Social Studies 201

Fall 2004

Problem Set 4

Due no later than 8:30 a.m., November 9, 2004

Note: If you complete the problem set by November 5, we will attempt to mark it by Tuesday am, November 9. Answers will be posted on the web site on Tuesday, November 9 by 8:30 a.m.

1. Religiosity and volunteer work. The cross-classification of Table 1 gives the number of respondents with each combination of hours of volunteer work and religiosity. If an individual is randomly selected from the set of 975 Saskatchewan respondents, answer the following.

a. What is the probability of selecting an individual who (i) performs less than 30 hours volunteer work; (ii) is very religious?

b. What is the probability of selecting an individual who is somewhat religious and performs 30-99 hours of volunteer work?

c. What is the probability of selecting an individual who is not very religious or not at all religious?

d. What is the probability of selecting an individual who is not very religious or volunteers for 30 to 99 hours per year?

e. What are the conditional probabilities of selecting an individual who volunteers less than 30 hours (i) given very religious, and (ii) given not at all religious?

f. Are the events (A) volunteering for 100 plus hours and (B) being very religious independent or dependent? Explain.

g. From Table 1 and the above, do the very religious appear to volunteer more than the others? Explain in a sentence or two.

Table 1. Cross-classification of annual hours of volunteer work by religiosity, Saskatchewan respondents. Number with each combination of characteristics

Religiosity	Annual hours of volunteer work			Total
	Less than 30	30 to 99	100 plus	
Very religious	48	43	96	187
Somewhat religious	152	190	212	554
Not very religious	63	47	48	158
Not at all religious	31	21	24	76
Total	294	301	380	975

Source: Statistics Canada. National Survey of Giving, Volunteering and Participating, 1997: Main file [machine readable data file]. Ottawa, ON: Statistics Canada. August 20, 1999.

2. Probability statements. From the following quotes, identify (a) two pairs of dependent events and (b) one pair of independent events.

> On average, 16% of employees felt that poor interpersonal relations were a source of stress at work. This compares with about 10% of primary industry workers and 21% of workers in health occupations who feel this way. The likelihood of feeling stressed at work as a result of poor interpersonal relationships did not vary significantly from the average in most occupations.
>
> Source: Cara Williams, "Stress at work," *Canadian Social Trends*, No. 70, Autumn 2003, p. 10.

> "Health of female immigrants declines over time: study." Most female immigrants to Canada say they're in good to excellent health shortly after their arrival, but this sense of positive health diminishes over time.
>
> Source: *Leader-Post*, Regina, October 28, 2004, p. F4.

> It has been observed that the longer immigrants live in Canada, the more their health resembles that of the Canadian-born population.
>
> Source: Statistics Canada, "Health status of Canada's Immigrants," *The Daily*, Sept. 19, 2002. http://www.statcan.ca/Daily/English/020919/d020919a.htm

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

a. What is the area between Z of 0 and Z of +1.75?
b. What is the area between Z of +0.7 and Z of +1.8?
c. What is the percentage of cases between Z = -0.8 and Z = +1.5?
d. What proportion of the area under the normal curve is to the right of Z = -0.87?
e. What is the area under the normal curve below Z = 1.27?
f. In a normally distributed population, what is the percentage of the population is more than one-half standard deviations from the mean?
g. What is the Z-value so that 0.40 of the area is less than this Z? (That is, what is the Z-value of the 40^{th} percentile?)
h. What are the Z-values so that there is 0.0425 of the area in each tail of the distribution lying beyond these Z-values, for a total of 0.085 in the two tails of the distribution?
i. In a standardized normal distribution, where is the seventieth percentile?
j. A "trimmed mean" can be defined as the mean when the largest 10% and the smallest 10% of the cases have been eliminated. In the standardized normal distribution, what are the Z-values for these trimming points?

4. Annual hours worked. For Saskatchewan respondents aged 15-24, the mean annual hours worked is 1420 hours and the standard deviation is 730 hours.

a. If hours worked is exactly normally distributed, what are the following?
   i. The proportion who work more than 2,250 hours annually.
   ii. The percentage who work less than 750 hours annually.
   iii. If there are 502 individuals, how many work between 750 and 1,750 hours?

b. Using Table 2 and the accompanying histogram, compare the results from the above with the frequency distribution and histogram for 15-24 year olds. Using these data, write a short note comparing the actual distribution of hours worked with that of a normal distribution.

5. Computer problems – use the file ssae.sav in the t:\students\public\201 folder.

a. Is pay normally distributed? Use *Analyze-Descriptive Statistics-Frequencies*, with options *Charts-Histograms-With Normal Curve* and *Statistics* (obtain mean and standard deviation), to obtain a frequency distributions of the variable hourly pay at your job (PAY – question 55). The frequency distribution table and the histogram, with the normal curve superimposed, should be available on the printout.

Determine the intervals within one standard deviation of the mean and within two standard deviations of the mean. From the frequency distribution table, obtain the percentage of cases that are within each of these intervals. Compare with the percentages of cases within one and standard deviations from the normal distribution table. Use these figures and the diagram on the printout to write a short note comparing the actual distribution of pay with that of a normal distribution.

b. Probabilities from a crosstabulation. Use *Analyze-Descriptive Statistics-Crosstabs* to obtain a cross-classification table of *V4 (Gay and lesbians married, from question 15)* by *VALUES (importance of religious and spiritual values, from question 41)*. If an individual is randomly selected from this table, what are the following probabilities?

i. The probability of selecting an individual who considers religious and spiritual values as somewhat or very important.

ii. The probability of selecting an individual who agrees (response 4 or 5) that gay and lesbian couples should be regarded as married.

iii. The probability of selecting an individual who strongly agrees that gay and lesbian couples should be regarded as married, given that the individual considers religious and spiritual values as (1) very important; (2) not at all important.

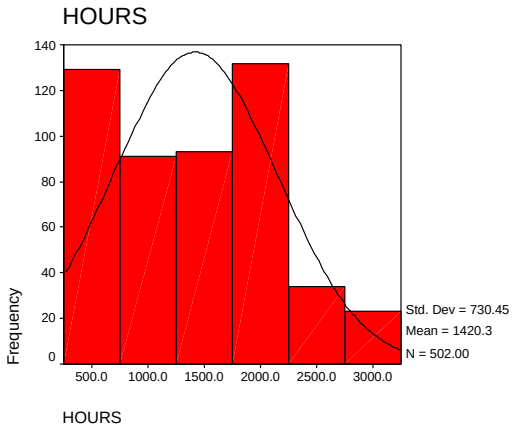
iv. Is the event of strongly disagreeing that gay and lesbian couples should be regarded as married independent of the event of religious and spiritual values being very important?

v. Attempt to find two events that are very close to being independent of each other.

Table 2. Annual hours worked at jobs by Saskatchewan respondents of different ages

Annual hours worked at jobs	X (hours in thousands of hours)	Number of respondents by age		
		All ages	Ages 15-24	Ages 35-44
250-750	0.5	264	129	39
750-1,250	1.0	259	91	44
1,250-1,750	1.5	340	93	74
1,750-2,250	2.0	1,134	132	380
2,250-2,750	2.5	232	34	63
2,750 plus	3.0	257	23	84
Total		2,486	502	684

Histogram of annual hours for 15-24 year olds, with normal distribution superimposed



Source: Statistics Canada. Survey of Labour and Income Dynamics (SLID), 1999: Person file [machine readable data file]. Release 1 Edition. Ottawa, ON: Statistics Canada. 4/16/2003.