University of Swaziland

Department of Statistics and Demography

Main Examination 2015/16

Course Title	:	**Introduction to Regression Analysis**
Course Code	:	**ST 304**
Time allowed	:	**2 (two) hours**
Requirements	:	**Scientific calculator and Statistical tables**
Instructions	:	**Answer any 4(four) questions.** **All questions carry equal marks**

Do not open the paper until you are granted permission to do so by the invigilator

Question 1

Use the SPSS output below to determine the presence of multicollinearity

a) Find the Variance Inflation Factor (VIF) and comment on the correlation of the two predictor variables. [10]

b) What are the effects of multicollinearity on; [10]
   i) the t- statistic
   ii) the sample coefficients

c) How would you deal with the problem of multicollinearity here? [5]

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Experience, Age[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Salary

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.879[a]	.773	.708	5.48362

a. Predictors: (Constant), Experience, Age

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	715.110	2	357.555	11.891	.006[a]
	Residual	210.490	7	30.070		
	Total	925.600	9			

a. Predictors: (Constant), Experience, Age

b. Dependent Variable: Salary

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	36.117	20.964		1.723	.129
	Age	-.014	.795	-.013	-.018	.986
	Experience	.890	.741	.892	1.202	.269

a. Dependent Variable: Salary

Coefficient Correlations[a]

Model			Experience	Age
1	Correlations	Experience	1.000	-.970
		Age	-.970	1.000
	Covariances	Experience	.549	-.571
		Age	-.571	.632

a. Dependent Variable: Salary

Question 2

The following data was taken from a sample of 10 observations.

Y	X
0.41	40
0.74	60
0.69	60
0.7	66
0.88	78
1.13	105
1.07	98
1.04	98
1.06	105
1.15	119

Given that the sample least square line is $\hat{y} = b_0 + b_1 x$ and

$f(b_0,b_1) = \sum(y - b_0 - b_1 x)^2$,

Where, b_0 = intercept and b_1 = slope

a) Use your knowledge of calculus to derive the formula for; [10]
   i) the intercept
   ii) the slope

b) Find the value of the intercept [5]
c) Find the value of the slope [5]
d) Find the sample correlation coefficient (r) and comment on the relationship between the two variables. [5]

Question 3

The following regression equation was calculated from data set of 20 observations:
$\hat{Y}=1.0 + 2.5X_1 + 4.0X_2$. The value of MSR and MSE are 0.465 and 0.004, respectively. The standard deviations of the regression coefficients of X_1 and X_2 are 0.26 and 0.21respectively.

i) Test the hypothesis that at least one of the regression coefficients is not equal to zero, use $\alpha = 0.05$ [10]
ii) Test H_0: $\beta_1 = 0$ versus H_1: $\beta_1 \neq 0$ Use $\alpha = 0.05$ [7.5]
iii) Test H_0: $\beta_2 = 0$ versus H_1: $\beta_2 \neq 0$ Use $\alpha = 0.05$ [7.5]

Question 4

Highway miles per gallon (mpg) for large pickups range from 14 to 23. Typically, the more horsepower and the larger the engine size, the lower the mpg. A random sample of eight pickups is selected.

Pickup	MPG	Horse-power	Engine Size
Chevrolet Avalanche	15	285	5.3
Chevrolet Silverado 2500	16	300	6.0
Dodge Ram 1500	19	215	3.7
F-150Regular Cab	21	202	4.2
GMC1500Regular Cab	23	200	4.3
GMC Sierra Denali	14	325	6.0
Lincoln Blackwood	17	300	5.4
Toyota Tundra Regular	15	190	3.4

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.785[a]	.616	.462	2.35150

a. Predictors: (Constant), Engine size, Horse power

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.352				
	Residual	27.648				
	Total	72.000				

a. Predictors: (Constant), Engine size, Horse power

b. Dependent Variable: MPG

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	25.114	4.267		5.885	.002
	Horse power	-.115	.050	-1.984	-2.292	.071
	Engine size	4.463	2.724	1.418	1.638	.162

a. Dependent Variable: MPG

Given the partial SPSS output above:

i) Find the 95% confidence intervals for the regression coefficients of a model using horsepower and engine size to predict mpg. [15]

ii) Interpret these two confidence intervals. [10]

Question 5

Excel output summaries from a sample of 10 homes are given below, where

Y = Home size in hundreds of square metres
X_1 = Annual income in thousands of Emalangeni
X_2 = Family size
X_3 = Combined years of formal education (beyond high school) for all household wage earners

SUMMARY OUTPUT (Initial Model)

Regression Statistics	
Multiple R	....
R Square	0.905265
Adjusted R Square	....
Standard Error	2.035454
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	....	237.5416	....	....	0.001792
Residual	....	....	4.143075		
Total	....	262.4	....		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	5.656717	2.834204	....	0.092956
Income	0.193878	0.0877	....	0.069076
Family size	2.338108	0.907791	....	0.042017
Education	-0.16277	0.244071	....	0.529633

a) Find the missing values for the Regression Statistics and the ANOVA tables in the initial model [5]

b) Using $\alpha = 0.10$, are all the predictor variables necessary in the model? [5]

SUMMARY OUTPUT (Final Model)

Regression Statistics	
Multiple R	....
R Square	0.898243
Adjusted R Square	0.869169
Standard Error	....
Observations	10

ANOVA

	df	*SS*	*MS*	*F*
Regression	2	235.6989	117.8495	30.89559
Residual	7	26.7011	3.814442	
Total	9	262.4		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	5.091053	2.594836	1.961994	0.090553
Income	0.164853	0.073055	2.25654	0.05863
Family size	2.65694	0.740463	3.588214	0.008878

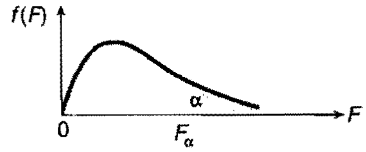
c) Find the missing values for the Regression Statistics in the final model [5]

d) Carry out an F-test for the final model [5]

e) Construct the 90% confidence interval for the two parameters [5]

End of Examination

TABLE A.7a Percentage points of the F distribution. $\alpha = .10$.

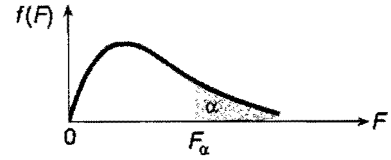


Numerator Degrees of Freedom

Denominator Degrees of Freedom

ν_2 \ ν_1	1	2	3	4	5	6	7	8
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67

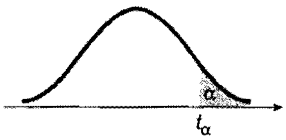
TABLE A.7b Percentage points of the F distribution. $\alpha = .05$.



Numerator Degrees of Freedom

Denominator Degrees of Freedom ν_2 \ ν_1	1	2	3	4	5	6	7	8
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94

Critical values of *t*.



Degrees of Freedom	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
35	1.306	1.690	2.030	2.438	2.724
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

TABLE A.9a Critical values for the Durbin–Watson DW statistic. $\alpha = .05$.

n	$k=1$ d_L	$k=1$ d_U	$k=2$ d_L	$k=2$ d_U	$k=3$ d_L	$k=3$ d_U	$k=4$ d_L	$k=4$ d_U	$k=5$ d_L	$k=5$ d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika*, 1951, *30*, 159–178. Reproduced by permission of the *Biometrika* trustees.