Principles of Inferential Statistics in Medicine

Midterm Exam – 513–607A, October 22^{nd}, 1998.

1. Suppose that the random variable X follows a Normal distribution with mean 5 and variance 25, so that we have $X \sim N(\mu = 5, \sigma^2 = 25)$.

(a) Calculate $Pr\{X \leq 0\}$
(b) Calculate $Pr\{-1 \leq X \leq 0\}$

2. Suppose that the true rate of asthma in the poulation of Quebec children aged 4 to 12 years old is 15%. If you decide to sample Quebec children between 4 and 12 years old at random, and stop as soon as you encounter your first child with asthma, what is the probability you will stop at the sixth child? (Hint: This means that you will first have 5 children without asthma.)

3. In a study of household environment and allergies, you have gathered the following information about cats in the household:

number of cats	percent of households
No cats	50%
Exactly one cat	40%
Exactly two cats	9%
Exactly three cats	1%

The number of households with four or more cats was very small, so is ignored in this study.

(a) What is the average numbers of cats in Quebec households?
(b) What is the variance of the number of cats in Quebec households?

4. Suppose that you would like to measure the effect of stress on workers in a manufacturing plant. Two different diagnostic tests of stress are available, and their properties are listed below:

	Stress Test Alpha	Stress Test Delta
sensitivity	60%	75%
specificity	95%	90%

(a) Which of the two tests will generate a larger number of false positive test results?
(b) Suppose that the prevalence of stress is 5%. Using Stress Test Alpha, what is the probability that a person with a positive stress test will actually be stressed?

5. A study has measured cholesterol values on a random sample of 101 adults over the age of 50 in Montreal. They found the mean cholesterol level to be 203 mg/dl, with a standard deviation of 47 mg/dl. Calculate a 95% confidence interval for the true mean cholesterol in this population.

6. An epidemiologic study wishes to investigate the relationship between calcium intake during childhood and bone mineral density (BMD). The following data are collected:

	High calcium intake group	Low calcium intake group
mean BMD	1.0	0.96
SD	0.1	0.15
Sample size	101	201

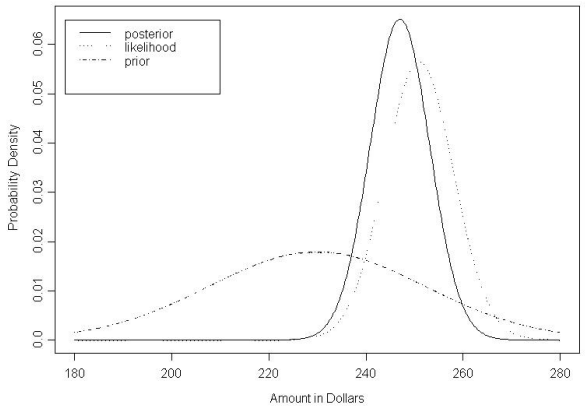
Test the null hypothesis that there is no difference in mean BMD for high versus low calcium intake populations against two-sided alternative. State the null and alternative hypotheses, carry out the test, provide a p-value, and state your conclusion. Assume that the variances in the two groups are not equal.

7. Suppose that you wish to plan a study to investigate two different drugs for blood pressure lowering (call them Drug A and Drug B). Suppose that you expect Drug A to lower blood pressure by about 10 mmHg on average, while the other drug may be slightly better, lowering blood pressure by an average of 13 mmHg. You would therefore like to estimate the true difference in the blood pressure lowering abilities of these drugs to within ± 1 mmHg with a 95% confidence interval. If the standard deviation of the drop in pressure is expected to be 10 mmHg in both groups (the standard deviations are assumed equal), what sample size will be required in each group?

8. Suppose that the true efficacy of a certain headache remedy is 90%.

(a) If five persons with headaches all take this remedy, what is the probability that all five will be cured of their headaches?
(b) If 1000 persons all take this remedy, what is the probability that 920 or more will be cured of their headaches?

9. A survey is designed to estimate the average yearly amount spent on alternative medicines in Canadian households. Data on 1000 households are collected, and a Bayesian analysis is used to analyse the data. A tri-plot of the data is given below:



Explain carefully what each curve in the triplot represents.

10. The SF-36 questionnaire is a series of eight scales, each measuring an aspect of quality of life. One of the scales measures overall General Health. Differences of 10 points or more on this scale are considered to be clinically important, with higher scores indicating better health than lower scores. Suppose that in a clinical trial of rheumatology patients, Drug A is compared to Drug B. The average score on the General Health scale in the Drug A group is 65, while the average score in the Drug B group is 77. A 95% confidence interval for the difference in SF-36 General Health (Drug B scores – Drug A scores) is calculated, and it is found to be (6, 18). Based on this information, would you prefer to use Drug A or Drug B? How certain are you of this decision?