

# Stat 140: Probability Examples

## Bayes Rule

### AIDS Testing

The ELISA test for AIDS is used in the screening of blood donations. As with most medical diagnostic tests, the ELISA test is not infallible. If a person actually carries the AIDS virus, experts estimate that this test gives a positive result 97.7% of the time. (This number is called the *sensitivity* of the test.) If a person does not carry the AIDS virus, ELISA gives a negative result 92.6% of the time (the *specificity* of the test). Recent estimates are that 0.5% of the American public carries the AIDS virus (the *base rate* with the disease).

**(a) When the American Red Cross receives blood donations, they test them for HIV/AIDs. Suppose a blood donation from a particular person has tested positive. Given this information, how likely do you think it is that the person actually carries the AIDS virus? Don't do any calculations for now, just make a guess.**

SOLUTION:

Let's define some relevant events:

A = the event that a (randomly selected) person carries the AIDs virus B = the event that a (randomly selected) person tests positive for AIDs

**(b) From the description above, determine the following (note that there are no calculations to do here - just matching up numbers in the text to probabilities of events):**

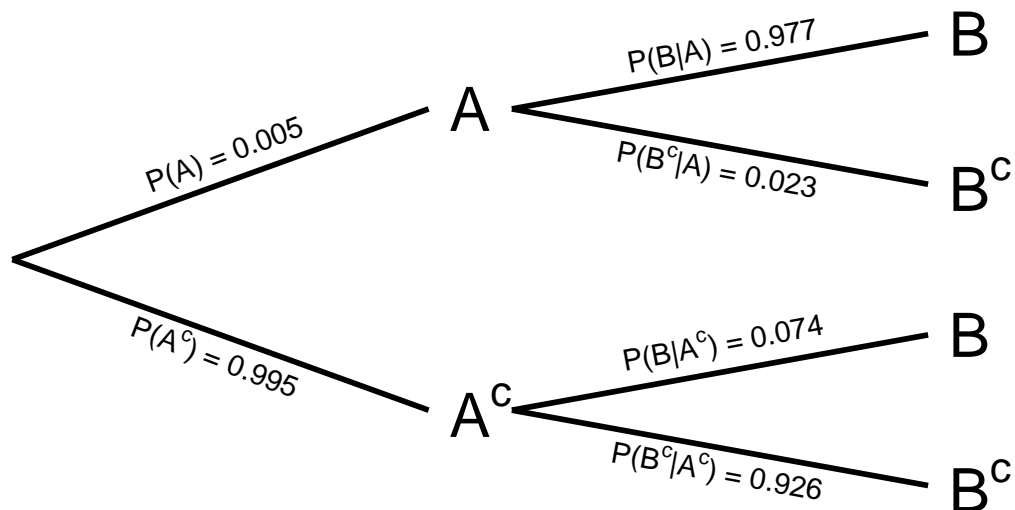
$$P(A) = 0.005$$

$$P(B|A) = 0.977$$

$$P(B^c|A^c) = 0.926$$

**(c) Calculate the probability that someone who has tested positive for AIDs actually has AIDs.**

Always make a diagram! (next page)



We want to calculate  $P(\text{person has AIDS, given that they tested positive for AIDS})$ . In symbols, this is  $P(A|B)$ . We can calculate this using Bayes' Rule as follows:

$$\begin{aligned}
 P(A|B) &= \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)} \\
 &= \frac{0.005 \times 0.977}{0.005 \times 0.977 + 0.995 \times 0.074} \\
 &= 0.062
 \end{aligned}$$

# Email Spam

According to Wikipedia, about 80% of all email that is sent is spam ([https://en.wikipedia.org/wiki/Email\\_spam](https://en.wikipedia.org/wiki/Email_spam)). According to this article in Wired magazine, as of 2015, Google's spam filter for Gmail correctly identifies 99.9% of all spam emails as spam, and only incorrectly identifies 0.05% of non-spam emails as spam (<https://www.wired.com/2015/07/google-says-ai-catches-99-9-percent-gmail-spam/>).

Define A to be the event that an email is a spam message, and B to be the event that an email is identified as spam by Gmail's filters.

(a) From the description above, determine the following (note that there are no calculations to do here - just matching up numbers in the text to probabilities of events):

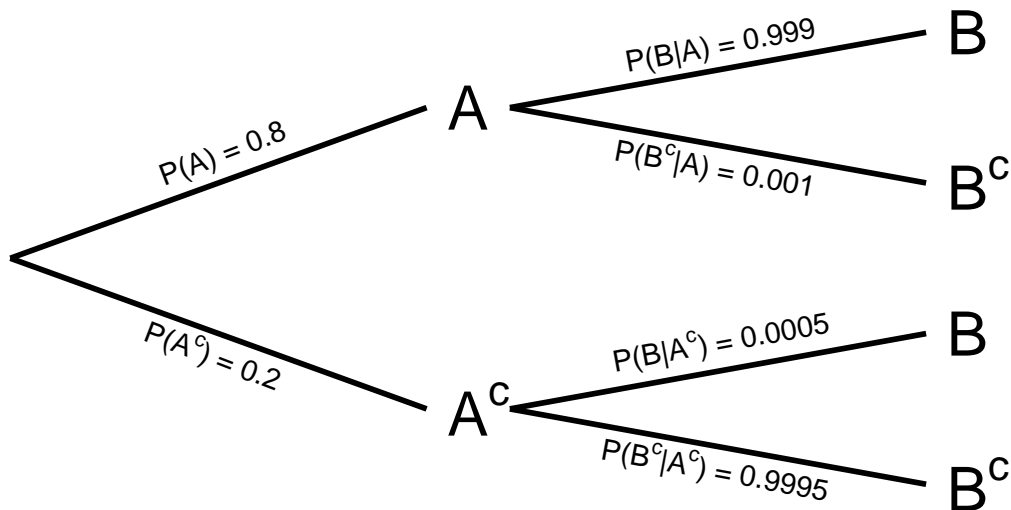
$$P(A) = 0.8$$

$$P(B|A) = 0.999$$

$$P(B|A^c) = 0.0005$$

(b) Find the probability that the next email that makes it to my inbox will be a spam message. That is, find  $P(A|B^c)$ .

Always make a diagram!



The usual statement of Bayes' Rule is  $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)}$ . However, you can switch around the letters and complements to get different variations on this rule. Here's one variation:  $P(A|B^c) = \frac{P(A)P(B^c|A)}{P(A)P(B^c|A) + P(A^c)P(B^c|A^c)}$ . I got this by taking the original version of Bayes' Rule, and replacing all of the B's with  $B^c$ .

From Bayes' Rule,

$$\begin{aligned} P(A|B^c) &= \frac{P(A)P(B^c|A)}{P(A)P(B^c|A) + P(A^c)P(B^c|A^c)} \\ &= \frac{0.8 \times 0.001}{0.8 \times 0.001 + 0.2 \times 0.9995} \\ &= 0.004 \end{aligned}$$

## Drunk Driving

Police often set up roadblocks where drivers are asked a few questions to allow the officer to judge whether or not the person may have been drinking. If the officer does not suspect a problem, drivers are released to go on their way. Otherwise, drivers are detained for a Breathalyzer test that will determine whether or not they will be arrested. The police say that based on the brief initial stop, trained officers can make the correct decision 80% of the time. Suppose the police operate a sobriety checkpoint after 9:00pm on a Saturday night, a time when national traffic safety experts suspect that about 12% of drivers have been drinking.

**(a) You are stopped at the checkpoint and, of course, have not been drinking. what's the probability that you are detained for further testing?**

Let's define the events  $A$ : the driver was drinking; and  $B$ : the police decide to detain them.

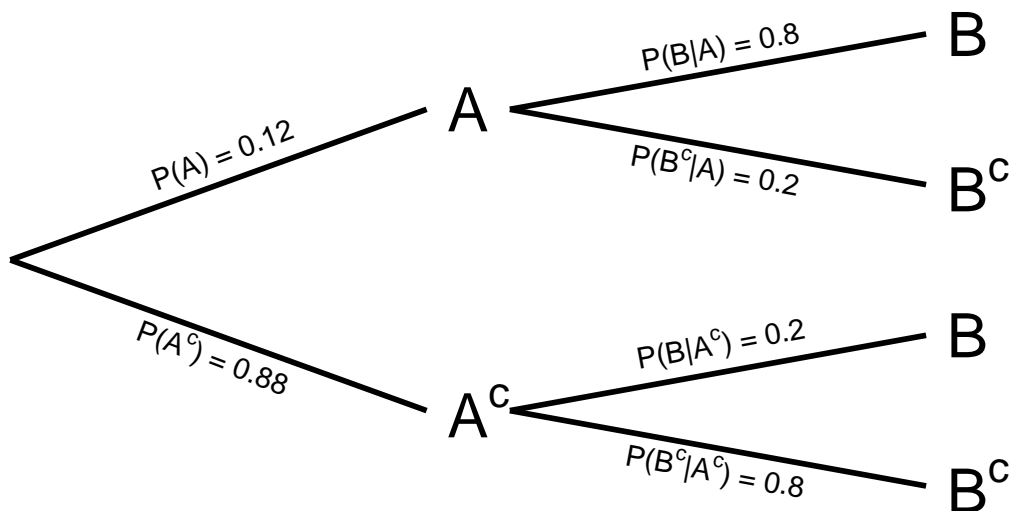
$$P(A) = 0.12$$

$$P(B|A) = 0.8$$

$$P(B^c|A^c) = 0.8$$

I got the lower two because if the driver has been drinking, the “correct decision” is to detain the driver. So  $P(B|A) = 0.8$ . Similarly, if the driver has not been drinking, the correct decision is to not detain the driver. So  $P(B^c|A^c) = 0.8$ .

Always make a diagram!



The probability that you are detained given that you have not been drinking is  $P(B|A^c) = 0.2$ , which you can read from the tree diagram.

**(b) What's the probability that any given driver will be detained?**

This is  $P(B)$ . To find  $P(B)$ , note from the tree diagram that there are two ways of getting to “B”: by going through  $A$  (the driver had been drinking), or by going through  $A^c$  (the driver had not been drinking). These are disjoint, so  $P(B) = P(A \text{ and } B) + P(A^c \text{ and } B)$ . You can also convince yourself of this by drawing a Venn diagram.

To find  $P(A \text{ and } B)$ , multiply the two probabilities on the branches leading to  $A$ , and from there to  $B$ :

$$P(A \text{ and } B) = P(A)P(B|A) = 0.12 \times 0.8 = 0.096$$

Similarly,

$$P(A^c \text{ and } B) = P(A^c)P(B|A^c) = 0.88 \times 0.2 = 0.176$$

Putting this all together, we get  $P(B) = 0.096 + 0.176 = 0.272$

**(c) What's the probability that a driver who is detained has actually been drinking?**

We want to calculate  $P(\text{the driver has been drinking, given that they were detained}) = P(A|B)$ .

We can calculate this using Bayes' Rule as follows:

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)} \\ &= \frac{0.12 \times 0.8}{0.12 \times 0.8 + 0.88 \times 0.2} \\ &= 0.353 \end{aligned}$$

**(d) What's the probability that a driver who was released had actually been drinking?**

We want to calculate  $P(\text{the driver has been drinking, given that they were released}) = P(A|B^c)$ .

We can calculate this using Bayes' Rule as follows:

$$\begin{aligned} P(A|B^c) &= \frac{P(A)P(B^c|A)}{P(A)P(B^c|A) + P(A^c)P(B^c|A^c)} \\ &= \frac{0.12 \times 0.2}{0.12 \times 0.2 + 0.88 \times 0.8} \\ &= 0.033 \end{aligned}$$