Slide 2.1

Chapter 2: Probability

- Probability underlies statistical inference - the drawing of conclusions from a sample of data.
- If samples are drawn at random, their characteristics (such as the sample mean) depend upon chance.
- Hence to understand how to interpret sample evidence, we need to understand chance, or probability.

The definition of probability

- The probability of an event *A* may defined in different ways:
  - The frequentist view: the proportion of trials in which the event occurs, calculated as the number of trials approaches infinity
  - The subjective view: someone’s degree of belief about the likelihood of an event occurring

Some vocabulary

- an experiment: an activity such as tossing a coin, which has a range of possible outcomes.
- a trial: a single performance of the experiment.
- the sample space: all possible outcomes of the experiment. For a single toss of a coin the sample space is {Heads, Tails}.

Probabilities

- With each outcome in the sample space we can associate a probability (calculated according to either the frequentist or subjective view)
- Pr(Heads) = 1/2
  Pr(Tails) = 1/2
- This is an example of a probability distribution (more detail in Chapter 3)

Rules for probabilities

- $0 \leq \Pr(A) \leq 1$
- $\sum p = 1$, summed over all outcomes
- $\Pr(\text{not-}A) = 1 - \Pr(A)$

Examples: picking a card from a pack

- The probability of picking any one card from a pack (e.g. King of Spades) is 1/52. This is the same for each card
- Summing over all cards: 1/52 + 1/52 + ... = 1
- Pr(not-King of Spades) = 51/52 = 1-Pr(King of Spades)

Compound events

- Often we need to calculate more complicated probabilities:
  - what is the probability of drawing any Spade?
  - what is the probability of throwing a ‘double six’ with two dice?
  - what is the probability of collecting a sample of people whose average IQ is greater than 100?
- These are compound events.

Rules for calculating compound probabilities

1. The addition rule: the **'or'** rule

- Pr(*A* or *B*) = Pr(*A*) + Pr(*B*)
  The probability of rolling a five or six on a single roll of a die is
  Pr(5 or 6) = Pr(5) + Pr(6) = 1/6 + 1/6 = 1/3

1	2	3	4	5	6

A slight complication...

- If *A* and *B* can simultaneously occur, the previous formula gives the wrong answer...
  - Pr(King or Heart) = 4/52 + 13/52 = 17/52 ✖
- This double counts the King of Hearts
- 16 dots highlighted

	A	K	Q	J	10	9	8	7	6	5	4	3	2
Spades	•	•	•	•	•	•	•	•	•	•	•	•	•
Hearts	•	•	•	•	•	•	•	•	•	•	•	•	•
Diamonds	•	•	•	•	•	•	•	•	•	•	•	•	•
Clubs	•	•	•	•	•	•	•	•	•	•	•	•	•

A slight complication... (continued)

- We therefore subtract the King of Hearts:
  - Pr(King or Heart) = 4/52 + 13/52 - 1/52 = 16/52
- The formula is therefore
  - Pr(*A* or *B*) = Pr(*A*) + Pr(*B*) - Pr(*A* and *B*)
- When *A* and *B* cannot occur simultaneously, Pr(*A* and *B*) = 0

The multiplication rule

- When you want to calculate Pr(*A* and *B*):
- Pr(*A* and *B*) = Pr(*A*) × Pr(*B*)
- The probability of obtaining a double-six when rolling two dice is
  - Pr(Six and Six) = Pr(Six) × Pr(Six)
    = 1/6 × 1/6 = 1/36.

Another slight complication: independence

- Pr(drawing two Aces from a pack of cards, without replacement)...
- If the first card drawn is an Ace (P = 4/52), that leaves 51 cards, of which 3 are Aces.
- The probability of drawing the second Ace is 3/51, different from the probability of drawing the first Ace. They are not independent events. The probability changes.
- Pr(two Aces) = $4/52 \times 3/51 = 1/221$

Conditional probability

- 3/51 is the probability of drawing an Ace **given** that an Ace was drawn as the first card
- This is the conditional probability and is written Pr(Second Ace | Ace on first draw)
- To simplify notation write as Pr(A2|A1)
- This is the probability of event A2 occurring, given A1 has occurred

Conditional probability (continued)

- Consider Pr(A2| not-A1)...
- A ‘not-Ace’ is drawn first, leaving 51 cards of which 4 are Aces
- Hence Pr(A2| not-A1) = 4/51
- So Pr(A2| not-A1) $\neq$ Pr(A2|A1)
- They are not independent events

The multiplication rule again

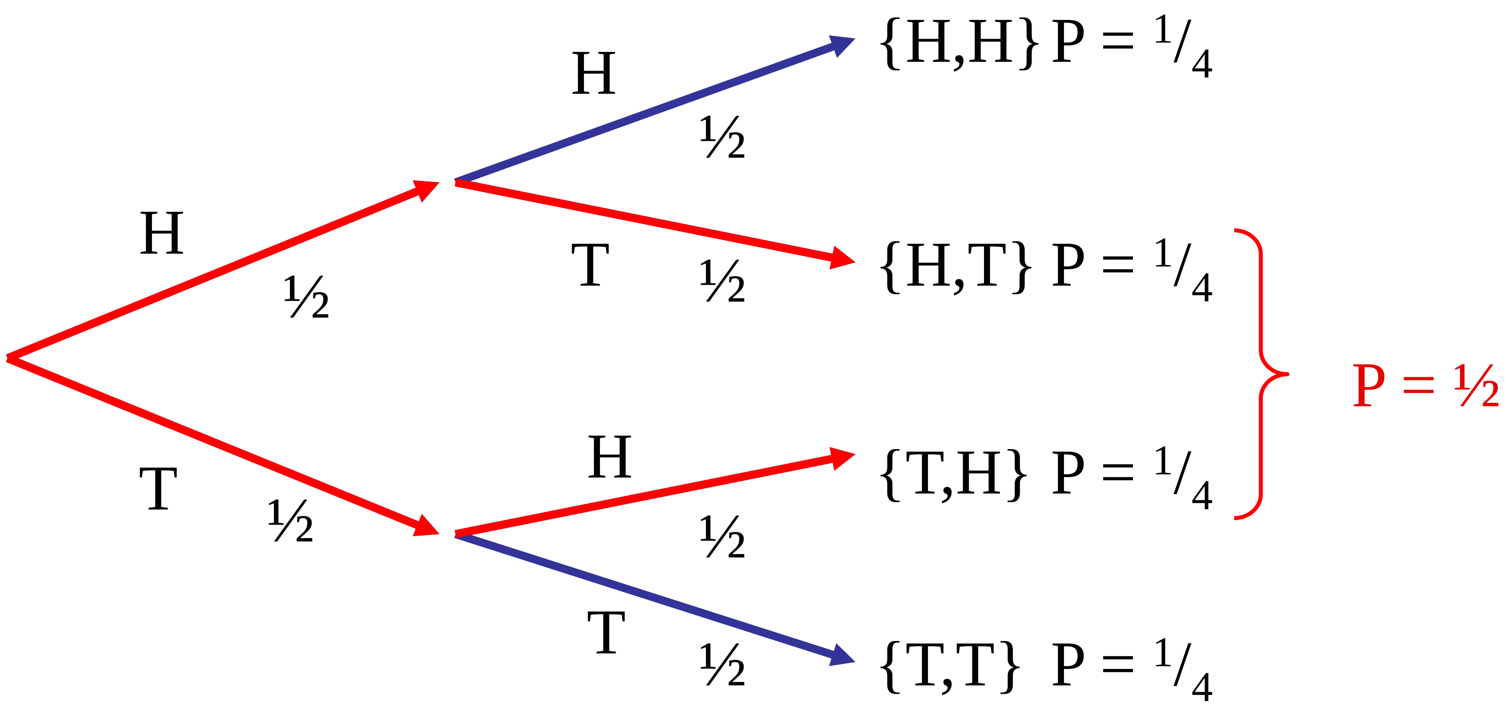
- The general rule is
  - $\Pr(A \text{ and } B) = \Pr(A) \times \Pr(B|A)$
- For independent events
  - $\Pr(B|A) = \Pr(B|\text{not-}A) = \Pr(B)$
- and so
  - $\Pr(A \text{ and } B) = \Pr(A) \times \Pr(B)$

Combining the rules

- Pr(1 Head in two tosses)...

- ...= Pr([H and T] or [T and H])
  = Pr([H and T]) + Pr([T and H])
  = [1/2 × 1/2] + [1/2 × 1/2])
  = 1/4 + 1/4 = 1/2

The tree diagram



But it gets complicated quickly...

- Pr(3 Heads in 5 tosses)?
- Pr(30 Heads in 50 tosses)?
- How many routes? Drawing takes too much time, we need a formula...

The Combinatorial formula

- The combinatorial formula gives the number of routes through the diagram:

$$nCr = \frac{n!}{r!(n-r)!}$$

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 1$$

- E.g. $5C3 = \frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \times (2 \times 1)} = 10$

n is the number of trials, *r* is the number of ‘successes’ required. Note the formula is not that difficult to evaluate, many terms cancel out.

The Combinatorial formula (continued)

- We can write the probability of 1 Head in 2 tosses as the probability of a head and a tail (in that order) times the number of possible orderings.
- Pr(2 Heads) = $\frac{1}{2} \times \frac{1}{2} \times 2C1 = \frac{1}{4} \times 2 = \frac{1}{2}$
- We can formalise this in the Binomial distribution...

The Binomial distribution

- If we define P as the probability of Heads, and hence (1-P) is the probability of Tails, we can write
  - $\Pr(1\text{ Head}) = P^1 \times (1-P)^1 \times 2C1$
- or, in general
  - $\Pr(r\text{ Heads}) = P^r \times (1-P)^{(n-r)} \times nCr$
- This is the formula for the Binomial distribution
- The Binomial distribution applies in circumstances where the underlying experiment has only two possible outcomes, e.g. success or failure.

Example

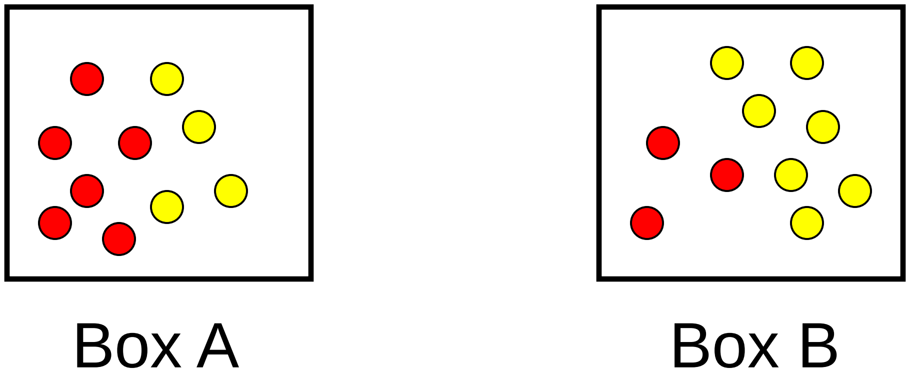
- P(3 heads in 5 tosses):
  - $n = 5$, $r = 3$, $P = ½$

- $\Pr(3 \text{ Heads}) = P^r \times (1-P)^{(n-r)} \times nCr$

$$= ½^3 \times (1 - ½)^2 \times 5C3$$
$$= {}^1/_8 \times {}^1/_4 \times 5!/(3! \times 2!)$$
$$= 10/32$$

Bayes' Theorem

- A ball is drawn at random from one of the boxes below. It is red.



- Intuitively, it seems more likely to have come from Box A. But what is the precise probability? Bayes theorem answers this question.

Solution

- We require Pr(A|R). This can be written:

$$\Pr(A \mid R) = \frac{\Pr(A \,\text{and}\, R)}{\Pr(R)}$$

- Expanding top and bottom we have:

$$\Pr(A \mid R) = \frac{\Pr(R \mid A) \times \Pr(A)}{\Pr(R \mid A) \times \Pr(A) + \Pr(R \mid B) \times \Pr(B)}$$

- We now have the answer in terms of probabilities we can evaluate.

Solution (continued)

- Hence we obtain:

$$\Pr(A \mid R) = \frac{6/10 \times 0.5}{6/10 \times 0.5 + 3/10 \times 0.5} = \frac{2}{3}$$

- There is a 2/3 probability the ball was taken from Box A.
- A similar calculation yields Pr(*B*|*R*) = $^1/_3$
- These are the posterior probabilities. The prior probabilities were ½, ½.

Prior and posterior probabilities

- Prior probabilities: Pr(A), Pr(B)
- Likelihoods: Pr($R|A$), Pr($R|A$)
- Posterior probabilities: Pr($A|R$), Pr($B|R$)

$$\text{posterior probability} = \frac{\text{likelihood} \times \text{prior probability}}{\sum(\text{likelihood} \times \text{prior probability})}$$

Table of likelihoods and probabilities

	Prior probabilities	Likelihoods	Prior $\times$ likelihood	Posterior probabilities
A	0.5	0.6	0.30	0.30/0.45 = 2/3
B	0.5	0.3	0.15	0.15/0.45 = 1/3
Total			0.45	

- The posterior probabilities are calculated as $^2/_3$ and $^1/_3$, as before.

Summary

- Probability underlies statistical inference
- There are rules (e.g. the multiplication rule) for calculating probabilities
- Independence simplfies the rules
- These rules lead on to probability distributions such as the Binomial
- Bayes theorem tells us how to update probabilities in the light of evidence