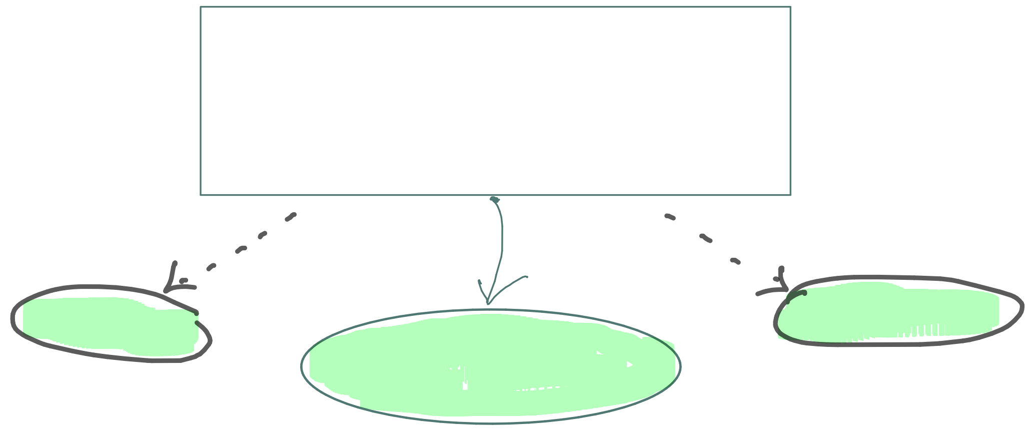
Module 3: Sampling Data



T6: Understanding Chance
T7: Chance Variability (The Box Model)
T8: Sample Surveys

LO6: Use the box model to describe chance & chance variability, including sample surveys & the CLT.

Why study chance?

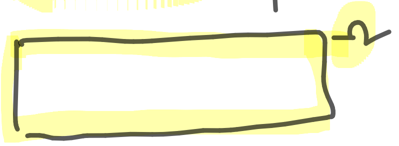
① The Prosecutor's Fallacy

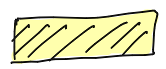
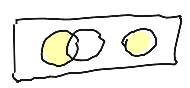
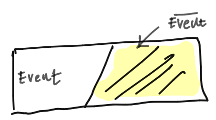
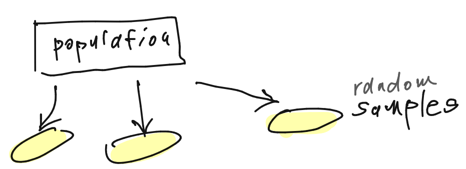
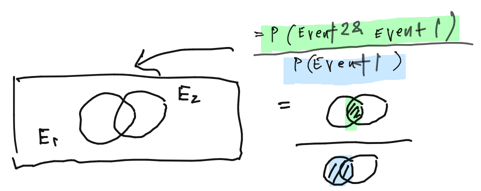
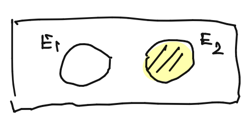
$$P(\text{innocent} \mid \text{evidence matches}) \neq P(\text{evidence matches} \mid \text{innocent})$$

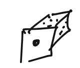
② The relationship between population & samples

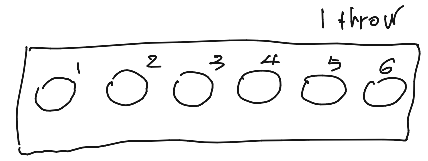
Chance = P(Event) = % of times an event is likely to occur, if a process is repeated long-term.

Sample space = all the possible outcomes of an event



	P(Event)	Picture of Sample Space
Impossible	0	
Certain	1	
Possible	≤ 1	
Opposite/complement	1 − P(Event)	 Event, Event
Random	same for every collection of objects	 population, random samples
Conditional	P(Event 2 | Event 1)	 E₁, E₂, P(Event 2 & event 1) / P(Event 1) =
(i) Dependent	≠ P(Event 2)	sampling without replacement
(ii) Independent	= P(Event 2)	sampling with replacement
Mutually Exclusive	0	 E₁, E₂

Example: Throw a fair die once. 



	P(Event)
Impossible	P(throwing a 7)
Certain	P(throwing a number less than 7)
Possible	P(throwing a 1)
Opposite/complement	P(not throwing a 2)
Random	2 throws: P(throwing a 6) ← as it's "fair"
Conditional	2 throws: P(throwing a 6 on 2nd throw | got a 6 on 1st throw)
(i) Dependent	
(ii) Independent	
Mutually Exclusive	1 throw: P(getting a 1 and a 6)

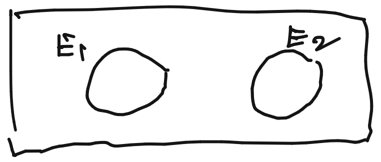
1 throw

Classic FAQs

FAQ 1 What's the difference between mutually exclusive & independence?

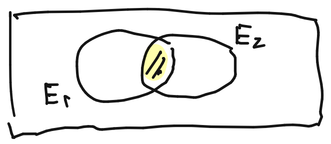
Mutually exclusive = the occurence of Event 1 prevents Event 2 occuring

("no overlap")

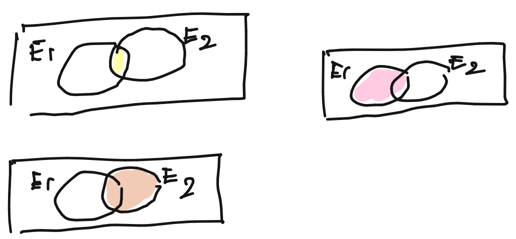


$P(\text{Event } 1 \cap \text{Event } 2) = 0$

Independence = the occurence of Event 1 does not change the chance of Event 2 occuring

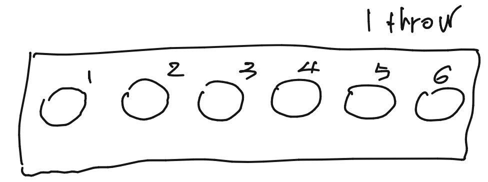


$$P(\text{Event } 2 \mid \text{Event } 1) = \frac{P(\text{Event } 1 \cap \text{Event } 2)}{P(\text{Event } 1)} = P(\text{Event } 2)$$

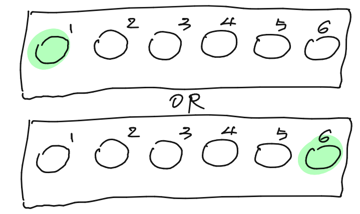


Example: Throw a fair die once.

Mutually exclusive

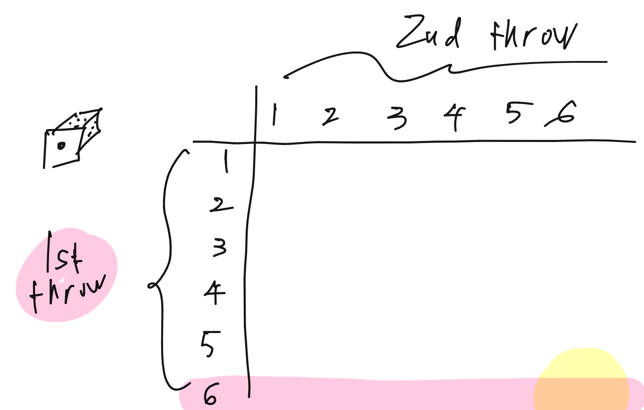


$P(\text{throw a } 1 \text{ and a } 6) = 0$



Example: Throw a fair die twice

Independence



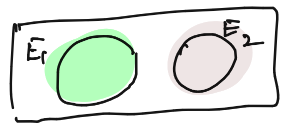
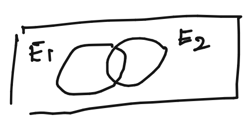
$$P(\text{throw a } 6 \text{ on 2nd throw} \mid \text{got a } 6 \text{ on 1st throw})$$

$$= \frac{P(\text{threw } 2 \text{ } 6\text{'s})}{P(\text{got a } 6 \text{ on 1st throw})}$$

$$= \frac{1}{6} = P(\text{throw a } 6 \text{ on 2nd throw})$$

So the 2 events are independent

FAQ 2. When do I add & multiply?

			When?
Addition Rule	P(at least 1 of 2 events occurs)	P(Event 1) + P(Event 2) 	mutually exclusive
Multiplication Rule	P(both events occur)	P(Event 1) × P(Event 2)	independent
		P(Event 1) × P(Event 2 | Event 1) 	dependent

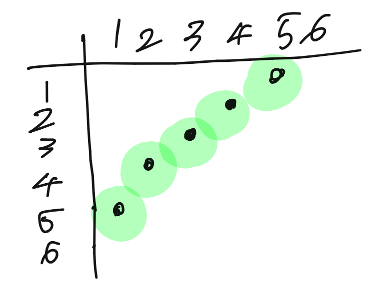
Simulation :)

the easy, quick way to estimate chance of event

Example: 2 fair dice are thrown. What is the chance of getting a total of 6?

1. Exact

$P(\text{total of } 6) = \frac{5}{36} \doteq 0.139$



2. Estimate

```
set.seed(1)
totals = sample(1:6, 1000, rep = T)
       + sample(1:6, 1000, rep = T)
table(totals)
```

R Output

2	3	4	5	6	7	8	9	10	11	12
28	50	79	112	148	167	150	105	82	58	21

$\frac{148}{1000}$

Binomial Model

1. Suppose we have n distinct objects in a row.

$$\underbrace{A\ B\ C\ \_\ \_\ \_\ \_}$$

The number of ways of arranging the objects is

$$n! = n(n-1)(n-2)\cdots(3)(2)(1) \qquad 0! = 1$$

2. Suppose we have n objects of 2 types, in a row.

$$\underbrace{A\ A\ B\ A\ \_\ \_\ \_}_{n} \qquad \begin{array}{l} A \times x \\ B \times n-x \end{array}$$

The number of ways of arranging the objects is

$$= \frac{n!}{x!\,(n-x)!}$$

$$= \binom{n}{x} \quad \text{"n choose x"}$$

$$\binom{n}{0} = 1$$

3. Suppose we have $\overset{\text{fixed}}{n}$ $\underbrace{\text{independent binary trials}}$

where $P(\text{Event}) = \underset{\text{fixed.}}{p}$.

either Event or $\overline{\text{Event}}$ occurs each time

$$P\left(\begin{array}{l}\text{observing}\\ \text{exactly}\\ x \text{ events}\end{array}\right) = \binom{n}{x} p^{x} (1-p)^{n-x} \qquad 0 \le x \le n$$

Example: We throw a fair dice 5 times. What is the chance of getting 4 6's. $P(6) = 1/6$

Let X = the # ("number") of 6's in 5 throws of a fair dice.

$$\sim \text{Bin}\left(n = 5,\ p = \tfrac{1}{6}\right)$$

So

$$P(X = 4) = \binom{n}{x} p^{x} (1-p)^{n-x}$$

$$= \binom{5}{4} \left(\frac{1}{6}\right)^{4} \left(\frac{5}{6}\right)^{1}$$

$$= \frac{5!}{4!\,1!} \left(\frac{1}{6}\right)^{4} \left(\frac{5}{6}\right)^{1}$$

$$\doteqdot 0.003$$

Quick way: dbinom$\left(4, 5, \frac{1}{6}\right)$