

**Social Studies 201****Winter 2001****Problem Set No. 3****Due Friday, March 2, 2001**

1. Explain which concept of probability (subjective, theoretical, or frequency) appears to be used in each of the following statements.

In Switzerland, a man was roughly twice as likely to be murdered in the 1880s as he is today. ... Finland had a particularly violent past: the chance of being murdered at the turn of the 19th century was higher than it is in many American cities today. (*The Economist*, October, 15, 1994, p. 21).

A professor allocates the grades of 50 students in a class according to a normal distribution with mean 65 and standard deviation 10. If a student is randomly selected from this class, the probability that the student's grade is above 85 is 0.0227.

A new scientific study of nearly 100,000 women found that, contrary to some earlier research, hair dye does not raise a woman's **chance** of contracting certain blood and lymph cancers. (*The Leader-Post*, October 6, 1994).

Health care will **likely** not dominate the debate, in part because of Mr. Klein's effective management of the issues, where he addressed every concern thrown at him point by point (*National Post*, February 13, 2001, p. A8).

Mr. Rock's father died of prostate cancer several years ago – a fact that doubled his own **risk** of getting the disease. (*National Post*, February 13, 2001, p. A1).

2. Identify (i) one set of independent events and (ii) and one set of dependent events in the following quote from *The Economist*, October 15, 1994, p. 114. Explain your reasoning.

... the damage that smoking does to health is even larger than previously thought: the habit reduces life expectancy, on average, by eight years rather than five. The good news is that quitting smoking at any age improves life expectancy. Quitting before you reach 35 reduces the risk of death to the level of the life-long abstinent.

3. Use the tables from problem 6 of Computer Problem 2 to answer the following.
- (a) Questions from cross-classification table of V8 by SEX. If an individual is randomly selected from the Table, what is the probability of
- Agreeing (response 4 or 5)?
  - Agreeing or being male?
  - Being female given a neutral response (3)?
  - Disagreeing (response 1 or 2), given that the respondent is male? Disagreeing given that the respondent is female?
  - From (iv), comment on the independence or dependence of events and how responses differ by sex of the individual.
- (b) Questions from cross-classification table of V8 by FEDVOTE. If an individual is randomly selected from the Table, what is the probability of
- Disagree (response 1 or 2) and support PC party?
  - Disagree or support PC party?
  - Disagree given NDP? Disagree given PC?
  - Somewhat disagreeing (response 2), given NDP support?
  - From (iii) and (iv), comment on the independence or dependence of the events.

4. Obtain the following using the standardized normal distribution:
  - (a) What is the area between  $Z = 0$  and  $Z = 1.85$ ?
  - (b) What is the area between  $Z = -2.23$  and  $Z = 0.57$ ?
  - (c) What is the proportion of cases above  $Z = -1.09$ ?
  - (d) What is the proportion of cases between  $Z = 1.12$  and  $Z = 2.12$ ?
  - (e) What is the percentage of cases that are within 1.5 standard deviations of the mean? More than 2.1 standard deviations from the mean?
  - (f) What is the 35th percentile? The 72nd percentile?
  - (g) What is the value of  $Z$  such that 61% of the cases are less than this value?
  - (h) What are the  $Z$ -values so that 0.03 of the area is in each tail of the distribution?
5. The 944 Saskatchewan respondents in Statistics Canada's General Social Survey, Cycle 11, 1996, reported a mean household income of \$38 thousand, with a standard deviation of \$27 thousand.
  - (a) Assuming that incomes are normally distributed, what is
    - i. the proportion of households in the sample with an income of \$80,000 plus?
    - ii. the probability that a household randomly selected from the sample has an income below \$20,000?
    - iii. the number of households in the sample with incomes between \$30,000 and \$50,000?
    - iv. the 75th percentile of income?
    - v. the interquartile range of incomes?
  - (b) Table 1 provides that income distribution from the sample of the 944 Saskatchewan households in the sample. From the results in (a) and Table 1, comment concerning how well the normal distribution fits that actual distribution of incomes as provided by the sample.

Table 1: Frequency and Percentage Distributions of Household Income, Saskatchewan, 1996

Household Income in Thousands of Dollars	Number of Households	Per Cent of Households
less than 10	74	7.8
10 - 20	212	22.5
20 - 30	159	16.8
30 - 40	154	16.3
40 - 50	106	11.2
50 - 60	84	8.9
60 - 80	77	8.2
80 plus	78	8.3
Total	944	100.0