Week 3
Basics of probability theory and Bayes' rule

Probability theory

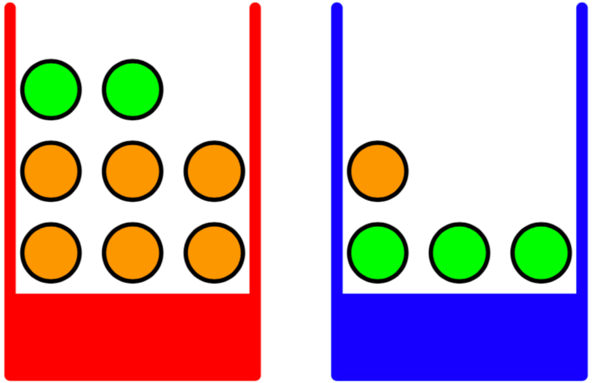
sum rule $$p(X) = \sum_Y p(X,Y)$$

product rule $$p(X,Y) = p(Y|X)p(X).$$

$$p(x) = \int p(x,y)\,\mathrm{d}y$$

$$p(x,y) = p(y|x)p(x).$$

Probability theory



Suppose we pick the red box 40% of the time and we pick the blue box 60% of the time, and that when we remove an item of fruit from a box we are equally likely to select any of the pieces of fruit in the box.

$$\begin{aligned} p(B=r) &= 4/10 \\ p(B=b) &= 6/10 \end{aligned}$$

$$p(B=r)+p(B=b)=1$$

$$\begin{aligned} p(F=a|B=r) &= 1/4 \\ p(F=o|B=r) &= 3/4 \\ p(F=a|B=b) &= 3/4 \\ p(F=o|B=b) &= 1/4. \end{aligned}$$

$$p(F=a|B=r)+p(F=o|B=r)=1$$

$$p(F=a|B=b)+p(F=o|B=b)=1.$$

$$\begin{aligned} p(F=a) &= p(F=a|B=r)p(B=r)+p(F=a|B=b)p(B=b) \\ &= \frac{1}{4}\times\frac{4}{10}+\frac{3}{4}\times\frac{6}{10}=\frac{11}{20} \end{aligned}$$

$$p(F=o)=1-11/20=9/20.$$

Bayes' rule

product rule $p(X, Y) = p(Y|X)p(X).$

$$\downarrow$$

$$p(X, Y) = p(Y, X)$$

$$\downarrow$$

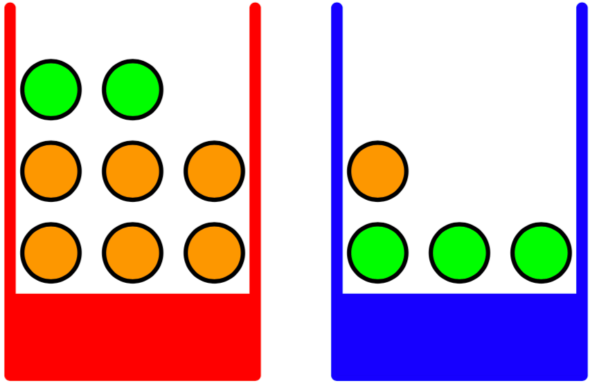
$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

$$\downarrow$$

$$p(X) = \sum_Y p(X|Y)p(Y).$$

Using sum rule, the denominator can be expressed in terms of the quantities appearing in the numerator. We can view the denominator as being the normalization constant required to ensure that the sum of the conditional probability on the left-hand side of over all values of Y equals one.

Bayes' rule



Suppose instead we are told that a piece of fruit has been selected and it is an orange, and we would like to know which box it came from.

$$p(B = r|F = o) = \frac{p(F = o|B = r)p(B = r)}{p(F = o)} = \frac{3}{4} \times \frac{4}{10} \times \frac{20}{9} = \frac{2}{3}.$$

$$p(B = b|F = o) = 1 - 2/3 = 1/3.$$

Bayesian probability

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})p(\mathbf{w})}{p(\mathcal{D})} \qquad p(\mathcal{D}) = \int p(\mathcal{D}|\mathbf{w})p(\mathbf{w})\,\mathrm{d}\mathbf{w}.$$

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

Conjugate: posterior distributions having the same functional form as the prior.

https://seeing-theory.brown.edu/index.html

Exercise

Exercise (homework)

A diagnostic test has a probability 0.95 of giving a positive result when applied to a person suffering from a certain disease, and a probability 0.10 of giving a (false) positive when applied to a non-sufferer. It is estimated that 0.5 % of the population are sufferers. Suppose that the test is now administered to a person about whom we have no relevant information relating to the disease (apart from the fact that he/she comes from this population). Calculate the following probabilities:

(a) that the test result will be positive;

(b) that, given a positive result, the person is a sufferer;

(c) that, given a negative result, the person is a non-sufferer;

(d) that the person will be misclassified.

Exercise

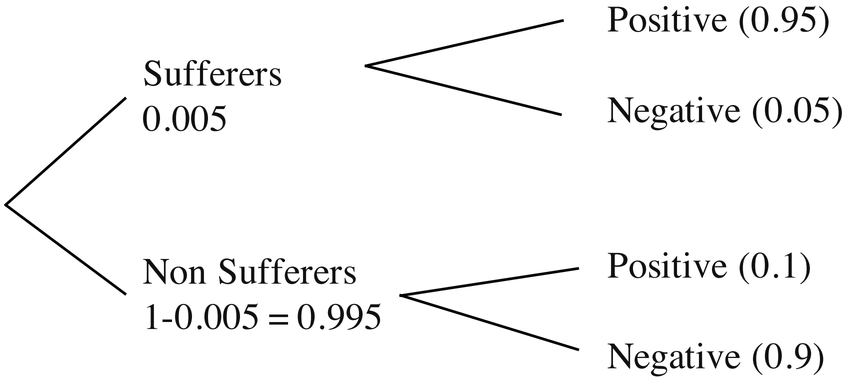
A diagnostic test has a probability 0.95 of giving a positive result when applied to a person suffering from a certain disease, and a probability 0.10 of giving a (false) positive when applied to a non-sufferer. It is estimated that 0.5 % of the population are sufferers. Suppose that the test is now administered to a person about whom we have no relevant information relating to the disease (apart from the fact that he/she comes from this population). Calculate the following probabilities:

(a) that the test result will be positive;

(b) that, given a positive result, the person is a sufferer;

(c) that, given a negative result, the person is a non-sufferer;

(d) that the person will be misclassified.



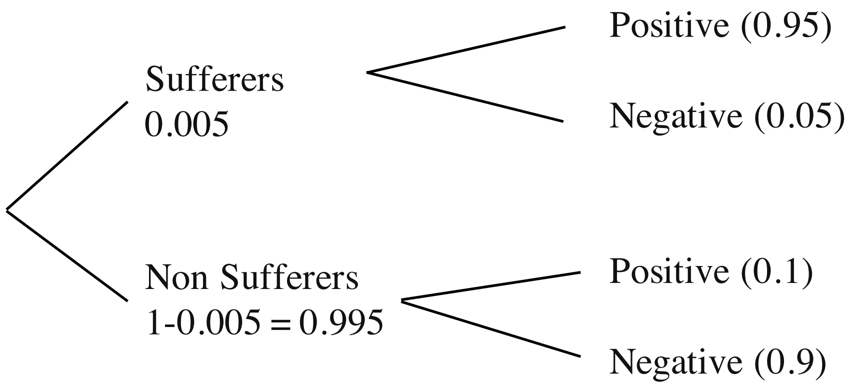
Exercise

(a) that the test result will be positive;

(b) that, given a positive result, the person is a sufferer;

(c) that, given a negative result, the person is a non-sufferer;

(d) that the person will be misclassified.



(a) $P(\text{positive}) = 0.005*0.95 + 0.995*0.1 = 0.10425$

(b) $P(\text{suffer} \mid \text{positive}) = P(\text{positive} \mid \text{suffer}) * P(\text{suffer}) / P(\text{positive}) = 0.95*0.005/0.10425 = 0.0455$

(c) $P(\text{non suffer} \mid \text{negative}) = P(\text{negative} \mid \text{non suffer}) * P(\text{non suffer}) / P(\text{negative}) = 0.9*0.995/(1-0.10425) = 0.9997$

(d) $P(\text{misclassified}) = P(\text{positive} \cap \text{non suffer}) + P(\text{negative} \cap \text{suffer}) = 0.1*0.995 + 0.05*0.005 = 0.09975$