Probability: Joint, Conditional, Marginal, and Bayes' Theorem

Mark Andrews

Division of Psychology
Nottingham Trent University

What is probability?

- Probability is a means to quantify uncertainty.
- If a variable can take on more than one value, probability can be used to describe the certainty that is will take each one of its possible values.
- Probabilities must lie between zero and one.
- The sum of the probabilities of all values of a variable must equal one.

What is probability? Notation

- If X is a variable with possible values $\{x_1, x_2 \ldots x_k \ldots x_K\}$, then

$$P(X = x_k) \tag{1}$$

gives the probability that X takes the value x_k.
- The rules of probability require that

$$0 \leqslant P(X = x_k) \leqslant 1 \quad \forall x_k \tag{2}$$

and

$$\sum_{k=1}^{K} P(X = x_k) = 1. \tag{3}$$

Simple probability rules

- If two values are mutually exclusive, then the probability of either happening is the *sum* of their individual probabilities, i.e.

$$P(X = x_k \text{ or } X = x_j) = P(X = x_k) + P(X = x_j). \tag{4}$$

- If two different variables are *independent*, the joint probability of their outcomes is equal to the product of their individual outcomes, i.e. if X and Y are independent then

$$P(X = x_k \text{ and } Y = y_l) = P(X = x_k) \times P(Y = y_l). \tag{5}$$

Conditional, joint and marginal probability

- We can discuss at least three different types of probability distributions, referred to as *joint*, *marginal* or *conditional* probabilities.
- There are fundamental relationships between them.
- We can illustrate these concepts by looking at survival rates of males and females on *RMS Titanic*.

Frequency distribution of survival rates on Titanic

- The following table shows the number of men and women who died or survived on the *Titanic*:

	Men	Women
Perished	1352	109
Survived	338	316

.

- Altogether, there were 2115 onboard.

Conditional and marginal probabilities using frequencies

- From the numbers

	Men	Women
Perished	1352	109
Survived	338	316

,

we are able to answer:

1. What is the overall probability of dying?
2. What is the overall probability of being a man, or a woman?
3. What is the probability of dying given the person is a man, or a woman?
4. What is the probability of being a man given a person survived?

From frequencies to probabilities

- We can convert the frequencies

	Men	Women
Perished	1352	109
Survived	338	316

,

to the probabilities

	Men	Women
Perished	.64	.05
Survived	.16	.15

,

by dividing each frequency by the total number onboard, i.e. 2115.

Joint probability tables

- The table

	Men	Women
Perished	.64	.05
Survived	.16	.15

,

  is a *joint probability* table.
- It provides the probability for every combination of the two variables.
- It is just another probability distribution, like what we have met already.
- Each element in the table lies between 0 and 1 and together they must sum to 1.

Marginal probabilities from joint probabilities

- From the table

	Men	Women
Perished	.64	.05
Survived	.16	.15

  ,

  what is the overall probability of dying?
- Recall that each element in the table provides the probability for a combination of values of the two variables, e.g. the probability of dying and being male is .64.
- Following the rules of probability, we calculate the probability of dying as follows:

$$\begin{aligned} \mathrm{P}(\text{Perished}) &= \mathrm{P}(\text{Male \& Perished}) + \mathrm{P}(\text{Female \& Perished}), && (6) \\ &= .64 + .05 = .69 && (7) \end{aligned}$$

Conditional probabilities from joint probabilities

- From the table

	Men	Women
Perished	.64	.05
Survived	.16	.15

,

what is the probability of dying if the person is a man?

- First, the probability of being a man onboard *Titanic* is

$$\begin{aligned} \mathrm{P}(\mathrm{Male}) &= \mathrm{P}(\mathrm{Male} \ \& \ \mathrm{Perished}) + \mathrm{P}(\mathrm{Male} \ \& \ \mathrm{Survived}), & (8) \\ &= .64 + .16 = .8\ . & (9) \end{aligned}$$

Conditional probabilities from joint probabilities (cont'd)

- Then, what is the probability of being a man and dying? This is simply .64.
- Putting these together: 80% of the ship's total was male and 64% of the total were men who died. That means the fraction of men who died is 64/80, i.e.

$$\mathrm{P}(\text{Perished}|\text{Male}) = \frac{\mathrm{P}(\text{Male \& Perished})}{\mathrm{P}(\text{Male})}, \tag{10}$$

$$= \frac{.64}{.8} = .8 \tag{11}$$

Calculating conditional probabilities from frequencies

- Notice how we calculate the conditional probabilities.
- For example, the probability of dying if the person is a man is

$$
\begin{aligned}
\mathrm{P}(\text{Perished}|\text{Male}) &= \frac{\#(\text{Male \& Perished})}{\#(\text{Male})}, &(12)\\
&= \frac{\#(\text{Male \& Perished})}{\#(\text{Male \& Perished}) + \#(\text{Male \& Survived})}, &(13)\\
&= \frac{1352}{1352 + 338} = .8\,. &(14)
\end{aligned}
$$

Conditional, joint and marginal probability: Rules

- Conditional probability is the probability distribution of a variable when the value of another variable is known, i.e.

$$P(X = x_k | Y = y_l)$$

is read as the "the probability that X is x_k *given* Y is y_l".

- Conditional probability can be derived from the joint probability as follows:

$$P(X = x_k | Y = y_l) = \frac{P(X = x_k, Y = y_l)}{P(Y = y_l)}$$

- Marginal probability can also be derived from the joint probability as follows:

$$P(Y = y_l) = \sum_{k=1}^{K} P(X = x_k, Y = y_l).$$

Conditional, joint and marginal: More Rules

- As we have

$$P(X = x_k | Y = y_l) = \frac{P(X = x_k, Y = y_l)}{P(Y = y_l)}$$

  then we also have

$$P(X = x_k, Y = y_l) = P(X = x_k | Y = y_l) P(Y = y_l),$$

- Likewise, given that

$$P(Y = y_l | X = x_k) = \frac{P(Y = y_l, X = x_k)}{P(X = x_k)}$$

  then

$$P(Y = y_l, X = x_k) = P(Y = y_l | X = x_k) P(X = x_k).$$

Putting the rules together

- We have seen

$$P(X = x_k, Y = y_l) = P(X = x_k | Y = y_l) P(Y = y_l),$$

…and

$$P(Y = y_l, X = x_k) = P(Y = y_l | X = x_k) P(X = x_k).$$

- As $P(Y = y_l, X = x_k) = P(X = x_k, Y = y_l)$, then

$$P(Y = y_l | X = x_k) P(X = x_k) = P(X = x_k | Y = y_l) P(Y = y_l).$$

Putting the rules together (cont'd)

$$
\begin{aligned}
P(Y = y_l, X = x_k) &= P(X = x_k, Y = y_l), \\
P(Y = y_l | X = x_k) P(X = x_k) &= P(X = x_k | Y = y_l) P(Y = y_l), \\
P(Y = y_l | X = x_k) &= \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{P(X = x_k)}
\end{aligned}
$$

And we arrive at Bayes' Theorem

- We have

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{P(X = x_k)}. \tag{15}$$

- As $P(X = x_k)$ is

$$\begin{aligned} P(X = x_k) &= \sum_{l=1}^{L} P(X = x_k, Y = y_l), & (16) \\ &= \sum_{l=1}^{L} P(X = x_k | Y = y_l) P(Y = y_l). & (17) \end{aligned}$$

- Therefore

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{\sum_{l'=1}^{L} P(X = x_k | Y = y_{l'}) P(Y = y_{l'})}$$

- This is Bayes's Theorem.