Conditional probability

Enrico Bibbona, DISMA, Politecnico di Torino
enrico.bibbona@polito.it

Joint and Conditional Probability, 3.6

We call the joint probability of two events E and F the probability of their intersection

$$\mathbb{P}(E \cap F).$$

This represents the probability that both conditions defining the events E and F are satisfied.

Joint and Conditional Probability, 3.6

In many examples, calculating the probability of an event E, is not obvious, but becomes much easier is some further information is provided Example:

- We have two urns containg red and blue balls in different proportions (e.g urn 1 has 3R, 2B while urn two has 1R, 11B)
- An urn is selected uniformly at random and then one ball is extracted
- Calculating the probability that the extected ball is red is greatly simplified if we know which is the urn out of which it gets extracted.

Conditional Probability, 3.6

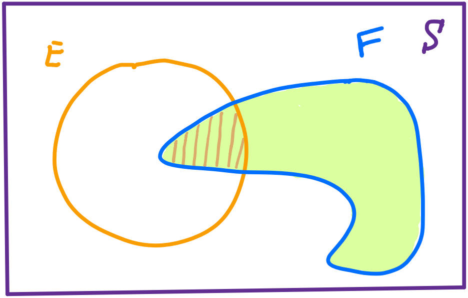
$\mathbb{P}(E \cap F) = \mathbb{P}(E|F) \cdot \mathbb{P}(F)$

We define the conditional probability of E given F as the following quantity

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Intuitively, it is as if we restrict the sample space to only the outcomes contained in F and, among these, we calculate the probability of E.

Graphically, it can be viewed as a ratio of areas (see figure).



$\mathbb{P}(E|F) = \dfrac{\quad}{\quad} = \dfrac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$

Example, 3.6.1

A package contains: 5 faulty transistors, 10 defective transistors, and 25 good transistors. The defective ones work correctly for a while and then break.

- A transistor is selected at random. It initially works. What is the probability that it is one of the good ones?

Example, 3.6.1

A package contains: 5 faulty transistors, 10 defective transistors, and 25 good transistors. The defective ones work correctly for a while and then break.

- A transistor is selected at random. It initially works. What is the probability that it is one of the good ones?

Solution

$$\mathbb{P}(\text{good}|\text{not faulty}) = \frac{\mathbb{P}(\text{good and not faulty})}{\mathbb{P}(\text{not faulty})} = \frac{\mathbb{P}(\text{good})}{\mathbb{P}(\text{not faulty})}$$
$$= \frac{\frac{25}{40}}{\frac{35}{40}} = \frac{25}{35} = \frac{5}{7}$$

Law of Total Probability (Probability by conditioning), 3.7

Let's assume that the state space can be partitioned as a union of disjoint events $F_1, \cdots, F_n$, with the assumption that all the F_i are mutually disjoint.

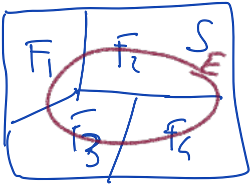
$S = \bigcup_{i=1}^{n} F_i \qquad \forall i \neq j \quad F_i \cap F_j = \emptyset$

Example

Thinking of an insurance application, imagine that all the insured individuals (sample space S) can be divided into two categories:

- prone to accidents (event I)
- not prone to accidents (event NI)

Thus, we have $S = I \cup NI$ and $I \cap NI = \emptyset$



Law of Total Probability

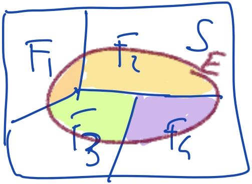
For any event E, $\mathbb{P}(E)$ can be calculated if the conditional probabilities $\mathbb{P}(E|F_i)$ are known, using the following property:

$$\mathbb{P}(E) = \sum_i \mathbb{P}(E|F_i)\mathbb{P}(F_i)$$

The book refers to this property as the *probability by conditioning*.

Law of Total Probability, 3.7

$$E = (E \cap F_1) \cup (E \cap F_2) \cup (E \cap F_3) \cup (E \cap F_4)$$



$$\mathbb{P}(E) = \sum_{i=1}^{4} \mathbb{P}(E \cap F_i) = \sum_i \mathbb{P}(E|F_i) \cdot \mathbb{P}(F_i)$$

$$\mathbb{P}(E|F_i) = \frac{\mathbb{P}(E \cap F_i)}{\mathbb{P}(F_i)}$$

$$\Downarrow$$

$$\mathbb{P}(E \cap F_i) = \mathbb{P}(E|F_i) \cdot \mathbb{P}(F_i)$$

The proof of the previous property is straightforward, in fact:

$$S = \bigcup_i F_i \qquad E = \bigcup_i (E \cap F_i) \text{ (disjoint)}$$

$$\mathbb{P}(E) \overset{iii}{=} \sum_i \mathbb{P}(E \cap F_i) = \sum_i \mathbb{P}(E|F_i)\mathbb{P}(F_i)$$

by the definition of conditional probability.

Example, 3.7.1

Let's return to the insurance application. Suppose the probability of having at least one accident in a year is 0.4 for customers in A (prone to accidents), and 0.2 for those in A^c (not prone). Now, suppose that 30% of the population is in A.

- What is the probability that a new insured customer, about whom nothing is known, has at least one accident in their first year of driving?

Example, 3.7.1

Let's return to the insurance application. Suppose the probability of having at least one accident in a year is 0.4 for customers in I (prone to accidents), and 0.2 for those in NI (not prone). Now, suppose that 30 % of customers are in I

- What is the probability that a new insured customer, about whom nothing is known, has at least one accident in their first year of driving?

$$\mathbb{P}(A) = \mathbb{P}(A \mid I) \cdot \mathbb{P}(I) + \mathbb{P}(A \mid NI) \cdot \mathbb{P}(NI)$$

$$= 0.4 \cdot 0.3 + 0.2 \cdot 0.7 =$$

Bayes' Theorem, 3.7

There are cases where it may be useful to calculate the conditional probability of an event E given F, knowing the probability in reverse, i.e., $\mathbb{P}(F|E)$. To do this, we can use the following Bayes' formula:

$$\mathbb{P}(E/F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(F|E)\mathbb{P}(E)}{\mathbb{P}(F)}$$

$$\mathbb{P}(E \cap F) = \mathbb{P}(E/F) \cdot \mathbb{P}(F)$$

The proof is immediate, since by the definition of conditional probability:

$$\mathbb{P}(F|E) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)}$$

$$\mathbb{P}(F|E)\mathbb{P}(E) = \mathbb{P}(F \cap E) = \mathbb{P}(E|F)\mathbb{P}(F)$$

and by dividing by $\mathbb{P}(F)$, we obtain the result.

$$\mathbb{P}(E \cap F) = \mathbb{P}(F|E) \cdot \mathbb{P}(E)$$

Exercise, 3.7.5

A medical test is 99% effective in detecting a disease when it is present. The *false positives* (healthy individuals being classified as sick by the test) occur 1% of the time. If the prevalence of this disease in the population is 0.5%, calculate the probability of being sick given that a positive test result is obtained.

Exercise, 3.7.5

A medical test is 99% effective in detecting a disease when it is present. The *false positives* (healthy individuals being classified as sick by the test) occur 1% of the time. If the prevalence of this disease in the population is 0.5%, calculate the probability of being sick given that a positive test result is obtained.

Observation

Tests like this are very common (e.g., PSA for prostate cancer, tri-test in pregnancy...), and you will see that the result of this calculation is surprising. Normally, one would think such a test is very reliable. However, a positive result should not cause excessive alarm. This example highlights the non-intuitive nature of probability and its relevance in real-life applications.

Exercise, 3.7.5

P(TP|S)

A medical test is 99% effective in detecting a disease when it is present. The *false positives* (healthy individuals being classified as sick by the test) occur 1% = P(TP|H) of the time. If the prevalence of this disease in the population is 0.5%, calculate the probability of being sick given that a positive test result is obtained.

P(S)

Solution

Let $\overset{S}{\cancel{M}}$ be the event that the person is sick (and $\overset{H}{\cancel{S}}$ for healthy), TP the event of a positive test, and TN for a negative test.

$$\mathbb{P}(\overset{S}{\cancel{M}}|TP) = \frac{\mathbb{P}(TP|\overset{S}{\cancel{M}})\mathbb{P}(\overset{S}{\cancel{M}})}{\mathbb{P}(TP)} = \frac{0.99 \times 0.005}{\mathbb{P}(TP)} = \frac{0.00495}{\mathbb{P}(TP)}$$

To calculate $\mathbb{P}(TP)$, we use the law of total probability:

$$\mathbb{P}(TP) = \mathbb{P}(TP|\overset{S}{\cancel{M}})\mathbb{P}(\overset{S}{\cancel{M}}) + \mathbb{P}(TP|\overset{H}{\cancel{S}})\mathbb{P}(\overset{H}{\cancel{S}}) = 0.99 \times 0.005 + 0.01 \times 0.995 = 0.0149$$

The probabilities calculated this way always correspond to the ratio between one of the terms from the law of total probability and the total sum.

$$\mathbb{P}(\overset{S}{\cancel{M}}|TP) = \frac{\mathbb{P}(TP|\overset{S}{\cancel{M}})\mathbb{P}(\overset{S}{\cancel{M}})}{\mathbb{P}(TP|\underset{S}{\cancel{M}})\mathbb{P}(\underset{S}{\cancel{M}}) + \mathbb{P}(TP|\underset{H}{\cancel{S}})\mathbb{P}(\cancel{S}) \, H} = \frac{0.00495}{0.0149} \approx 0.33$$

Independent Events, 3.8

Two events E and F are said to be independent if knowing that F occurs does not affect the probability of E, that is:

$$\frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)} = \mathbb{P}(E|F) = \mathbb{P}(E)$$

By the definition of conditional probability, this is equivalent to:

A_1, A_2, A_3
all independent

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F) \text{ and } \mathbb{P}(F|E) = \mathbb{P}(F)$$

NB $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3)$

Example

In the colored dice roll on the third slide of this group, we reasoned as follows:

- Let two dice be rolled, one blue and one red. Let B and R be the respective outcomes.
- If it is known that $B = 6$, then $B - R \geq 3$ if and only if $R \leq 3$, and the probability of $R \leq 3$ can be easily calculated (do it!).

To do this correctly, in the second point, I should have calculated $\mathbb{P}(R \leq 3 | B = 6)$. However, the outcome of the blue die can reasonably be assumed to be independent of the outcome of the red die (as long as they are not magnetized!), so $\mathbb{P}(R \leq 3 | B = 6)$ and $\mathbb{P}(R \leq 3)$ coincide, and our reasoning is valid.

Exercises

At the end of Chapter 3 in the textbook, you will find many exercises. The recommendation is to solve a few, focusing on those that do NOT require complicated applications of combinatorial calculations.

Exercises

$$\mathbb{P}(Acc) = \mathbb{P}(Acc \mid G) \cdot \mathbb{P}(G) + \mathbb{P}(Acc \mid Av) \cdot \mathbb{P}(Av) + \mathbb{P}(Acc \mid B) \cdot \mathbb{P}(B)$$

$$= 0.05 \times 0.2 + 0.15 \times 0.5 + 0.3 \times 0.3$$

Suppose that an insurance company classifies people into one of three classes: good risks, average risks, and bad risks.
Their records indicate that the probabilities that good, average, and bad risk persons will be involved in an accident over a 1-year span are, respectively, .05, .15, and .30.

1. If 20 percent of the population are "good risks," 50 percent are "average risks," and 30 percent are "bad risks," what proportion of people have accidents in a fixed year?
2. If policy holder A had no accidents in 1987, what is the probability that he or she is a good (average) risk?

$$\mathbb{P}(A \text{ is } G \mid no\ Acc) = \frac{\mathbb{P}(no\ Acc \mid G) \cdot \mathbb{P}(G)}{\mathbb{P}(no\ Acc)}$$

$$\mathbb{P}(no\ Acc \mid G) = 0.95 = 1 - \mathbb{P}(Acc \mid G)$$

Exercises

A certain organism possesses a pair of each of 5 different genes (which we will designate by the first 5 letters of the English alphabet). Each gene appears in 2 forms (which we designate by lowercase and capital letters). The capital letter will be assumed to be the dominant gene in the sense that if an organism possesses the gene pair xX, then it will outwardly have the appearance of the X gene. For instance, if X stands for brown eyes and x for blue eyes, then an individual having eithergene pair XX or xX will have brown eyes, whereas one having gene pair xx will be blue-eyed. The characteristic appearance of an organism is called its phenotype, whereas its genetic constitution is called its geno- type. (Thus 2 organisms with respective genotypes aA, bB, cc, dD, ee and AA, BB, cc, DD, ee would have different genotypes but the same pheno- type.) In a mating between 2 organisms each one contributes, at random, one of its gene pairs of each type. The 5 contributions of an organism (one of each of the 5 types) are assumed to be independent and are also independent of the contributions of its mate. In a mating between or- ganisms having genotypes aA, bB, cC, dD, eE, and aa, bB, cc, Dd, ee, what is the probability that the progeny will phenotypically (and then genotipically) resemble

MUM

aA, bB, cC, dD, eE,

$\downarrow A$

DAD

aa, bB, cc, Dd, ee

1) Calculate the probability that the progeny will phenotipically resemble the first (second) parent.

$$P_1 = \mathbb{P}_p(\text{1st gene is either } aA \text{ or } AA,\ \text{2nd gene } bB \text{ or } BB,\ \text{3rd gene is either } cC \text{ or } CC,\ \text{4th gene is } dD \text{ or } DD,\ \text{5th } eE \text{ or } EE) = ?$$

By independence

$$P_1 = \mathbb{P}_p(aA \cup AA) \cdot \mathbb{P}_p(bB \cup BB) \cdot \mathbb{P}_p(cC \cup CC) \cdot \mathbb{P}_p(dD \cup DD) \cdot \mathbb{P}(eE \cup EE)$$

$$= \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{3}{4} \cdot 1 = \frac{9}{2^6}$$

$$\mathbb{P}_p(aA \cup AA) = \mathbb{P}(aA \cup AA \mid Mum = aA,\ Dad = aa)$$

$$\begin{array}{ll} M & aA \\ D & aa \end{array} \quad \left\{ \begin{array}{cc} M & D \\ a & a \\ a & a \\ A & a \\ A & a \end{array} \right.$$

$$\mathbb{P}(aA \cup AA) = \frac{1}{2}$$

$$\begin{array}{ll} M & bB \\ D & bB \end{array} \quad \left\{ \begin{array}{cc} M & D \\ b & b \\ b & B \\ B & b \\ B & B \end{array} \right.$$

$$P(bB \cup BB) = \frac{3}{4}$$

M cC

D cc

$\rightarrow$ same as for the first gene $\mathbb{P}(cC \cup CC) = \frac{1}{2}$

M dD

D dD

$\mathbb{P}(dD \cup DD) = \frac{3}{4}$

M eE

D EE

$\mathbb{P}(eE \cup EE) = 1$

since Dad for sure transmits E

Exercises

Three prisoners are informed by their jailer that one of them has been chosen at random to be executed, and the other two are to be freed. Pris- oner A asks the jailer to tell him privately which of his fellow prisoners will be set free, claiming that there would be no harm in divulging this information because he already knows that at least one of the two will go free. The jailer refuses to answer this question, pointing out that if A knew which of his fellow prisoners were to be set free, then his own probability of being executed would rise from 13 to 12 because he would then be one of two prisoners. What do you think of the jailer's reasoning?

Thank you for the attention

POLITECNICO
DI TORINO