Homework 2 Output

Emil Hafeez (eh2928)

10/3/2020

Problem 1

I use the shorthand $P(T) = P(TestPositive)$, and $P(T-) = P(TestNegative)$, and similar for $P(D)$ where D is the disease at hand.

Part 1

Given that "in a group of 113 patients with prostatic cancer, 79 have a positive diagnosis" and that "in a group of 217 individuals without prostatic cancer, 10 have a positive diagnosis,", we have the following table.

	With Prostatic Cancer	Without Prostatic Cancer
Test Positive (T)	79	10
Test Negative (T-)	34	207
Total	113	217

To calculate the sensitivity and the specificity of the test, we take the sensitivity to be equal to $P(Test|Disease)$ and the specificity to be $P(T-|D-)$. For the sensitivity, equates to $\frac{79}{113}$ and for the specificity to be $\frac{207}{217}$, which is to say the sensitivity equals 0.6991 or 69.91% and the specificity equals 0.9539 or 95.39%.

Part 2

In this other hypothetical scenario, it will not be enough to use only the data provided by this new test being developed to assess its test characteristics like sensitivity and specificity. This is because determining sensitivity and specificity requires reference to a gold-standard test, so that we can compare the accuracy of the new test data to a standard for which the "truth" is known; with just the new test data available, this investigation cannot be made.

Part 3

a) In this example, the $P(T+|D+)$ is given as 0.8 and the $P(T-|D-)$ as 0.95. The $P(D)$ is 0.5. Therefore, we can use Bayes' Theorem and then apply the LTP to the denominator to determine the following $P(D|T)$.

$$P(T|D) = \frac{P(T|D)*P(D)}{P(T)} = \frac{P(T|D)*P(D)}{P(T)} = \frac{P(T|D)*P(D)}{P(T|D)P(D)+(P(T|D-)P(D-)} = \frac{(0.8)(0.5)}{(0.8)(0.5)+(0.95)(0.5)} = 0.4571.$$

This is known as the positive predictive value (PPV).

b) We now calculate the same test characteristic as above in Part 3a, but now we decrease the pre-test probability of the disease to 0.1 instead of 0.5.

$$P(T|D) = \frac{P(T|D)*P(D)}{P(T)} = \frac{P(T|D)*P(D)}{P(T)} = \frac{P(T|D)*P(D)}{P(T|D)P(D)+(P(T|D-)P(D-)} = \frac{(0.8)(0.1)}{(0.8)(0.1)+(0.95)(0.9)} = 0.0856.$$

Problem 2

Part 1 We are given that $Probability(Prediabetic) = P(D) = 0.345$, and it follows from the co,plement that $P(D) + P(D^c) = 1$ such that that $P(D^c) = (1 - P(D)) = (1 - 0.345) = 0.655$

Part 2 The probability that less than (not less than or equal to) 10 of these 50 patients are prediabetic is given by the following equation. We observe that the scenario follows a binomial distribution $X \sim Bin(50, 0.0345)$ where X is a random variable denoting the number of patients that are prediabetic among the 50 randomly selected group. We observe the binomial distribution given that there are a fixed number of trials, trials are independent, there are only two possible outcomes (success or failure on each draw), and the probability of success if fixed for each trial.

Where n is the number of patients total, p is the probability of success, and $1 - p$ is the probability of failure, such that $n = 20, p = 0.345, (1 - p) = 0.655$.

$P(X < 10) = \sum_{x=10}^{20} \frac{10!}{x!(10-x)!}(0.345)^x(1 - 0.345)^{10-x}$

We can use r to compute the code "pbinom(9, 50, 0.345)". We use 9 rather than 10 because we are computing less than 10, not equal to 10.

Part 3 The probability that 34.5% or $\frac{17}{50}$ of these patients are prediabetic is given by the expected value of the binomial distribution, where $E(X) = np$ where $E(X) = \frac{17}{50}$ and $n = 50$. Therefore $p = \frac{E(X)}{n} = \frac{17/50}{50} = 0.34$

10 of these 50 patients are prediabetic is given by the following equation.

Part 4 In some conditions, the Binomial distribution can be approximated by the Poisson distribution: when n is large (> 100) and when the probability of success is small $p < 0.01$. These conditions are not met, and so it would not be appropriate to use this approximation method.

Problem 3

Part a)

Here, the incidence of uveal melanoma in the US can be assumed to follow a Poisson distribution characterizing the number of occurrences of an event (uveal melanoma) given that the events occur one at a time (one case a time per person, no one person can get it twice simultaneously), and the number of expected events is given as a constant.

$P(X = x) = f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, 2, \ldots, n$

Part b)