Topic 6

Conditional Probability and Independence

Conditional Probability

Outline

Definition

The Multiplication Principle

The Law of Total Probability

Introduction

Toss a fair coin 3 times. Let winning be "at least two heads out of three"

HHH	HHT	HTH	HTT
THH	THT	TTH	TTT

If we now know that the first coin toss is heads, then only the top row is possible and we would like to say that the probability of winning is

$$\frac{\#(\text{outcome that result in a win and also have a heads on the first coin toss})}{\#(\text{outcomes with heads on the first coin toss})}$$

$$= \frac{\#\{\text{HHH, HHT, HTH}\}}{\#\{\text{HHH, HHT, HTH, HTT}\}} = \frac{3}{4}.$$

We can take this idea to create a formula in the case of equally likely outcomes for the statement the conditional probability of A given B.

$$P(A|B) = \text{the proportion of outcomes in } A \text{ that are also in } B = \frac{\#(A \cap B)}{\#(B)}.$$

Definition

We can turn this into a more general statement using only the probability, P, by dividing both the numerator and the denominator in this fraction by $\#(\Omega)$.

$$P(A|B) = \frac{\#(A \cap B)/\#(\Omega)}{\#(B)/\#(\Omega)} = \frac{P(A \cap B)}{P(B)}$$

We thus take this version of the identity as the general definition of conditional probability for any pair of events A and B as long as the denominator $P(B) > 0$.

Exercise. Pick an event B so that $P(B) > 0$. Define, for every event A,

$$Q(A) = P(A|B).$$

Show that Q satisfies the three axioms of a probability. In words, *a conditional probability is a probability*.

Introduction

Roll two dice. The event {first roll is 3} is indicated.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Then $P\{\text{sum is 8}|\text{first die shows 3}\} = 1/6$, and $P\{\text{ sum is 8}|\text{first die shows 1}\} = 0$

Exercise. Roll two 4-sided dice. With the numbers 1 through 4 on each die, the value of the roll is the number on the side facing downward. Assume equally likely outcomes, find $P\{\text{sum is at least 5}\}$, $P\{\text{first die is 2}\}$ and $P\{\text{sum is at least 5}|\text{first die is 2}\}$.

The Multiplication Principle

The defining formula for conditional probability can be rewritten to obtain the multiplication principle, $P(A \cap B) = P(A|B)P(B)$.

$$\begin{aligned} P\{\text{ace on first 2 cards}\} &= P\{\text{ace on 2nd card}|\text{ace on 1st card}\}P\{\text{ace on 1st card}\} \\ &= \frac{3}{51} \times \frac{4}{52} = \frac{1}{17} \times \frac{1}{13} \end{aligned}$$

We can continue this process to obtain a chain rule:

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

Thus, $P\{\text{ace on first 3 cards}\}$

$$\begin{aligned} &= P\{\text{ace on 3rd card}|\text{ace on 1st and 2nd card}\}P\{\text{ace on 2nd card}|\text{ace on 1st card}\} \\ &\quad \times P\{\text{ace on 1st card}\} = \frac{2}{50} \times \frac{3}{51} \times \frac{4}{52} = \frac{1}{25} \times \frac{1}{17} \times \frac{1}{13}. \end{aligned}$$

The Multiplication Principle

Exercise. For an urn with b blue balls and g green balls, find

- the probability of green, blue, green (in that order)
- the probability of green, green, blue (in that order)
- $P\{$exactly 2 out of 3 are green$\}$
- $P\{$exactly 2 out of 4 are green$\}$

To answer the final part, appropriately modify the the first three parts above.

The Law of Total Probability

A partition of the sample space Ω is a finite collection of pairwise mutually exclusive events

$$\{C_1, C_2, \ldots, C_n\}$$

whose union is Ω.

Thus, every outcome $\omega \in \Omega$ belongs to *exactly* one of the C_i. In particular, distinct members of the partition are mutually exclusive. ($C_i \cap C_j = \emptyset$, if $i \neq j$)

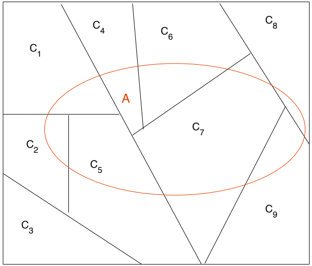

Figure: A partition of Ω into $n = 9$ events.

An event A can be written as the union

$$(A \cap C_1) \cup \cdots \cup (A \cap C_n)$$

of mutually exclusive events.

The Law of Total Probability

If we know, for each state, the fraction of the population whose ages are from 18 to 25, then we cannot just average these values to obtain this fraction over the whole country.This method fails because it give equal weight to California and Wyoming.

The law of total probability says that we should weigh these conditional probabilities by the probability of residence in a each state and then sum over the states.

Let $\{C_1, C_2, \ldots, C_n\}$ be a partition of Ω chosen so that $P(C_i) > 0$ for all i. Then, for any event A,

$$P(A) = \sum_{i=1}^{n} P(A|C_i)P(C_i).$$

$$\begin{aligned} P(A) &= P((A \cap C_1) \cup \cdots \cup (A \cap C_n)) = P(A \cap C_1) + \cdots + P(A \cap C_n) \\ &= P(A|C_1)P(C_1) + \cdots + P(A|C_n)P(C_n) \end{aligned}$$

NB. For the partition $\{C, C^c\}$, $P(A) = P(A|C)P(C) + P(A|C^c)P(C^c)$.