Queen Mary
University of London

January Examination Period 2023

ECN224 (Econometrics 1) ***Duration: 2 hours***

YOU ARE NOT PERMITTED TO READ THE CONTENTS OF THIS QUESTION PAPER UNTIL INSTRUCTED TO DO SO BY AN INVIGILATOR

Answer ALL questions

Calculators are permitted in this examination. Please state on your answer book the name and type of machine used. Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately.

It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms, it will be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

EXAM PAPERS MUST NOT BE REMOVED FROM THE EXAM ROOM

Examiner: Cristina Gualdani

The questions below are based on a dataset containing the characteristics of 243 students. In particular, we will use the following variables:

- *gpa*: grade point average. This is the average of all the grades taken in school. The higher *gpa*, the better the school performance.
- *drink*: how often and heavily a student drinks alcohol on a scale from 0 to 33. The higher *drink*, the more a student consumes alcohol.
- *study*: average hours of study per week.
- *athlete*: dummy variable equal to 1 if a student is an athlete, and 0 otherwise.
- *allnight1*: dummy variable equal to 1 if a student never studies all night, and 0 otherwise.
- *allnight2*: dummy variable equal to 1 if a student sometime studies all night, and 0 otherwise.
- *allnight3*: dummy variable equal to 1 if a student frequently studies all night, and 0 otherwise.
- *allnight4*: dummy variable equal to 1 if a student very frequently studies all night, and 0 otherwise.

Note that *allnight1, allnight2, allnight3,* and *allnight4* define mutually exclusive categories. This means that, for each individual, *allnight1*=1 or *allnight2*=1 or *allnight3*=1 or *allnight4*=1. Each individual has one and only one of these four dummies taking value one.

The Appendix contains tables with the critical values of the standard normal distribution and the F distribution.

Question 1

Answer the following questions.

a) Do you expect the correlation between *gpa* and *drink* to be positive, negative, or zero? Why? Which graph can allow you to visualise if such a correlation has the conjectured sign in the sample?

[6 marks]

b) We conduct a simple OLS regression of *gpa* on *drink*. The regression output is reported in Table 1. Interpret the OLS coefficient of *drink*. Compute the OLS estimate

of the intercept, knowing that the sample average of *gpa* is 2.808 and the sample average of *drink* is 19.107. Interpret the OLS estimate of the intercept.

[10 marks]

Table 1

```
> model <- lm(gpa~drink)
> summary(model)

Call:
lm(formula = gpa ~ drink)

Residuals:
     Min       1Q   Median       3Q      Max
-1.33574 -0.31269  0.00093  0.29783  1.13740

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  [redacted]   0.088900  [redacted]  [redacted]
drink       -0.019215   0.004448   -4.32
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

[redacted]
[redacted]
F-statistic: 18.66 on 1 and 216 DF,  p-value: 2.382e-05
```

c) Which measure of fit would you compute to assess how well the estimated regression line describes the data? Compute such a measure using the fact the Explained Sum of Squares (ESS) is 3.638 and the Total Sum of Squares (TSS) is 45.752. Interpret the measure of fit just obtained.

[6 marks]

d) Consider the results in Table 1. Using both the critical value and the p-value, test whether the coefficient of *drink* is statistically significant at 1% and 5%. Formally, which null hypothesis and alternative hypothesis are you testing?

[8 marks]

e) We re-run the OLS regression of *gpa* on *drink* using heteroskedasticity-robust standard errors. The regression output is reported in Table 2. First, explain the meaning of heteroskedasticity. Second, provide the OLS estimates of the intercept and of the coefficient of *drink* under heteroskedasticity-robust standard errors. Third, test whether the coefficient of *drink* is statistically significant at 5% and compare the result with your answer to d).

[10 marks]

Table 2

```
> coeftest(model, vcov = vcovHC(model, type = "HC1"))

t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept)          0.0919482 34.4764 < 2.2e-16 ***
drink        0.0046777
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Question 2

Answer the following questions.

a) We enrich the model of Question 1 by now regressing *gpa* on *drink, study, athlete, allnight1, allnight2, allnight3*. We use heteroskedasticity-robust standard errors. We include the intercept. The regression output is reported in Table 3. Interpret each OLS coefficient. (Hint: *allnight1, allnight2, allnight3, allnight4* define mutually exclusive categories and the baseline category of the regression is *allnight4*).

[12 marks]

Table 3

```
> model2 <- lm(gpa~drink+study+athlete+allnight1+allnight2+allnight3)
> coeftest(model2, vcov = vcovHC(model2, type = "HC1"))

t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept)  3.0363133  0.2173432 13.9701 < 2.2e-16 ***
drink       -0.0133408  0.0046577 [redacted]
study        0.0080064  0.0032210  2.4857  0.013733 *
athlete     -0.2041497  0.0939723 -2.1724  0.030975 *
allnight1   -0.0464453 [redacted]
allnight2   -0.1073747
allnight3   -0.0483904
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

b) We now regress *gpa* on *drink, study, athlete, allnight1, allnight2, allnight3, allnight4.* We exclude the intercept. The regression output is reported in Table 4. Is the “no perfect multicollinearity” assumption violated? Interpret the OLS coefficients of *allnight1, allnight2, allnight3, allnight4.* Does the economic interpretation of these estimates change with respect to Table 3? Explain.

[12 marks]

Table 4

```
> model3 <- lm(gpa~0+drink+study+athlete+allnight1+allnight2+allnight3+allnight4)
> coeftest(model3, vcov = vcovHC(model3, type = "HC1"))

t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
drink       [redacted]
study       0.0080064  0.0032210  2.4857  0.013733 *
athlete    -0.2041497  0.0939723 -2.1724  0.030975 *
allnight1   2.9898680 [redacted] 27.2335 [redacted]
allnight2   2.9289386            27.2148
allnight3   2.9879229            16.6668
allnight4   3.0363133            13.9701
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

c) Consider the results in Table 3. Is the coefficient of *drink* statistically significant at 1%? Would you expect the answer to change if we considered the results in Table 4? Why?

[10 marks]

d) Consider the results in Table 4. Are the coefficients of *allnight1, allnight2, allnight3, allnight4* statistically significant at 5%? Here you should conduct the four tests separately and not as a joint test. Calculate the standard errors of these coefficients. Calculate the 95% confidence intervals of these coefficients.

[16 marks]

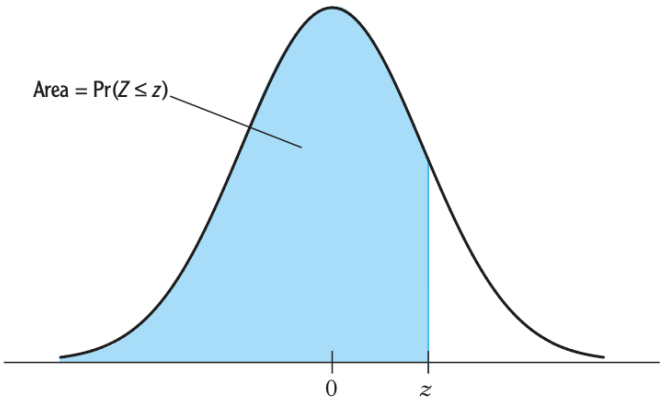
e) Consider the results in Table 4. After running the regression, we test whether the coefficients of *drink*, *study*, and *athlete* are jointly statistically significant at 5%. The F-statistic is 7.720. What is the outcome of the test? Specify the critical value. Formally, which null hypothesis and alternative hypothesis are you testing? Would it be correct to reject the null hypothesis if at least one of the t-statistics for *drink*, *study*, and *athlete* exceeded 1.96 in absolute value? Explain.

[10 marks]

End of Paper - An appendix of 3 pages follows

Appendix

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \leq z)$



	Second Decimal Value of z									
z	0	1	2	3	4	5	6	7	8	9
–2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
–2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
–2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
–2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
–2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
–2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
–2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
–2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
–2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
–2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
–1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
–1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
–1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
–1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
–1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
–1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
–1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
–1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
–1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
–1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
–0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

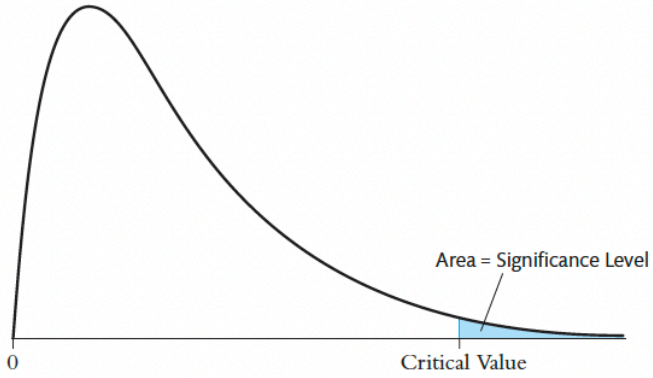
(Table 1 continued)

(Table 1 continued)

	Second Decimal Value of *z*									
z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
–0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
–0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
–0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
–0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
–0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
–0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
–0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
–0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
–0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

This table can be used to calculate $\Pr(Z \leq z)$ where Z is a standard normal variable. For example, when $z = 1.17$, this probability is 0.8790, which is the table entry for the row labeled 1.1 and the column labeled 7.

TABLE 4 Critical Values for the $F_{m,\infty}$ Distribution



	Significance Level		
Degrees of Freedom	**10%**	**5%**	**1%**
1	2.71	3.84	6.63
2	2.30	3.00	4.61
3	2.08	2.60	3.78
4	1.94	2.37	3.32
5	1.85	2.21	3.02
6	1.77	2.10	2.80
7	1.72	2.01	2.64
8	1.67	1.94	2.51
9	1.63	1.88	2.41
10	1.60	1.83	2.32
11	1.57	1.79	2.25
12	1.55	1.75	2.18
13	1.52	1.72	2.13
14	1.50	1.69	2.08
15	1.49	1.67	2.04
16	1.47	1.64	2.00
17	1.46	1.62	1.97
18	1.44	1.60	1.93
19	1.43	1.59	1.90
20	1.42	1.57	1.88
21	1.41	1.56	1.85
22	1.40	1.54	1.83
23	1.39	1.53	1.81
24	1.38	1.52	1.79
25	1.38	1.51	1.77
26	1.37	1.50	1.76
27	1.36	1.49	1.74
28	1.35	1.48	1.72
29	1.35	1.47	1.71
30	1.34	1.46	1.70

This table contains the 90^{th}, 95^{th}, and 99^{th} percentiles of the $F_{m,\infty}$ distribution. These serve as critical values for tests with significance levels of 10%, 5%, and 1%.

End of Appendix

End of Examination/ Cristina Gualdani