

## 10.10: Reversing a Conditional Probability- Bayes' Rule

In many cases, we know  $P(A|B)$  but we really want to know  $P(B|A)$ . This commonly occurs in medical screening, where we know  $P(\text{positive test result}|\text{disease})$  but what we want to know is  $P(\text{disease}|\text{positive test result})$ . For example, some doctors recommend that men over the age of 50 undergo screening using a test called prostate specific antigen (PSA) to screen for possible prostate cancer. Before a test is approved for use in medical practice, the manufacturer needs to test two aspects of the test's performance. First, they need to show how *sensitive* it is – that is, how likely is it to find the disease when it is present:  $\text{sensitivity} = P(\text{positive test}|\text{disease})$ . They also need to show how *specific* it is: that is, how likely is it to give a negative result when there is no disease present:  $\text{specificity} = P(\text{negative test}|\text{no disease})$ . For the PSA test, we know that sensitivity is about 80% and specificity is about 70%. However, these don't answer the question that the physician wants to answer for any particular patient: what is the likelihood that they actually have cancer, given that the test comes back positive? This requires that we reverse the conditional probability that defines sensitivity: instead of  $P(\text{positive test}|\text{disease})$  we want to know  $P(\text{disease}|\text{positive test})$ .

In order to reverse a conditional probability, we can use *Bayes' rule*:

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)}$$

Bayes' rule is fairly easy to derive, based on the rules of probability that we learned earlier in the chapter (see the Appendix for this derivation).

If we have only two outcomes, we can express Bayes' rule in a somewhat clearer way, using the sum rule to redefine  $P(A)$ :

$$P(A) = P(A|B) * P(B) + P(A|\neg B) * P(\neg B)$$

Using this, we can redefine Bayes's rule:

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A|B) * P(B) + P(A|\neg B) * P(\neg B)}$$

We can plug the relevant numbers into this equation to determine the likelihood that an individual with a positive PSA result actually has cancer – but note that in order to do this, we also need to know the overall probability of cancer in the person, which we often refer to as the *base rate*. Let's take a 60 year old man, for whom the probability of prostate cancer in the next 10 years is  $P(\text{cancer}) = 0.058$ . Using the sensitivity and specificity values that we outlined above, we can compute the individual's likelihood of having cancer given a positive test:

$$\begin{aligned} P(\text{cancer}|\text{test}) &= \frac{P(\text{test}|\text{cancer}) * P(\text{cancer})}{P(\text{test}|\text{cancer}) * P(\text{cancer}) + P(\text{test}|\neg \text{cancer}) * P(\neg \text{cancer})} \\ &= \frac{0.8 * 0.058}{0.8 * 0.058 + 0.3 * 0.942} = 0.14 \end{aligned}$$

That's pretty small – do you find that surprising? Many people do, and in fact there is a substantial psychological literature showing that people systematically neglect *base rates* (i.e. overall prevalence) in their judgments.

This page titled 10.10: Reversing a Conditional Probability- Bayes' Rule is shared under a [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/) license and was authored, remixed, and/or curated by [Russell A. Poldrack](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.