Math 107

Applications of Contingency tables – Medical Testing – Day 1

Not all medical testing in society is completely effective. It is possible for a medical test to give a positive result and the individual not have the given disease. These results can have implications on public policy. One aspect of this is increasing costs of medical testing when the results are not terribly effective. Another public policy aspect is drug testing – certainly a false positive drug test can have serious repercussions for an individual. You might be surprised that probability plays a big role in medical and drug testing. In order to understand the specifics of this important applications we need to understand the concept of conditional probability

Conditional Probability is the probability of an event occurs given that another has occurred. We have a special notation:
P(A|B) is said to be the probability of A given B

The best way to understand conditional probability is to look at some examples:

Example:

Much has been made in the recent presidential election concerning the turnout by different groups of voters and the outcome of the election. Consider the table given below of voters from a region in the U.S.

	Caucasian	African American	Latino	Asian
Voted for Obama	10,825	4,562	3516	4812
Voted for Romney	10,516	986	1245	3602

a) What is the probability that a person voted for Obama given they are Latino?

b) What is the probability Person is Caucasian given they voted for Romney?

c) What is the probability that person voted for Obama given they are African American?

d) What is the probability a person is African American given they voted for Romney?

Example:

	Has Skin Cancer	No Skin Cancer
Visits Tanning Parlor frequently	12500	1560
Does not visit tanning parlor frequently	3000	14,800

a) What is the probability a person selected from this chart visits a tanning parlor frequently?
b) What is the probability that a person selected from this chart has skin cancer
c) What is the probability that a person selected from this chart has skin cancer given they visit a tanning parlor frequently?
d) What is the probability that a person selected from this chart visits a tanning parlor frequently given they have skin cancer?

SENSITIVITY and SPECIFICITY

There is a specific application of conditional probability related to medical testing. An accurate medical test should detect the disease when it is present and give negative results when it is not present. The results fall into four categories

TRUE POSITIVES People who test positive and have the disease	**FALSE POSITIVES** People who test positive but do not have the disease
FALSE NEGATIVE People who test negative but do have the disease	**TRUE NEGATIVES** People who test negative and do not have the disease

The **sensitivity** of a test is the probability that the test will detect the disease in person who does have the disease

Sensitivity = $\frac{\text{True positives}}{\text{True positives plus false negatives}}$

The **specificity** of a test is the probability that the test will give a negative result for a person who does not have the disease

$$\textbf{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

Sensitivity and Specificity measure how accurate the test will give desired results. There are two other conditional probabilities that are also useful

Among a population of individuals, **the positive predictive value (PPV)** of a test is the probability that a person in the population who tests positive actually has the disease.

$$PPV = \frac{\text{True positive}}{\text{All positives}} = \frac{\text{True Positives}}{\text{True positives} + \text{False positives}}$$

Among a population of individuals, **the negative predictive value (NPV)** of a test is the probability that a person who tests negative actually does not have the disease

$$NPV = \frac{\text{True Negatives}}{\text{All Negatives}} = \frac{\text{True Negatives}}{\text{True negatives} + \text{False negatives}}$$

One more definition that is critical is The **prevalence** of a disease in a given population is the percentage of the population that has the disease.

Example: Hepatitis C is an infection caused by a virus that attacks the liver and leads to inflammation. Here is a table of results from the general population concerning a test for Hep C

	Has Hep C	No Hep C
Test positive	9	99
Test negative	1	891

a) According to this test what is the prevalence of Hep C?

b) What is the sensitivity and specificity of this test?

c) What is the PPV and NPV?

d) What conclusions can you draw from these results?

Example: Hepatitis C is an infection caused by a virus that attacks the liver and leads to inflammation. Here is a table of results for a population of older and former injection drug users concerning a test for Hep C

	Has Hep C	No Hep C
Test positive	720	20
Test negative	80	180

a) According to this test what is the prevalence of Hep C in this population?

b) What is the sensitivity and specificity of this test?

c) What is the PPV and NPV?

d) What conclusions can you draw from these results? As a policy maker what would you suggest given these two examples concerning Hep C?

Example: Crohn's disease is a chronic inflammatory bowel disease that affects the lining of the digestive tract. Here is a table of results for the general population of the United States for this disease (numbers are in millions)

	Has Crohns	No Crohns
Test positive	0.46	31.30
Test negative	0.12	288.12

a) According to this test what is the prevalence of Crohn's in this population?

b) What is the sensitivity and specificity of this test?

c) What is the PPV and NPV?

d) What conclusions can you draw from these results? As a policy maker what would you suggest given these two examples concerning Crohn's disease?

Example: HIV , while a fatal disease, is fairly rare with only 0.4% of the general population who has the disease. Suppose there are 320 million individuals living in the United States. Also assume the AIDS test that is generally used (there are many) has a sensitivity and specificity of 99%

a) Fill in the Chart Below (yes you have enough information)

	Has HIV	No HIV	Totals
Tests Positive			
Tests Negative			
Totals			

b) Calculate the PPV and NPV –what can you conclude about these results?