MONASH University

Information Technology

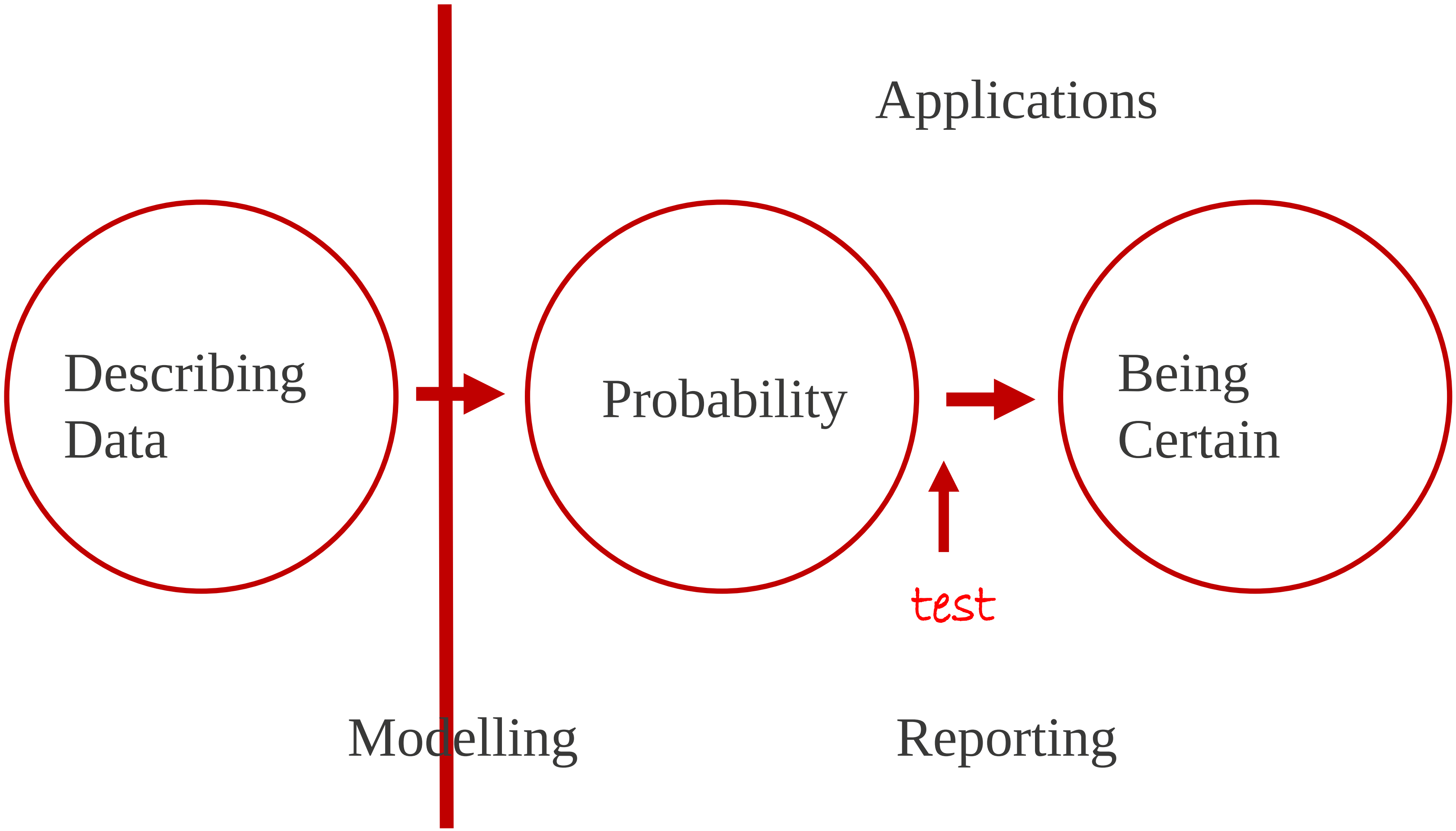
# FIT1006
# Business Information Analysis

## Lecture 9
## Introduction to Probability

# Topics covered:

- The probability of an event.
- Set notation and set operations for probability.
- Probability distributions.
- The mean and variance of a probability distribution.

# Course outline: Progress report



# Gambling on our future

As a nation we love to bet and, for too many of us, gambling infects how we manage our finances as well, writes **David Wilson**.

Research shows Australians bet and lose more than any other people. A report, published in *The Economist* last year, pegs Australian gambling losses at a world-beating $1144 per resident.

impulsive, which it seems is human nature.

"People live in the here-and-now, so if you know your next pay cheque is $2423, which is a mine worker's average weekly income, you are more likely to give in to impulse," he says. He does not judge, but rather advises accountability: being more frugal to become financially free.

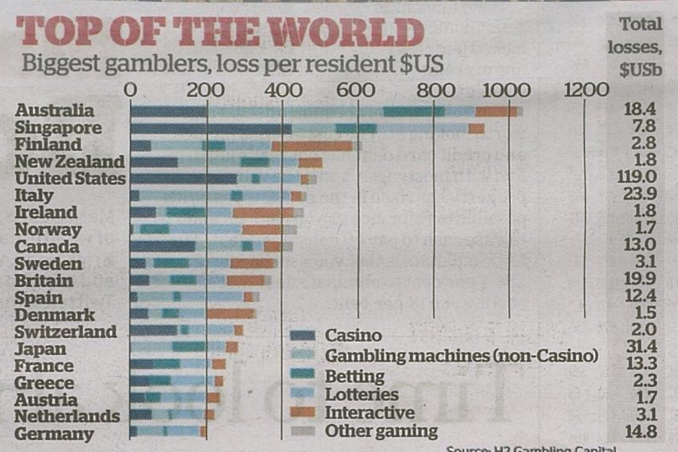
For starters, he advises, shred the credit card. The feeling of entering a department store knowing you have $400 to live off for the week is very different from when you have a $10,000 credit limit, he says, adding that the secret is to remove the ability to make impulse buys. Understanding that you can own your house in 10 years not 30, through thrift, does the trick.

Wealth psychology researcher Rik Schnabel's take is that Australians have an unfortunate saver or earner mindset. They would benefit from an investor view, Schnabel says, adding that a saver seeks bargains to conserve cash, although saving means spending.


gambling, Schnabel says, noting the two are polarities.

Worse still, Australians have a history of real gambling and will bet on almost anything, he says, arguing the trait was conditioned into the psyche by the traditional "two-up" game. Certainly, he says, the convict heritage breeds contempt for the rich and authority.

Encouraged by a "she'll be right" attitude, anti-elite battlers happily gamble, unable even to grasp the language of investing and wealth, according to Schnabel. In fact, he says, battlers shun tycoons, worried about estranging their low-income ... who think the rich are




**Clockwise from main: The traditional "two-up" game, shown in movie *Wake in Fright*, is part of the Australian psyche; problem gambling may affect 5 million Australians; big toys can hobble your investment intentions; race day can spike risky spending.**
Photos: Tamara Voninski, Nic Walker

offer, the more irresponsible Australians seem.

Financial planner Peter Horsfield also reckons Australians are reckless – especially when speculating on property with borrowed funds. Horsfield partly blames market forces – the tax deductions and historically low

narcissism. Warning how easily a "cross-collateralised" empire built on multiple mortgages can unwind, he quotes Warren Buffett on the perils of leverage.

"To make money they didn't have and didn't need, they risked what they did have and did need."

Fiscal recklessness can spike

their day, ME found in its survey last year of 1000 Australian adults. The findings suggest Australians effectively gamble on gambling.

According to the government, half a million Australians risk becoming, or are, problem gamblers. A single person's actions jeopardise the lives of up to 10 others, so 5 million

https://flux.qa (Feed code: SJ6KGV)

# Question 1 – Motivating Problem

There are 30 of us in a room. What is the probability that at least two of us have the same birthdate (disregard the year)? (Assume 365 days in the year)

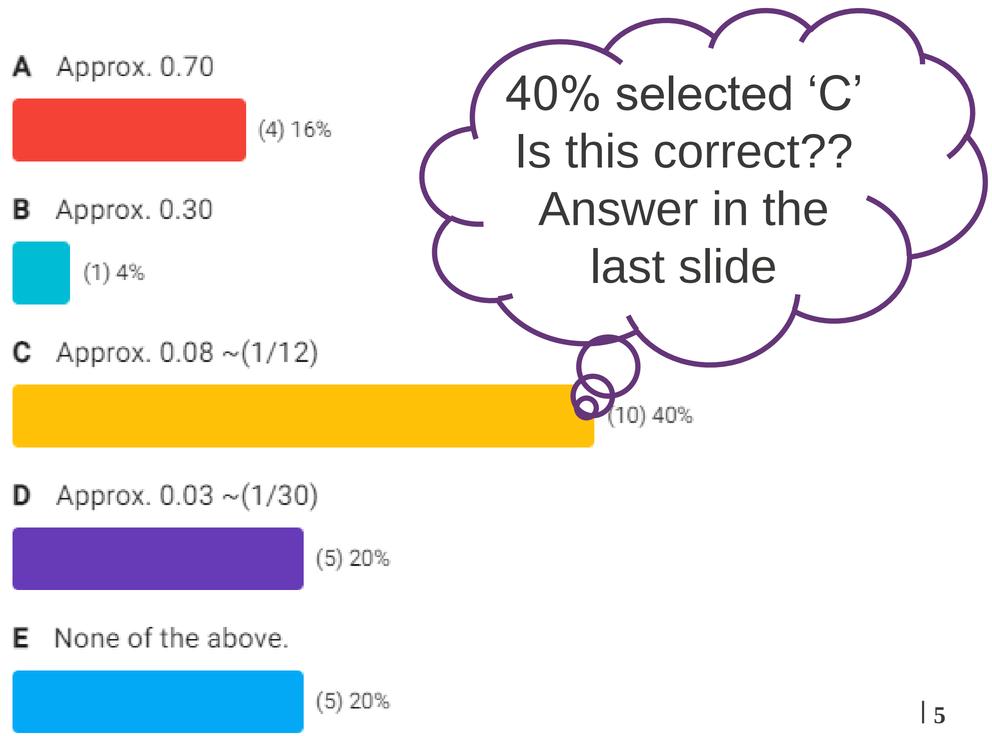
A. $\approx 0.70$

B. $\approx 0.30$

C. $\approx 0.08 \sim (1/12)$

D. $\approx 0.03 \sim (1/30)$

E. None of the above.



# Some Terminology

- Experiment (some activity leading to an)
- Outcome (a well defined result)
- Event (a collection of outcomes)
- Random Variable (value of the event)
- Sample Space (set of possible values of R.V.)
- Probability (chance R.V. assumes a certain value)

https://flux.qa (Feed code: SJ6KGV)

## Question 2

Two dice are thrown and the numbers appearing on the upper faces are observed. The figure below shows a(n):

A. Outcome
B. Event
C. Random Variable
✓D. Sample Space
E. Probability

| | | | | | |
|---|---|---|---|---|---|
| 11 | 12 | 13 | 14 | 15 | 16 |
| 21 | 22 | 23 | 24 | 25 | 26 |
| 31 | 32 | 33 | 34 | 35 | 36 |
| 41 | 42 | 43 | 44 | 45 | 46 |
| 51 | 52 | 53 | 54 | 55 | 56 |
| 61 | 62 | 63 | 64 | 65 | 66 |

https://flux.qa (Feed code: SJ6KGV)

# Question 3

Two dice are thrown and the numbers appearing on the upper faces are observed. Which event is more likely?

---

A. Both numbers are even 9

B. Sum is greater than 9 6

✓ C. At least one 6 is thrown 11

D. A double is thrown 6

| | | | | | |
|---|---|---|---|---|---|
| 11 | 12 | 13 | 14 | 15 | 16 |
| 21 | 22 | 23 | 24 | 25 | 26 |
| 31 | 32 | 33 | 34 | 35 | 36 |
| 41 | 42 | 43 | 44 | 45 | 46 |
| 51 | 52 | 53 | 54 | 55 | 56 |
| 61 | 62 | 63 | 64 | 65 | 66 |

# Determining Probability

- Objective Probability
  - Limit of relative frequency - based on an analysis of repeated outcomes.
  - Logical deduction - (previous slide) - based on an analysis of all outcomes. (Sometimes called the Classical Method)
  - Empirically - based on historical observations.
- Subjective Probability
  - Based on personal assessment using intuition or judgement. This is a method often employed for business decision making.

[https://flux.qa](https://flux.qa) (Feed code: SJ6KGV)

# Question 4

A potentially biased coin is tossed repeatedly and the number of heads appearing is divided by the number of tosses: (1 = head, 0 = tail). The long-run probability of throwing heads with this coin is:

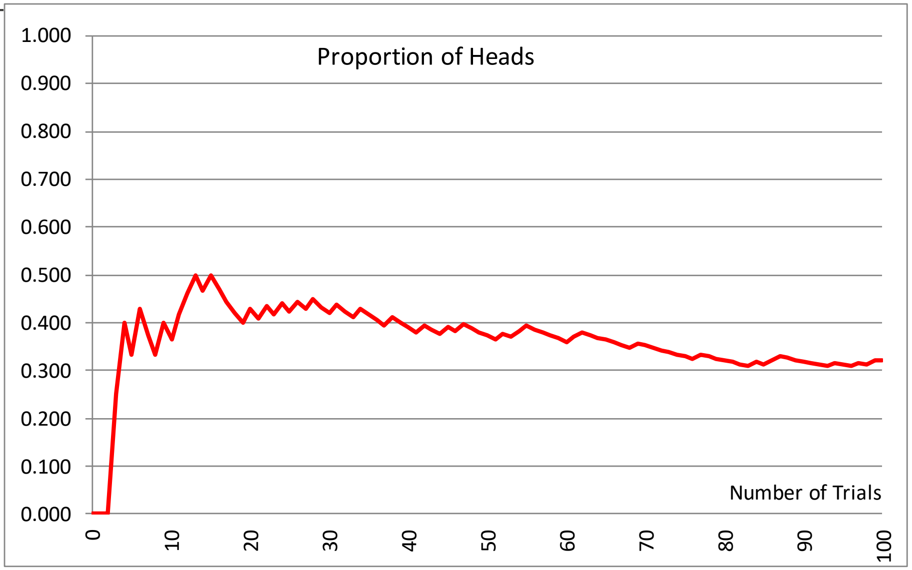
A. 0.5

B. 0.4

✓ C. 0.3

D. 0.0

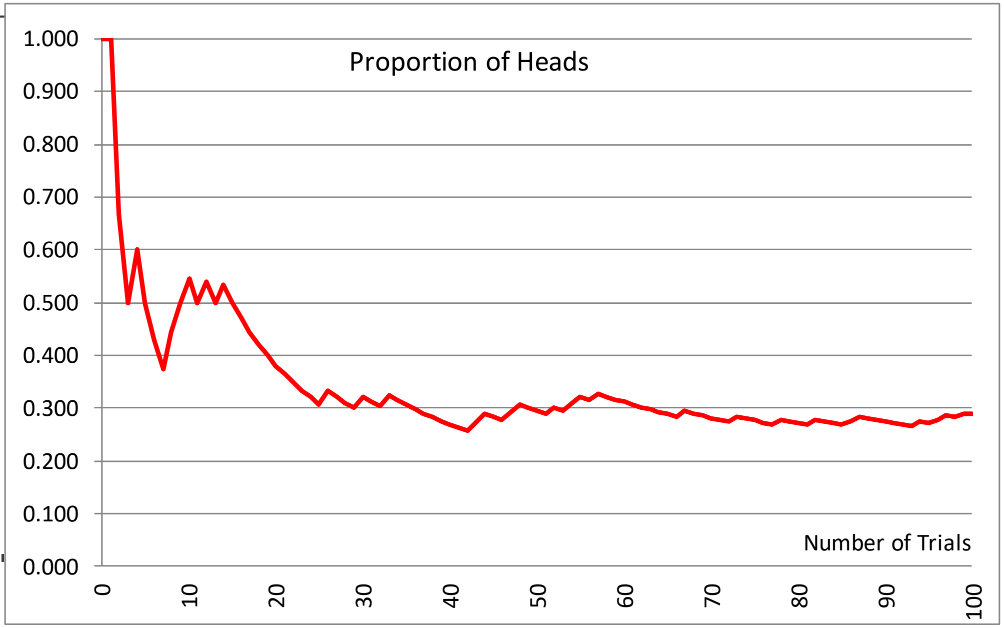
E. None of the above.



https://flux.qa (Feed code: SJ6KGV)

# Question 5

A potentially biased coin is tossed repeatedly and the number of heads appearing is divided by the number of tosses: (1 = head, 0 = tail). The long-run probability of throwing heads with this coin is:




A. 1.0

B. 0.5

✓ C. 0.3

D. 0.0

E. None of the above.

https://flux.qa (Feed code: SJ6KGV)

## Question 6

For the previous questions, how many tosses would it require before you were confident of your estimate of the probability of throwing heads with a potentially biased coin?

---

A. 10

B. 100

C. 1000

D. 10,000

E. More than 10,000.

We will revisit this question again in week 11 and you'll learn enough to work this out – whether you want to be 95% confident or 99% confident.

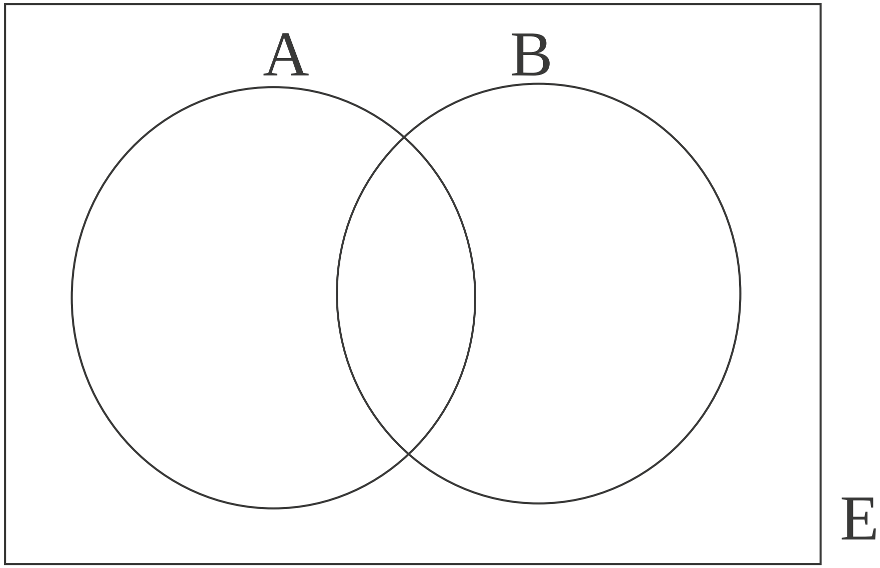
# Sets

- We define a <u>set</u> as a collection of objects, often related by some common property, and often use them to describe events or sample spaces.
- Notation. Use { } to mean *'the set of'*
- Let A represent the outcomes of throwing a die, then A = {1, 2, 3, 4, 5, 6}
- Let Y represent the set of even numbers, then

  Y = {…, -4, -2, 0, 2, 4, 6, …}
- By convention we refer to the set of all outcomes, or the sample space as the *Universal Set (E)*.

# Venn Diagrams

- Named after John Venn, these diagrams are a standard way of representing sets.
- We can show disjoint sets, intersection $\cap$, union $\cup$, complementary sets etc.



https://flux.qa (Feed code: SJ6KGV)

## Question 7

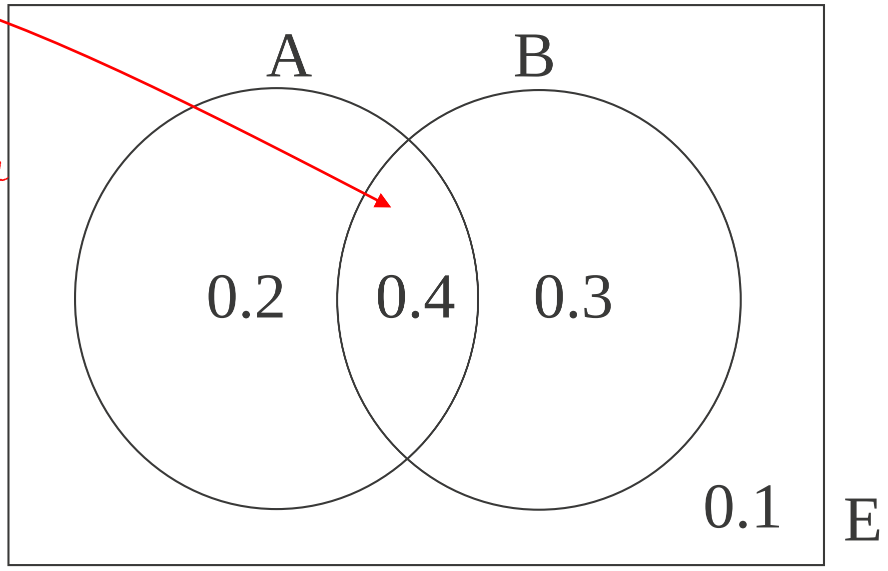
Using the Venn diagram below which probability is greatest?

A. $P(A \cap B)$ $= 0.4$

B. $P(\overline{A})$ $= 1 - (0.2 + 0.4) = 0.4$

✓ C. $P(B)$ $= (0.3 + 0.4) = 0.7$

D. $P(\overline{A \cap B})$ $= 1 - 0.4 = 0.6$



https://flux.qa (Feed code: SJ6KGV)

## Question 8

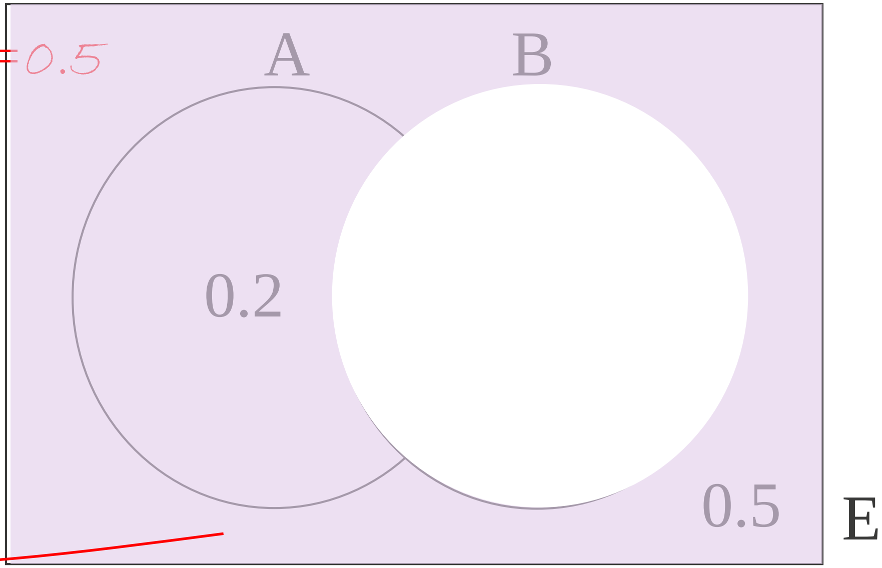
Using the Venn diagram below which probability is greatest?

A. $P(A \cup B)$ $= 0.2+0.2+0.1 = 0.5$

B. $P(A)$ $= 0.2+0.2=0.4$

C. $P(B)$ $= (0.2+0.1)=0.3$

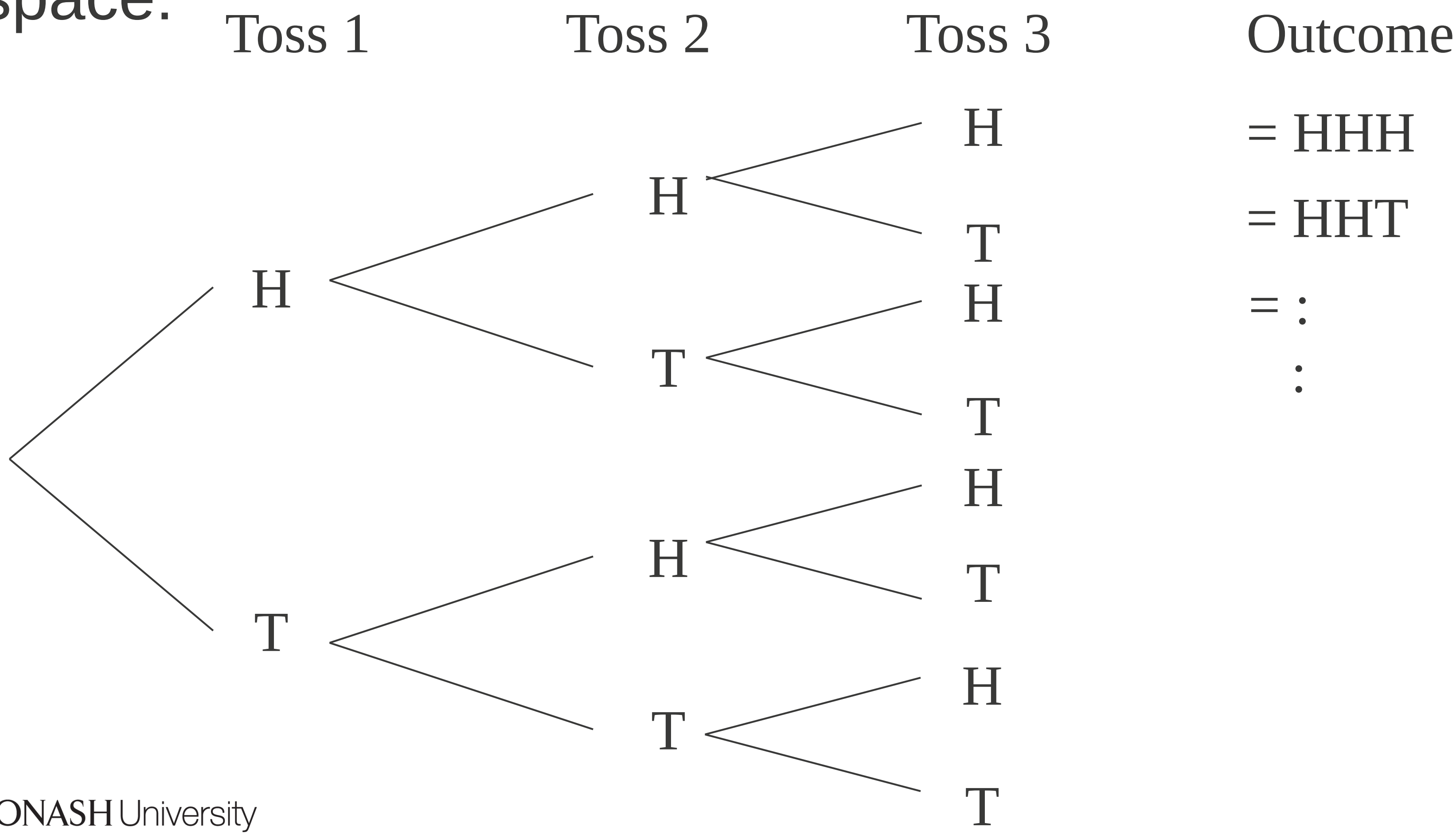
✓ D. $P(A \cup \overline{B})$ $= 0.7$




Shaded region

# Probability Tree

- Three coins are tossed. Complete the sample space:



# Probability Distribution

- Three coins are tossed.
- Let X denote the number of heads showing.
- The probability of each outcome, e.g. {H, H, T} is?

  $= 0.5 \times 0.5 \times 0.5 = 0.125$

- Why? → Each event is independent of the previous
- If events A and B are independent then

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

# Probability Distribution

Collating all the results for each x gives the following probability distribution of X.

| x | P(x) |
|---|---|
| 0 | 0.125 |
| 1 | 0.375 |
| 2 | 0.375 |
| 3 | 0.125 |

→ P(TTT) = 0.5 x 0.5 x 0.5 = 0.125

→ P(1 Head) = P(HTT) + P(THT) + P(TTH)

→ P(2 Heads) = P(HHT) + P(HTH) + P(THH)

→ P(HHH) = 0.5 x 0.5 x 0.5 = 0.125

# Expected Value

- If a random variable $X$, can take on a range of values, $x_i$, each with a certain probability of occurring, $P(x_i)$ then the Expected Value, $E(X)$, is the sum of each outcome multiplied by its probability of occurring.

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(x_i)$$

- The Expected Value is the outcome we 'expect' to obtain on average...

# Variance of a random variable

- In the same way that the variance of statistical data is the 'average' squared deviations, we can calculate the variance and standard deviation of a random variable.

$$Var(X) = \sigma^2 = E(X-\mu)^2 = \sum_{i=1}^{n}(x_i - \mu)^2 P(x_i)$$

Giving the computational formula of:

$$= \sum_{i=1}^{n} {x_i}^2 P(x_i) - \mu^2$$

MONASH University

# Group Activity

The probability distribution for the number of heads showing when 3 coins are tossed is

| x | P(x) |
|---|---|
| 0 | 0.125 |
| 1 | 0.375 |
| 2 | 0.375 |
| 3 | 0.125 |

Calculate the expected value

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(x_i)$$

| x | P(x) | xP(x) |
|---|---|---|
| 0 | 0.125 | |
| 1 | 0.375 | |
| 2 | 0.375 | |
| 3 | 0.125 | |
| | Sum | |

Calculate the expected value of $X^2$

| x | P(x) | $x^2$ | $x^2P(x)$ |
|---|---|---|---|
| 0 | 0.125 | | |
| 1 | 0.375 | | |
| 2 | 0.375 | | |
| 3 | 0.125 | | |
| | | Sum | |

Calculate the mean and variance of X

# Calculate the Expected Value (= Mean)

The probability distribution for the number of heads showing when 3 coins are tossed is

| x | P(x) |
|---|---|
| 0 | 0.125 |
| 1 | 0.375 |
| 2 | 0.375 |
| 3 | 0.125 |

Calculate the expected value

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(x_i)$$

| x | P(x) | xP(x) |
|---|---|---|
| 0 | 0.125 | 0 |
| 1 | 0.375 | 0.375 |
| 2 | 0.375 | 0.75 |
| 3 | 0.125 | 0.375 |
| | Sum | 1.5 |

Expected value or Mean

# Calculate the Variance of x

The probability distribution for the number of heads showing when 3 coins are tossed is

| x | P(x) |
|---|---|
| 0 | 0.125 |
| 1 | 0.375 |
| 2 | 0.375 |
| 3 | 0.125 |

First calculate the expected value of $x^2$:

| x | P(x) | $x^2$ | $x^2P(x)$ |
|---|---|---|---|
| 0 | 0.125 | 0 | 0 |
| 1 | 0.375 | 1 | 0.375 |
| 2 | 0.375 | 4 | 1.5 |
| 3 | 0.125 | 9 | 1.125 |
| | | Sum | 3 |

Expected value of $x^2$

$$\text{Variance of x} = \sum_{i=1}^{n} x_i^2 P(x_i) - \mu^2$$

$$= 3 - (1.5)^2 = 0.75$$

# Motivating problem

- There are 30 of us in a room. What is the probability that at least two of us have the same birthdate (disregard the year)?  (Assume 365 days in the year)

| Person | Event | Sample Space | Probability | Joint Prob | Interpretation |
|---|---|---|---|---|---|
| 1 | Their birthday is day A | 365/365 | 1.000 | 1.000 | P(one person has birthday on any day) |
| 2 | Their birthday is different to A - say B | 364/365 | 0.997 | 0.997 | P(two people have birthdays on diff days) |
| 3 | Their birthday is different to A and B - say C | 363/365 | 0.995 | 0.992 | P(three people have birthdays on diff days) |
| 4 | … not days A & B & C but D | 362/365 | 0.992 | 0.984 | … |
| 5 | … | … | 0.989 | 0.973 | |
| 6 | … | … | 0.986 | 0.960 | |
| 27 | … | … | 0.929 | 0.373 | |
| 28 | … | … | 0.926 | 0.346 | |
| 29 | … | … | 0.923 | 0.319 | |
| 30 | … | … | 0.921 | 0.294 | Prob that 30 people have birthdays on different days |

- Why is joint probability multiplied: P(A)*P(B)*P(C) etc.?
- 1 – P(30 people have b'days on different days) = ?

$= 1 - 0.294 = 0.706 \rightarrow$ so, answer to Q1 is 'A'

# Reading/Questions

- Reading:
  - $7^{th}$ Ed Sections 6.1, 7.1 – 7.3.
- Questions:
  - $7^{th}$ Ed Questions 6.2, 6.5, 6.17, 7.4, 7.5, 7.11, 7.15, 7.24, 7.25.