UNIVERSITY OF LIMERICK
OLLSCOIL LUIMNIGH
COLLEGE OF HUMANITIES

END OF TERM ASSESSMENT PAPER
MODULE CODE: SO5041 **SEMESTER**: Autumn 2012
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: Yes

---

1 Two types of t-test:

(i) The long-run average score on a test for incoming computer science students has been 68.4. The test is given to 10 new students, and the researchers want to ask the question whether this group differs from the norm. With the following data, test the hypothesis that the new group is different.

```
  1.   70
  2.   73
  3.   79
  4.   79
  5.   68
  6.   79
  7.   75
  8.   68
  9.   69
 10.   87
```

(ii) The table below is based on a random sample of workers, male and female, who report their daily travel time in minutes. Interpret the table in terms of the hypothesis that gender has an effect on travel time. Draw conclusions, explaining your reasoning (source: BHPS wave 18).

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    male |     319    27.44514     1.32498     23.6649    24.83831    30.05198
  female |     399    21.06266    .9187926    18.35287    19.25636    22.86895
---------+--------------------------------------------------------------------
combined |     718    23.89833    .7876325    21.10502    22.35199    25.44467
---------+--------------------------------------------------------------------
    diff |            6.382484    1.568205                3.303654    9.461315
------------------------------------------------------------------------------
    diff = mean(male) - mean(female)                              t =   4.0699
Ho: diff = 0                                     degrees of freedom =      716

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0001          Pr(T > t) = 0.0000
```

2 Distributions

(i) If sleep time is normally distributed, with a mean of 470 minutes and a standard deviation of 16 minutes, carry out the following calculations:

a. What proportion of people sleep more than 500 minutes
b. What proportion of people sleep more than 425 minutes
c. What proportion of people sleep more than between 430 and 460 minutes?
d. What amount of sleep time represents the level that 75% of people sleep less than or equal to?

(ii) Given a sampling distribution that is normal with mean μ and standard deviation (standard error) σ

a. What is the chance of getting a value outside the range $\mu \pm 2\sigma$ if the null hypothesis is true?
b. What is the chance of getting a value above $\mu + 2\sigma$?

(iii) If the null hypothesis is false, can you tell what the probability of getting a sample value above $\mu + 2\sigma$? Why?

3 The following table presents data from the British Household Panel Survey on subjective mental wellbeing. The dependent variable is males' GHQ score (range 0-36; higher scores imply lower mental wellbeing) and the explanatory variable is his spouse's score on the same variable.

```
. reg rhlghq1 spghq if rsex==1

      Source |       SS       df       MS              Number of obs =    3244
-------------+------------------------------           F(  1,  3242) =  165.05
       Model |  3651.26205     1  3651.26205           Prob > F      =  0.0000
    Residual |  71721.8354  3242  22.1227129           R-squared     =  0.0484
-------------+------------------------------           Adj R-squared =  0.0481
       Total |  75373.0974  3243  23.2417815           Root MSE      =  4.7035

------------------------------------------------------------------------------
     rhlghq1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       spghq |   .1941548   .0151128    12.85   0.000     .1645231    .2237865
       _cons |    8.35618   .1977541    42.26   0.000     7.968445    8.743916
------------------------------------------------------------------------------
```

(i) Report the regression equation and interpret it in words
(ii) Predict husband's GHQ for a spouse's value of 5
(iii) Predict husband's GHQ for a spouse's value of 25
(iv) Interpret the table to test the hypothesis that spouse's GHQ affects the husband's GHQ
(v) Report and interpret the R2
(vi) Discuss the possible reasons that spouses' GHQ scores could be related. How likely is it that this is a causal effect, with wife's score determining husband's score?
(more over page)

(vii) The second table uses the husband's own GHQ measured three years earlier to predict his current GHQ: compare the regression equation and the R2. How do the two analyses differ, and why might this be so?

```
. reg rhlghq1 ohlghq1 if rsex==1

      Source |       SS       df       MS              Number of obs =    3244
-------------+------------------------------           F(  1,  3242) =  962.88
       Model |  17259.7251     1  17259.7251           Prob > F      =  0.0000
    Residual |  58113.3723  3242  17.9251611           R-squared     =  0.2290
-------------+------------------------------           Adj R-squared =  0.2288
       Total |  75373.0974  3243  23.2417815           Root MSE      =  4.2338

------------------------------------------------------------------------------
     rhlghq1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ohlghq1 |   .5021801   .0161836    31.03   0.000      .470449    .5339111
       _cons |   5.497606   .1823538    30.15   0.000     5.140066    5.855146
------------------------------------------------------------------------------
```

4 Essay question: Discuss the importance of representativity and generalisability in quantitative social research.

5 The following two tables are drawn from the European Social Survey, and represent couples born in the 1940s and in the 1970s (percentage is row) classified according to their level of education.

(i) Describe the main features of the 1940s table

(ii) In what ways are the 1970s couples different?

(iii) What sorts of social processes are likely to be behind the general pattern in the tables?

(iv) Using the data provided with the 1940s table, carry out a statistical test for association between wife's and husband's educational level

** 1940s cohort

Wife's status	Husband's status: Low	Intermedi	Complete	Third lev	Total
Low	153	111	35	52	351
	43.59	31.62	9.97	14.81	100.00
Intermediate 2nd	92	80	38	73	283
	32.51	28.27	13.43	25.80	100.00
Complete 2nd	20	20	26	36	102
	19.61	19.61	25.49	35.29	100.00
Third level	34	34	29	173	270
	12.59	12.59	10.74	64.07	100.00
Total	299	245	128	334	1,006
	29.72	24.35	12.72	33.20	100.00

Pearson chi2(9) = 213.4555   Pr = 0.000

** 1970s cohort

Wife's status	Husband's status: Low	Intermedi	Complete	Third lev	Total
Low	21	20	10	8	59
	35.59	33.90	16.95	13.56	100.00
Intermediate 2nd	37	120	34	52	243
	15.23	49.38	13.99	21.40	100.00
Complete 2nd	15	43	77	53	188
	7.98	22.87	40.96	28.19	100.00
Third level	12	85	63	354	514
	2.33	16.54	12.26	68.87	100.00
Total	85	268	184	467	1,004
	8.47	26.69	18.33	46.51	100.00

Pearson chi2(9) = 337.1388   Pr = 0.000

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