Solutions Chapter 3 - Probability - the nuts and bolts of Bayesian inference

Alexandra Hotti

September 2019

Assignment 3.1 Messy probability density

Suppose that a probability density is given by the following function:

$$\begin{cases} 1, & \text{if } 0 \leq X < 0.5 \\ 0.2, & \text{if } 0.5 \leq X < 1 \\ 0.8(X-1), & \text{if } 1 \leq X < 2 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Assignment 3.1.1:

Demonstrate that the above density is a valid probability distribution.

Answer: A probability distribution is valid if it fullfills two conditions. First, it should sum to 1. Second, all the values of the distribution need to be real and non negative.

Between the x range: [0, 1] the distribution take on constant positive value and in between [1, 2] the values range between [0, 0.8]. Thus, all values are both real and positive.

Next we make sure that the distribution sums to 1 by calcuating the area:

$$area = 1 \cdot 0.5 + 0.2 \cdot 0.5 + 1 \cdot 0.5 \cdot 0.8 = 0.5 + 0.1 + 0.4 = 1$$

Q.E.D

Assignment 3.1.2:

What is the probability that $0.2 \leq X \leq 0.5$?

Answer: $Pr(0.2 \leq X \leq 0.5) = 0.3 \cdot 1 = 0.3$

Assignment 3.1.3:

Find the mean of the distribution.

Answer:

$$\mathbb{E} = \int_0^1 X \cdot f(X) = \int_0^{0.5} XdX + \int_{0.5}^1 0.2 \cdot XdX + \int_1^2 0.8(X-1)dX \approx 0.867$$

Assignment 3.1.4:

What is the median of the distribution?

Answer: The median is located where at the point where the probability of X is less then 0.5. This occurs when: $P(X < 0.5)$, i.e. at $X < 0.5$.

Assignment 3.2 - Keeping it discrete

Suppose that the number of heads obtained X in a series of N coin flips is described by a binomial distribution:

$$Pr(X = K|\theta) = \binom{N}{K}\theta^K(1-\theta)^{N-K} \tag{2}$$

where θ is the probability of obtaining a heads on any particular throw.

Assignment 3.2.1:

Suppose that $\theta = 0.5$ (that is, the coin is fair). Calculate the probability of obtaining 5 heads in 10 throws

Answer: In equation (3) the number of coin throws corresponds to N, the number of heads to K and lastly the fairness of the coin corresponds to our θ. We input theses values and get:

$$Pr(X = 10|\theta = 0.5) = \binom{10}{5}0.5^5(1-0.5)^{10-5} \approx 0.25 \tag{3}$$

Assignment 3.2.2:

Calculate the probability of obtaining fewer than 3 heads.

Answer: This is given by:

$$P(X < 3|\theta = 0.5) = \sum_{i=0}^{2} P(X = i|\theta = 0.5) = \binom{10}{0}0.5^0(1-0.5)^{10-0} + \binom{10}{1}0.5^1(1-0.5)^{10-1} + \binom{10}{2}0.5^2(1-0.5)^{10-2} =$$

$$0.5^{10} + 10 \cdot 0.5^{10} + 45 \cdot 0.5^{10} = 0.0546875$$

Assignment 3.2.3:

Find the mean of this distribution..

Answer: This is approximated by:

$$\mathbb{E} = \theta \cdot N = 0.5 \cdot 10 = 5$$

3.3 Continuously confusing

Suppose that the time that elapses before a particular component on the Space Shuttle fails can be modelled as being exponentially distributed:

$$p(t|\lambda) = \lambda e^{-\lambda t} \tag{4}$$

where $\lambda > 0$ is a rate parameter.

Assignment 3.3.1:

Show that the above distribution is a valid probability density.

Answer: Since the distribution is continuous and unbound in the upperbound for t we integrate in the following manner:

$$p(t|\lambda) = \int_0^\infty \lambda e^{-\lambda t} = [-e^{-\lambda t}]_0^\infty$$

Assignment 3.3.2:

Find the mean of this distribution.

Answer:

$$\mathbb{E} = \int_0^\infty t \cdot \lambda e^{-\lambda t} = \frac{1}{\lambda}$$

Assignment 3.3.3:

Suppose that $\lambda = 0.2$ per hour. Find the probability that the component fails in the first hour of flight.

Answer: Assuming that the parameter t has hours as its units:

$$P(0 \leq t < 1|\lambda = 0.2) = \int_0^1 (0.2e^{-0.2t})dt = [-e^{-0.2t}]_0^1 = 1 - e^{-0.2} = 0.181269... \approx 0.18$$

Assignment 3.3.5:

What is the probability that the component fails during the second hour given that it has survived the first?

Answer: We rewrite the conditional probability using that: $p(A|B) = \frac{p(A,B)}{p(B)}$

$$P(1 < t \leq 2|\lambda = 0.2, t > 1) = \frac{P(1 < t \leq 2, \lambda = 0.2, t > 1)}{P(\lambda = 0.2, t > 1)}$$

Now we want to use the memoryless property of the exponential distribution, i.e. that the past has no bearing on its future behavior. Every instant is like the beginning of a new random period, which has the same distribution regardless of how much time has already elapsed. Now using this memoryless property we get:

$$\frac{P(1 < t \leq 2, \lambda = 0.2, t > 1)}{P(\lambda = 0.2, t > 1)} = \frac{P(1 < t \leq 2|\lambda = 0.2)P(\lambda = 0.2)}{P(t > 1|\lambda = 0.2)P(\lambda = 0.2)} = \frac{P(1 < t \leq 2|\lambda = 0.2)}{P(t > 1|\lambda = 0.2)} \tag{5}$$

Now, we can calculate eq. (5) by using eq. (4). The numerator becomes:

$$P(1 < t \leq 2|\lambda = 0.2) = \int_1^2 (0.2e^{-0.2t})dt = [-e^{-0.2t}]_1^2 = -e^{-0.4} + e^{-0.2} \approx 0.148 \tag{6}$$

The denominator becomes:

$$P(t > 1|\lambda = 0.2) = \int_1^\infty (0.2e^{-0.2t})dt = [-e^{-0.2t}]_1^\infty = -e^{-\infty} + e^{-0.2} \approx 0.819 \tag{7}$$

Thus, we finally get:

$$\frac{P(1 < t \leq 2|\lambda = 0.2)}{P(t > 1|\lambda = 0.2)} = \frac{0.148}{0.819} \approx 0.181$$

Assignment 3.3.6:

Show that the probability of the component failing during the (n + 1)th hour given that it has survived n hours is always 0.18.

Answer:

$$P(n \leq t \leq n+1|t > n, \lambda = 0.2) = \frac{P(n \leq t \leq n+1, t > n, \lambda = 0.2)}{P(t > n, \lambda = 0.2)} = \frac{P(n \leq t \leq n+1|\lambda = 0.2)P(\lambda = 0.2)}{P(t > n|\lambda = 0.2)P(\lambda = 0.2)} =$$

$$\frac{P(n \leq t \leq n+1|\lambda = 0.2)}{P(t > n|\lambda = 0.2)}$$

$$P(n \leq t \leq n+1|\lambda = 0.2) = \int_n^{n+1} (0.2e^{-0.2t})dt = [-e^{-0.2t}]_n^{n+1} = -e^{-0.2(n+1)} + e^{-0.2n} = -e^{-0.2} \cdot e^{-0.2n} + e^{-0.2n} = \tag{8}$$

$$-e^{-0.2} \cdot e^{-0.2n} + e^{-0.2n} = e^{-0.2n}(-e^{-0.2} + 1)$$

$$P(t > n|\lambda = 0.2) = \int_n^\infty (0.2e^{-0.2t})dt = [-e^{-0.2t}]_n^\infty = e^{-0.2n} \tag{9}$$

Using (9) and (8) gives us:

$$\frac{P(n \leq t \leq n+1|\lambda = 0.2)}{P(t > n|\lambda = 0.2)} = \frac{e^{-0.2n}(-e^{-0.2} + 1)}{e^{-0.2n}} = -e^{-0.2} + 1 = 0.18126924692$$

Q.E.D.

Planet Scrabble 3.5

On a far-away planet suppose that people's names are always two letters long, with each of these letters coming from the 26 letters of the Latin alphabet. Suppose that there are no constraints on individuals' names, so they can be composed of two identical letters, and there is no need to include a consonant or a vowel.

Problem 3.5.1.

How many people would need to be gathered in one place for there to be a 50% probability that at least two of them share the same name?

Answer: The probability that at least two people sharing the same name in a room could be calculated as:

$$P(\text{At least 2 people share the same name}) = 1 - P(\text{At least 2 people do not share the same name})$$

Then we compute $P(\text{At least 2 people do not share the same name})$. We know that that the alphabet contains 26 letters and there are no restrictions in the manner that these letter can be combined. I.e. we get: $26 \cdot 26 = 676$. When the first person goes into the room, there are 676 possible names "not taken". When the next person goes into the room there are 675 possible names "not taken", and so on. Using this logic we get:

$$P(\text{At least 2 people do not share the same name}) = \frac{676}{676} \cdot \frac{675}{676} \cdot \frac{674}{676} \cdot \frac{673}{676} \cdot \frac{672}{676} \cdot \frac{671}{676} \cdot \frac{670}{676} \cdot ... \tag{10}$$

The denominator in eq. (10) is generalized as eq. (11). Where k is the number of people in the room. The numerator is generalized in eq. (12).

$$\frac{1}{676^k} \tag{11}$$

$$676(676-1) \cdot ... \cdot (676-k+1) = \frac{676(676-1) \cdot ... \cdot 2 \cdot 1}{(676-k)(676-k-1) \cdot ... \cdot 1} = \frac{676!}{(676-k)!} \tag{12}$$

By combining eq. (11) and (12) we get (13).

$$P(\text{At least 2 people do not share the same name}) = \frac{1}{676^k} \cdot \frac{676!}{(676-k)!} \tag{13}$$

In the assignment we are asked to find the k for which eq. (13) is equal to 0.5. This gives us:

$$0.5 = \frac{1}{676^k} \cdot \frac{676!}{(676-k)!}$$

$$k = 30.88481298... \approx 31$$

This means that approximately 31 people have to be in the room for there to be a 50 % probability that at least two of them share the same name.

Problem 3.5.2.

Suppose instead that the names are composed of three letters. Now how many people would need to be gathered in one place for there to be a 50% probability that at least two of them share the same name?

Answer: Now there are $26^3 = 17576$ possible name combinations. Therefore, we simply use this number in (13) instead of 676, set the probability equal to 0.5 and solve for k. This gives us $k \approx 157$

Game theory 3.6

A game show presents contestants with four doors: behind one of the doors is a car worth $1000; behind another is a forfeit whereby the contestant must pay $1000 out of their winnings thus far on the show. Behind the other two doors there is nothing. The game is played as follows:

- The contestant chooses one of four doors.
- The game show host opens another door, always to reveal that there is nothing behind it.
- The contestant is given the option of changing their choice to one of the two remaining unopened doors.
- The contestant's final choice of door is opened, to their delight (a car!), dismay (a penalty), or indifference (nothing).

Assuming that:

- the contestant wants to maximise their expected wealth, and
- the contestant is risk-averse, what is the optimal strategy for the contestant?

Answer: Based on the description of the doors we get the following data:

- 50 % of the doors contain nothing and gives us $ 0.
- 25 % of the doors hide a car worth $ 1000 (kind of a cheap car if you ask me ;)).
- 25 % of the doors hide a penalty of $ 1000.

To maximize the contestants wealth, he or she should aim for getting the car.

Now, there are three possibilities for the door which the contestant picked:

- **You picked an empty door.** Now the two unopened doors contain the car and the penalty. If you stay you neither gain nor lose anything. If you switch you can either win or lose big time, with a probability of 0.5.
- **You picked the car.** Now the two unopened doors contain an empty door and the penalty. If you stay you get the car. If you switch lose money or get nothing, with a probability of 0.5.
- **You picked the penalty.** Now the two unopened doors contain an empty door and the car. If you stay you get the penalty. If you switch you get the car or nothing, with a probability of 0.5.

If the contestant simply stays, he or she has no information when making his or hers decision and therefore we get the following distribution:

Table 1: Distribution for the different outcomes if the contestant does not switch.

p(car)	p(nothing)	p(penalty)
25 %	50 %	25 %

If the contestant instead chooses to switch, he or she has access to more information and the situation changes. We calculate the probability for each outcome based on the additional information given to the contestant.

First, if the contestant switches, he or she can get nothing in three different ways:

$$p(penalty) = p(\text{switching from empty door to penalty}) + p(\text{switching from car to penalty}) +$$
$$p(\text{switching from penalty to penalty}) = 0.5 \cdot 0.5 + 0.25 \cdot 0.5 + 0.25 \cdot 0 = 0.375$$

$$p(car) = p(\text{switching from empty door to car}) + p(\text{switching from car to car}) +$$

$$p(\text{switching from penalty to car}) = 0.5 \cdot 0.5 + 0.25 \cdot 0 + 0.25 \cdot 0.5 = 0.375$$

$$p(nothing) = p(\text{switching from empty door to empty door}) + p(\text{switching from car to empty door}) +$$

$$p(\text{switching from penalty to empty door}) = 0.5 \cdot 0 + 0.25 \cdot 0.5 + 0.25 \cdot 0.5 = 0.25$$

We summarize these results in table 2.

Table 2: Distribution for the different outcomes if the contestant does switch.

p(car)	p(nothing)	p(penalty)
37.5 %	25 %	37.5 %

From the two tables above, one quickly realises that the expected value from both switching and staying is \$ 0. However, by switching we decrease the probability of getting nothing and increase the probability of losing money and getting the car. Since, both outcomes have the same return and the contestant is risk averse her or she should stick with the first door.

Blood doping in cyclists 3.7

Suppose, as a benign omniscient observer, we tally up the historical cases where professional cyclists either used or did not use blood doping, and either won or lost a particular race. This results in the probability distribution shown in Table 3.

Table 3: The historical probabilities of behaviour and outcome for professional cyclists.

	Lost	Won
Clean	0.70	0.05
Doping	0.15	0.10

Problem 3.7.1.

What is the probability that a professional cyclist wins a race?

Answer:

$$Pr(X = win) = \sum_{\alpha \in \{clean, doping\}} Pr(X = win, Y = \alpha) \tag{14}$$

$$Pr(X = win) = Pr(X = win, Y = clean) + Pr(X = win, Y = Doping) = 0.05 + 0.10 = 0.15$$

Problem 3.7.2.

What is the probability that a cyclist wins a race, given that they have cheated?

Answer:

$$Pr(X = win|Y = cheated) = \frac{Pr(X = win, Y = cheated)}{Pr(Y = cheated)} = \frac{0.10}{0.25} = 0.4 \tag{15}$$

Problem 3.7.3.

What is the probability that a cyclist is cheating, given that they win?

Answer:

$$Pr(Y = cheated|X = win) = \frac{Pr(X = win, Y = cheated)}{Pr(X = win)} = \frac{0.10}{0.15} = 0.666... \tag{16}$$

Now suppose that drug testing officials have a test that can accurately identify a blood-doper 90% of the time. However, it incorrectly indicates a positive for clean athletes 5% of the time.

Problem 3.7.4.

If the officials want to maximize the proportion of people correctly identified as dopers (i.e. keep down the proportion of false positives), should they test all the athletes or only the winners?

Answer: From the information above, one can decipher that:

Pr(+ | doped) = 90%

Pr(- | doped) = 10%

Pr(+ | clean) = 5%

Pr(- | clean) = 95%

Now we want decide whether we should only test the winners or every cyclist. We can decide this by comparing: $Pr(Y = \text{cheated}|X = \text{winners}, Z = +)$ and $Pr(Y = \text{cheated}|X = \text{all cyclists}, Z = +)$.

We rewrite the first expression:

$$Pr(Y = \text{cheated}|X = \text{winners}, Z = +) = \frac{Pr(Y = \text{cheated}, X = \text{winners}, Z = +)}{Pr(X = \text{winners}, Z = +)} =$$

$$\frac{Pr(Y = \text{cheated}, Z = +|X = \text{winners})Pr(X = \text{winners})}{Pr(Z = +|X = \text{winners})Pr(X = \text{winners})} =$$

$$\frac{Pr(Y = \text{cheated}, Z = +|X = \text{winners})Pr(X = \text{winners})}{Pr(Z = +|X = \text{winners})Pr(X = \text{winners})} = \frac{Pr(Y = \text{cheated}, Z = +|X = \text{winners})}{Pr(Z = +|X = \text{winners})} \tag{17}$$

The numerator can be rewritten as:

$$Pr(Y = \text{cheated}, Z = +|X = \text{winners}) = \frac{Pr(X = \text{winners}|Y = \text{cheated}, Z = +)Pr(Y = \text{cheated}, Z = +)}{Pr(X = \text{winners})}$$

Since we test the cyclists that win after they have won, winning is conditionally independent of a positive test result.

$$Pr(Y = \text{cheated}, Z = +|X = \text{winners}) = \frac{Pr(X = \text{winners}|Y = \text{cheated})Pr(Y = \text{cheated}, Z = +)}{Pr(X = \text{winners})} =$$

$$Pr(Y = \text{cheated}, Z = +|X = \text{winners}) = \frac{Pr(X = \text{winners}|Y = \text{cheated})Pr(Y = \text{cheated}, Z = +)Pr(Y = \text{cheated})}{Pr(X = \text{winners})}$$

Now we plug in the numbers:

$$Pr(Y = \text{cheated}, Z = +|X = \text{winners}) = \frac{0.9 \cdot 0.4 \cdot 0.25}{0.15} = 0.6 \tag{18}$$

Now we have an expression for the numerator and can rewrite the denominator in (17) as:

$$Pr(Z = +|X = \text{winners}) = Pr(+, \text{cheated}|\text{winners}) + Pr(+, \text{clean}|\text{winners}) =$$

$$\frac{Pr(+, \text{cheated}, \text{winners})}{Pr(\text{winners})} + \frac{Pr(+, \text{clean}, \text{winners})}{Pr(\text{winners})} =$$

$$\frac{Pr(+|\text{cheated}, \text{winners})Pr(\text{cheated}|\text{winners})Pr(\text{winners})}{Pr(\text{winners})} + \frac{Pr(+|\text{clean}, \text{winners})Pr(\text{clean}|\text{winners})Pr(\text{winners})}{Pr(\text{winners})} =$$

$$Pr(+|\text{cheated}, \text{winners})Pr(\text{cheated}|\text{winners}) + Pr(+|\text{clean}, \text{winners})Pr(\text{clean}|\text{winners})$$

The test result is conditionally independent of whether we won the race, therefore we finally get:

$$Pr(+|\text{winners}) = Pr(+|\text{cheated})Pr(\text{cheated}|\text{winners}) + Pr(+|\text{clean})Pr(\text{clean}|\text{winners}) \tag{19}$$

We have all the information we need to calculate (19):

$$Pr(+|\text{winners}) = 0.9 \cdot \frac{2}{3} + 0.05 \cdot \frac{1}{3} \approx 0.62 \tag{20}$$

Using the result from (20) and (18), we get:

$$Pr(Y = \text{cheated}|X = \text{winners}, Z = +) = \frac{0.6}{0.62} \approx 0.97 \tag{21}$$

So, now we start investigating expression number two, i.e.: $Pr(Y = \text{cheated}|X = \text{all cyclists}, Z = +)$.

$$Pr(Y = \text{cheated}|X = \text{all cyclists}, Z = +) = Pr(Y = \text{cheated}|Z = +) = \tag{22}$$

Using Bayes's theorem gives us:

$$Pr(\text{cheated}|+) = \frac{Pr(+|\text{cheated})Pr(\text{cheated})}{Pr(+)} = \frac{Pr(+|\text{cheated})Pr(\text{cheated})}{Pr(+|cheated)Pr(cheated) + Pr(+|clean)Pr(clean)}$$

Plugging in the numbers gives us:

$$Pr(\text{cheated}|+) = \frac{0.9 \cdot 0.25}{0.9 \cdot 0.25 + 0.05 \cdot 0.75} \approx 0.86 \tag{23}$$

So, comparing the numbers for testing every cyclist in (23) and only testing the winners as in (21), one can conclude that is is better to only test the winners. As doping would increase ones chances of winning, it makes sense that this group would contain many dopers.

3.8 Breast cancer revisited

Suppose that the prevalence of breast cancer for a randomly chosen 40-year-old woman in the UK population is about 1%. Further suppose that mammography has a relatively high sensitivity to breast cancer, where in 90% of cases the test shows a positive result if the individual has the disease. However, the test also has a rate of false positives of 8%.

Problem 3.8.1.

Show that the probability that a woman tests positive is about 9%.

Answer:
The above information gives us the following probabilities:

- p(cancer) = 1%
- p(not cancer) = 99%
- p(+|cancer) = 90%
- p(-|cancer) = 10%
- p(+|not cancer) = 8%
- p(-|not cancer) = 92%

Now, we want to show that $p(+) \approx 9\%$. We rewrite the expression using marginalization as:

$$p(+) = p(+|\text{cancer})p(\text{cancer}) + p(+|\text{not cancer})p(\text{not cancer}) = 0.9 \cdot 0.01 + 0.08 \cdot 0.99 = 0.0882 \approx 8.8\%$$

Problem 3.8.2.

A woman tests positive for breast cancer. What is the probability she has the disease?

Answer: Bayes' theorem gives us:

$$p(c|+) = \frac{p(+|c)p(c)}{p(+)} = \frac{0.9 \cdot 0.01}{0.0882} = 0.10204081632 \approx 10\%$$

Problem 3.8.3.

Draw a graph of the probability of having a disease, given a positive test, as a function of (a) the test sensitivity (true positive rate) (b) the false positive rate, and (c) the disease prevalence. Draw graphs (a) and (b) for a rare (1% prevalence) and a common (10% prevalence) disease. What do these graphs imply about the relative importance of the various characteristics of medical tests?

Answer:

The first task is to plot the probability of having cancer given a positive test: $p(c|+)$ as a function of the true positive rate, i.e. $p(+|c)$. Disease prevalence is given by: $p(c)$ and false positives by: $p(+|\text{not cancer})$. All of these variables are taken into consideration in 24. All of these variables are probabilities, so they will vary in the range: [0,1].

$$p(cancer|+) = \frac{p(+|cancer)p(cancer)}{p(+)} = \frac{p(+|cancer)p(cancer)}{p(+|\text{cancer})p(\text{cancer}) + p(+|\text{not cancer})p(\text{not cancer})} \tag{24}$$

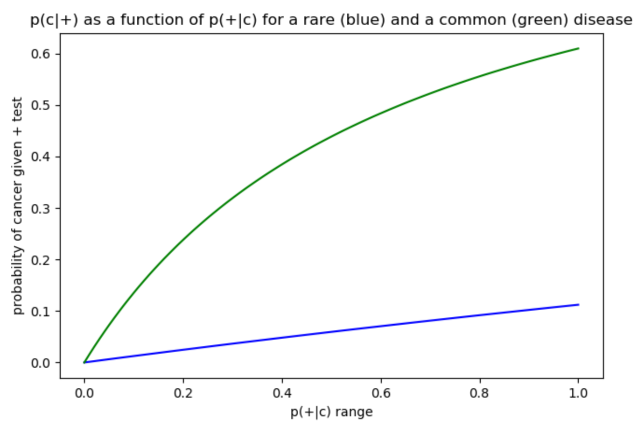

Figure 1: Probability of having cancer, given a positive test as a function of the true positive rate for a common (green) and a rare (blue) disease.

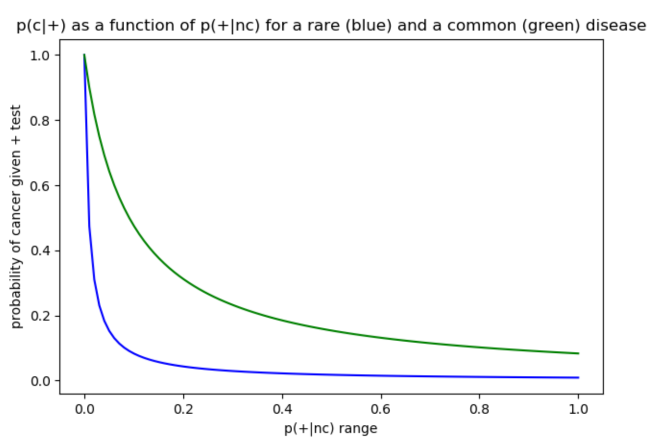

Figure 2: Probability of having cancer, given a positive test as a function of the false positive rate for a common (green) and a rare (blue) disease.

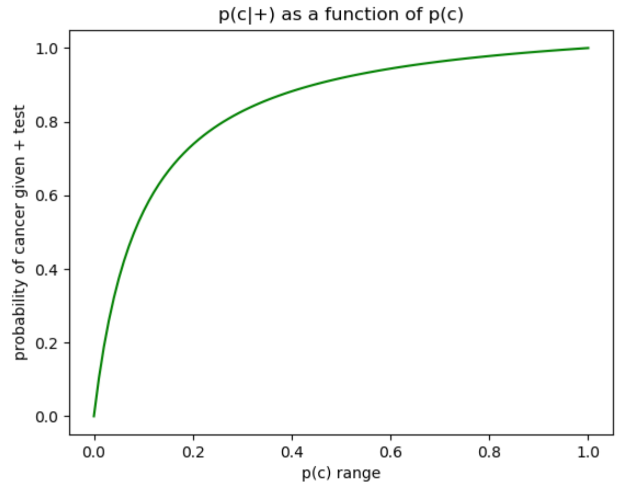

Figure 3: Probability of having cancer, given a positive test as a function of the disease prevalence.

Above in Figure 1 and 2 we can see that for the rare disease there is a small dependence with the true positive rate (blue) relative to the false positive rate (green). Therefore, it is more essential to keep down false positives. If we study equation (1) we see that if we are dealing with a rare disease the denominator will mostly be affected by the right hand term, i.e. $p(+|\text{not cancer})p(\text{not cancer})$. Since $p(\text{not cancer})$ becomes large and $p(\text{cancer})$ becomes small and therefore we become more dependent on the false positives, i.e. $p(+|\text{not cancer})$.

If we are dealing with a common disease instead the relative size of $p(\text{cancer})$ and $p(\text{not cancer})$ becomes smaller, see the green lines. Therefore, we are less concerned about false positives for common diseases.

Lastly, looking at how $p(c|+)$ vary as the disease gets more common along the x-axis, one can see that the test gets more accurate for common diseases.

Problem 3.8.4.

Assume the result of a mammography is independent when retesting an individual (probably a terrible assumption!). How many tests (assume a positive result in each) would need to be undertaken to ensure that the individual has a 99% probability that they have cancer?

Answer: We start by looking at the probability of cancer after taking two positive tests:

$$P(c|+,+) = \frac{P(+,+|c)P(c)}{P(+,+)} \tag{25}$$

Since the test results are independent of each other we know that: $P(A,B) = P(A)P(B)$. Also, we marginalize $P(+,+)$. Which gives us:

$$P(c|+,+) = \frac{P(+|c)P(+|c)P(c)}{P(+,+)} = \frac{P(+|c)P(+|c)P(c)}{P(+,+|c)P(c) + P(+,+|nc)P(nc)} = \frac{P(+|c)P(+|c)P(c)}{P(+|c)P(+|c)P(c) + P(+|nc)P(+|nc)P(nc)} =$$

$$1 + \frac{P(+|c)P(+|c)P(c)}{P(+|nc)P(+|nc)P(nc)} = \frac{P(+|c)^2P(c)}{P(+|c)^2P(c) + P(+|nc)^2P(nc)} \tag{26}$$

The general equation becomes:

$$P(c|+_1,..,+_n) = \frac{P(+|c)^nP(c)}{P(+|c)^nP(c) + P(+|nc)^nP(nc)} \tag{27}$$

Now, we calculate how many tests are needed to have a 99 % risk of having cancer.

$$0.99 = \frac{0.9^n \cdot 0.01}{0.9^n \cdot 0.01 + 0.08^n \cdot 0.99} \tag{28}$$

Thus, $n \approx 3.79704$. The patient should thus take 4 positive tests to be 99% sure that she has breast cancer.