Social Studies 201 – Fall 2006
Problem Set 5
Due December 4, 2006

1. Number of close friends. The data in Table 1 represent Saskatchewan respondents in Statistics Canada's General Social Survey of 2003, a survey that examined issues of social engagement of Canadians. Use the data in Table 1 to answer the following.

a. Obtain the 90% interval estimate for the mean number of close friends for Saskatchewan adults in the (i) 15-24 age group and (ii) the 75 plus age group.

b. For each of the same two age groups as in a. (15-24 and 75 plus) obtain the proportion of respondents that has two or less close friends. Then construct 95% interval estimates for the true proportion with two or less close friends for all those in each age group.

c. Among all respondents the mean is 6.3 close friends. Test whether the mean number of close friends for (i) all those of age 45-54 differs from 6.3 and (ii) all those of age 75 plus exceeds 6.3. Use 0.05 significance.

d. One-quarter of those in all age groups have two or less close friends. For all those of ages 15-24, test whether the proportion with two or less close friends is less than one-quarter. Use 0.10 significance.

e. From Table 1 and the results of a. through d., briefly comment on the differences in the number of close friends by age group.

Table 1. Frequency distributions and statistics of Saskatchewan respondents by number of close friends, 2003, ages 15-24. 45-54, and 75 plus

Number of close friends	Frequency for respondents of		
	Age 15-24	Age 45-54	Age 75 plus
None	2	10	16
1 or 2	28	33	22
3 to 5	76	57	39
6 to 10	52	48	29
11 to 20	19	20	19
More than 20	3	5	9
Total	180	173	134
Mean	6.38	6.48	7.35
Standard deviation	5.04	5.95	7.73

Source: Adapted from Statistics Canada. General Social Survey of Canada, 2003. Cycle 17: Social Engagement [machine readable data file]. 1st Edition. Ottawa, ON: Statistics Canada [publisher and distributor] 10/1/2004.

Questions 2 and 3 refer to survey results reported in the April 14 and 15, 2005 editions of the *Leader-Post*. Further details about the environment and Saskatchewan Learning surveys were obtained from http://www.executive.gov.sk.ca/polling.htm on April 15, 2005, where the full survey results were posted.

2. Sample sizes and margin of error

a. Opinion about climate change in Table 2 are measured on a five-point scale from 1 (not at all important) to 5 (very important). How large a sample size would be required to estimate the true mean for Saskatchewan correct to within plus or minus 0.1, seventeen times out of twenty? How large a sample size would be required to estimate the true mean for rural areas correct to within 0.25 points, with probability 0.98?

b. Following are statements about sample sizes and margins of error for the three surveys reported in the *Leader-Post*. Verify each of the following statements by using the formula for sample size or interval estimation. Show your calculations.

i. Total sample of 615 individuals would be interviewed randomly with a 95% confidence in the results; the maximum margin of error would be ±3.9%.
(Source: Arcas and Saskatchewan Environment. Environmental Issues – Public Opinion Survey)

ii. The total sample size and the corresponding margin of error of results were as follows: Sample size: 1,001 ± 3.1 percent, 19 times out of 20
(Pulse Research Ltd., Opinion Poll. Prepared for: Saskatchewan Learning, January 2005)

iii. The poll consists of 602 telephone interviews conducted with Regina residents aged 18 or older between April 8 and April 11. The results from a sample of this size are considered to be accurate to within plus or minus 4 percentage points, 19 times out of 20.
(Report on a Sigma Analytics poll on City of Regina tax hike. *Leader-Post*, April 15, 2005, p. A2)

c. In the survey in quote i., dealing with environmental issues, up to one-third of the 615 respondents were undecided about how to respond to some questions. If one-third of these respondents did not respond on a particular question, what would be the margin of error for this reduced number of respondents?

3. Concern about climate change. In the poll about environmental issues, respondents were asked to state the importance of climate change (global warming). Respondents were asked to respond on the five-point scale given in the first column of Table 2.

Table 2. Saskatchewan respondents' views concerning the importance of climate change (global warming); urban and rural respondents

Response to issue of climate change	Per cent response by region	
	Urban	Rural
Very important (5)	35.3	34.0
Important (4)	24.0	20.3
Neutral (3)	24.2	26.3
Not important (2)	6.4	11.1
Not at all important (1)	10.1	8.3
Total	100.0	100.0
Sample size	n = 395	n = 220
Mean	3.68	3.61
Standard deviation	1.29	1.19

Source: Saskatchewan Environment – Environment Issues – Public Opinion Survey, March 2005, from http://www.executive.gov.sk.ca/polling.htm, obtained April 15, 2005.

Use the data in Table 2 to answer the following:

a. Obtain interval estimates for the true mean response to the issue of climate change for Saskatchewan (i) urban residents and (ii) rural residents.

b. Obtain an interval estimate for the proportion of all urban residents in the province who consider the issue very important or important (i.e. all who responded 4 or 5).

c. Do more than one-half of rural residents in the province consider the issue of climate change very important or important (i.e. all those with response 4 or 5)?

d. What do you conclude about urban/rural similarities or differences as a result of the data in Table 3 and the results in a. through c.?

4. Downsizing and job loss. The data in Tables 3 and 4 come from a study of job loss due to downsizing at the Wells Fargo Bank in the United States. The data in Table 3 represent responses from employees six months after the downsizing. Assume that these results represent a random sample of bank employees.

a. The authors of the article state that the "results suggest that age did influence both the likelihood of being laid off and subsequent income and employment." Using the chi-square test, what do you conclude about the relationship between age of employee and labour force status. Do the test results support the authors' conclusions?

b. Table 3 was devised so one-third of respondents were in each of the three age groups. The authors state "the cohort of older workers was less likely to be currently employed than the two cohorts of younger workers, more likely to be unemployed, and more likely to have left the labour force." For each of the three age groups, obtain 95% interval estimates for the proportion reporting unemployed or not in labour force. From these results and the data in Table 4, does the authors' statement appear correct?

Table 3. Cross-classification of employees' labour force status and age, frequency and expected frequencies

LFS * AGE Crosstabulation

			AGE			
			under 37	37-44	over 44	Total
LFS	employed at bank	Count	44	47	45	136
		Expected Count	45.0	45.5	45.5	136.0
	employed elsewhere	Count	43	35	28	106
		Expected Count	35.1	35.5	35.5	106.0
	unemployed	Count	3	7	13	23
		Expected Count	7.6	7.7	7.7	23.0
	not in labour force	Count	9	11	14	34
		Expected Count	11.3	11.4	11.4	34.0
Total		Count	99	100	100	299
		Expected Count	99.0	100.0	100.0	299.0

Table 4. Chi-square test results for the data in Table 4

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11.025[a]	6	.088
Likelihood Ratio	11.231	6	.082
Linear-by-Linear Association	1.905	1	.167
N of Valid Cases	299		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.62.

Source: Jill Quadagno et al., "Downsizing and the Life-course Consequences of Job Loss: The Effects of Age and Gender on Employment and Income Security," from Victor W. Marshall et al., *Restructuring Work and the Life Course*, University of Toronto Press, Toronto, 2001. All data and quotes from pp. 310 and 311.