

Principles of Inferential Statistics in Medicine

Midterm Exam – 513–607A, October 28nd, 1999.

1. Suppose that a new blood pressure lowering drug has been developed. The drug is given to a group of mildly hypertensive patients. Average blood pressure in this sample of patients is found to be 96 mmHg before the drug is administered, and 88 mmHg after one month on the drug. A 95% confidence interval for the average before – after blood pressure is (1, 15).
 - (a) What do you conclude about the clinical effectiveness of the drug?
 - (b) If the sample size for this study was $n = 25$, what is your estimate of the sample standard deviation of the before – after differences?
2. Plasma clotting time can be affected by applying different treatments to the plasma. If a standard treatment is applied average clotting time is 10 minutes. A new treatment is being tested that may have different clotting times. Samples from 20 patients are treated, and it is found that $\bar{x} = 8$, and $s^2 = 9$. Test the null hypothesis that the clotting times are equivalent versus the alternative that they are different. State the null and alternative hypotheses, carry out the test, provide a p -value, and state your conclusion.
3. There is currently much ongoing research concerning antibodies in rheumatoid arthritis. Among all patients with rheumatoid arthritis, it is found that 40% have marker A, while 60% have marker B. Curiously, these markers seem to be present independently in rheumatoid arthritis patients, so that knowing whether Marker A was present or not gives no information about whether marker B was present or not. What is the probability that at least one of these markers would be present in a randomly selected rheumatoid arthritis patient?
4. The prevalence of disease D in a certain population is known to be 20%. Further, a certain diagnostic test for diagnosing disease D in this population is known to have a positive predictive value of 50%. If the specificity is 80%, what is the sensitivity of the diagnostic test for detecting disease D?
5. Consider three independent random variables, X_1 , X_2 , and X_3 . Suppose that the distributions of these three random variables are as follows:

$$\begin{aligned}X_1 &\sim N(\mu = 0, \sigma^2 = 1) \\X_2 &\sim N(\mu = 4, \sigma^2 = 2)\end{aligned}$$

$$X_3 \sim N(\mu = -4, \sigma^2 = 3)$$

State the distribution of $X_1 + 2 \times X_2 - X_3$, and provide the mean and variance of this distribution.

6. Calculate the following probabilities:

(a) Suppose that a random variable X has distribution

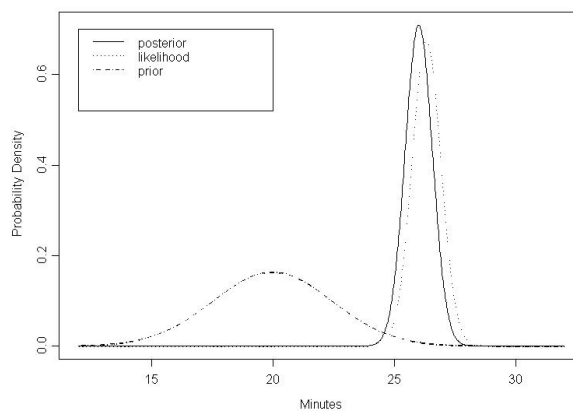
$$X \sim N(\mu = 5, \sigma^2 = 25)$$

Find $\Pr\{X \geq 10\}$.

(b) Suppose that a second random variable Y follows a t distribution with 30 degrees of freedom. Find $\Pr\{Y \leq -1.055\}$.

7. In a certain hospital, 90% of all knee surgeries are classified as “successful”. If 200 such operations are performed, what is the probability that at least 190 of them will be classified as successful?

8. In a fitness test, subjects between 20 and 30 years old are asked to jog until they are tired, and record the length of time in minutes that they could jog before tiring. Data on 10 subjects 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, both in general terms and in the context of this example.

9. A study of the number of patients per room in a certain hospital gave the following results:

Number of patients per room	Number of rooms
1	30
2	30
3	40

The total number of rooms in this hospital is 100. What is the average number of patients per room? What is the variance of the number of patients per room?

10. A study is planned to estimate the average cholesterol levels of Canadians. A random sample will be obtained. It is estimated that the average cholesterol level is 6.0 mmol/L , and the standard deviation of this number is 1.8 mmol/L . What sample size would be required so that the total width of a 95% confidence interval would be 0.1 mmol/L ?