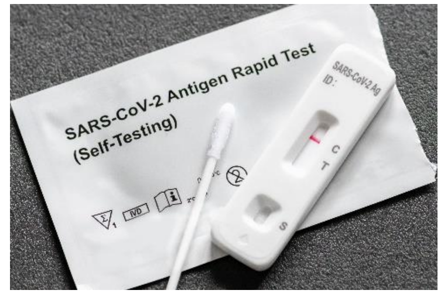
Chapter 6: Measuring Test Accuracy

Investigation: Rapid antigen tests are immunoassays that detect the presence of a specific viral antigen. They can be convenient for at home use without a clinical professional, but they are not always highly accurate due to the fact that 1) they can't detect small viral loads like PCR tests, and 2) human error in actual implementation and reading of the results.



In the following fictional study, 3,000 people agreed to be tested for the SARS-CoV-2 virus (the virus that leads to COVID-19) during a time of relatively high viral transmission. These people were asked to follow the instructions and identify whether their test was positive or negative.

These people were then tested with a highly reliable PCR test to determine how well the antigen test performed in identifying SARS-CoV-2. The results are in the following table.

	Verified Virus	Verified No Virus	Totals
Antigen test Negative	15 (False Negative)	2,818 (True Negative)	2,833
Antigen test Positive	80 (True Positive)	87 (False Positive)	167
Totals	95	2,905	3,000

- Let's start by thinking about probabilities for simple events involving one grouping at a time.
  - Let's define "Virus" be the event that someone truly has the SARS-CoV-2 virus.
  - Let "+" be the event that someone gets a positive test result.

What proportion of these residents were truly infected with the virus?

P(Virus) = 95/3000

What proportion of these residents received a positive test result?

P(+) = 167/3000

- Now, let's consider probabilities involving compound events.
  - ∪ **stands for "union."** A ∪ B would be every possibility where at least one of these two events A and B are true.
  - ∩ **stands for "intersection,"** A ∩ B would be every possibility where BOTH A and B are true.

What proportion of these residents were either truly infected with the virus, or had a positive test result?

P(Virus ∪ +) = (15+80+87)/3000

What proportion of these residents got a true positive?

P(Virus ∩ +) = 80/3000

- Finally, let's consider this investigation if we examine conditional probabilities.
  - This would be the probability of one event given known information about another
  - Notation might be something like P(B|A), which would be read "What is the probability of B given that A is true?"
  - We can represent this in a formula. P(B|A) = $\frac{P(A \cap B)}{P(A)} = \frac{\text{(Probability of both A and B happening)}}{\text{(Probability of A happening)}}$

Test Accuracy Estimated from Data

- In the context of test effectiveness, there are 4 different conditional probabilities we might examine
  - **Positive Predictive Value (PPV):** Percentage of positive test results that are correct
    - *Probability of a true result **given** the test result is "positive"*
  - **Negative Predictive Value (NPV):** Percentage of negative test results that are correct
    - *Probability of a true result **given** the test result is "negative"*
  - **Sensitivity:** Percentage of people who have the "condition" who correctly test positive
    - *Probability of a true result **given** the person has "condition of interest"*
  - **Specificity:** Percentage of people who don't have the "condition" who correctly test negative
    - *Probability of a true result **given** the person does not have the "condition of interest"*

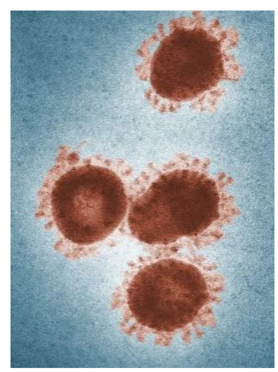
Practice: For each question, identify which measure is being asked for and estimate it from the data

What proportion of people verified with the SARS-CoV-2 virus received a positive test result?

P(+ | virus) = 80/95

What proportion of people with a negative test result were correctly identified?

P(no virus | -) = 2818/2833



Confidence Intervals for Test Accuracy

- Each measure is technically a proportion. Therefore, we can create a confidence interval for each proportion following the same method we've used before

Z-Interval for π: $\hat{p} \pm Z_{\alpha/2} * \sigma_{\hat{p}}$ where $\sigma_{\hat{p}} \approx \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

What would be our 95% confidence interval for the sensitivity of this test?

Point estimate:

SE:

Margin of error:

Remember that assumptions still apply for z-intervals here. In particular…

- We are assuming that the distribution of possible sample proportions is normally distributed
  - **The 10/10 Rule**: This is a reasonable assumption to make if our sample has **at least 10 of each response (e.g., >=10 "yes" and >=10 "no" responses)**
  - When this isn't true, the distribution of $\hat{p}$ might have some skew, or be too discrete to reasonably use a normal approximation.

Test Accuracy Derived from Theoretical Probabilities

- To enrich your conceptual understanding of test effectiveness, let's dig more deeply into the idea of unions, intersections, and conditional events from a "theoretical" probability perspective.
- …But let's start with a more fun example first!

Example: At an ice cream stand, customers have the choice of paying extra for a cone, or simply choosing a bowl. Furthermore, they have the choice of purchasing two scoops or just one scoop. https://setosa.io/conditional/

Let's call A the event that a customer pays extra for a cone, *and A' that they choose a bowl.*

Let's call B the event that a customer pays extra for 2 scoops*, and B' that they choose one scoop.*

In context, what would be the event A ∪ B? What color(s) are represented by that event?

<u>All of the red, purple and blue. All customers who purchase a cone or 2 scoops (or both)</u>

In context, what would be the event A ∩ B? What color(s) are represented by that event?

<u>Just the purple. All customers who purchase a cone with 2 scoops</u>

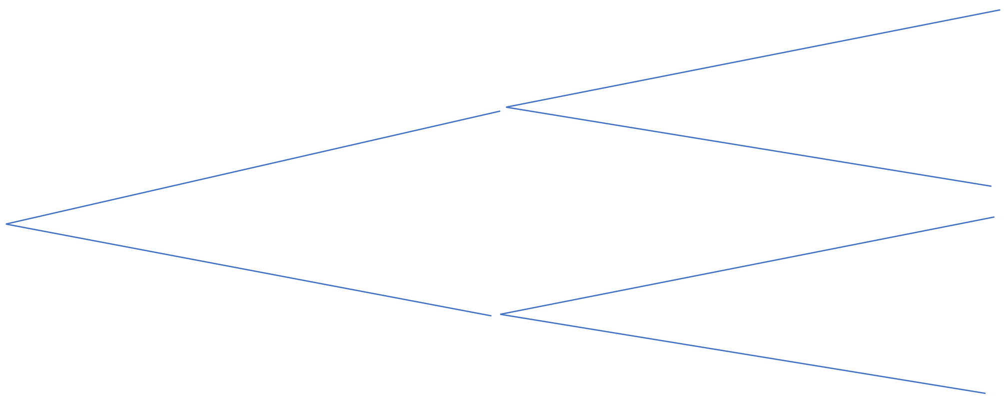
How would we find P(B | A)? And how would we write this in context to this example?

<u>Of those who purchase a cone, what proportion also purchase 2 scoops? 0.1/0.2</u>

Let's say that it's actually 40% of customers who purchase a cone, 65% who purchase 2 scoops, and 15% are purchasing both. Now, what is P(B | A)?

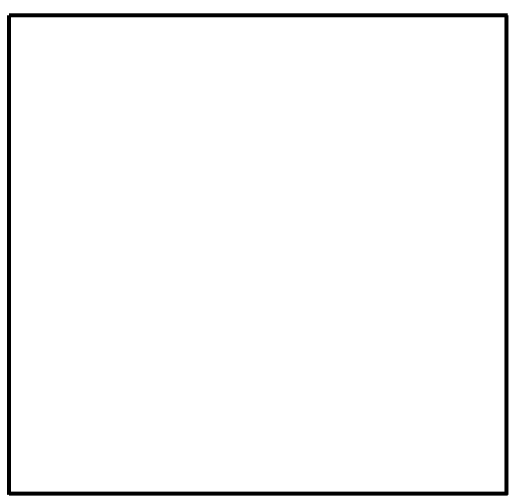
<u>0.15/0.4 = 0.375</u>

- **Tree Diagrams** are a great way to visualize compound events when wanting to identify conditional probabilities.
  - The first set of stems represent a simple event and its complement, and each of these probabilities should add up to 1
  - The second set of stems represent conditional probabilities given the stem they are linked to. Each **grouping** should represent possible conditions and should also add up to 1
  - *To save some extra calculation, the P(B | A') = 0.833*



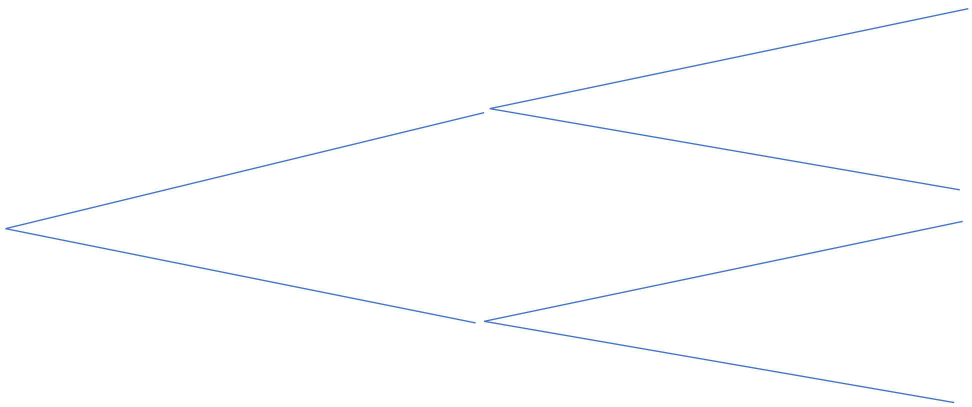
- Now, multiply down each branch to find the intersection probabilities.

- **Mosaic Plot** *(if time).* Everyone can be categorized as being in exactly one of four possible intersections, and we can represent those probabilities proportionally in the area below.



Investigation revisited: Let's return to the antigen test again and think about what results we might see with the following assumed information:

- Let's say that among people **with** the virus, **82%** of them will actually report a **positive** antigen test.
- Let's say that among people **without** the virus, **4%** of them will incorrectly report a **positive** antigen test
- Let's also assume that the **true incidence rate** in this community is **3%.**
- Let's finally assume that there are no inconclusive results (everyone identifies their test result as positive or negative).



What is the specificity of this test?

What's the probability that someone in this community gets an incorrect test result? *(hint: there are two ways to get an incorrect test result. What do they sum to?)*

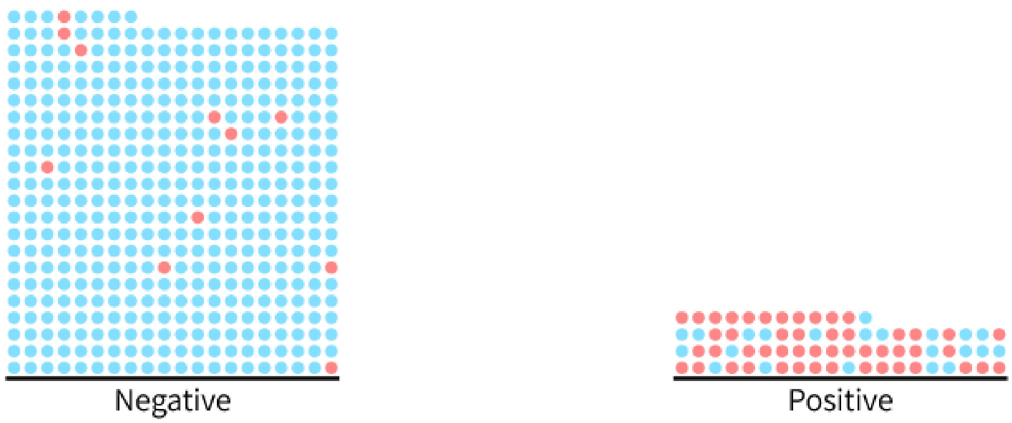
What is the positive predictive value of this test?

If time…look at the following simulation: https://seeing-theory.brown.edu/bayesian-inference/index.html#section1.

Let's say you were to get a positive test result and you'd like to know the true likelihood you have the virus. How does this probability change depending on the true rate of infection among those tested?

Visualizing Test Accuracy

- What proportion of the positive test results are correct? **<u>PPV</u>**
- What proportion of the negative test results are correct? **<u>NPV</u>**
- What proportion of the **red** (darker) dots are in the right place? **<u>Sensitivity</u>**
- What proportion of the **blue** (lighter) dots are in the right place? **<u>Specificity</u>**

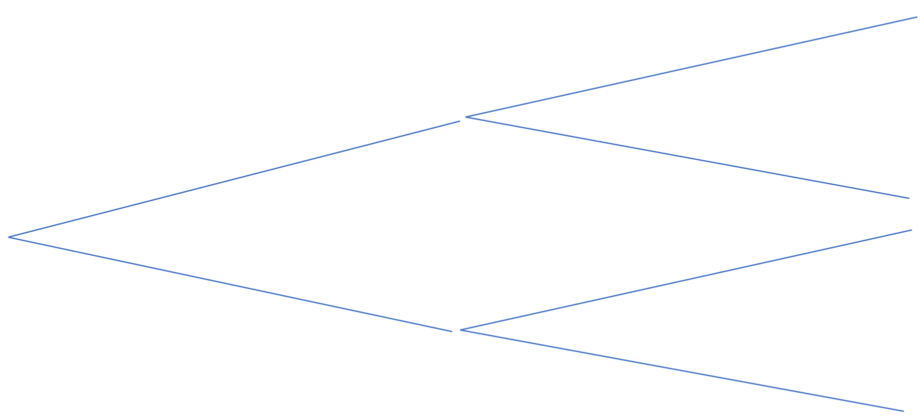


Chapter 6 Additional Practice

Practice: A test is used to screen for the presence of a particular drug in one's blood system.

- Let's assume for this example that 8% of the people we are screening truly have drugs in their system.
- Let's also assume that this drug test will produce a negative for 12% of users who truly have drugs in their system.
- Let's also assume it will produce a negative for 96% of users who truly don't have drugs in their system.

Use this information to fill in the tree diagram below



What is the test's sensitivity?

What is the test's specificity?

What is the test's negative predictive value *(if we assume an 8% incidence rate)*?

Now, let's assume we didn't know the true sensitivity or specificity of this test. We gather a sample of 500 people, of which 8% of them are known to have used drugs in the past 24 hours. We see the following results.

	Positive Test Result	**Negative Test Result**	**Totals**
Subject used drugs	31	9	40
Subject did not use drugs	24	436	460
Totals	55	445	500

Based on the sample data provided, what would be our estimate for specificity? And what would be the 95% margin of error on our estimate?

If the true proportion of test-takers who used drugs were to go down from 8%, would you expect the positive predictive value to get better or get worse?

Chapter 6 Learning Goals

After this chapter, you should be able to…

- Distinguish a true positive, true negative, false positive, and false negative in context
- Calculate the estimated probability of a simple event from frequency table
- Calculate the estimated probability of an event union or intersection from a frequency table
- Calculate the estimated conditional probability of an event from a frequency table
- Distinguish between sensitivity, specificity, positive predictive value, and negative predictive value, both in context, and from a frequency table
- Calculate a 95% confidence interval (or 95% margin of error) for a measure of test effectiveness
- Use a tree diagram to represent how one categorical variable's likelihood conditionally relates to the outcome of another categorical variable.
- Derive the probability of an intersection of events from a tree diagram
- Derive sensitivity, specificity, positive predictive value, and negative predictive value from a tree diagram.
- Reason about how the positive of negative predictive value might change based on the incidence rate of the population being tested.