UNIVERSITY OF JOHANNESBURG
UNIVERSITEIT VAN JOHANNESBURG

DEPARTMENT OF ECONOMICS AND ECONOMETRICS

JUNE/JULY SUPPLEMENTARY EXAMINATION 2014

Course / *Kursus*: Econometrics / *Ekonometrie* 2A01
Paper / *Vraestel*: Statistics for Economists
Statistiek vir Ekonome
Examiners / *Eksaminatore*: Dr J Muteba Mwamba
Mr JJ Kouadio

Time: **180 min**
Marks: **130**

Instructions:

1. ***Answer all questions***
2. ***This paper consists of 15 pages***
3. ***An Excel workfile is provided with 2 sheets. Use the Excel file provided for Section E and F.***
4. ***Please round off all quantitative answers to 4 decimal places***
5. ***Some formulae and tables are provided in Appendix***

Initials & Surname: ______________________________

Student number: ______________________________

SECTION/AFDELING	*TOTAL* / TOTAAL	*MARK* / PUNTE
A	*28*	
B	*33*	
C	*17*	
D	*22*	
E	*10*	
F	*20*	
TOTAL	*130*	

SECTION A: *Theory*

[28 marks]

Answer the following theory by providing a clear and correct answer. (2 marks each)

1. *What is the difference between continuous and discrete random variable?*

2. *Distinguish mutually exclusive events from independent events?*

3. *Give two most important properties of probability measures?*

4. *Enumerate three measures of central tendency that are equal when the distribution is symmetric?*

5. *What is an outlier observation?*

6. *What is another name for the alternate hypothesis?*

7. *What is the third step in hypothesis testing?*

8. *What is a residual in regression analysis?*

9. *What do we call a statement about the value of a population parameter?*

10. *As the sample size increases, the curve of the t-distribution approaches the ____ ?*

11. *If the absolute value of the computed value of the test statistic exceeds the critical value of the test statistic, what is our decision?*

12. *What is the difference between the null hypothesis and the alternative one?*

13. *What is a type I error in hypothesis testing?*

14. *What is the difference between a statistic and a parameter?*

SECTION / AFDELING B: *Probability*

[33 marks]

Answer the following questions and show all your calculations by using Excel or formulas.

1. $P(A_1) = 0.20$, $P(A_2) = 0.40$ and $P(A_3) = 0.40$. $P(B_1|A_1) = 0.25$, $P(B_1|A_2) = 0.05$ and $P(B_1|A_3) = 0.10$.

Use Bayes' theorem to determine $P(A_3|B_1)$.

[3 marks]

---

2. 2. *Let* $P(x)$ *be a probability distribution defined by;*

[9 marks]

$$P(x) = \begin{cases} \frac{x}{3}; \text{ for } x = 0, 1, \text{ and } 2, \\ 0 \text{ otherwise} \end{cases}$$

(a) *Show that* $P(x)$ *is indeed a probability distribution*

(b) *Calculate the mean*

(c) *Calculate the standard deviation*

---

3. *In southern California, a growing number of persons pursuing a teaching credential are choosing paid internships over traditional student teaching programs. A group of eight candidates for three local teaching positions consisted of five candidates who had enrolled in paid internships and three candidates who had enrolled in traditional student teaching programs. Assume that all eight candidates are equally qualified for the position. Let x represent the number of internship-trained candidates who are hired for these three positions.*

a) *Does x have a binomial distribution or a hypergeometric distribution? Support your answer.*

[2 marks]

----

b) *Find the probability that three internship-trained candidates are hired for these positions.*

[2 marks]

----

c) *What is the probability that none of the three hired was internship-trained?*

[2 marks]

----

d) *Find* $P(x \leq 1)$.

[2 marks]

----

4. *A pollster randomly selected 4 of 10 available people. How many different groups of 4 are possible?*

marks] *[3*

----

5. *The diameters of Douglas firs grown at a Christmas tree farm are normally distributed with a mean of 4 inches and a standard deviation of 1.5 inches.*

a) *What proportion of the trees will have diameters less than 3 inches?* *[3 marks]*

---

b) *Your Christmas tree stand will expand to a diameter of 6 inches. What proportion of the trees will not fit in your Christmas tree stand?* *[3 marks]*

---

6. *A salesperson makes six calls. Each call is a success or a failure, with the probability of success equal to 0.3.*

a) *What is the probability that the salesperson will make 3 successful calls?* *[2 marks]*

---

b) *What is the probability that the salesperson will make at least 2 successful calls?* *[2 marks]*

SECTION / AFDELING C: ANOVA **[17 marks]**

1. *Name the three assumptions of the ANOVA technique.* *[3 marks]*

---

2. *What are the criteria for the large sample test of means?* *[2 marks]*

---

3. *What are the characteristics of the F-distribution?* *[5 marks]*

---

4. *A company has four manufacturing plants, and company officials want to determine whether there is a difference in average age of workers at the four locations. The following data are the ages of randomly selected workers at each plant. Perform a one-way ANOVA to determine whether there is a significant difference in the mean average ages of the workers at the four plants. Use $\alpha = 0.01$ and show all the step in your solution.* *[7 marks]*

Plant 1	Plant 2	Plant 3	Plant 4
29	32	25	27
27	33	24	24
30	31	24	26
27	34	25	
	30	26	
	28		

<u>SECTION D:</u> *Confidence Interval* **[22 marks]**

1. A study is being conducted in a company that has 800 engineers. A random sample of 50 of these engineers reveals that the average sample age is 34.30 years. Historically the population standard deviation of the age of the company's engineers is approximately 8 years. Construct a 98% confidence interval to estimate the average age of all the engineers in this company.

[5 marks]

2. The owner of a large equipment company wants to make a rather quick estimate of the average number of days a piece of ditch digging equipment is rented out per person per time. The company has records of all rentals, but the amount of time required to conduct an audit of all accounts would be prohibitive. The owner decides to take a random sample of rental invoices. Fourteen different rentals of ditch diggers are selected randomly from the files, yielding the following data. Use these data to construct a 99% confidence interval to estimate the average number of days that a ditch digger is rented and assume that the number of days per rental is normally distributed in the population:

3 1 3 2 5 1 2 1 4 2 1 3 1 1 *[5 marks]*

3. *A study of 87 selected companies with a telemarketing operation revealed that 39% of the sampled companies had used telemarketing to assist them in order processing. Using this information, construct a 95% confidence interval of the population proportion of telemarketing companies that use telemarketing operation to assist them in order processing?* *[5 marks]*

4. *A sample of 87 professional working women showed that the average amount paid annually into a private pension fund per person was $3343, with a sample standard deviation of $1226. A sample of 76 professional working men showed that the average amount paid annually into a private pension fund per person was $5568, with a sample standard deviation of $1716. A women's activist group wants to "prove" that women do not pay as much per year as men into private pension funds. If they use an alpha of 0.001 and these sample data, will they be able to reject a null hypothesis that women annually pay the same as or more than men into private pension funds? (Show all steps).*

[7 marks]

SECTION / AFDELING E: *Descriptive Statistics* **[10 marks]**

Use the data in provided in the Excel file named: "EKM2A JUNE 2014 SUP EXAM data.xls"

The final marks (in percentage) for a group of Econometrics 2A students are given. Use the necessary descriptive statistics and a histogram to answer the following questions:

1. *Comment on the central tendency of the given marks by interpreting the relevant statistics.* *[3 marks]*

2. *Use the empirical rule to comment on the distribution of this sample.* *[3 marks]*

3. *Interpret the first class of the histogram by referring to the bin, the frequency and the cumulative frequency and what exactly each of those means for this specific sample.* *[3 marks]*

4. *Above which value was the marks 75 % of the time?* *[1 mark]*

SECTION / AFDELING F: *Regression* **[20 marks]**

Use the data provided in the Excel file named: "EKM2A JUNE 2014 SUP EXAM data.xls"

Consider Keynes' marginal propensity to consume (the rate of consumption for one unit change in income = > 0 and <1):

1. *What is the dependent?* *[2 marks]*

2. *What is the independent variable?* *[4 marks]*

3. Run the regression model; report and interpret the coefficient of independent variable?

[8 marks]

----

4. Report and interpret the coefficient of determination? *[3 marks]*

----

5. Is the independent variable statistical significant? *[3 marks]*

------------------Good Luck------------------

Appendix: Some Formulas

GEOMETRIC MEAN $GM = \sqrt[n]{(X_1)(X_2) \cdots (X_n)}$ [3–4]

POPULATION VARIANCE $\sigma^2 = \frac{\Sigma(X - \mu)^2}{N}$ [3–8]

ARITHMETIC MEAN OF GROUPED DATA $\overline{X} = \frac{\Sigma fM}{n}$ [3–12]

where:

- $\overline{X}$ is the designation for the sample mean.
- M is the midpoint of each class.
- f is the frequency in each class.
- fM is the frequency in each class times the midpoint of the class.
- ΣfM is the sum of these products.
- n is the total number of frequencies.

STANDARD DEVIATION, GROUPED DATA $s = \sqrt{\frac{\Sigma f(M - \overline{X})^2}{n - 1}}$ [3–13]

where:

- s is the symbol for the sample standard deviation.
- M is the midpoint of the class.
- f is the class frequency.
- n is the number of observations in the sample.
- $\overline{X}$ is the designation for the sample mean.

GENERAL RULE OF MULTIPLICATION $P(A \text{ and } B) = P(A)P(B|A)$ [5-6]

BAYES' THEOREM $P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}$ [5–7]

MULTIPLICATION FORMULA Total number of arrangements $= (m)(n)$ [5–8]

PERMUTATION FORMULA $${}_nP_r = \frac{n!}{(n-r)!}$$ [5–9]

where:

n is the total number of objects.
r is the number of objects selected.

Combination : $nCr = \dfrac{n!}{r!(n-r)!}$

MEAN OF A PROBABILITY DISTRIBUTION $$\mu = \Sigma[xP(x)]$$ [6–1]

VARIANCE OF A PROBABILITY DISTRIBUTION $$\sigma^2 = \Sigma[(x-\mu)^2 P(x)]$$ [6–2]

BINOMIAL PROBABILITY FORMULA $$P(x) = {}_nC_x\, \pi^x (1-\pi)^{n-x}$$ [6–3]

where:

C denotes a combination.
n is the number of trials.
x is the random variable defined as the number of successes.
π is the probability of a success on each trial.

MEAN OF A BINOMIAL DISTRIBUTION $$\mu = n\pi$$ [6–4]

VARIANCE OF A BINOMIAL DISTRIBUTION $$\sigma^2 = n\pi(1-\pi)$$ [6–5]

HYPERGEOMETRIC DISTRIBUTION $$P(x) = \frac{({}_SC_x)({}_{N-S}C_{n-x})}{{}_NC_n}$$ [6–6]

where:

N is the size of the population.
S is the number of successes in the population.
x is the number of successes in the sample. It may be 0, 1, 2, 3, ...
n is the size of the sample or the number of trials.
C is the symbol for a combination.

POISSON DISTRIBUTION	$P(x) = \frac{\mu^x e^{-\mu}}{x!}$	[6–7]

where:

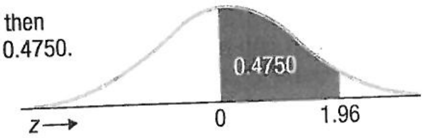
μ (mu) is the mean number of occurrences (successes) in a particular interval.
e is the constant 2.71828 (base of the Napierian logarithmic system).
x is the number of occurrences (successes).
$P(x)$ is the probability for a specified value of x.

$$z = \frac{X - \mu}{\sigma}$$

Appendix B: Tables

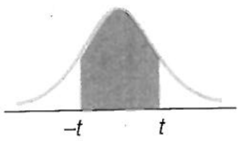
B.1 Areas under the Normal Curve

Example:
If $z = 1.96$, then
$P(0 \text{ to } z) = 0.4750$.



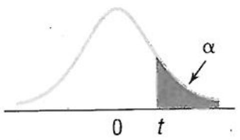
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

B.2 Student's *t* Distribution

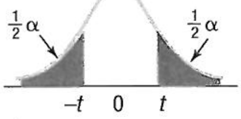


Confidence interval

Left-tailed test



Right-tailed test



Two-tailed test

	Confidence Intervals, *c*					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test, α					
df	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test, α					
	0.200	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.768
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
31	1.309	1.696	2.040	2.453	2.744	3.633
32	1.309	1.694	2.037	2.449	2.738	3.622
33	1.308	1.692	2.035	2.445	2.733	3.611
34	1.307	1.691	2.032	2.441	2.728	3.601
35	1.306	1.690	2.030	2.438	2.724	3.591

	Confidence Intervals, *c*					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test, α					
df	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test, α					
	0.200	0.10	0.05	0.02	0.01	0.001
36	1.306	1.688	2.028	2.434	2.719	3.582
37	1.305	1.687	2.026	2.431	2.715	3.574
38	1.304	1.686	2.024	2.429	2.712	3.566
39	1.304	1.685	2.023	2.426	2.708	3.558
40	1.303	1.684	2.021	2.423	2.704	3.551
41	1.303	1.683	2.020	2.421	2.701	3.544
42	1.302	1.682	2.018	2.418	2.698	3.538
43	1.302	1.681	2.017	2.416	2.695	3.532
44	1.301	1.680	2.015	2.414	2.692	3.526
45	1.301	1.679	2.014	2.412	2.690	3.520
46	1.300	1.679	2.013	2.410	2.687	3.515
47	1.300	1.678	2.012	2.408	2.685	3.510
48	1.299	1.677	2.011	2.407	2.682	3.505
49	1.299	1.677	2.010	2.405	2.680	3.500
50	1.299	1.676	2.009	2.403	2.678	3.496
51	1.298	1.675	2.008	2.402	2.676	3.492
52	1.298	1.675	2.007	2.400	2.674	3.488
53	1.298	1.674	2.006	2.399	2.672	3.484
54	1.297	1.674	2.005	2.397	2.670	3.480
55	1.297	1.673	2.004	2.396	2.668	3.476
56	1.297	1.673	2.003	2.395	2.667	3.473
57	1.297	1.672	2.002	2.394	2.665	3.470
58	1.296	1.672	2.002	2.392	2.663	3.466
59	1.296	1.671	2.001	2.391	2.662	3.463
60	1.296	1.671	2.000	2.390	2.660	3.460
61	1.296	1.670	2.000	2.389	2.659	3.457
62	1.295	1.670	1.999	2.388	2.657	3.454
63	1.295	1.669	1.998	2.387	2.656	3.452
64	1.295	1.669	1.998	2.386	2.655	3.449
65	1.295	1.669	1.997	2.385	2.654	3.447
66	1.295	1.668	1.997	2.384	2.652	3.444
67	1.294	1.668	1.996	2.383	2.651	3.442
68	1.294	1.668	1.995	2.382	2.650	3.439
69	1.294	1.667	1.995	2.382	2.649	3.437
70	1.294	1.667	1.994	2.381	2.648	3.435

(continued)

B.2 Student's *t* Distribution (*concluded*)

df	Confidence Intervals, *c*					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test, α					
	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test, α					
	0.200	0.10	0.05	0.02	0.01	0.001
71	1.294	1.667	1.994	2.380	2.647	3.433
72	1.293	1.666	1.993	2.379	2.646	3.431
73	1.293	1.666	1.993	2.379	2.645	3.429
74	1.293	1.666	1.993	2.378	2.644	3.427
75	1.293	1.665	1.992	2.377	2.643	3.425
76	1.293	1.665	1.992	2.376	2.642	3.423
77	1.293	1.665	1.991	2.376	2.641	3.421
78	1.292	1.665	1.991	2.375	2.640	3.420
79	1.292	1.664	1.990	2.374	2.640	3.418
80	1.292	1.664	1.990	2.374	2.639	3.416
81	1.292	1.664	1.990	2.373	2.638	3.415
82	1.292	1.664	1.989	2.373	2.637	3.413
83	1.292	1.663	1.989	2.372	2.636	3.412
84	1.292	1.663	1.989	2.372	2.636	3.410
85	1.292	1.663	1.988	2.371	2.635	3.409
86	1.291	1.663	1.988	2.370	2.634	3.407
87	1.291	1.663	1.988	2.370	2.634	3.406
88	1.291	1.662	1.987	2.369	2.633	3.405
89	1.291	1.662	1.987	2.369	2.632	3.403
90	1.291	1.662	1.987	2.368	2.632	3.402
91	1.291	1.662	1.986	2.368	2.631	3.401
92	1.291	1.662	1.986	2.368	2.630	3.399
93	1.291	1.661	1.986	2.367	2.630	3.398
94	1.291	1.661	1.986	2.367	2.629	3.397
95	1.291	1.661	1.985	2.366	2.629	3.396
96	1.290	1.661	1.985	2.366	2.628	3.395
97	1.290	1.661	1.985	2.365	2.627	3.394
98	1.290	1.661	1.984	2.365	2.627	3.393
99	1.290	1.660	1.984	2.365	2.626	3.392
100	1.290	1.660	1.984	2.364	2.626	3.390
120	1.289	1.658	1.980	2.358	2.617	3.373
140	1.288	1.656	1.977	2.353	2.611	3.361
160	1.287	1.654	1.975	2.350	2.607	3.352
180	1.286	1.653	1.973	2.347	2.603	3.345
200	1.286	1.653	1.972	2.345	2.601	3.340
∞	1.282	1.645	1.960	2.326	2.576	3.291