Introduction to Probability

Combining Probabilities

Events rarely occur in isolation. Usually we are interested in a combination or compound of events; for example

- The probability that two sections of a factory will be understaffed on the same day
- The probability of having a car accident today, given that you have had a car accident in the last five years.

We will look at two laws of probability for combining events

- The Addition Law
- The multiplication Law

Introduction to Probability

Conditional Probability The conditional probability of an event is the probability that an event A occurs given that another event B has already occurred. This type of probability is calculated by restricting the sample space that were working with to only the set B.

The formula for conditional probability can be rewritten using some basic algebra. Instead of the formula:

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

Introduction to Probability

Probability trees The setting out of solutions to problems requiring the manipulation of the probabilities of mutually exclusive and independent events can sometimes be helped by the use of probability tree diagrams. These have useful applications in decision theory.

The best choice of probability tree structure often depends upon the question and the natural order in which events like A and B above occur.

Introduction to Probability

Histograms A histogram is constructed from a frequency table. The intervals are shown on the X-axis and the number of scores in each interval is represented by the height of a rectangle located above the interval. A histogram of the response times from the dataset Target RT is shown below.

Introduction to Probability

Cumulative Distribution Function

The cumulative distribution function (c.d.f.) of a discrete random variable X is the function F(t) which tells you the probability that X is less than or equal to t. So if X has p.d.f. $P(X = x)$, we have:

$$F(t) = P(X \leq 1)$$

In other words, for each value that X can be which is less than or equal to t, work out the probability that X is that value and add up all such results.

Introduction to Probability

Example

In the above example where the die is thrown repeatedly, lets work out $P(X \leq t)$ for some values of t.

$P(X \leq 1)$ is the probability that the number of throws until we get a 6 is less than or equal to 1. So it is either 0 or 1.

- $P(X = 0) = 0$
- $P(X = 1) = 1/6$.
- Hence $P(X \leq 1) = 1/6$

Similarly, $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$
$= 0 + 1/6 + 5/36 = 11/36$

Conditional Probability

Suppose *B* is an event in a sample space *S* with $P(B) > 0$. The probability that an event *A* occurs once *B* has occurred or, specifically, the conditional probability of A given *B* (written $P(A|B)$), is defined as follows:

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

- This can be expressed as a multiplication theorem

$$P(A \cap B) = P(A|B) \times P(B)$$

- The symbol | is a vertical line and does not imply division.
- Also $P(A|B)$ is not the same as $P(B|A)$.

Remark: The Prosecutor's Fallacy , with reference to the O.J. Simpson trial.

Independent Events

Events A and B in a probability space *S* are said to be independent if the occurrence of one of them does not influence the occurrence of the other.

More specifically, *B* is independent of A if $P(B)$ is the same as $P(B|A)$. Now substituting $P(B)$ for $P(B|A)$ in the multiplication theorem from the previous slide yields.

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

We formally use the above equation as our definition of independence.

Mutually Exclusive Events

- Two events are mutually exclusive (or disjoint) if it is impossible for them to occur together.
- Formally, two events *A* and *B* are mutually exclusive if and only if $A \cap B = \varnothing$

Consider our die example

- Event A = 'observe an odd number' = $\{1,3,5\}$
- Event B = 'observe an even number' = $\{2,4,6\}$
- $A \cap B = \varnothing$ (i.e. the empty set), so *A* and *B* are mutually exclusive.

Addition Rule

The addition rule is a result used to determine the probability that event *A* or event *B* occurs or both occur. The result is often written as follows, using set notation:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- $P(A)$ = probability that event *A* occurs.
- $P(B)$ = probability that event *B* occurs.
- $P(A \cup B)$ = probability that either event *A* or event *B* occurs, or both occur.
- $P(A \cap B)$ = probability that event *A* and event *B* both occur.

Remark: $P(A \cap B)$ is subtracted to prevent the relevant outcomes being counted twice.

Addition Rule (Continued)

For mutually exclusive events, that is events which cannot occur together:
$P(A \cap B) = 0$. The addition rule therefore reduces to

$$P(A \cup B) = P(A) + P(B)$$

Addition Rule: Worked Example

Suppose we wish to find the probability of drawing either a Queen or a Heart in a single draw from a pack of 52 playing cards. We define the events Q = 'draw a queen' and H = 'draw a heart'.

- $P(Q)$ probability that a random selected card is a Queen
- $P(H)$ probability that a randomly selected card is a Heart.
- $P(Q \cap H)$ probability that a randomly selected card is the Queen of Hearts.
- $P(Q \cup H)$ probability that a randomly selected card is a Queen or a Heart.

Solution

- Since there are 4 Queens in the pack and 13 Hearts, so the $P(Q) = 4/52$ and $P(H) = 13/52$ respectively.

- The probability of selecting the Queen of Hearts is $P(Q \cap H) = 1/52$.

- We use the addition rule to find $P(Q \cup H)$:

$$P(Q \cup H) = (4/52) + (13/52) - (1/52) = 16/52$$

- So, the probability of drawing either a queen or a heart is $16/52 (= 4/13)$.

Multiplication Rule

The multiplication rule is a result used to determine the probability that two events, *A* and *B*, both occur. The multiplication rule follows from the definition of conditional probability.

The result is often written as follows, using set notation:

$$P(A|B) \times P(B) = P(B|A) \times P(A) \qquad (= P(A \cap B))$$

Recall that for independent events, that is events which have no influence on one another, the rule simplifies to:

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

Multiplication Rule

From the first year intake example, check that

$$P(E|F) \times P(F) = P(F|E) \times P(E)$$

- $P(E|F) \times P(F) = 0.58 \times 0.38 = 0.22$
- $P(F|E) \times P(E) = 0.55 \times 0.40 = 0.22$

Law of Total Probability

The law of total probability is a fundamental rule relating marginal probabilities to conditional probabilities. The result is often written as follows, using set notation:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

where $P(A \cap B^c)$ is probability that event A occurs and B does not.

Using the multiplication rule, this can be expressed as

$$P(A) = P(A|B) \times P(B) + P(A|B^c) \times P(B^c)$$

Law of Total Probability

From the first year intake example , check that

$$P(E) = P(E \cap M) + P(E \cap F)$$

with $P(E) = 0.40$, $P(E \cap M) = 0.18$ and $P(E \cap F) = 0.22$

$$0.40 = 0.18 + 0.22$$

Remark: M and F are complement events.

Bayes' Theorem

Bayes' Theorem is a result that allows new information to be used to update the conditional probability of an event.

Recall the definition of conditional probability:

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

Using the multiplication rule, gives Bayes' Theorem in its simplest form:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Probability: Worked Example

An electronics assembly subcontractor receives resistors from two suppliers: Deltatech provides 70% of the subcontractors's resistors while another company, Echelon, supplies the remainder.

1% of the resistors provided by Deltatech fail the quality control test, while 2% of the resistors from Echelon also fail the quality control test.

1. What is the probability that a resistor will fail the quality control test?
2. What is the probability that a resistor that fails the quality control test was supplied by Echelon?

Probability: Worked Example

Firstly, let's assign names to each event.

- D : a randomly chosen resistor comes from Deltatech.
- E : a randomly chosen resistor comes from Echelon.
- F : a randomly chosen resistor fails the quality control test.
- P : a randomly chosen resistor passes the quality control test.

We are given (or can deduce) the following probabilities:

- $P(D) = 0.70$,
- $P(E) = 0.30$.

Probability: Worked Example

We are given two more important pieces of information:

- The probability that a randomly chosen resistor fails the quality control test, given that it comes from Deltatech: $P(F|D) = 0.01$.
- The probability that a randomly chosen resistor fails the quality control test, given that it comes from Echelon: $P(F|E) = 0.02$.

Probability: Worked Example

The first question asks us to compute the probability that a randomly chosen resistor fails the quality control test. i.e. $P(F)$.

All resistors come from either Deltatech or Echelon. So, using the *law of total probability*, we can express $P(F)$ as follows:

$$P(F) = P(F \cap D) + P(F \cap E)$$

Probability: Worked Example

Using the ***multiplication rule*** i.e. $P(A \cap B) = P(A|B) \times P(B)$, we can re-express the formula as follows

$$P(F) = P(F|D) \times P(D) + P(F|E) \times P(E)$$

We have all the necessary probabilities to solve this.

$$P(F) = 0.01 \times 0.70 + 0.02 \times 0.30 = 0.007 + 0.006 = 0.013$$

Probability: Worked Example

- The second question asks us to compute probability that a resistor that fails the quality control test was supplied by Echelon.
- In other words; of the resistors that did fail the quality test only, what is the probability that a randomly selected resistor was supplied by Echelon?
- We can express this mathematically as $P(E|F)$.
- We can use *Bayes' theorem* to compute the answer.

Probability: Worked Example

Recall Bayes' theorem

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$$P(E|F) = \frac{P(F|E) \times P(E)}{P(F)} = \frac{0.02 \times 0.30}{0.013} = 0.46$$

More on probability

For this lecture and the next.

1. Contingency Tables
2. Conditional Probability: Worked Examples
3. Joint Probability Tables
4. The Multiplication Rule
5. Law of Total Probability
6. Bayes' Theorem
7. Exam standard Probability Question
8. Random Variables