CONCORDIA UNIVERSITY
SOCIOLOGY AND ANTHROPOLOGY DEPARTMENT
SOCI 213/4 AA & BB: STATISTICS II / PROF. HUSSEIN MERHI
FINAL EXAM / WINTER 2020
Date: Friday April 17, 2020 / Duration: 7:00-10:00 PM

I. A math teacher wanted to know why some of his students were performing better than the others. He collected information about each student's Math Score (Y), Number of Courses per Semester (X_1), Number of Days of Absence/semester (X_2), and Number of Hours of Study/Week (X_3). After processing the data, these statistics were obtained: ***(8 pts)***

Variable	$\overline{X}$	S	b
Y	65	10.3	------
X_1	3	1.2	-1.16
X_2	2	1.5	-1.76
X_3	5	1.9	2.89

1. Find the multiple regression equation and explain its meaning as well as the meaning of its slopes and constant. What score do you predict for someone with X_1=5 courses, X_2=4 days, & X_3=3 hrs ?
2. Which of the three independent variables has greater impact on the dependent variable ? Rank the variables in order in terms of their impact on Y.

II. You are asked to select a **sample of 450** students from the total high school population of a small city in order to find the proportion of students who are likely to drop out of school. This student population is divided on three high schools: **A=2200, B=1800, C=1200**. ***(8 pts)***

1. If you use stratified proportionate random sampling, how many students would you select from each high school ? Evaluate this procedure in terms of the sampling rules.
2. Using the proportionate random sample, the survey shows that the number of students who are likely to drop out in each high school sample is: **A=25, B=30, C=35**.
   a. Estimate the proportion of drop-out students for each high school.
   b. Estimate the proportion of drop-outs for the entire student population of the city.

III. A normally distributed sample (**N=122**) has a mean number of years of education **$\overline{X}$ =10 years and a standard deviation S=4 years**. ***(4 pts)***
1. What proportion of scores fall between the mean and a score of 12 ?
2. What score would someone need to be in the top 5% of the sample distribution ?

IV. A researcher wanted to use the sample statistics above (**question III**) to estimate the mean of the population from which the sample was taken. ***(4 pts)***
1. What is the probability of selecting a sample with a mean greater than 11 ?
2. Construct a 90% confidence interval for the population mean.

V. The average length of stay in a hospital for the entire population of a city is 5.4 days. A new research on a sample of 125 patients obtains a sample mean of 4.9 days and a sample standard deviation of 2.1. ***(4 pts)***

Conduct a two-tailed test to determine whether the sample mean is significantly different from the population mean at $\alpha = .05$.

VI. We selected a sample of 20 persons equally divided into 4 marital categories: never married, married, divorced, and widowed and ranked them from 0 (Extremely dissatisfied) to 35 (Extremely satisfied). Use the ANOVA test to determine if there's a significant variation in happiness between the categories of marital status. ***(6 pts)***

Degree of happiness by Marital status, N=20			
Never married	**Married**	**Divorced**	**Widowed**
X	**X**	**X**	**X**
23	19	16	5
30	35	5	6
20	15	9	4
20	26	10	5
27	30	5	0

VII. A sample of 310 individuals was classified according to Alcohol consumption and Type of accident in the following way: ***(6 pts)***

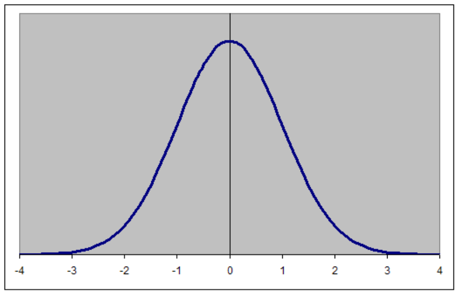
Type of accident	**Amount of alcohol consumption**			
	Little	**Moderate**	**Much**	**All**
Minor	78	47	21	146
Major	42	63	59	164
Total	120	110	80	N=310

Use the **Chi square** test to find the probability of independence between the two variables at a level of significance of **.05**.

VIII. Here are a population and a sample income distributions: ***(6 pts)***

Income distribution ($000)	**Population %**	**Sample**
Less than 10	10	40
10 to less than 20	20	125
20 to less than 40	35	200
40 to less than 80	20	115
80 or more	15	20
	100	500

Use **the Goodness-of-fit** test to determine whether the two distributions are significantly different at a level of significance of **.01**.



STANDARD NORMAL Z-DISTRIBUTION

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

T-DISTRIBUTION						
$One-tail\ \alpha=$	.10	.05	.025	.01	.005	.0005
$Two-tail\ \alpha=$	.20	.10	.05	.02	.01	.001
df 1	3.078	6.314	12.71	31.82	63.66	636.6
2	1.886	2.920	4.303	6.965	9.925	31.60
3	1.638	2.353	3.182	4.541	5.841	12.92
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
40	1.303	1.684	2.021	2.423	2.704	3.551
50	1.295	1.676	2.009	2.403	2.678	3.496
60	1.296	1.671	2.000	2.390	2.660	3.460
80	1.292	1.664	1.990	2.374	2.639	3.416
100	1.290	1.660	1.984	2.364	2.626	3.390
1000	1.282	1.646	1.962	2.330	2.581	3.300
∞	1.282	1.64	1.960	2.326	2.576	3.291

F Distribution (p=.05)

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	249.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.61	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.50	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	1.89	1.62
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.75	1.52	1.00

Chi Square Distribution									
df \ α	**.99**	**.75**	**.50**	**.25**	**.10**	**.05**	**.025**	**.010**	**.005**
1	0.00016	0.10153	0.45494	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944
2	0.02010	0.57536	1.38629	2.77259	4.60517	5.99146	7.37776	9.21034	10.59663
3	0.11483	1.21253	2.36597	4.10834	6.25139	7.81473	9.34840	11.34487	12.83816
4	0.29711	1.92256	3.35669	5.38527	7.77944	9.48773	11.14329	13.27670	14.86026
5	0.55430	2.67460	4.35146	6.62568	9.23636	11.07050	12.83250	15.08627	16.74960
6	0.87209	3.45460	5.34812	7.84080	10.64464	12.59159	14.44938	16.81189	18.54758
7	1.23904	4.25485	6.34581	9.03715	12.01704	14.06714	16.01276	18.47531	20.27774
8	1.64650	5.07064	7.34412	10.21885	13.36157	15.50731	17.53455	20.09024	21.95495
9	2.08790	5.89883	8.34283	11.38875	14.68366	16.91898	19.02277	21.66599	23.58935
10	2.55821	6.73720	9.34182	12.54886	15.98718	18.30704	20.48318	23.20925	25.18818
11	3.05348	7.58414	10.34100	13.70069	17.27501	19.67514	21.92005	24.72497	26.75685
12	3.57057	8.43842	11.34032	14.84540	18.54935	21.02607	23.33666	26.21697	28.29952
13	4.10692	9.29907	12.33976	15.98391	19.81193	22.36203	24.73560	27.68825	29.81947
14	4.66043	10.16531	13.33927	17.11693	21.06414	23.68479	26.11895	29.14124	31.31935
15	5.22935	11.03654	14.33886	18.24509	22.30713	24.99579	27.48839	30.57791	32.80132
16	5.81221	11.91222	15.33850	19.36886	23.54183	26.29623	28.84535	31.99993	34.26719
17	6.40776	12.79193	16.33818	20.48868	24.76904	27.58711	30.19101	33.40866	35.71847
18	7.01491	13.67529	17.33790	21.60489	25.98942	28.86930	31.52638	34.80531	37.15645
19	7.63273	14.56200	18.33765	22.71781	27.20357	30.14353	32.85233	36.19087	38.58226
20	8.26040	15.45177	19.33743	23.82769	28.41198	31.41043	34.16961	37.56623	39.99685
21	8.89720	16.34438	20.33723	24.93478	29.61509	32.67057	35.47888	38.93217	41.40106
22	9.54249	17.23962	21.33704	26.03927	30.81328	33.92444	36.78071	40.28936	42.79565
23	10.19572	18.13730	22.33688	27.14134	32.00690	35.17246	38.07563	41.63840	44.18128
24	10.85636	19.03725	23.33673	28.24115	33.19624	36.41503	39.36408	42.97982	45.55851
25	11.52398	19.93934	24.33659	29.33885	34.38159	37.65248	40.64647	44.31410	46.92789
26	12.19815	20.84343	25.33646	30.43457	35.56317	38.88514	41.92317	45.64168	48.28988
27	12.87850	21.74940	26.33634	31.52841	36.74122	40.11327	43.19451	46.96294	49.64492
28	13.56471	22.65716	27.33623	32.62049	37.91592	41.33714	44.46079	48.27824	50.99338
29	14.25645	23.56659	28.33613	33.71091	39.08747	42.55697	45.72229	49.58788	52.33562
30	14.95346	24.47761	29.33603	34.79974	40.25602	43.77297	46.97924	50.89218	53.67196

FORMULAS

Regression

Simple: $Y = a + bX$ $\qquad a = \overline{Y} - b\overline{X}$

Multiple: $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots\ldots. + b_k X_k$

$$a = \overline{Y} - [b_1 \overline{X}_1 + b_2 \overline{X}_2 + b_3 \overline{X}_3 + \ldots. + b_k \overline{X}_k]$$

$$\beta = b\left(\frac{S_x}{S_y}\right) \qquad R^2 = 1 - \frac{SSE}{SST}$$

Sampling

$$K = \frac{population}{sample}$$

Sampling distribution & the Normal Curve

$$Z = \frac{X - \mu}{\sigma} \qquad Z = \frac{X - \overline{X}}{S} \qquad Z = \frac{X - \mu}{\sigma_{\overline{X}}}$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} \qquad \sigma_{\overline{X}} = \frac{S}{\sqrt{N-1}}$$

$$X = \mu \pm Z\sigma \qquad X = \overline{X} \pm ZS \qquad X = \mu \pm Z\sigma_{\overline{X}}$$

Estimation

$Limits\ \ for\ \mu = \overline{X} \pm Z\sigma_{\overline{X}}$ **or** $Limits\ \ for\ \mu = \overline{X} \pm t\sigma_{\overline{X}}$

$Limits\ \ for\ \hat{p} = p \pm Z\sigma_p$ **or** $Limits\ \ for\ \hat{p} = p \pm t\sigma_p$

Hypothesis Testing

$$Z_s = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} \quad \text{or} \quad t_s = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

$$p = X / N \qquad \sigma_p = \sqrt{\frac{pq}{N}}$$

$$Z_s = \frac{p - \hat{p}}{\sigma_p} \quad \text{or} \quad t_s = \frac{p - \hat{p}}{\sigma_p}$$

ANOVA test

n_1 = (K – 1) ***n_2= (N – K)***

$$\overline{X} = \frac{\sum X}{N}$$

$$SST = \sum X^2 - N_T \overline{X}_T^2 \qquad SSB = \sum N_k (\overline{X}_k - \overline{X}_T)^2$$

SSW = SST - SSB

MSW = SSW / n_2

MSB = SSB / n_1

F ratio = MSB / MSW

Chi Square tests

$$EF = \frac{C \, x \, R}{N} \qquad EF = \frac{N \; x \; P}{100}$$

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$df = (C-1)(R-1) \qquad df = K - 1$$