

DEPARTMENT OF STATISTICS AND DEMOGRAPHY

MAIN EXAMINATION, 2008/9

COURSE TITLE: INTRODUCTION TO REGRESSION ANALYSIS

COURSE CODE: ST 304

TIME ALLOWED: TWO (2) HOURS

INSTRUCTION: ANSWER ANY FOUR QUESTIONS
ALL QUESTIONS CARRY EQUAL MARKS (15 MARKS)

SPECIAL REQUIREMENTS: SCIENTIFIC CALCULATORS AND STATISTICAL TABLES

**DO NOT OPEN THIS PAGE UNTIL PERMISSION HAS BEEN GRANTED BY THE
INVIGILATOR**

Question 1

A car dealer lists the following used car prices for a certain car with age X measured in years and selling price Y measured in thousands of Emalangeni.

X	1	2	3	4	5	6	7	8	9	10
Y	2.45	1.8	2.0	2.0	1.7	1.2	1.15	0.69	0.6	0.47

- Determine the equation of the Least Squares regression line.
- Construct a 95% confidence interval for the slope of the regression line.
- Determine the predicted value for the average selling price of a five-year-old model and construct a 95% confidence interval.

(5+5+5 marks)

Question 2

In a study of factors thought to be related to admission patterns of a large general hospital, the administrator collected the following data on ten communities in the hospital's catchment's area:

Persons admitted per 1000 population during study period (Y)	Index of availability of other health services (X_1)	Index of indigence (X_2)
61.6	6.0	6.3
53.2	4.4	5.5
65.5	9.1	3.6
64.9	8.1	5.8
72.7	9.7	6.8
52.5	4.8	7.9
50.2	7.6	4.2
44.0	4.4	6.0
53.8	9.1	2.8
53.5	6.7	6.7

Given that:

$$\sum X_1^2 = 525.73, \sum X_1 X_2 = 374.31, \sum Y^2 = 33349.92, \sum X_2^2 = 331.56, \sum X_1 Y = 4104.32, \sum X_2 Y = 3185.94$$

- Obtain the regression equation of Y on X_1 and X_2 .
- Predict the admission population when $X_1=11.5$ and $X_2=5$, using the fitted regression model.

(10+5 marks)

Question 3

(a) Given the following results from the regression model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_i$

t	1	2	3	4	5	6	7	8	9
$Y - \hat{Y}$	-0.418	-0.350	0.507	-0.374	-0.181	0.652	0.256	0.342	-0.434

Obtain the estimate of ρ and test at 5% level of significance for positive correlation given that $n = 50$. state clearly the hypotheses, decision rule and conclusion. **(15 marks)**

Question 4

- (a) Explain the meaning of heteroscedasticity. What happens to the OLS estimators in a simple linear regression model in the presence of heteroscedasticity, if all other assumptions hold? **(5 marks)**
- (b) When and how would the method of Generalised Least Squares (GLS) help tackle the problem of Heteroscedasticity? **(5 marks)**
- (c) For a simple regression model $y_i = \beta_0 + \beta_1 X_i + \mu_i$, fitted to data with 10 sample values, the values of X_i and the respective absolute values $|e_i|$ of the residuals are:

X_i	12.1	21.4	18.7	21.7	12.5	10.4	20.8	10.2	16.0	12.0
$ e_i $	1.03	1.24	0.20	0.22	0.26	0.59	0.83	0.10	0.06	0.03

Perform the Spearman's test of Heteroscedasticity.

(5 marks)

Question 5

Data concerning the weekly fuel consumption (Y), average hourly temperature (X1) and chill index (X2) at a certain factory has been observed for the last eight years.

Week (i)	Average Hourly temperature	Chill Index	Fuel Consumption
1	28	18	12.4
2	28	14	11.7
3	32.5	24	12.4
4	39	22	10.8
5	45.9	8	9.4
6	57.8	16	9.5
7	58.1	1	8
8	62.5	0	7.5

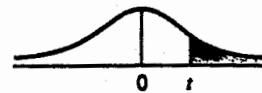
- (a) Fit the regression model on the fuel consumption data.
(b) Test for the significance of the coefficients at 5% level of significance.

(8+7 marks)

END OF EXAM!!

Table VIII The *t* Distribution Table[†]

The entries in the table give the critical values of *t* for the specified number of degrees of freedom and areas in the right tail.



<i>df</i>	Area in the Right Tail under the <i>t</i> Distribution Curve					
	.10	.05	.025	.01	.005	.001
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
31	1.309	1.696	2.040	2.453	2.744	3.375
32	1.309	1.694	2.037	2.449	2.738	3.365
33	1.308	1.692	2.035	2.445	2.733	3.356
34	1.307	1.691	2.032	2.441	2.728	3.348
35	1.306	1.690	2.030	2.438	2.724	3.340
36	1.306	1.688	2.028	2.434	2.719	3.333
37	1.305	1.687	2.026	2.431	2.715	3.326
38	1.304	1.686	2.024	2.429	2.712	3.319
39	1.304	1.685	2.023	2.426	2.708	3.313
40	1.303	1.684	2.021	2.423	2.704	3.307
∞	1.282	1.645	1.960	2.326	2.576	3.090

[†]This table is an abbreviated version of Table VIII that appears in Appendix C. This table goes up to 40 degrees of freedom. For degrees of freedom from 41 to 70, use Table VIII of Appendix C.

TABLE A5 A Durbin-Watson Table: Values of d_{LW} and d_{UL}

n	k=1		k=2		k=3		k=4		k=5	
	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}
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	1.88
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	1.60	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.61	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

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika* 30 (1951), 159-178. Reproduced by permission of the Biometrika Trustees.

TABLE A6 A Durbin-Watson Table: Values of d_{LW} and d_{UL}

n	k=1		k=2		k=3		k=4		k=5	
	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}	d_{LW}	d_{UL}
15	.81	1.07	.70	1.25	.59	1.46	.49	1.70	.39	1.96
16	.84	1.09	.74	1.25	.63	1.44	.53	1.66	.44	1.90
17	.87	1.10	.77	1.25	.67	1.43	.57	1.63	.48	1.85
18	.90	1.12	.80	1.26	.71	1.42	.61	1.60	.52	1.80
19	.93	1.13	.83	1.26	.74	1.41	.65	1.58	.56	1.77
20	.95	1.15	.86	1.27	.77	1.41	.68	1.57	.60	1.74
21	.97	1.16	.89	1.27	.80	1.41	.72	1.55	.63	1.71
22	1.00	1.17	.91	1.28	.83	1.40	.75	1.54	.66	1.69
23	1.02	1.19	.94	1.29	.86	1.40	.77	1.53	.70	1.67
24	1.04	1.20	.96	1.30	.88	1.41	.80	1.53	.72	1.66
25	1.05	1.21	.98	1.30	.90	1.41	.83	1.52	.75	1.65
26	1.07	1.22	1.00	1.31	.93	1.41	.85	1.52	.78	1.64
27	1.09	1.23	1.02	1.32	.95	1.41	.88	1.51	.81	1.63
28	1.10	1.24	1.04	1.32	.97	1.41	.90	1.51	.83	1.62
29	1.12	1.25	1.05	1.33	.99	1.42	.92	1.51	.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	.94	1.51	.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	.96	1.51	.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	.98	1.51	.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.18	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.59
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.60
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

Source: From J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika* 30 (1951), 159-178. Reproduced by permission of the Biometrika Trustees.