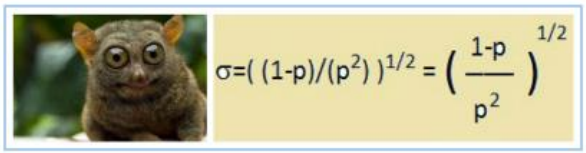
Final exam:

Emphasis: everything after the study break in Feb….

Course Lectures

Class lectures;

In order to use our lecture time more effectively, I will be posting lecture slides the day prior to each lecture (likely, the Monday evening prior to class). I strongly recommend that you print up these presentations and bring them along to class (for note taking).



PASSWORD protected.

Week 1

Week 2

Week 3

Week 4

Week 5

Week 6

STUDY WEEK

Week 8

Grades for midterm here

Week 10

Week 11

Week 12

Week 13 ***(APLIA ASSIGNMENTS DUE)***

Week 14 (2nd PROBLEMS ASSIGNMENT DUE)

Review carefully all slides after the FEB study break including the APLIA assignments AND 2nd problem solving assignment..

PLEASE RETURN EVERYTHING ON COMPLETION OF THIS EXAM, INCLUDING THIS EXAM, YOUR SCANTRON AND EXAM BOOKLET

NAME _____ **STUDENT #** _____

KING'S UNIVERSITY COLLEGE DEPARTMENT OF SOCIOLOGY DECEMBER

Introduction to Statistics
Final Examination
Sociology 2205 (Section 575)
Don Kerr

General Instructions:

Read the questions carefully and be sure you are clear as to what's required before attempting to answer.
Show all of your work.
Refer to the formulas provided on the last two pages to help you answer the questions.

You may use your calculators, but no additional notes.

20 multiple choice questions..

PART 1 involves multiple choice, and is worth 30 points.
PART 2 involves problem solving and calculations. Of the 8 problems, you must solve 7 (10 points each, for a total of 70 points).

A few things to think about:

Level of measurement is always relevant!! (nominal/ordinal/interval-ratio)
You should always be considering this, in answering questions.
-> level of measurement involved is relevant in deciding which statistics/formula to work with..

Association
Strength of the association?
If both or one variable is nominal -> cramer's v, phi, lambda (and also Maximum difference method)
  If both are ordinal -> gamma
  If both are interval -> Pearson's R

If both variables are "ordinal" or both are "interval ratio", we can speak of the direction of a relationship (positive or negative)..

The word "significance" is important, and it always involves some sort of formal test (always, 5 steps)

Either:
-> association (cross tab)
  If at least one variable is "nominal" -> X^2 test of independence
  If both are "ordinal" -> use the Gamma based test of significance.
  If both varaibles are "interval/ratio" use the regression based test of significance

Also, don't forget the stuff we covered prior to X^2 test of independence (REVIEW Week 8 slides)
 -> Tests of difference, 2 samples (means & proportions) NOTE: we won't deal with 1 sample tests on the final..

Proportion: $p = \frac{f}{N}$

Percentage change $= \left(\frac{f_2 - f_1}{f_1}\right) \times 100$

$\overline{X} = \frac{\Sigma(X_i)}{N}$

$Z = \frac{X_i - \overline{X}}{s}$

$Z(\text{obtained}) = \frac{\overline{X} - \mu}{s/\sqrt{n-1}}$

Percentage: % $= \left(\frac{f}{N}\right) \times 100$

Ratio = f_1 / f_2

$s = \sqrt{\frac{\Sigma(X_i - \overline{X})^2}{N}}$

$v = 1 - \left(\frac{f_m}{n}\right)$

$Z(\text{obtained}) = \frac{P_s - P_u}{\sqrt{P_u(1-P_u)/n}}$

c.i. $= \overline{X} \pm Z\left(\frac{s}{\sqrt{n-1}}\right)$

c.i. $= P_s \pm Z\sqrt{\frac{P_u(1-P_u)}{n}}$

$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$

$t(\text{obtained}) = \frac{\overline{X} - \mu}{s/\sqrt{n-1}}$

$Z(\text{obtained}) = \frac{(\overline{X}_1 - \overline{X}_2)}{\sigma_{\overline{x}-\overline{x}}}$

$\sigma_{\overline{x}-\overline{x}} = \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}$

$f_e = \frac{\text{Row marginal X Column marginal}}{N}$

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

$t(\text{obtained}) = \frac{(\overline{X}_1 - \overline{X}_2)}{\sigma_{\overline{x}-\overline{x}}}$

$\sigma_{\overline{x}-\overline{x}} = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$

$Z(\text{obtained}) = \frac{(P_{s1} - P_{s2})}{\sigma_{p-p}}$

$\sigma_{p-p} = \sqrt{P_u(1-P_u)}\sqrt{\frac{n_1+n_2}{n_1 n_2}}$

$P_u = \frac{n_1 P_{s1} + n_2 P_{s2}}{n_1 + n_2}$

$\phi = \sqrt{\frac{\chi^2}{N}}$

$V = \sqrt{\frac{\chi^2}{(N)(\min r - 1,\ c - 1)}}$

$\lambda = \frac{E_1 - E_2}{E_1}$

$y = a + bx$

$b = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{n\Sigma X^2 - (\Sigma X)^2}$

$t_{obtained} = r\sqrt{\frac{n-2}{1-r^2}}$

$G = \frac{N_s - N_d}{N_s + N_d}$

$z(obtained) = G\sqrt{\frac{N_s + N_d}{N(1-G^2)}}$

$a = \overline{Y} - b\overline{X}$

$r = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[n\Sigma X^2 - (\Sigma X)^2][n\Sigma Y^2 - (\Sigma Y)^2]}}$

Confidence Level	Alpha (α)	α/2	Z score
90%	0.10	0.05	± 1.65
95%	0.05	0.025	± 1.96
99%	0.01	0.005	± 2.58

a. Is there a gender gap in use of the Internet? Random samples of men and women have been questioned about the average number of minutes they spend each week on the Internet for any purpose. Is the difference significant?

Women	Men
$\bar{X}_1 = 55$	$\bar{X}_2 = 60$
$s_1 = 2.5$	$s_2 = 2.0$
$N_1 = 520$	$N_2 = 515$

LECTURE 8

We are working with two samples:

A sample of King's students
N=150
12% smoke

A sample of Brock students
n= 125
20% smoke

Are King's students significantly less likely to smoke than Brock students?

The following table demonstrates the crosstabulation of marital status & where people work. Is there a significant association between the two variables? If so, using the appropriate measure of association, how strong is the relationship? What can you tell me about the pattern of the relationship??

	Marital Status			
Do You Work from Home?	Married/ Common-law	Separated/ Divorced/ Widowed	Single	Totals
Yes, but not regularly	2	0	3	5
Yes, 1–2 days each week	8	4	0	12
Yes, 3 or more days each week	0	3	3	6
No	10	8	4	22
Totals	20	15	10	45

Are the two variables significantly associated with each other?
What's the direction/pattern of the relationship?
Is there a strong association?

	Low	Medium	High
1.Strongly Disagree	24	20	18
2. Disagree	22	22	33
3. Agree	21	22	18
4. Strongly Agree	20	7	22

Calculate an appropriate measure of association. What does this tell us in terms of direction and strength of the association?

Is the association significant?

b. There is some evidence that people's involvement in their communities (membership in voluntary organizations, participation in local politics, and so forth) has been declining, and television has been cited as the cause. Do the data below support the idea that TV is responsible for the decline?

Hours of Community Service	Television Viewing			
	Low	Moderate	High	Totals
Low	5	10	18	33
Moderate	10	12	10	32
High	15	8	7	30
Totals	30	30	35	95

For each situation, compute and interpret the appropriate measure of association. Describe relationships in terms of the strength and pattern or direction.

Consider crime as dependent variable

What can you tell me about this relationship? Is it significant??

a. For 10 Canadian cities data has been gathered on rates of property crime (theft, breaking and entering, fraud) per 100,000 population and the percentage of people who are new immigrants (arrived in Canada within the past five years). Are the variables related? i.e. is the relationship significant?

City	Property Crime Rate	Percent Immigrants
A	15	10
B	12	18
C	20	9
D	17	11
E	16	15
F	10	20
G	17	9
H	13	22
I	9	10
J	7	15