# *Probability Basics*

Mark Andrews

# *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 $\{x_1, x_2 \dots x_k \dots x_K\}$, 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

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

# *Two 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

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

$$P(X = x_k | Y = y_l) \tag{6}$$

  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:

$$P(X = x_k | Y = y_l) = \frac{P(X = x_k, Y = y_l)}{P(Y = y_l)} \tag{7}$$

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

$$P(X = x_k) = \sum_{l=1}^{L} P(X = x_k, Y = y_l). \tag{8}$$

# Bayes Theorem

- So far, we have seen that

$$P(Y = y_l | X = x_k) = \frac{P(Y = y_l, X = x_k)}{P(X = x_k)} \tag{9}$$

  and

$$P(Y = y_l, X = x_k) = P(X = x_k | Y = y_l) P(Y = y_l). \tag{10}$$

- Given this, we can always write

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{P(X = x_k)}. \tag{11}$$

- This is the essence of Bayes' theorem. If we know $P(X = x_k | Y = y_l) P(Y = y_l)$ we can calculate $P(Y = y_l | X = x_k)$, assuming we also know $P(Y = y_l)$ and $P(X = x_k)$.

# *Conditional, joint and marginal probability*

- There 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?

# *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[1]

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

[1]My probability notation here is less exact.

# *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:

$$\mathrm{P}(\text{Perished}) = \mathrm{P}(\text{Male \& Perished}) + \mathrm{P}(\text{Female \& Perished}), \tag{15}$$

$$= .64 + .05 = .69 \tag{16}$$

# *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

$$\begin{aligned} \mathrm{P}(\text{Male}) &= \mathrm{P}(\text{Male \& Perished}) + \mathrm{P}(\text{Male \& Survived}), && (17)\\ &= .64 + .16 = .8\,. && (18) \end{aligned}$$

# *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.

$$
\begin{aligned}
P(\text{Perished}|\text{Male}) &= \frac{P(\text{Male \& Perished})}{P(\text{Male})}, && (19)\\
&= \frac{.64}{.8} = .8 && (20)
\end{aligned}
$$

# *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.

$$P(X = x_k | Y = y_l) \tag{21}$$

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:

$$P(X = x_k | Y = y_l) = \frac{P(X = x_k, Y = y_l)}{P(Y = y_l)} \tag{22}$$

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

$$P(X = x_k) = \sum_{l=1}^{L} P(X = x_k, Y = y_l). \tag{23}$$

# Conditional, joint and marginal: More Rules

- As we have

$$P(X = x_k | Y = y_l) = \frac{P(X = x_k, Y = y_l)}{P(Y = y_l)} \tag{24}$$

then we also have

$$P(X = x_k, Y = y_l) = P(X = x_k | Y = y_l) P(Y = y_l), \tag{25}$$

- Likewise, given that

$$P(Y = y_l | X = x_k) = \frac{P(Y = y_l, X = x_k)}{P(X = x_k)} \tag{26}$$

then

$$P(Y = y_l, X = x_k) = P(Y = y_l | X = x_k) P(X = x_k). \tag{27}$$

# *Putting the rules together*

- We have seen
$$P(X = x_k, Y = y_l) = P(X = x_k | Y = y_l) P(Y = y_l), \tag{28}$$
- ... and
$$P(Y = y_l, X = x_k) = P(Y = y_l | X = x_k) P(X = x_k). \tag{29}$$
- As $P(Y = y_l, X = x_k) = P(X = x_k, Y = y_l)$,
-
$$P(Y = y_l, X = x_k) = P(X = x_k, Y = y_l), \tag{30}$$
$$P(Y = y_l | X = x_k) P(X = x_k) = P(X = x_k | Y = y_l) P(Y = y_l) \tag{31}$$

# *Putting the rules together (cont'd)*

$$P(Y = y_l, X = x_k) = P(X = x_k, Y = y_l), \tag{32}$$

$$P(Y = y_l | X = x_k) P(X = x_k) = P(X = x_k | Y = y_l) P(Y = y_l), \tag{33}$$

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{P(X = x_k)} \tag{34}$$

## *And we arrive at Bayes' Theorem*

- We have

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{P(X = x_k)}. \tag{35}$$

- As $P(X = x_k)$ is

$$\begin{aligned} P(X = x_k) &= \sum_{l=1}^{L} P(X = x_k, Y = y_l), & (36) \\ &= \sum_{l=1}^{L} P(X = x_k | Y = y_l) P(Y = y_l). & (37) \end{aligned}$$

- Therefore

$$P(Y = y_l | X = x_k) = \frac{P(X = x_k | Y = y_l) P(Y = y_l)}{\sum_{l'=1}^{L} P(X = x_k | Y = y_{l'}) P(Y = y_{l'})} \tag{38}$$

# *Expected Value*

- Consider a variable X that takes on discrete values $\{1, 2, 3\}$ with probabilities $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\}$.
- Here are a set of 25 random draws from that distribution,

$$1, 1, 1, 3, 1, 1, 1, 2, 1, 3, 2, 1, 3, 2, 1, 1, 3, 1, 1, 3, 2, 1, 1, 2, 1.$$

- The mean of this sample is 1.6, which we obtain by

$$\bar{x} = \frac{1}{25}(1 + 1 + 1 + 3 + 1 \ldots + 1 + 2 + 1).$$

# *Expected Value (cont'd)*

- If we count the number of ones, two, threes in the sample, we see there are 15 ones, 5 twos and 5 threes. This means $\bar{x}$ can also be written as

$$\bar{x} = \frac{1}{25}(15 \times 1 + 5 \times 2 + 5 \times 3), \tag{39}$$

$$= \frac{15}{25} \times 1 + \frac{5}{25} \times 2 + \frac{5}{25} \times 3, \tag{40}$$

$$= 1.6. \tag{41}$$

- In other words, the sample mean is identical to sum of the values times their probability of occurrence in the sample.

# *Expected Value (cont'd)*

- By the law of large numbers, we can be guaranteed that if we keep on drawing samples, these fractions will approach the true probabilities of the variable's values, so that in the limit the average of a large sample will be

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 3 = 1.75.$$

  This value of 1.75 is equal to the true mean or expectation or expected value of the variable X.

# Expected Utility

- On a roulette wheel, there are 38 pockets, 18 of which are red, 18 are black and 2 are green.
- On each play, you pay \$1 to play. You choose either red or black. If your chosen colour comes up, you win \$2, otherwise you gain nothing and lose your initial \$1.
- Your chance of winning is $18/38$, and your chance of losing is $20/38$.
- If you win, you win \$1. If you lose, you lose \$1.
- Your expected utility from playing this game is the calculated by adding the probability of winning times 1 to the probability of losing times $-1$, or

$$\frac{18}{38} \times 1 + \frac{20}{38} \times -1 = -\frac{2}{38} \approx -.0526. \tag{42}$$