**BIOS 4120: Introduction to Biostatistics**
**Breheny**

# Lab #6

*Note: We will not be using R for this lab.*

Today's lab is about probability. You have learned some very useful properties about probabilities in class, now it's time to provide some examples of when these properties come into play. Then we'll talk a little bit about conditional probabilities, and when they come in handy. Last we will do an example or two of Bayes' rule, learning how to calculate sensitivity and specificity along the way.

## 1 Probability Properties

- **Addition Rule:** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Complement Rule:** $P(A^C) = 1 - P(A)$
- **Multiplication Rule:** $P(A \cap B) = P(A)P(B|A)$ or $P(A \cap B) = P(B)P(A|B)$
    - *If A,B are independent:* $P(A \cap B) = P(A)P(B)$
- **Law of Total Probability:** $P(A) = P(A \cap B) + P(A \cap B^C)$
- **Bayes' Rule:** $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^C)P(B|A^C)}$
    - *More on this later*

## 2 Probability Question Examples

a) Conditional Probability Card Game – Guess my card using the rules of 20 questions! (i.e. is it an 8? Is it a heart? Etc.)

b) In a class of 101 students, 30 are freshmen and 41 are sophomores. Find the probability that a student picked from this group at random is either a freshman or a sophomore.

c) Suppose two 6-sided dice are rolled. What is the probability that I roll two sixes? What assumption(s) are you making about the dice?

d) A different class, still with 101 students, has 40 juniors. Also, the class has 50 females, 22 of whom are juniors. Find the probability that a student picked from this group at random is either a junior or a female.

e) Blind man's bluff: I'm going to show you my card, then shuffle the cards again (with my card included) and deal one to you. You have to determine if your card will be higher or lower than mine. Assume aces are high. What is the probability that your card is greater than or equal to my card? Would you bet that your card is greater than or equal to mine?

f) (I will pass out cards now). Check your card. Write down whether or not you won this round of blind man's bluff (i.e. whether or not your card beat my card).

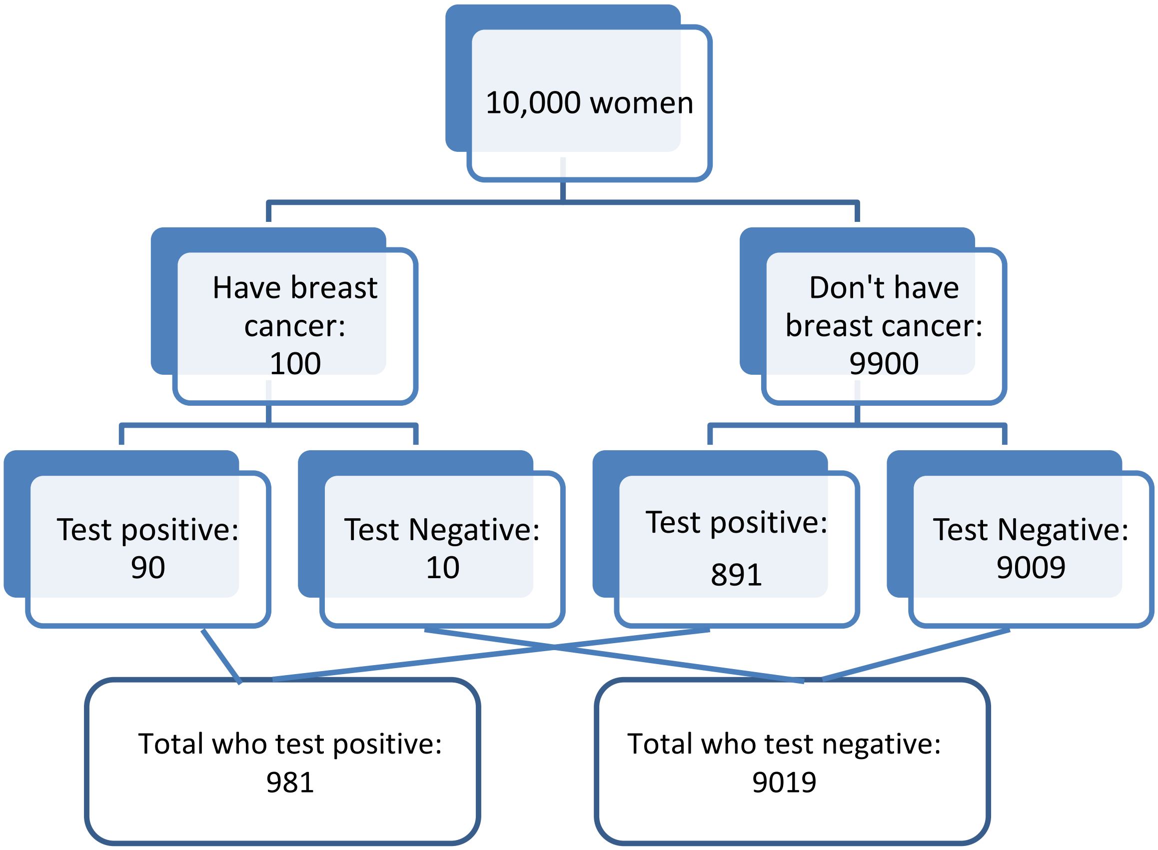
# 3 Screening/Diagnostic Test Examples

**Example 1:**
Scenario (from Gigerenzer 2007): The prevalence of breast cancer in women is 1%. A mammography has sensitivity of 90% (90% of women with breast cancer will test positive) and a false positive rate of 9%.

Of the four choices below, which best describes what you would tell a patient with a positive mammogram?

a) It is not certain you have breast cancer, yet the probability is about 81%.
b) Out of 10 women who test positive as you did, about 9 have breast cancer.
c) Out of 10 women who test positive as you did, only about 1 has breast cancer.
d) The chance you have breast cancer is about 1%.

- **Probability tree**




**Now let's look at the probability a person has the disease given that they test positive.**

$$P(cancer|+) = \frac{90}{90+891} = \frac{90}{981} = 0.092$$

- **Sensitivity, specificity, prevalence**
  - **Prevalence:** P(cancer) = 0.01
  - **Sensitivity:** P(+ | cancer) = 0.9
  - **False Positive rate:** P(+ | no cancer) = 0.09
  - **Specificity**: P(- | no cancer) = 1-FPR = 1-0.09 = 0.91

**Example 2:**

The following data are taken from a study investigating the use of a technique called radionuclide ventriculography as a diagnostic test for detecting coronary artery disease. Assume the participants were randomly selected from the population (Exercise 6.15)

| **Test** | **Disease** | | **Total** |
|---|---|---|---|
| | **Present (D)** | **Absent ($D^c$)** | |
| + | 302 | 80 | 382 |
| - | 179 | 372 | 551 |
| **Total** | 481 | 452 | 933 |

- **Find P(+ | D)**
  - Otherwise known as the *sensitivity*

- **Find P(D)**
  - Otherwise known as *prevalence*

- **Find P(+ | $D^c$)**
  - Otherwise known as the *false positive rate*

- **Find P(- | $D^c$)**
  - Otherwise known as the *specificity*

- **Find P( D | +)** (from the table and with Baye's Theorem)
  - Otherwise known *predictive value positive*