
X2            1    21234267
X3            1     5713693
X4            1     7137818
X2X4          1      415213

Unusual Observations
Obs         X1           Y         Fit      SE Fit    Residual    St Resid
 23       1.00       1900        3364         441       -1464       -2.71R
 24       1.00       6360        4511         440        1849        3.42R

R denotes an observation with a large standardized residual

Durbin-Watson statistic = 2.07

Predicted Values for New Observations

New Obs     Fit     SE Fit         95.0% CI             95.0% PI
1          1170        259   (     640;    1700)  (    -349;    2689)

Values of Predictors for New Observations

New Obs        X1        X2        X3        X4      X2X4
1            1.00       400      64.0      5.00      2000
```

Best Subsets Regression: Y versus X1; X2; X2X4

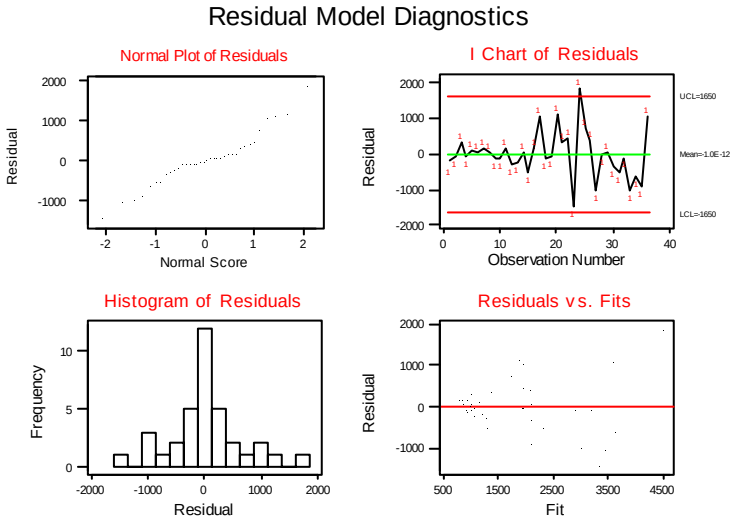
Response is Y

Vars	R-Sq	R-Sq(adj)	C-p	S	X1	X2	X4
1	67.7	66.8	0.3	684.26			X
1	42.8	41.2	25.2	910.58		X	
2	68.0	66.0	2.0	691.92		X	X
2	67.8	65.8	2.2	693.89	X		X
3	68.0	65.0	4.0	702.25	X	X	X

Correlations: Y; X1; X2; X3; X4

	Y	X1	X2	X3
X1	0.072			
	0.678			
X2	0.655	-0.020		
	0.000	0.910		
X3	0.691	0.045	0.658	
	0.000	0.795	0.000	
X4	0.819	0.083	0.761	0.708
	0.000	0.632	0.000	0.000

Cell Contents: Pearson correlation
P-Value



Given this output and your knowledge of multiple regression, answer the following;

a. The slope of the *Speed* variable is	
b. Is the relationship between *RAM* and *Hard drive* significant? Why?	H_o: H_A:
	Decision:

c. Are the independent variables significant in explaining the variation in the *Price*? Why?	H_o: H_A:
	Decision:
d. Do *Speed* and *Hard drive* interact on varying the value of the *Price*? Why?	H_o: H_A:
	Decision:
e. Check the assumptions of the multiple regression	
f. What will be the *Price* of a computer *including* the monitor, has a *Speed* of 400 Mhz, a *RAM* of 64 MB's and a *Hard drive* capacity of 5 GB's?	
g. A 95% CI for the average Price of computers with the specs in (f) is	
h. The two variables that have no severe multicolinearity are	
i. A 99% CI for the slope of the *Hard drive capacity* of the computer is	
j. The percentage of variation in *Price* explained by the indep. variables is	
k. The estimated variance of the regression model is	
l. Using the best subset regression option, what is the best group of indep. variables that explain the variation in the *Price*? And has a C-p value of	

Q4. The U.S. Golf Association undertook a study of two brands of golf balls with the objective to see whether there is a consistency in the distance (in feet) that the two golf ball brands will fly off the tee. To conduct the test, the U.S.G.A. uses a robot named "Iron Byron," which swings the club at the same speed and with the same swing pattern each time it is used. The following data reflect sample data for a random sample of balls of each brand.

											Σx	Σx^2
Brand A	234	236	230	227	234	233	228	229	230	238	2319	537895
Brand B	240	236	241	236	239	243	230	239	243	240	2387	569913

Using 2.5% sig. level, do you think that the variation in brand A is not greater than that in brand B?

Hypotheses:	H_o: H_A:
Check the assumptions:	1. 2.
Test statistic =	
Critical value =	
Decision rule:	
Decision:	
Conclusion:	

Q5. A maker of toothpaste is interested in testing whether the proportion of adults (over age 18) who use their toothpaste and have no cavities within a six-month period is any different than the proportion of children (18 and under) who use the toothpaste and have no cavities within a six-month period. To test this, they have selected a sample of adults and a sample of children randomly from the population of those customers who use their toothpaste. The following results were observed.

	Adults	Children
Sample Size	**100**	**200**
Number with 0 cavities	**83**	**165**

Using the p-value approach and a significance level of 0.05, do you think that adults are different than children?

Hypotheses:	H_o: H_A:
Check the assumptions:	1. 2. 3.
Test statistic =	
P-value =	
Decision rule:	
Decision:	
Conclusion:	

Q6. Referring to question (2) above, if the manufacturer of brand A balls thinks that he average flying distance for the balls is approximately 235 feet. Do the data provide sufficient evidence to support the claim of the manufacturer? Use 2% significance level.

Hypotheses:	H_o: H_A:
Check the assumptions:	1. 2. 3.
Test statistic =	
Critical value =	
Decision rule:	
Decision:	
Conclusion:	

Q7.

a. The following table represents the total sales of a big company (In millions of S.R.)

Years	1997	1998	1999	2000
Sales	1.5	2.01	2.25	2.35

Use the year 1997 as the base year to find the simple index value for the year 1999, and interpret its value

$I_{1999} =$	Interpretation:

b. The following table represents the expenses of a big university for three years (In millions of S.R.)

Year	Salaries	Lab Materials	Housing	Maintenance
1999	2	0.15	0.5	0.75
2000	2.1	0.16	0.45	0.6
2001	2.3	0.2	0.55	0.8

Calculate an unweighted aggregate price index for the year 2001 using 1999 as the base year, and interpret its value.

$I_{2001} =$	Interpretation:

c. The following values represent advertising rates paid by a regional catalog retailer that advertises either on TV or in newspaper (In S.R.)

Year	TV Ad.	%on TV Ad.	Newspapers Ad.
1	1050	30	1400
2	1085	35	1470
3	1115	35	1610
4	1330	45	2240

I. Find the Paashe index for the year 3 using year 1 as a base year, and interpret its value

$I_3 =$	Interpretation:

II. Find the Laspeyres index for the year 4 using year 1 as a base year, and interpret its value

$I_4 =$	Interpretation:

d. Al-Riyadh bank has two major branches, one in Jeddah, and the other in Al-Riyadh. The manger of the bank wants to evaluate the number of new clients in each quarter, the bank considered four years 1999 – 2002. The seasonal index for each quarter is given below

Quarter	1	2	3	4
Seasonal Index	1.0323	0.9236	1.0823	0.9745

I. Normalize the seasonal index values

Quarter	1	2	3	4
Normal Seasonal Index values				

II. Suppose that the seasonally unadjusted forecast for the 3rd quarter of 2002 is 263.6149, find the adjusted forecasted value for the 3rd quarter of 2002.

Adjusted forecasted value =

III. Find the desesonalized value of $y_t = 190$ if it was in the 2nd quarter of the year 1999.

Desesonalized value =

e. The following data represents the total sales for a certain market for the previous four years (In 100,000 S.R.)

Year	Sales	Forecasted value F_t
2001	780	780
2002	815	780
2003	795	787
2004	820	788.6

I. Using the values in the above table, find the single exponential smoothing forecast value for the year 2005 using $\alpha = 0.20$

The forecasted value for the year 2005 =

II. Using $\alpha = 0.20$ and $\beta = 0.25$, find the double exponential smoothing forecast value for the year 2002, given that the fitted line equation for the sales is $\hat{y}_t = 777.50 + 10\,t$, $y_1 = 780$

C_0 =	T_0 =
C_1 =	T_1 =
F_2 =	

III. If MAD for the single and double exponential smoothing methods are 16.445, 11.246 respectively, which method is better? Why?

Answer: Reason:

With Our Best Wishes