Problem Wk.10.1.4: Operations on Conditional Distributions

Part 1: Bayesian Update

In many cases, we will have an original distribution over some random variable, *P(A)* and then get some evidence that a related random variable *B* has value *b*. These two random variables are typically related through a conditional distribution describing the probability of the evidence given the variable of interest, *P(B | A)*. The quantity we're interested in is *P(A | B = b)*. We can compute it by constructing the joint distribution *P(A, B)* and then conditioning it on *B = b*.

We'll use this method to compute a distribution *P(Disease | Test = 'posTest')*

1. What is the result of conditionalizing the joint distribution *P(Disease, Test)* from the previous problem on `Test = 'posTest'`?

   Enter the probabilities below; use 6 decimal digits of precision.

   `DDist('disease': [ ], 'noDisease': [ ] )`

Part 2: Total Probability

One more common operation on distributions is sometimes called the *law of total probability*:

$$P(B) = \sum_a P(B \mid A = a) P(A = a)$$

One way to think about it is that, starting with some information about *A*, *P(A)* and knowing how *B* depends on *A*, *P(B | A)*, we can summarize what we know about *B* in *P(B)*.

Recalling that

- *P(posTest | disease) = 0.98*, and that
- *P(posTest | noDisease) = 0.05*,

  and that `Disease` is defined as follows

  `Disease = DDist({'disease' : 0.0001, 'noDisease' : 0.9999}),`

1. what is the result of applying the law of total probability to get *P(Test)* given *P(Test | Disease)* and *P(Disease)* as given before.

   Enter the probabilities below; use 6 decimal digits of precision.

   `DDist('posTest': [ ], 'negTest': [ ] )`