

UNIVERSITY OF SWAZILAND

FINAL EXAMINATION PAPER 2014

TITLE OF PAPER : LINEAR STATISTICAL METHODS
COURSE CODE : ST204
TIME ALLOWED : 2 (TWO) HOURS
**REQUIRMENTS : STATISTICAL TABLES
AND CALCULATOR**
**INSTRUCTIONS : ANSWER ANY 4 (FOUR) QUESTIONS.
ALL QUESTIONS CARRY EQUAL MARKS.**

**THIS PAPER IS NOT TO BE OPENED UNTIL PERMISSION HAS
BEEN GRANTED BY THE INVIGILATOR**

QUESTION ONE.

[4 + 4 + 3 + 3 + 7 + 4 marks]

- 1.1 State the Simple Linear Regression Model with distribution of error terms unspecified.
- 1.2 State the Simple Linear Regression Model when distribution of error terms is normal.
- 1.3 Explain the similarities and differences of the above two models.
- 1.4 Explain why normality assumption is important in the above regression model in part 1.2.
- 1.5 State the estimated (fitted) regression function and discuss the important properties of the fitted regression line.
- 1.6 Explain why we consider the scope of the model includes (or does not include) the zero value for the independent variable (x) in interpreting the meaning of the parameters β_0 and β_1 . Use the numerical values of $\hat{\beta}_0 = 10$ and $\hat{\beta}_1 = 3$ in your explanation of the meaning of the parameters β_0 and β_1 .

QUESTION TWO.

[2 + 2 + 4 + 1 + 6 + 6 + 4 marks]

A production manager has compared the dexterity test scores of five assembly-line employees with hourly productivity (units produced per hour). The following two tables are a part of the output; obtained by running the model, $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ using SPSS:

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	234.000	1	234.000	30.789	.012
	Residual	22.800	3	7.600		
	Total	256.800	4			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	Constant	19.200	7.669		2.504	.087
	x	3.000	.541	.955	5.549	.012

- 2.1. Find the fitted regression line.
- 2.2. What would be expected productivity of an employee who scored 15 on the test?
- 2.3. State the null and alternative hypotheses for the Goodness-of-Fit test of the said model and explain clearly the conclusion of the test.
- 2.4. What is the estimated value of σ^2 ?
- 2.5. Test $\beta_0 = 20$ against $\beta_0 < 20$ at $\alpha = 0.01$.
- 2.6. Test $\beta_1 = 2$ against $\beta_1 > 2$ at $\alpha = 0.05$.
- 2.7. Compute the coefficient of correlation, r and explain the nature and strength of the relationship between dependent and independent variables.

QUESTION THREE.

[4 + 2 + 2 + 4 + 2 + 1 + 6 + 4 marks]

- 3.1 a. Define Coefficient of Correlation and Coefficient of Determination.
 b. Discuss how we use above two measures in describing linear association between the independent and dependent variables.
- 3.2 An educator wants to see how the number of absences a student in her class (of seven students) has affected the student's final grade. The following quantities have been calculated from the data:

$$\sum x = 38 \quad \sum y = 482 \quad \sum xy = 2776 \quad \sum x^2 = 338 \quad \sum y^2 = 39526$$

- a. Identify the response variable and the predictor variable.
 b. Fit the regression line, $Y_i = \beta_0 + \beta_1 X_i$.
 c. Interpret the estimated values of β_0 and β_1 .
 d. What would be the expected final mark if a student misses seven classes?
 e. Construct a 99% confidence interval for β_1 .
 f. Compute the coefficient of determination and interpret the value.

QUESTION FOUR.

[5 + 6 + 3 + 5 + 4 + 2 marks]

- 4.1 Compare Regression Analysis and Analysis of Variance in terms of their similarities and differences.
 4.2 State and compare the *Cell Means Model* and *Factor Effects Model* for single factor studies.
 4.3 An industrial sales manager, testing the effectiveness of three different sales presentations, randomly selected a presentation to be used when making the next sales call on 14 customers. The number of units purchased as a result of each sales call are shown below:

<u>Presentation 1</u>	<u>Presentation 2</u>	<u>Presentation 3</u>
7	13	14
9	10	15
11	11	12
7	13	10
	14	13

- a. Identify the dependent variable, factor studied and factor levels.
 b. Complete the computation of the ANOVA table.
 c. Use the 0.05 level in determining whether the three sales presentations could be equally effective. Clearly state all the steps in the test between the hypotheses and the conclusion.
 d. Which sales presentation would you prefer? Explain.

QUESTION FIVE.

[8 + 1 + 2 + 2 + 4 + 2 + 2 + 2 + 2 marks]

- 5.1 State the Cell Means Model for two-factor studies with equal sample sizes and its important features.
- 5.2 A company wishes to test the effectiveness of its advertising. A product is selected, and two types of ads are written; one is serious and one is humorous. Also the ads run on both mediums of advertising; television and radio. Sixteen potential customers are selected and assigned randomly to one of the four groups. After seeing or listening to the ad, each customer is asked to rate its effectiveness on a scale of 1 to 20 and the data were analyzed using SPSS. The following ANOVA table is a part of the output from that analysis:

ANOVA TABLE

Source of Variation	Sum of Squares	df	Mean Square	F
Between treatments	186.189			
Factor A	10.563			
Factor B	175.563			
A X B	0.063			
Within treatments	66.250			
Total	252.439			

- What is size of the sample used in this experiment?
- What are the treatments in this experiment?
- Identify the Factor A and the Factor B.
- Complete the ANOVA Table.

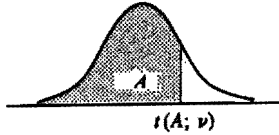
For the following tests:

State only the decision and explain the conclusions (based on F-test). You do not need to write all steps of F-test. Use $\alpha = 0.01$ for all tests.

- Test whether the effectiveness of the advertising is influenced by the types of ads.
- Test whether the effectiveness of the advertising is influenced by the mediums of advertising.
- Test whether the effectiveness of the advertising is influenced by the interactions between the types of ads and the mediums of advertising.
- Test whether the effectiveness of the advertising is influenced by the treatments.

TABLE A.2 Percentiles of the *t* Distribution

Entry is $t(A; \nu)$ where $P\{t(\nu) \leq t(A; \nu)\} = A$



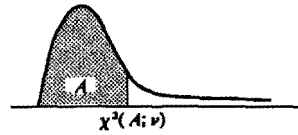
ν	A						
	.60	.70	.80	.85	.90	.95	.975
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

TABLE A.2 (concluded) Percentiles of the *t* Distribution

ν	A						
	.98	.985	.99	.9925	.995	.9975	.9995
1	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	2.197	2.336	2.528	2.661	2.845	3.153	3.849
21	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	2.123	2.250	2.423	2.542	2.704	2.971	3.551
60	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	2.054	2.170	2.326	2.432	2.576	2.807	3.291

TABLE A.3 Percentiles of the χ^2 Distribution

Entry is $\chi^2(A; \nu)$ where $P\{\chi^2(\nu) \leq \chi^2(A; \nu)\} = A$

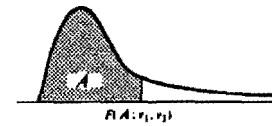


ν	A									
	.005	.010	.025	.050	.100	.900	.950	.975	.990	.995
1	0.0*393	0.0*157	0.0*982	0.0*393	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2

Source: Reprinted, with permission, from C. M. Thompson, "Table of Percentage Points of the Chi-Square Distribution," *Biometrika* 32 (1941), pp. 188-89.

TABLE A.4 Percentiles of the F Distribution

Entry is $F(A; \nu_1, \nu_2)$ where $P\{F(\nu_1, \nu_2) \leq F(A; \nu_1, \nu_2)\} = A$



$$F(A; \nu_1, \nu_2) = \frac{1}{F(1-A; \nu_2, \nu_1)}$$

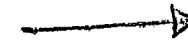


TABLE A.4 (continued) Percentiles of the F Distribution

Den. df	A	Numerator df								
		1	2	3	4	5	6	7	8	9
30	.50	0.466	0.709	0.807	0.858	0.890	0.912	0.927	0.939	0.948
	.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.975	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.995	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
.999	13.3	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	
60	.50	0.461	0.701	0.798	0.849	0.880	0.901	0.917	0.928	0.937
	.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.975	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.995	8.49	5.80	4.73	4.14	3.76	3.49	3.29	3.13	3.01
.999	12.0	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69	
120	.50	0.458	0.697	0.793	0.844	0.875	0.896	0.912	0.923	0.932
	.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.95	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
	.975	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.995	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
.999	11.4	7.32	5.78	4.95	4.42	4.04	3.77	3.55	3.38	
∞	.50	0.455	0.693	0.789	0.839	0.870	0.891	0.907	0.918	0.927
	.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	.975	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.995	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62
.999	10.8	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10	

TABLE A.4 (concluded) Percentiles of the F Distribution

Den. df	A	Numerator df								
		10	12	15	20	24	30	60	120	∞
30	.50	0.955	0.966	0.978	0.989	0.994	1.00	1.01	1.02	1.02
	.90	1.82	1.77	1.72	1.67	1.64	1.61	1.54	1.50	1.46
	.95	2.16	2.09	2.01	1.93	1.89	1.84	1.74	1.68	1.62
	.975	2.51	2.41	2.31	2.20	2.14	2.07	1.94	1.87	1.79
	.99	2.98	2.84	2.70	2.55	2.47	2.39	2.21	2.11	2.01
	.995	3.34	3.18	3.01	2.82	2.73	2.63	2.42	2.30	2.18
.999	4.24	4.00	3.75	3.49	3.36	3.22	2.92	2.76	2.59	
60	.50	0.945	0.956	0.967	0.978	0.983	0.989	1.00	1.01	1.01
	.90	1.71	1.66	1.60	1.54	1.51	1.48	1.40	1.35	1.29
	.95	1.99	1.92	1.84	1.75	1.70	1.65	1.53	1.47	1.39
	.975	2.27	2.17	2.06	1.94	1.88	1.82	1.67	1.58	1.48
	.99	2.63	2.50	2.35	2.20	2.12	2.03	1.84	1.73	1.60
	.995	2.90	2.74	2.57	2.39	2.29	2.19	1.96	1.83	1.69
.999	3.54	3.32	3.08	2.83	2.69	2.55	2.25	2.08	1.89	
120	.50	0.939	0.950	0.961	0.972	0.978	0.983	0.994	1.00	1.01
	.90	1.65	1.60	1.55	1.48	1.45	1.41	1.32	1.26	1.19
	.95	1.91	1.83	1.75	1.66	1.61	1.55	1.43	1.35	1.25
	.975	2.16	2.05	1.95	1.82	1.76	1.69	1.53	1.43	1.31
	.99	2.47	2.34	2.19	2.03	1.95	1.86	1.66	1.53	1.38
	.995	2.71	2.54	2.37	2.19	2.09	1.98	1.75	1.61	1.43
.999	3.24	3.02	2.78	2.53	2.40	2.26	1.95	1.77	1.54	
∞	.50	0.934	0.945	0.956	0.967	0.972	0.978	0.989	0.994	1.00
	.90	1.60	1.55	1.49	1.42	1.38	1.34	1.24	1.17	1.00
	.95	1.83	1.75	1.67	1.57	1.52	1.46	1.32	1.22	1.00
	.975	2.05	1.94	1.83	1.71	1.64	1.57	1.39	1.27	1.00
	.99	2.32	2.18	2.04	1.88	1.79	1.70	1.47	1.32	1.00
	.995	2.52	2.36	2.19	2.00	1.90	1.79	1.53	1.36	1.00
.999	2.96	2.74	2.51	2.27	2.13	1.99	1.66	1.45	1.00	

Source: Reprinted from Table 5 of Pearson and Hartley, *Biometrika Tables for Statisticians*, Volume 2, 1972, published by the Cambridge University Press, on behalf of The Biometrika Society, by permission of the authors and publishers.