Introduction to Probability

Lecture 5

Today

Marginal Distributions

Bayes' Theorem

Recap

Attendance: 57438842

Summary

Conditional Probability

Independent Events

$$P(\text{event}|\text{condition}) = \frac{P(\text{event occurs under given condition})}{P(\text{condition})}$$

or

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

If knowing B tells us nothing about A:

$$P(\mathrm{A}|\mathrm{B}) = P(A)$$

Law of Total Probability

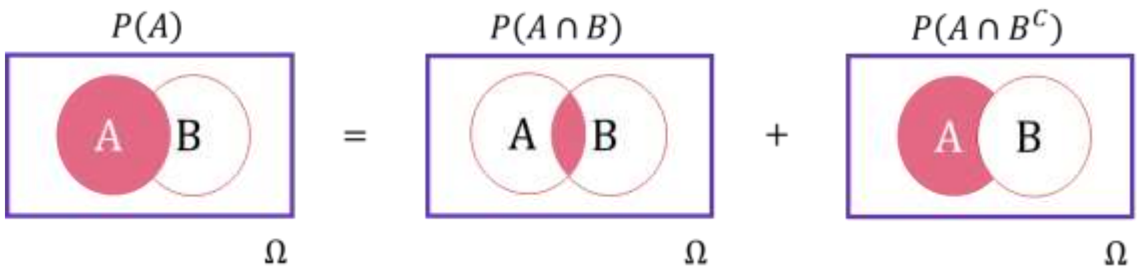
Total Probability

We can take an event A and break it into:

1. The piece of A in B
2. The piece of A not in B

Mathematically this means

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



For sets

$$|A| = |A \cap B| + |A \cap B^C|$$

Law of Total Probability

If there is a sequence of disjoint sets $B_1, B_2 \ldots B_N$ which "tile" A then

$$P(A) = \sum_{n=1}^{N} P(A \cap B_n)$$

The distribution $P(A)$ is called the **marginal distribution**.

Example

If $P(A \cap B) = 0.4$ and $P(A$

Total law:

$$P(A) = P(A \cap B) + P(A \cap B^C) = 0.4 + 0.2 = 0.6$$

Change of Notation

Change of Notation

From now on the **union** won't play a large role

Therefore

$$P(A \cap B) \rightarrow P(A, B)$$

This is called the **joint** distribution.

Change of Notation (2)

The previous equations are given by

$$P(A) = \sum_{n=1}^{N} P(A, B_n)$$

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

$$P(A, B) = \mathrm{P(A)P(B)}$$

Marginalising

Conditional Probability

Independence

Example

If the joint distribution $P(x, y)$ is given by the following

$P(x, y)$	$x = 0$	$x = 1$	$x = 2$
$y = 0$	${}^{2}/_{9}$	${}^{2}/_{9}$	${}^{2}/_{9}$
$y = 1$	${}^{1}/_{9}$	${}^{1}/_{9}$	${}^{1}/_{9}$

What is $P(y)$?

We use $P(y) = \sum_x P(x, y)$

$$P(y = 0) = \sum_x P(x, y = 0) = \frac{2}{9} + \frac{2}{9} + \frac{2}{9} = \frac{2}{3}$$

$$P(y = 1) = 1 - P(y = 0) = \frac{1}{3}$$

Class Examples

The following table represents the joint distribution of x and y

$P(x,y)$	$x = 0$	$x = 1$
$y = 0$	$1/6$	$1/12$
$y = 1$	$1/3$	$1/6$
$y = 2$	0	$1/4$

Calculate $\mathrm{P}(x)$ and $\mathrm{P}(y)$

Start with $P(x)$

$$P(x = 0) = \frac{1}{6} + \frac{1}{3} + 0 = \frac{1}{2} \rightarrow P(x = 1) = \frac{1}{2}$$

$$P(y = 0) = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}$$

$$P(y = 1) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

$$P(y = 2) = \frac{1}{4}$$

Law of Total Probability (2)

$$P(A,B) = P(A|B)P(B)$$

$$P(A) = \sum_{n=1}^{N} P(A,B_n) = \sum_{n=1}^{N} P(A|B_n)P(B_n)$$

Example:

If $P(A|B) = 0.2$, $P(A|\bar{B}) = 0.4$ and $P(B) = 0.1$ what is $P(A)$

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|\bar{B})P(\bar{B}) \\ &= 0.2 \times 0.1 + 0.4 \times (1 - 0.1) \\ &= 0.38 \end{aligned}$$

Example

3 machines manufacture components as follows

Machine	% of components made	% having defects
1	40	5
2	30	10
3	30	5

What is the probability a component has a defect?

We have $P(M_1) = 0.4$, $P(M_2) = 0.3$, $P(M_3) = 0.3$

$$P(d|M_1) = 0.05$$
$$P(d|M_2) = 0.1$$
$$P(d|M_3) = 0.05$$

$$P(d) = \sum_{i=1}^{3} P(d|M_i)P(M_i)$$
$$= 0.05 \times 0.4 + 0.1 \times 0.3 + 0.05 \times 0.3 = 0.065$$

Bayes' Theorem

Bayes' Theorem

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

We can use that

$$P(B) = \sum_{A} P(A, B) = \sum_{A} P(A|B)P(B)$$

So

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

Example

If $P(A|B) = \frac{1}{5}$, $P(A|B^C) = \frac{3}{10}$ and $P(B) = \frac{1}{2}$ what is $P(B|A)$?

Using Bayes' theorem

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

$$= \frac{\frac{1}{5} \times \frac{1}{2}}{\frac{1}{5} \times \frac{1}{2} + \frac{3}{10} \times \frac{1}{2}} = \frac{2}{5}$$

Example

There are two boxes:

1. Box A has 5 gold coins and 5 silver.
2. Box B has 5 gold coins and 10 silver.

If a box is picked at random, and a coin is then picked at random, what is the probability that the box chosen was A, given that the coin was silver?

Let g be the event cold coin

$$P(g|A) = \frac{1}{2} = P(\bar{g}|A)$$

$$P(g|B) = \frac{1}{3}; \; P(\bar{g}|B) = \frac{2}{3}$$

$$P(A) = P(B) = \frac{1}{2}$$

Then

$$P(A|\bar{g}) = \frac{P(\bar{g}|A)P(A)}{P(\bar{g}|A)P(A) + P(\bar{g}|B)P(B)}$$

$$= \frac{\frac{1}{2}}{\frac{1}{2} + \frac{2}{3}} = \frac{3}{7}$$

Example

You have two dice:

One is a biased die is three times as likely to be even as odd.

The other is fair.

If you pick the die at random and throw an odd number, what is the probability you have the biased die?

Let coin one (denoted c_1) be biased. Then

$$1 = 3p + 9p \rightarrow P(\text{odd}|c_1) = \frac{1}{4}$$

And $P(\text{odd}|c_2) = \frac{1}{2}$

$$P(c_1|\text{odd}) = \frac{P(\text{odd}|c_1)P(c_1)}{P(\text{odd}|c_1)P(c_1) + P(\text{odd}|c_2)P(c_2)}$$

$$= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{2}} = \frac{1}{3}$$

Class Example

Polygraphs are used to screen people. Each individual either tells the truth or lies. The (imperfect) polygraph returns 0 if it thinks the person is lying, or 1 for truthfulness, with the following probabilities:

$$P(0|\mathrm{Lies}) = 0.88$$
$$P(1|\mathrm{Truth}) = 0.86$$

People don't lie commonly, approximately 1% of the time.

What is the probability that the polygraph wrongly accuses someone of lying?

We have $\mathrm{P(Lies)} = 0.01$ so

$$\mathrm{P(Truth|0)}$$
$$= \frac{\mathrm{P(0|Truth)}P(\mathrm{Truth})}{\mathrm{P(0|Truth)}P(\mathrm{Truth}) + \mathrm{P(0|Lies)}P(\mathrm{Lies})}$$

$$= \frac{(1-0.86)\times(1-0.01)}{(1-0.86)\times(1-0.01) + 0.88\times 0.01}$$

$$= \frac{0.1386}{0.1474} \approx 0.94$$

Summary

Law of Total Probability

$$P(A) = \sum_B P(A, B) = \sum_B P(A|B)P(B)$$

Bayes' Theorem

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

Recap so far

Lecture 1: Introduction and Recap

1. Discrete probability is the long term frequency of an event.
2. The **sample space** (Ω) defines all our outcomes.
3. The **probability function** (P) assigns the probability to any event out (A) of the sample space.

The probability function satisfies $P(\Omega) = 1$, or:

$$1 = \sum_{x \in \Omega} P(x)$$

This is called **normalisation**.

Lecture 2: Combinatorics

The size of the sample space is given by

1. Unordered, with replacement: $|\Omega| = N^k$
2. Unordered, without replacement: $|\Omega| = {}^N P_k$
3. Ordered, without replacement: $|\Omega| = {}^N C_k$

The number of ways of placing k items with k_1 in group 1, k_2 in group 2 etc. up to group P, is given by the **multinomial coefficient**

$$\frac{k!}{k_1! k_2! \dots k_P!}$$

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Number of events in } A}{\text{Number of events in } \Omega}$$

Recap

Lecture 3: Foundations

We can manipulate our distribution using the underlying theory of sets. The basic law is:

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

in words: the probability of the event "A and B" equals the probability of A plus the probability of B, minus the probability of "A or B". If a sequence of events e_1, $e_2 \ldots e_N$ are disjoint then

$$P(e_1 \cup e_2 \ldots \cup e_N) = \sum_{n=1}^{N} P(e_n)$$

Lecture 4: Conditional Probability and Independence

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

$$P(A, B) \qquad \textbf{Joint Distribution}$$

$$P(A|B) = \frac{P(A, B)}{P(B)} \qquad \textbf{Conditional Distribution}$$

If events are conditionally independent then

$$P(A, B) = P(A)P(B)$$

For discrete distributions it can be represented in a **contingency table**.

Lecture 5: Marginal Distribution and Bayes Theorem

The marginal distribution is given by

$$P(A) = \sum_B P(A, B) \qquad \textbf{Marginal Distribution}$$

From which we can find

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_B P(B|A)P(A)} \qquad \textbf{Bayes' Formula}$$

Examples

Example

Given a part was defective from the following data, what is the probability it was manufactured by machine B?

Machine	Proportion of components made (%)	Proportion having defects (%)
A	20	1
B	50	2
C	30	3

We want $P(B|\text{defect})$

$$P(B|\text{defect}) = \frac{P(\text{defect}|B)P(B)}{P(\text{defect})}$$

$$P(\text{defect}) = \sum_{\text{machine}} P(\text{defect}|\text{machine})P(\text{machine})$$

$$= 0.01 \times 0.2 + 0.02 \times 0.5 + 0.03 \times 0.3 = 0.021$$

$$\rightarrow P(B|\text{defect}) = \frac{0.01}{0.021} = \frac{10}{21}$$

Example

A part is manufactured and then tested and are defective 30% of the time. The test identifies and removes defective parts with probability 0.9, non-defective parts with probability 0.2.

What is the probability a customer ends up with a defective part?

Letting d be defective, and r be removed. We have

$$P(r|d) = 0.9; P\left(r|\bar{d}\right) = 0.2; P(d) = 0.3$$

Then we are interested in

$$P(d|\bar{r}) = \frac{P(\bar{r}|d)P(d)}{P(\bar{r}|d)P(d) + P\left(\bar{r}|\bar{d}\right)P\left(\bar{d}\right)}$$

$$= \frac{(1-0.9)\times 0.3}{(1-0.9)\times 0.3 + (1-0.2)\times(1-0.3)}$$

$$= \frac{3}{59}$$

Class Example

A part is manufactured and then tested and are defective 30% of the time. The test identifies and removes defective parts with probability 0.9, non-defective parts with probability 0.2.

What is the probability the part is defective if it is removed?

Letting d be defective, and r be removed. We have

$$P(r|d) = 0.9; P(r|\bar{d}) = 0.2; P(d) = 0.3$$

Then we are interested in

$$P(d|r) = \frac{P(r|d)P(d)}{P(r|d)P(d) + P(r|\bar{d})P(\bar{d})}$$

$$= \frac{0.9 \times 0.3}{0.9 \times 0.3 + 0.2 \times (1 - 0.3)}$$

$$= \frac{27}{41}$$