

Social Studies 201**Winter 2005****Problem Set 4****Due no later than March 14, 2005**

Note: If you complete the problem set by Friday, March 11, we will attempt to have it marked by Tuesday morning, March 15. Answers will be posted on the web site on Monday afternoon, March 14.

1. Religiosity of Saskatchewan adults. Table 1 shows the relationship between the frequency of attendance of 1,069 Saskatchewan adults in the year 2001 and when these same respondents were age 15.

Table 1. Cross-classification of frequency of attendance at religious services for Saskatchewan adults in 2001 and when they were age 15

Frequency of attendance in 2001	Frequency of attendance at age 15			Total
	At least once a month (M15)	Several times a year (S15)	Not at all (N15)	
At least once a month (M)	420	47	43	510
Several times a year (S)	196	89	39	324
Not at all (N)	102	34	99	235
Total	718	170	181	1,069

Source: Adapted from Statistics Canada, General Social Survey, 2001. Cycle 15: Family History (Main File) [machine readable data file]. 3rd edition. Ottawa, ON: Statistics Canada 7/2/2003.

- a. If an individual is randomly selected from the respondents in Table 2, what is the
 - i. Probability of attending at least once a month when age 15?
 - ii. Probability of attending several times a year in 2001?
 - iii. Probability that they either were or are regular attenders, at least once a month.
 - iv. Probability the individual has never attended religious services?
 - v. Conditional probability of attending not at all in 2001, given attendance of at least once a month at age 15.
 - vi. Conditional probability of attending several times a year at age 15, given attending at least once a month in 2001.
- b. Are the events of attending several times a year at age 15 (S15) and the event of not attending in 2001 (N) independent of each other?
- c. Explain whether those who attended at least once a month at age 15 more or less likely than those who attended several times a year at age 15 to have the same attendance pattern in 2001.

2. Explanations of probability

a. The *Leader-Post* of November 8, 2004 (pp. A1-A2) reported on a survey of 2,100 Canadians which asked “Who was prime minister of Canada during the Second World War?” Possible responses were Mackenzie King, Lester Pearson, Wilfrid Laurier, and Louis St. Laurent. The report stated:

More than half of those over [age] 60 correctly identified King, with 16 per cent naming Pearson. Respondents aged 45 to 59 were twice as likely to choose King over Pearson. ... But more than a third of the 30-to-44 demographic – 35 per cent – named Pearson as our wartime PM, while only 25 per cent got it right.

Identify a pair of dependent events from this quote

b. A study by Douglas L. Palmer stated that “prejudice and racism seems to increase in direct parallel to the numbers of immigrants.” At the same time, “negative feelings toward immigrants do not connect with unemployment rates.” Explain how independence and dependence are involved in this quote. (From www.sfu.ca/~riim/riim-info.webarchive/msg00542.html)

3. Standardized normal distribution. For the standardized normal distribution,

- a. What is the area between Z of 0 and Z of +1.43?
- b. What is the area between Z of +0.4 and Z of +1.6?
- c. What is the percentage of cases between $Z = -0.72$ and $Z = +0.87$?
- d. What proportion of the area under the normal curve is to the left of $Z = 1.25$?
- e. What is the area under the normal curve with a value larger than $Z = -1.33$?
- f. In a normally distributed population, what is the percentage of the population is more than one and one-half standard deviations from the mean?
- g. What is the Z-value so that 0.35 of the area is less than this Z? (That is, what is the Z-value of the 35th percentile?)
- h. What are the Z-values so that there is 0.04 of the area in each tail of the distribution lying beyond these Z-values, for a total of 0.08 in the two tails of the distribution?
- i. In a standardized normal distribution, where is the sixty-second percentile?
- j. The 80 per cent range can be defined as the range so that only the extreme twenty per cent of cases (ten per cent in each end of the distribution) are eliminated. In the standardized normal distribution, what are the Z-values for the eighty per cent range?

4. Distribution of years of education

a. The mean years of education for the 3,862 Saskatchewan adults in the Survey of Labour and Income Dynamics, 2000 was 12.4 years and the standard deviation was 2.1 years. If the variable years of education is normally distributed, obtain the following.

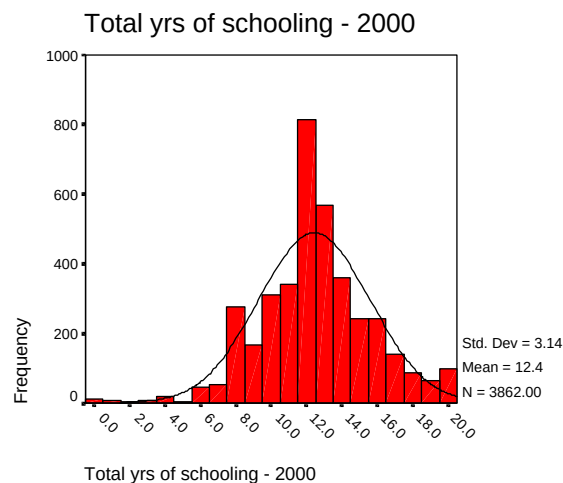
- Proportion of adults with less than 12 years of education.
- Proportion of adults with at least 9 but less than 12 years of education.
- Percentage of adults with more than 15 years of education.
- The interquartile range for years of education?

b. Use Table 3 and Figure 1 for this part. Compare the results you obtained in part a. with the actual percentages from Table 3 and the histogram of Figure 1 (with the normal curve superimposed on the frequency distribution). From this comparison, explain whether the variable “years of education” appears normally distributed or not.

Table 2. Percentage distribution of years of education of Saskatchewan adults, 2000

Years of education	Percentage
Less than 9	11.3
9 to under 12	21.4
12	24.0
13 to under 16	27.5
16 plus	15.8
Total	100.0

Figure 1. Histogram of distribution of years of education of Saskatchewan adults, 2000, with normal curve superimposed on the histogram



Source: Data from Statistics Canada. Survey of Labour and Income Dynamics (SLID), 2000. Person file [machine readable data file]. Ottawa, ON. Statistics Canada. 7/16/2003.

5. Computer problems – use the file *ssae.sav* in the folder *t:\students\public\201*. Hand in the SPSS output and the answers to the following.

a. Use *Analyze-Descriptive Statistics-Frequencies*, with options *Charts-Histograms-With Normal Curve* and *Statistics* (mean and standard deviation), to obtain frequency distributions, statistics, and histograms of the hours per week worked at jobs (*jobhours*). The frequency distribution table, statistics, and the histogram, with the normal curve superimposed, should appear on the printout.

Count the percentage of cases that are within one standard deviation of the mean of job hours; within two standard deviations of the mean of job hours. Compare with the percentages of cases within one and two standard deviations from the table of the normal distribution (Appendix H). Use these figures and the diagram on the printout to write a short note comparing the actual distribution of study hours with that of a normal distribution.

b. In Problem Set 3, you worked with the four variables *ued1* through *ued4*. Use *Analyze-Descriptive Statistics-Frequencies*, with options *Charts-Histograms-With Normal Curve* to obtain frequency distributions and histograms for these same four variables. Explain which of these four variables appears to have a distribution closest to that of a normal distribution and which is least like a normal distribution.

c. Crosstabulations and probability. Use *Analyze-Descriptive Statistics-Crosstabs* to obtain cross-classifications of (I) *ued4* by *sex* and (II) *ued4* by *pv* (provincial political preference). Use these tables to answer the following questions.

If one of the respondents in the table of part (I) is randomly selected, what are the following:

- i. Probability the respondent is a female?
- ii. Chance of selecting a respondent who somewhat or strongly agrees (response 4 or 5)?
- iii. Likelihood that the respondent is female and strongly disagrees?
- iv. Probability that the respondent strongly or somewhat disagrees (response 1 or 2), given that the respondent is female?
- v. Probability that the respondent strongly or somewhat disagrees (response 1 or 2), given that the respondent is male?

If an individual is randomly selected from the table of part (II), obtain the following.

- i. Probability of selecting an individual who supports no political party?
- ii. Probability of selecting an individual with a neutral response of 3?
- iii. The conditional probability of strongly agreeing, given each of the four provincial political preferences.

From the probabilities above, in a sentence or two explain which groups seem most in support of the statement that “University education should be made more accessible by lowering tuition.”