KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2016/2017

SECOND SEMESTER EXAMINATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN BUSINESS

BMS 902: BUSINESS STATISTICS II

DATE: THURSDAY 18TH MAY 2017 **TIME: 5.30 P.M. - 8.30 P.M.**

INSTRUCTIONS: Answer Question ONE and any other Three Questions

Question One (40 Marks)

a) Briefly explain five assumptions that are usually made about the unobservable error terms in the classical linear regression model (CLRM)? Why are these assumptions made? (12 marks)

b) A sample survey of seven families was taken and figures were obtained with respect to their annual savings Y (in '000), annual income X_1 (in '000) and family size X_2. The data is summarized in the table below:

Family	A	B	C	D	E	F	G
Annual Savings (Y)	5	4	3	2	6	7	9
Annual Income (X_1)	6	6	4	4	7	8	10
Family Size (X_2)	4	3	2	2	5	6	6

Required

i) Find the least squares regression equation of Y on X_1 and X_2. (10 marks)

ii) Estimate the annual savings of a family whose annual income is Shs.16,000 and family size is 4. (2 marks)

iii) Compute the coefficient of multiple determination and the adjusted value and interpret. (4 marks)

iv) Test the significance of the individual regression coefficient at the 5% level of significance (8 marks)

v) Test the overall significance of the regression at 5% level. (4 marks)

Question Two (20 marks)

a) What is meant by perfect multicollinearity in multiple regression analysis? (2 marks)

b) A researcher estimates the following model for stock market returns, but thinks that there may be a problem with it.

$$Y_t = 0.638 + 0.402X_{1t} - 0.891X_{2t} \quad \text{with } R^2 = 0.96, \text{ adj } R^2 = 0.89$$
$$(0.436) \quad (0.291) \quad (0.763)$$

i) By calculating the t - ratios, and considering their significance and by examining the value of R^2, suggest what the problem might be. (5 marks)

ii) How might you go about solving the perceived problem? (3 marks)

c) An Econometrician at Kenyatta University was performing a regression analysis, but realized that her model could be violating the OLS assumption of homoscedasticity.

i) What are the consequences of her just proceeding with her analysis? (3 marks)

ii) Discuss the white's test for this problem (4 marks)

iii) What can she do to remedy this problem? (3 marks)

Question Three (20 Marks)

a) List the conditions under which Durbin-Watson test is valid for no autocorrelation. (4 marks)

b) An econometrician suspects that the residuals of her model might be auto correlated. Explain the steps involved in testing this theory using the Durbin-Watson test. (6 marks)

c) The econometrician follows your guidance in part (c) and calculates a value for the Durbin-Watson statistic of 0.95. The regression has 60 quarterly observations and three explanatory variables (plus a constant term). Perform the test. What is your conclusion? (4 marks)

d) What are the advantages of panel data over cross-section or time series data? (4 marks)

e) Distinguish between balanced and unbalanced panels, giving examples of each. (2 marks)

Question Four (20 marks)

Study the Minitab regression output that follows.

Regression Analysis: Y versus X1, X2, X3, X4

Predictor	Coef	SE Coef	T	P
Constant	-55.93	24.22	-2.31	0.025
X1	**M**	0.02100	0.50	0.619
X2	-0.10720	0.03503	**P**	0.003
X3	0.57922	**Q**	7.59	0.000
X4	-0.8695	0.1498	-5.81	0.000

Analysis of Variance

Source	DF	SS	MS	F
Regression	4	**V**	4522.1	**W**
Residual Error	55	4479.7	81.4	
Total	59	22568.2		

Using the results from the two tables above,

i) Calculate the values of **M, P, Q, V** and **W** indicated in the cells in the tables. (10 marks)

ii) Write down the fitted regression function. (3 marks)

iii) Test for the significance of overall regression at the 5% level of significance. (3 marks)

iv) Determine the coefficient of multiple determination and interpret its value (4 marks)

Question Five (20 marks)

a) Explain why the linear probability model is inadequate as a specification for binary response variable estimation. (4 marks)

b) Explain the dummy variable trap. What are its consequences (4 marks)

c) Suppose that a researcher, using wage data on 250 randomly selected male workers and 280 female workers, estimates the OLS regression,

$$\widehat{Wage} = \underset{(0.23)}{12.52} + \underset{(0.36)}{2.12} \times male, \quad R^2 = 0.06, \; SER = 4.2$$

where *wage* is measured in \$/hour and *male* is a binary variable that is equal to 1 if the person is male and 0 if the person is a female. Define the wage gender gap as the difference in mean earnings between men and women.

i) What is the estimated gender gap? (2 marks)

ii) Is the estimated gender gap significantly different from zero? (2 marks)

iii) Construct a 95% confidence interval for the gender gap (2 marks)

iv) In the sample, what is the mean wage of women? Of men? (2 marks)

v) Another researcher uses the same data but regresses *wages* on *female*, a variable that is equal to 1 if the person is female and 0 if the person is male. What are the regression estimates calculated from this regression?

$Wage$ = ------- + -------$\times female$, R^2 = --------, SER = ------- (4 marks)

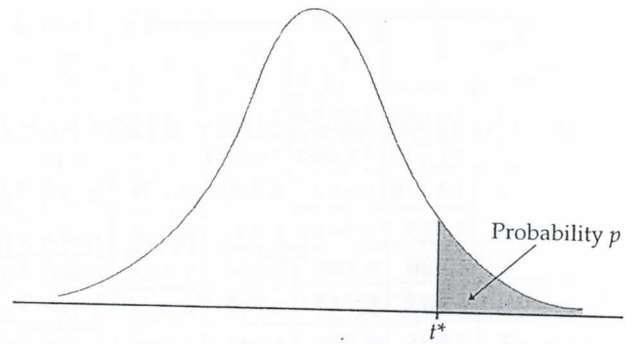


Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.

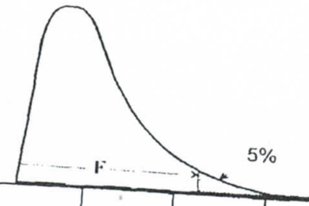
TABLE D

t distribution critical values

	Upper-tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

F-Table for 5%

Percentage points of the F-distribution: upper 5% points



df_{num} / df_{den}	1	2	3	4	5	6	7	8	9	10	15	20
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	245.95	248.01
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.429	19.446
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.703	8.660
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.858	5.803
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.619	4.558
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	3.938	3.874
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.511	3.445
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.218	3.150
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.006	2.936
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.845	2.774
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.719	2.646
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.617	2.544
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.533	2.459
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.463	2.388
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.403	2.328
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.352	2.276
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.308	2.230
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.269	2.191
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.234	2.155
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.203	2.124
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321	2.176	2.096
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297	2.151	2.071
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275	2.128	2.048
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255	2.108	2.027
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.089	2.007
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220	2.072	1.990
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204	2.056	1.974
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190	2.041	1.959
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177	2.027	1.945
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.015	1.932
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	1.924	1.839
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.836	1.748
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.750	1.659
100K	3.842	2.996	2.605	2.372	2.214	2.099	2.010	1.939	1.880	1.831	1.666	1.571

K (Multiply this value by 1000)

INVOLVEMENT IN ANY EXAMINATION IRREGULARITY SHALL LEAD TO DISCONTINUATION