Lecture 27: Basic probability theory

Nitin Saxena ⋆

IIT Kanpur

1 Conditional probability

Given two events A, B, the conditional probability of A given B is defined by,

$$P\left(A \,|\, B\right) := \frac{P(A \cap B)}{P(B)}\,.$$

Two events A, B are said to be *independent* if,

$$P(A \cap B) = P(A)P(B).$$

Prove that– A, B are independent events iff $P(A \,|\, B) = P(A)$ iff $P(B \,|\, A) = P(B)$.

Let us look at the application of these concepts. There was a survey conducted by the Health department in a hospital (with Asthma and Diabetes patients); it found that people who had diabetes did not have Asthma with higher probability as compared to the general population. This suggests that people who have Asthma have less likelihood of getting Diabetes.

It turns out that even if having Asthma and Diabetes are independent of each other, there will be a negative correlation between Asthma and Diabetic patients in a hospital.

In general, suppose A, B are two independent events. They may not be independent if we consider the conditional probabilities given $A \cup B$. That is, given event $A \cup B$, the events A, B might be negatively correlated (even if A, B are independent). This is known as *Berkson's Paradox*.

To make it more quantitative, consider a sample of 1000 balls. We know that 100 of them are red, 50 of them are shiny (they may be red or not) and 5 of them are red and shiny. The probability of being shiny is $1/20$, and the probability of a shiny ball being red is $5/100 = 1/20$. Hence being red and being shiny are independent.

Say, we pick only the balls which are red *or* shiny. (Note- there are $100 + 50 - 5 = 145$ such balls.) Then the probability that a ball is shiny is $50/145$ but a shiny ball being red remains at $5/100 = 1/20$. This shows that a shiny ball is mostly not red; thus, the two events have become *negatively* correlated.

Exercise 1. Let A be the event that the ball is shiny and B be the event that the ball is red. Prove that $P\left(A \,|\, A \cup B\right) > P(A)$. Convince yourself that this is the reason why events A and B seem to be negatively related, once we restrict to $A \cup B$.

Exercise 2. Convince yourself that the same thing happened when the health department surveyed *in a hospital* in the above example.

Let D, A, H be the events of having diabetes, asthma, or being in a hospital respectively. Assume $H = D \cup A$. If $P(D \,|\, H) > P(D) = P(D \,|\, A)$ then clearly, $P(D \,|\, H) \neq P(D \,|\, H \cap A)$. The latter simply means that the two illnesses are related in a hospital, but independent in the world!

⋆ Edited from Rajat Mittal's notes.

2 Bayes theorem

Suppose a scientific theory predicts that there will be a solar eclipse on 17th Oct., 2016 with high probability. If we observe that there is a solar eclipse on 17th Oct., what is the probability that the theory is correct?

Such problems are called *hypothesis testing*. The issue is that we are not 100% sure of the "scientific theory" and want to still use it to make predictions.

Let A be the event that the scientific theory is true and B be the event that solar eclipse happens on 17th Oct. So we know the conditional probability $P(B \mid A)$ and are interested in the conditional probability $P(A \mid B)$.

Bayes' theorem gives an answer to such meta-problems with a very clean expression.

Theorem 1 (Bayes 1700s). *Let A and B be two events with A^C, B^C denoting their complement. Then the conditional probability $P(A \mid B)$ is given by,*

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)} .$$

Note 1. The denominator is mostly obtained using the *disjoint-sum* formula $P(B) = P(B \mid A) \cdot P(A) + P(B \mid A^c) \cdot P(A^c)$.

Proof. The proof follows from the definition, and symmetry, of the conditional probability. □

Let us take some everyday examples and see how Bayes' theorem can be applied in various settings.

- Suppose there is a test for early detection of cancer and a study shows that it is very successful. If a person has cancer then the test diagnoses it with probability 0.9. If a person does not have cancer then the test correctly diagnoses with probability 0.9. Suppose a person is tested and the test shows that the person has cancer, what is the probability that the person actually has cancer?
  A naïve guess would be that the test works in both cases, so it seems pretty accurate. Hence, the person has cancer with very high probability. Let us try to study it using Bayes' theorem. Let A be the event that a person has cancer and B be the event that the test outputs that the person has cancer. Then,

  $$P(A \mid B) = \frac{0.9P(A)}{0.9P(A) + 0.1P(A^c)} .$$

  We realize that the information given to us is incomplete. We need, at least, the ratio between $P(A)$ and $P(A^c)$ (this is known as the *base rate*). Say 1% of the general population has cancer. Then,

  $$P(A \mid B) = \frac{0.9 \times 1}{0.9 \times 1 + 0.1 \times 99} \approx 0.1 .$$

  This shows that the base rate matters a lot in this calculation and should not be ignored (as it badly amplifies the error rate of 10%.).
  What if the base-rate is 50% ?

- In Mumbai, 90% of the taxis are black and the rest are white. It was observed by *Times of India* that white taxi drivers are very rash and are 5 times more likely to be involved in an accident as compared to a black taxi.

  Exercise 3. You are told that recently there was an accident involving a taxi, what is the probability that the taxi was white?

  Let W be the event that a taxi is white, and A be the event that there is a taxi accident. We know that $P(W \mid A) = P(W)P(A \mid W)/P(A) = 10 \times P(A \mid W)/(10 \times P(A \mid W) + 90 \times P(A \mid W^c))$ $= 5/(5 + 9 \times 1) = 5/14 \approx 36\%$.

References

1. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.