UNIVERSITY OF LIMERICK
OLLSCOIL LUIMNIGH
COLLEGE OF HUMANITIES

END OF TERM ASSESSMENT PAPER
MODULE CODE: SO5041 **SEMESTER**: Autumn 2010
MODULE TITLE: Quantitative Research Methods I (MA Sociology) **EXAM DURATION**: Two hours
LECTURER: Dr. Brendan Halpin **% OF TOTAL MARKS**: 40%
EXTERNAL EXAMINER: Prof. Chris Whelan

INSTRUCTIONS TO CANDIDATES:

- Answer four questions: all questions carry ten points each
- Calculators allowed: Yes
- Dictionaries allowed: No

---

1 (i) If student spending on entertainment is normally distributed with a mean of €98.25 per week, and standard deviation €17.34, what proportion of students spend more than €140?
(ii) How many spend more than €65?
(iii) How many between €80 and €105?
(iv) What spending level marks the 80th percentile?
(v) Calculate a 95% confidence interval around the mean spending, using the extra information that the sample size was 900. Interpret the confidence interval.
(vi) Of the 900 students, 324 had a part-time job. Calculate a 99% confidence interval around the proportion with a part-time job. Interpret the confidence interval.

2 In the Irish Survey of Sexual Health and Relationships, respondents were asked the age of their first sexual partner.
This Stata output shows a regression with age of partner (at the time of the first encounter) as the dependent variable, and age of respondent (now) as the explanatory variable.

```
      Source |       SS       df       MS              Number of obs =    6912
-------------+------------------------------           F(  1,  6910) = 1082.40
       Model |  20295.2464     1  20295.2464           Prob > F      =  0.0000
    Residual |  129564.125  6910  18.7502352           R-squared     =  0.1354
-------------+------------------------------           Adj R-squared =  0.1353
       Total |  149859.371  6911  21.6841805           Root MSE      =  4.3302

------------------------------------------------------------------------------
  agepartner |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .1304632   .0039655    32.90   0.000     .1226896    .1382367
       _cons |   16.07907    .160198   100.37   0.000     15.76503     16.3931
------------------------------------------------------------------------------
```

(i) Write out the regression equation
(ii) Conduct a formal test of the hypothesis that age has an effect on age-of-partner
(iii) What is the predicted age-of-partner for
  - someone currently aged 20?
  - someone currently aged 45?

(iv) Write down the R^2 value and explain what it means
(v) What sort of research question could this regression analysis be addressing? What can it tell us about the Irish population?

Table 1: Mean number of sex partners over past five years, by educational qualification and gender

Highest education qualification	Gender	
	Male	Female
Primary	1.4	0.8
Group, Junior Certificate or equivalent	2.4	1.2
Leaving Certificate or equivalent	3.2	1.5
Post-Leaving Cert Diploma/Certificate	2.9	1.5
Third level (university, I.T s)	3.3	1.9

Table 2: Mean number of sex partners over past five years, by educational qualification and age group

Highest education qualification	Gender		
	Male	Female	
Primary	1.4	0.8	1.1
	(262)	(304)	(566)
Group, Junior Certificate or equivalent	2.4	1.2	1.7
	(544)	(656)	(1,200)
Leaving Certificate or equivalent	3.2	1.5	2.2
	(859)	(1,168)	(2,027)
Post-Leaving Cert Diploma/Certificate	2.9	1.5	2.0
	(337)	(612)	(949)
Third level (university, I.T s)	3.3	1.9	2.5
	(1,180)	(1,510)	(2,690)
Total	2.9	1.5	2.1
	(3,182)	(4,250)	(7,432)

Note: Numbers of cases in parentheses

3 Table 1 reports the mean number of sexual partners over the past five years, broken down by educational qualification and gender, from the Irish Survey of Sexual Health and Relationships.

(i) What patterns of association do you see in Table 1? Why do you think these patterns might arise?

(ii) Table 2 shows a similar table, broken down by age group as well as qualification. Describe the main features of the table. Does the picture from this table change your mind about the picture youd describe from Table 1?

(iii) The χ^2 for the table of qualification by age group (i.e., the table of numbers in parentheses in Table 2) is 1261.0. Conduct a test of the null hypothesis that age group and qualification are independent. Explain your answer.

4 Write a short essay on the practical strengths and weaknesses of quantitative methods for social research.

Table 3: Economics and Sociology average grades for 10 students

Student	Economics	Sociology
1	1.8	1.8
2	4.0	2.9
3	1.8	1.8
4	4.0	4.0
5	3.1	1.8
6	2.2	3.2
7	3.8	1.8
8	3.2	2.4
9	3.6	1.8
10	4.0	1.8

5 (i) Ten students on the Economics & Sociology programme are chosen at random, and their average grades on modules offered by each department are calculated – see Table 3.

a. Conduct a test the hypothesis that performance in the two departments differs

b. Interpret your findings

(ii) Using information from the Student Records System, a test for equality between the average Economics grade (across all Economics students) and the average Sociology grade (across all Sociology students) was carried out:

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     ECO |     172    2.659302    .0727711    .9543839    2.515657    2.802948
     SOC |     130        2.84    .0441929    .5038764    2.752563    2.927437
---------+--------------------------------------------------------------------
combined |     302    2.737086    .0458294    .7964302    2.646899    2.827273
---------+--------------------------------------------------------------------
    diff |           -.1806977    .0921236               -.3619879    .0005926
------------------------------------------------------------------------------
    diff = mean(ECO) - mean(SOC)                                  t =  -1.9615
Ho: diff = 0                                     degrees of freedom =      300

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0254         Pr(|T| > |t|) = 0.0507          Pr(T > t) = 0.9746
```

Interpret the output.

(iii) What are the important differences between the two tests?

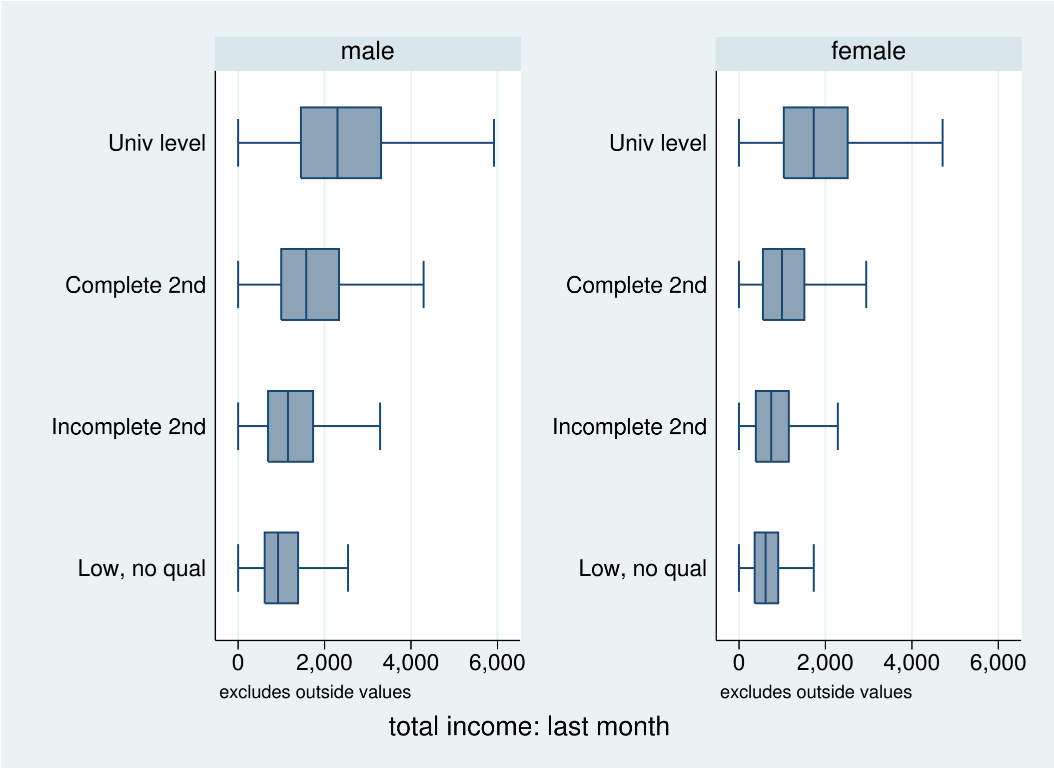

Figure 1: Boxplots of monthly income, by educational level and gender

6 Figure 1 shows the distribution of monthly income, broken down by gender and educational qualification. Use the information presented in the figure to write a short report on the effects of education and gender on income. (Source: British Household Panel Survey, 2005 wave.)

Formulas and Tables

(a) Standard deviation:

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

(b) z-score: If X is drawn from a normal distribution, with mean μ, and standard deviation σ, its corresponding z-score is:

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

(c) Standard deviation for a proprortion, π:

$$\sigma_\pi = \sqrt{\pi(1-\pi)}$$

(d) Sample standard error, SE, depends on sample standard deviation, s, and sample size, n:

$$SE = \frac{s}{\sqrt{n}}$$

(e) Confidence interval around point estimate, ε, where z is the z-score for the required level of confidence, and SE the standard error (note: z-score may be derived from standard normal distribution or t-distribution, as appropriate):

$$\varepsilon \pm z \times SE$$

(f) Chi-squared statistic for a table,

$$X^2 = \sum \frac{(O-E)^2}{E}$$

where O is the observed value and E the expected value.

(g) Expected value under independence in a table:

$$E = \frac{rc}{T}$$

where r is the row total and c the column total for that cell, and T the grand total for the table.

(h) Predicted value from a bi-variate regression, where a is the constant and b the slope coefficient:

$$\hat{Y} = a + bx$$

(i) Standard error for comparing means of two sub-samples, whose variance may not be the same:

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where s_i is the standard deviation for group i, and n_i the number of cases in group i.

Table of the Standard Normal Distribution

Right tail (probability of $X > z$)

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.00	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.10	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.20	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.30	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.40	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.50	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.60	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.70	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.80	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
.90	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.00	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.10	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.20	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.30	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.40	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.50	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.60	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.70	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.80	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.90	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.00	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.10	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.20	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.30	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.40	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.50	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.60	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.70	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.80	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.90	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.00	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.10	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.20	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.30	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.40	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.50	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.60	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.70	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.80	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.90	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
4.00	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Table of Student's *t* Distribution

Two-tailed probability

Degrees of Freedom	Probability level (Area under both tails)				
	0.10	0.05	0.025	0.01	0.005
1	6.314	12.706	25.452	63.657	127.321
2	2.920	4.303	6.205	9.925	14.089
3	2.353	3.182	4.177	5.841	7.453
4	2.132	2.776	3.495	4.604	5.598
5	2.015	2.571	3.163	4.032	4.773
6	1.943	2.447	2.969	3.707	4.317
7	1.895	2.365	2.841	3.499	4.029
8	1.860	2.306	2.752	3.355	3.833
9	1.833	2.262	2.685	3.250	3.690
10	1.812	2.228	2.634	3.169	3.581
11	1.796	2.201	2.593	3.106	3.497
12	1.782	2.179	2.560	3.055	3.428
13	1.771	2.160	2.533	3.012	3.372
14	1.761	2.145	2.510	2.977	3.326
15	1.753	2.131	2.490	2.947	3.286
16	1.746	2.120	2.473	2.921	3.252
17	1.740	2.110	2.458	2.898	3.222
18	1.734	2.101	2.445	2.878	3.197
19	1.729	2.093	2.433	2.861	3.174
20	1.725	2.086	2.423	2.845	3.153
21	1.721	2.080	2.414	2.831	3.135
22	1.717	2.074	2.405	2.819	3.119
23	1.714	2.069	2.398	2.807	3.104
24	1.711	2.064	2.391	2.797	3.091
25	1.708	2.060	2.385	2.787	3.078
26	1.706	2.056	2.379	2.779	3.067
27	1.703	2.052	2.373	2.771	3.057
28	1.701	2.048	2.368	2.763	3.047
29	1.699	2.045	2.364	2.756	3.038
30	1.697	2.042	2.360	2.750	3.030
35	1.690	2.030	2.342	2.724	2.996
40	1.684	2.021	2.329	2.704	2.971
50	1.676	2.009	2.311	2.678	2.937
60	1.671	2.000	2.299	2.660	2.915
75	1.665	1.992	2.287	2.643	2.892
100	1.660	1.984	2.276	2.626	2.871
500	1.648	1.965	2.248	2.586	2.820
1000	1.646	1.962	2.245	2.581	2.813
100000	1.645	1.960	2.241	2.576	2.807

Table of the χ^2 distribution (chi-sq)

Values of the χ^2 statistic for various degrees of freedom and areas under the right tail

Degrees of	Area under right tail				
Freedom	0.100	0.050	0.025	0.010	0.005
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28.300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.144	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.289	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.652	40.646	44.314	46.928