# Stat 201: Introduction to Statistics

Standard 16: Probability – Medical Testing Terminology

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- Let Event…
  - D = Has Disease
  - $D^c$= Disease Free
  - Pos = Positive Test
  - Neg = Negative Test

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | **48** | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- A **true positive** is when a participant tests positive for the disease and does have it
- Here there are 48 true positives

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | **3921** | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- A **true negative** is when a participant tests negative for the disease and doesn't have it
- Here there are 3,921 true negatives

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | **1307** | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- A **false positive** is when a participant tests positive for the disease but doesn't have it
- Here there are 1,307 false positives

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | **6** | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- A **false negative** is when a participant tests negative for the disease but does have it
- Here there are 6 false negatives

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | **1355** | 3927 | **5282** |

- The probability a randomly selected participant had a positive test

$$\mathrm{P}(Pos) = \frac{Number\ of\ positive\ observations}{Total\ number\ of\ observations}$$

$$= \frac{1355}{5282} = .25653162$$

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | **54** |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | **5282** |

- The probability a randomly selected participant had Disease:

$$P(D) = \frac{Number\ of\ D\ observations}{Total\ number\ of\ observations} = \frac{54}{5282} = .0102234$$

# Example 2: Probability

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | **48** | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | **5282** |

- The probability a randomly selected participant had Disease given they tested positive

$$P(D|Pos) = \frac{P(D\&Pos)}{P(Pos)} = \frac{\left(\frac{48}{5282}\right)}{.25653162}$$

$$= .03542435$$

# Example 2: Probability

| | **Positive Test (Pos)** | **Negative Test (Neg)** | **Total** |
|---|---|---|---|
| **Has Disease (D)** | **48** | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | **1355** | 3927 | 5282 |

- As before we can rewrite this as:

$$P(D|Pos) = \frac{Number\ of\ D\&Pos\ Observations}{Number\ of\ Pos\ Observations}$$

$$= \frac{48}{1355} = .03542435$$

# Example 2: Probability

- $P(D|Pos) = .03542435$
- $P(D) = .0102234$

- Because $P(D|Pos) \neq \; P(D)$ events D and POS are **not independent** events

# Adjectives for Tests

- **Sensitivity** – Probability that a test detects a substance correctly by giving a positive test result
  - $Sensitivity = P(Pos|Substance)$
- **Specificity** – Probability that a test correctly does not detect a substance by giving a negative result
  - $Specificity = P(Neg|Substance^c)$

# Example 3: Adjectives for Tests

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | **48** | 6 | **54** |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | **5282** |

- $Sensitivity = P(Pos|D) = \frac{P(Pos\ \&\ D)}{P(D)} = \frac{\left(\frac{48}{5282}\right)}{\left(\frac{54}{5282}\right)} = \frac{48}{54} = .8889$

# Example 3: Adjectives for Tests

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| Has Disease (D) | **48** | 6 | **54** |
| Disease Free ($D^c$) | 1307 | 3921 | 5228 |
| Total | 1355 | 3927 | 5282 |

- As before we can rewrite this as:

$$Sensitivity = \frac{number\ of\ Pos\&D\ observations}{number\ of\ D\ observations} = \frac{48}{54} = .8889$$

# Example 3: Adjectives for Tests

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- $Specificity = P(Neg|D^c) = \frac{P(Neg \,\&\, D^c)}{P(D^c)} = \frac{\left(\frac{3921}{5282}\right)}{\left(\frac{5228}{5282}\right)} = \frac{3921}{5228} = .75$

# Example 3: Adjectives for Tests

| | Positive Test (Pos) | Negative Test (Neg) | Total |
|---|---|---|---|
| **Has Disease (D)** | 48 | 6 | 54 |
| **Disease Free ($D^c$)** | 1307 | 3921 | 5228 |
| **Total** | 1355 | 3927 | 5282 |

- As before we can rewrite this as:

$$Specificity = \frac{number\ of\ Neg\&D^c\ observations}{number\ of\ D^c\ observations}$$

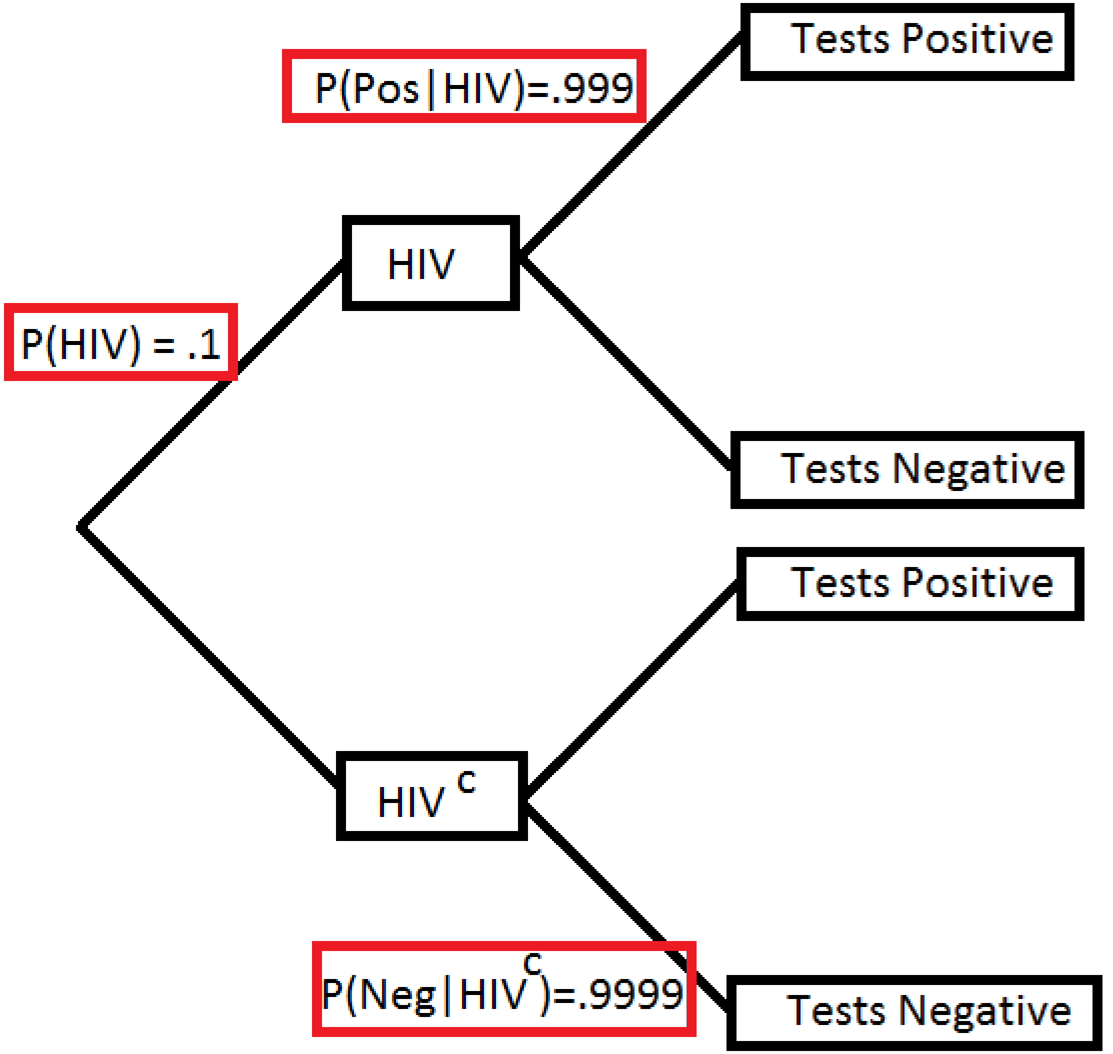
$$= \frac{3921}{5228} = .75$$

# Example 4: Tree Diagrams

- For a Western blot blood test the sensitivity is about .999 and the specificity is about .9999
  - $Sensitivity = P(Pos|HIV) = .999$
  - $Specificity = P(Neg|HIV^c) = .999$

- Consider a high risk group here 10% are truly HIV positive
  - P(HIV)=.1

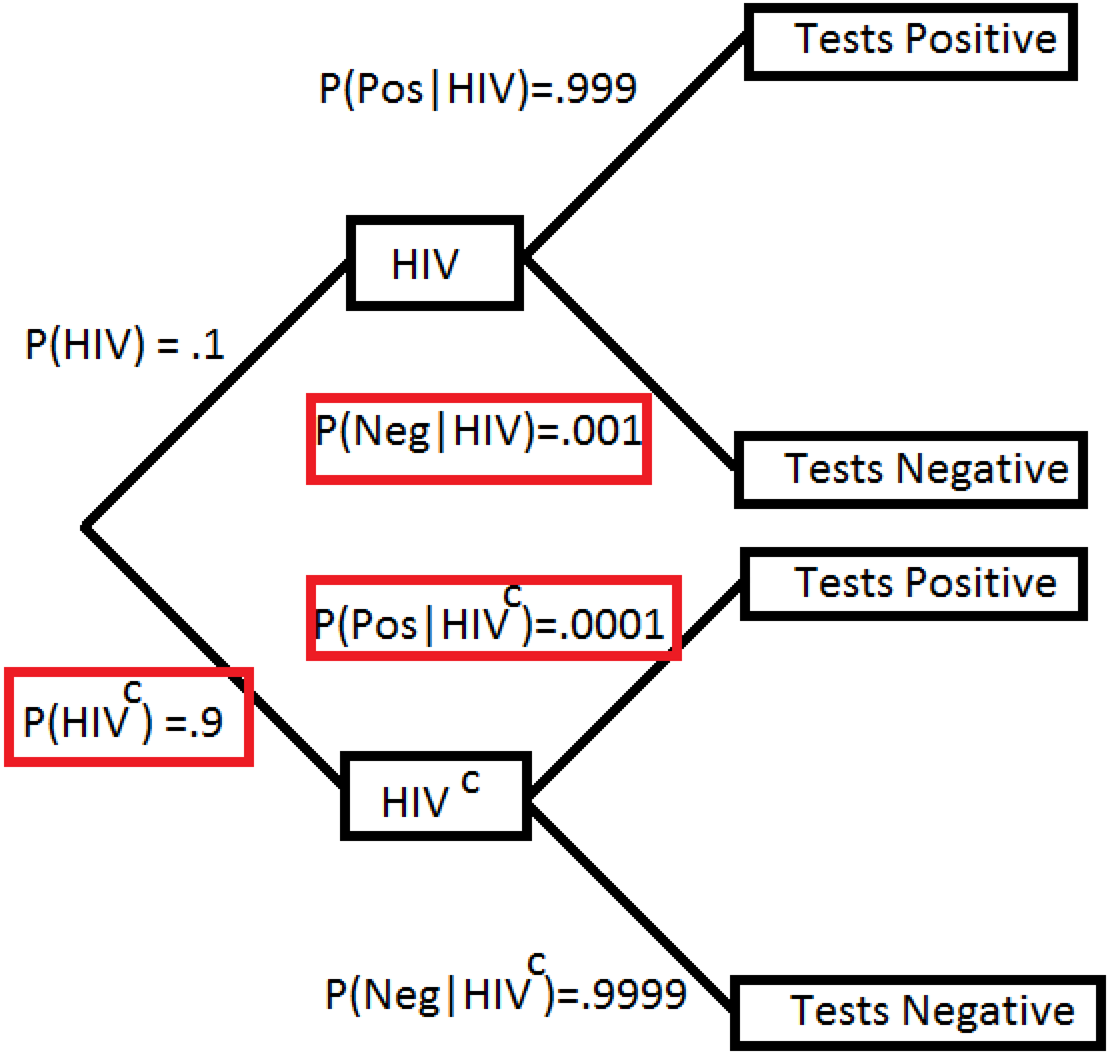
# Example 4: Tree Diagrams



# Example 4: Tree Diagrams

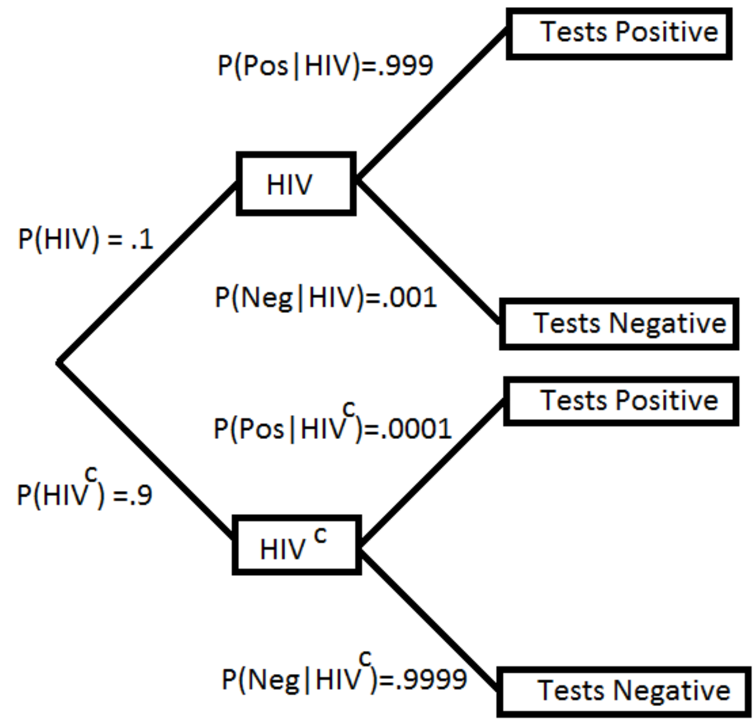
- Using Complement Rule we can find:
  - $P(Neg|HIV) = P(Pos^c|HIV) = 1 - P(Pos|HIV) = 1 - .999 = .001$
  - $P(Pos|HIV^c) = P(Neg^c|HIV^c) = 1 - P(Neg|HIV^c) = 1 - .9999 = .0001$
  - $P(HIV^c) = 1 - P(HIV) = 1 - .1 = .9$

# Example 4: Tree Diagrams



# Example 4: Tree Diagrams

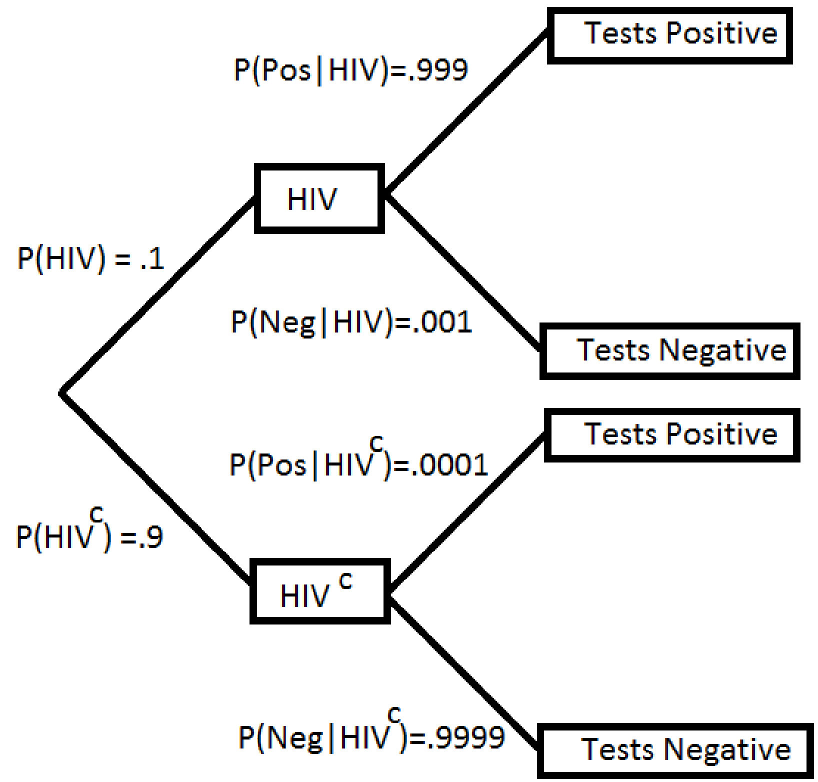
- Note, using the complement rule helped us find the rest of the probabilities
- The branches leaving any point should have probabilities that add to one



# Example 4: Tree Diagrams

- The probability that a randomly selected participant tests positive **and** has HIV:

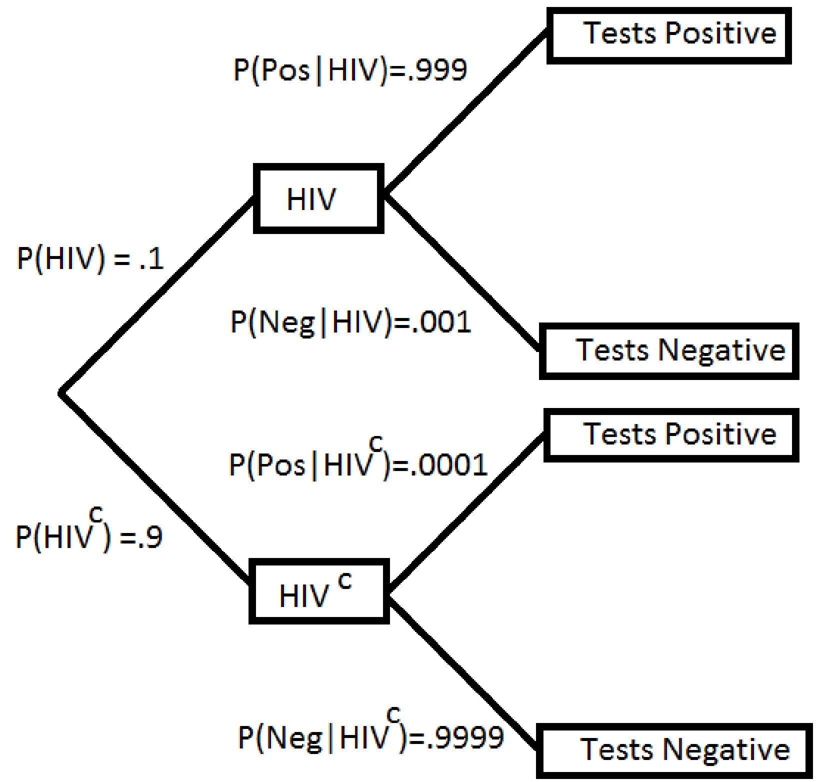
$$P\left(HIV \bigcap Pos\right)$$
$$= P(HIV) * P(Pos|HIV)$$
$$= .1 * .999 = .0999$$



# Example 4: Tree Diagrams

- The probability that a randomly selected participant tests negative **and** has HIV:

$$P\left(HIV \bigcap Neg\right)$$
$$= P(HIV)$$
$$* P(Neg|HIV)$$
$$= .1 * .001$$
$$= .0001$$



# Example 4: Tree Diagrams

- The probability that a randomly selected participant tests positive **and** doesn't have HIV:

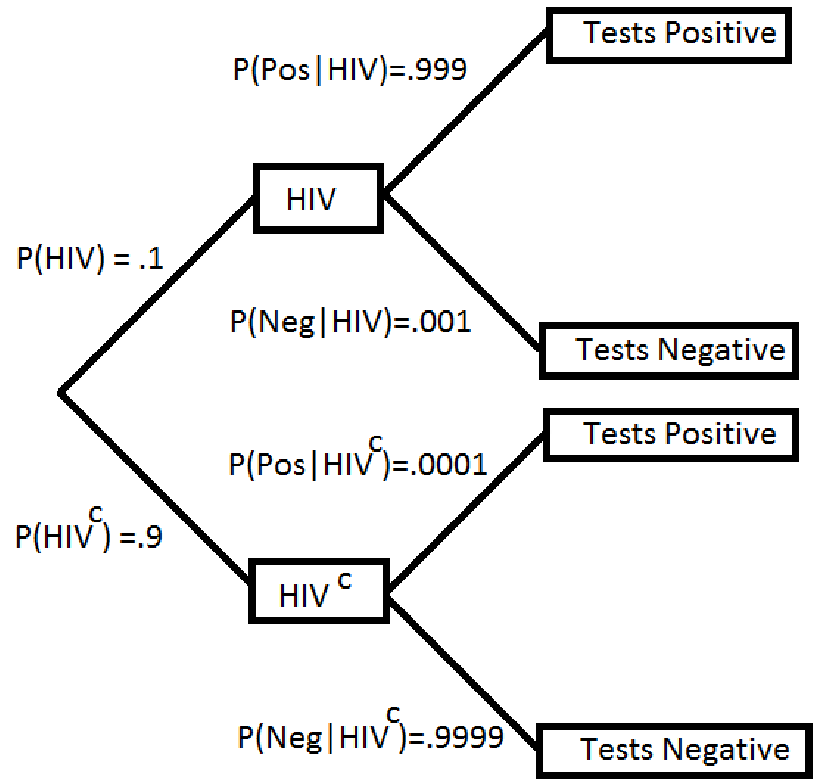
$$P\left(HIV^c \bigcap Pos\right)$$

$$= P(HIV^c)$$

$$* \mathrm{P}(\mathrm{Pos}|\mathrm{HIV}^c)$$

$$= .9 * .0001$$

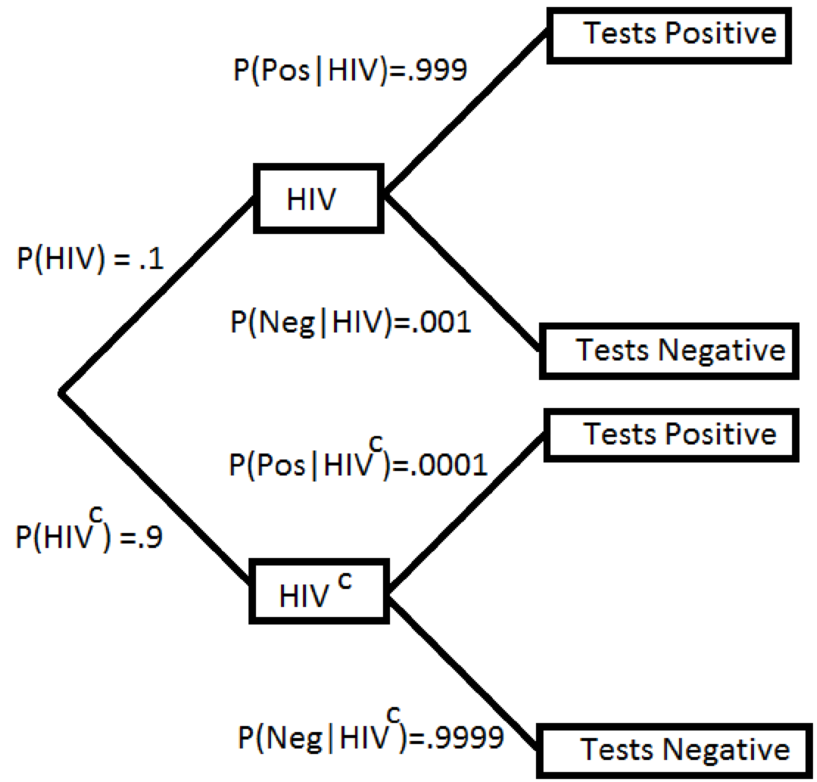
$$= .00009$$



# Example 4: Tree Diagrams

- The probability that a randomly selected participant tests negative **and** doesn't have HIV:

$$P\left(HIV^c \bigcap Neg\right)$$
$$= P(HIV^c) * \mathrm{P(Neg|HIV^c)}$$
$$= .9 * .9999$$
$$= .89991$$



# Example 4: True Positive

- The probability a randomly selected person has HIV given they tested positive
  - $P(HIV|Pos) = \frac{P(HIV \cap Pos)}{P(Pos)} = \frac{P(HIV \cap Pos)}{P(Pos|HIV^c)P(HIV^c)+P(Pos|HIV)P(HIV)} = \frac{.0999}{.0001*.9+.999*.1} = .999$
- Since this is very close to one, if you test positive, it is very likely that you have HIV

# Example 4: True Positive

- The probability a randomly selected person has HIV given they tested positive
  - $P(HIV|Neg) = \frac{P(HIV \cap Neg)}{P(Neg)} = \frac{.0001}{.9} = \ .0001$

- Since this is very close to zero, if you test negative, it is very unlikely that you have HIV