Probability Theory

Dr. Wan Nor Arifin

Biostatistics and Research Methodology Unit
Universiti Sains Malaysia

Last update: Oct 25, 2023

Outlines

- Introduction
- Basic Concepts
- Counting Method
- Conditional Probability

Expected outcomes

- Understand the basic concepts in probability
- Able to calculate probability by counting method
- Understand the concept of conditional probability and able to apply the concept to calculate related probability

Introduction

Introduction

Probability is…

"the chance that a given event will occur" (Merriam-Webster, 2022)

"a branch of mathematics that studies the possible outcomes of given events together with the outcomes' relative likelihoods and distributions" (Weisstein, 2022)

Range: **Impossible** 0 → 1 **Certain**

Classification

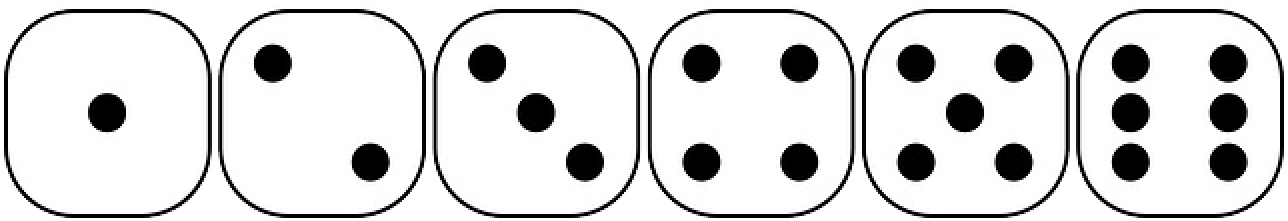
- Classical
- Frequentist
- Bayesian

Classification

- Classical
- Frequentist
- Bayesian

- Game of chance – flipping coin, rolling dice
- Finite number of possible outcomes

$$P(A) = \frac{N_A}{N}$$



Example: If a fair 6-sided die is rolled, probability of getting a 1 is

$$P(1) = \frac{1}{6}$$

Classification

- Classical
- Frequentist
- Bayesian

- Relative frequency of outcome after a number of repetition of random trials

$$P(x) \approx \frac{n_x}{n_t}$$

Example: Based on data collected over 200 years, it rained 15 out of 30 days in September. The probability of rain on 23 Sept 2022 is

$$P(\text{Rain on September 23}) = \frac{15}{30} = \frac{1}{2}$$

Classification

- Classical
- Frequentist
- Bayesian

- 1763 Thomas Bayes – Bayes' Theorem
- Updates prior knowledge (probability) in light of new data
- Will be introduced formally later in this lecture

Basic Concepts

Terms

- Experiment
- Sample Space
- Event
- Union
- Intersection
- Complement
- Disjoint

Terms

- Experiment

A situation for which the outcomes occur randomly

- Sample Space

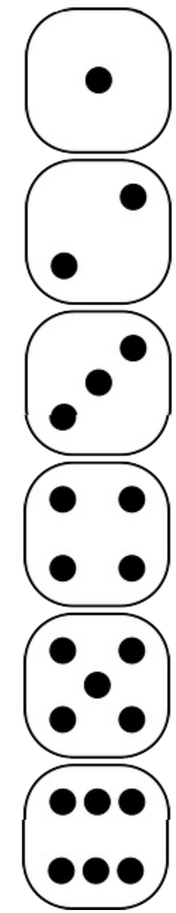
List of all possible outcomes of an experiment, Ω

- Event

A subset of the sample space

Sample space for a fair 6-sided die,

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$



Event A, odd numbers for a fair 6-sided die,

$$A = \{1,3,5\}$$

Terms

When either A or B or both occurs

$$A = \{1,2,3\}$$
$$B = \{3,4,5\}$$

$$A \cup B =$$

$$\{1,2,3,\cancel{3},4,5\} = \{1,2,3,4,5\}$$

- Union
- Intersection
- Complement
- Disjoint

Terms

- Union
- Intersection
- Complement
- Disjoint

When both A and B occurs

$$A = \{1,2,3\}$$
$$B = \{3,4,5\}$$

$$A \cap B =$$

$$\{\cancel{1,2,}3,\cancel{3},\cancel{4,5}\} = \{3\}$$

Terms

- Union
- Intersection
- Complement
- Disjoint

When A does NOT occur

$$\Omega = \{1,2,3,4,5,6\}$$
$$A = \{1,2,3\}$$

$$A^C =$$

$$\{\cancel{1,2,3,}4,5,6\} = \{4,5,6\}$$

Terms

- Union
- Intersection
- Complement
- Disjoint

When two events have no shared elements

$$A = \{1,2,3\}$$
$$C = \{4,5,6\}$$

$$A \cap B = \emptyset$$

$\emptyset$ is empty set

Properties

1. $P(A^c) = 1 - P(A)$

2. $P(\emptyset) = 0$

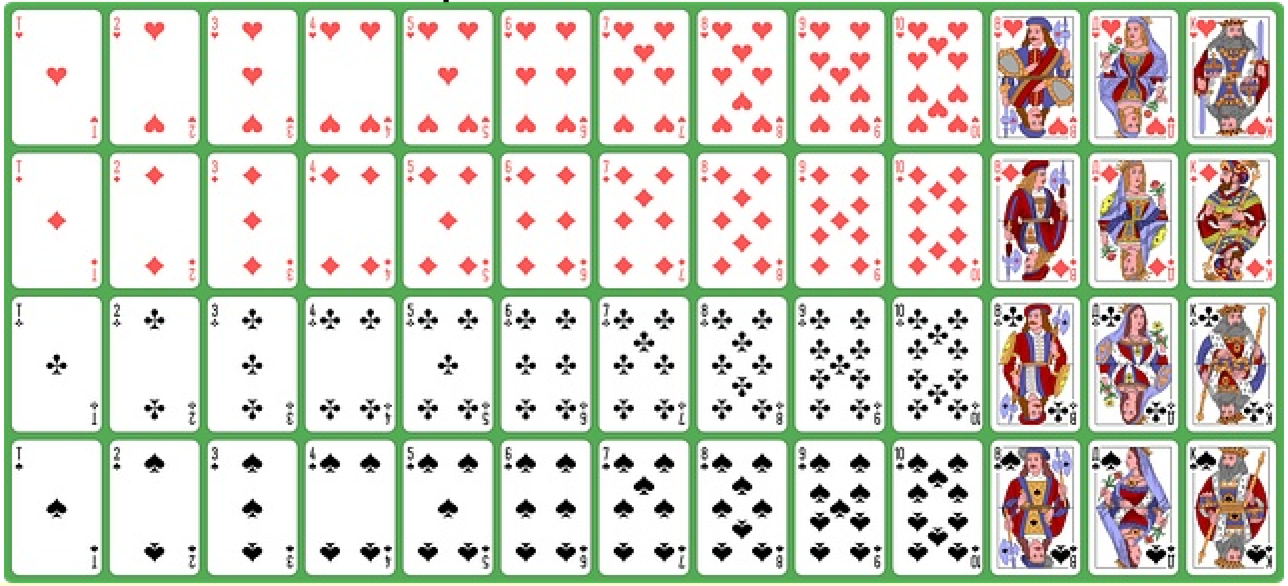
3. If $A \subset B$, then $P(A) \leqslant P(B)$

4. Addition Law $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Counting Methods

Multiplication Principle

Basic: If one experiment has *m* outcomes and another experiment has *n* outcomes, then there are *mn* possible outcomes for the two experiments.



Example: Playing cards have 13 face values (outcomes) per suit and 4 suits (experiments). Thus

$$13 \times 4 = 52 \text{ face values}$$

Multiplication Principle

Extended: If there are p experiments and the 1st has n_1 possible outcomes, the 2nd n_2, …, and the pth n_p possible outcomes, then there are a total of

$$n_1 \times n_2 \times \ldots \times n_p$$

possible outcomes for the p experiments.

Example: A fair 10 cent coin is thrown 4 times, each with two possible outcomes (hibiscus, congkak), thus

$$2 \times 2 \times 2 \times 2 = 2^4 = 16$$ possible outcomes

{HHHH, CCCC, HHHC, ...}

Permutation

For a set of size *n* and a sample of size *r*, the number of different ordered samples **without** replacement:

$$^{n}P_{r} = \frac{n!}{(n-r)!} = n(n-1)(n-2)\ldots(n-r+1)$$

Example: If the same number cannot appear twice, how many different ways to arrange number 0 – 9 to form a 3 digits number sequence?

Sample size, *r* = 3; Number of elements, *n* = 10

$$^{n}P_{r} = \frac{n!}{(n-r)!} = \frac{10!}{7!} = 10.9.8 = 720 \text{ ways}$$

10	9	8

Permutation

For a set of size n and a sample of size r, there are

$$n^r$$

different ordered samples **with** replacement

Example: How many different ways to arrange number 0 – 9 to form a 3 digits number sequence?

Sample size, r = 3; Number of elements, n = 10

$$n^r = 10^3 = 1000 \text{ ways}$$

Combination

The number of unordered samples of *r* objects from *n* objects **without** replacement:

Binomial coefficient

$$^{n}C_{r}=\binom{n}{r}=\frac{n(n-1)(n-2)\ldots(n-r+1)}{r!}=\frac{n!}{(n-r)!}\times\frac{1}{(r)!}=\frac{n!}{(n-r)!\,r!}$$

Example: If the same color cannot appear again, how many combinations of 2 colors out of 3 colors are possible?

$r = 2$; $n = 3$

$$^{n}C_{r}=\binom{n}{r}=\frac{n!}{(n-r)!\,r!}=\frac{3!}{(3-2)!\,2!}=\frac{3.2!}{1!\,2!}=\frac{3}{1}=3$$

RG
RB
GB

Combination

The number of unordered samples of *r* objects from *n* objects **with** replacement:

$$\frac{(r+n-1)!}{r!(n-1)!}$$

Example: How many combinations of 2 colors out of 3 colors are possible?

$r = 2; n = 3$

$$\frac{(r+n-1)!}{r!(n-1)!}=\frac{(2+3-1)!}{2!(3-1)!}=\frac{4!}{2!2!}=\frac{4.3}{2!}=6$$

RR	RG
GG	RB
BB	GB

Combination

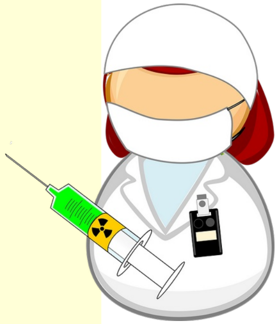
The number of ways that n objects can be grouped into r classes with n_i in the ith class:

Multinomial coefficient

$$\binom{n}{n_1 n_2 \ldots n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}$$

Example: A group of 7 eligible subjects in a clinical trial is allocated into 3 groups with size of 3, 2 and 2. How many ways the allocation could be done?

$$\binom{7}{3\,2\,2} = \frac{7!}{3!2!2!} = \frac{7.6.5.4.3!}{3!2!2!} = \frac{7.6.5.4}{2.1.2.1} = 210$$



Marginal, Joint and Conditional Probabilities

Marginal Probability

- Probability when the numerator is the marginal total of a table (subset)
- Probability of event A,

$$P(A)=\frac{n_A}{n_A+n_{A^C}}$$

Table in count

New test	ELISA		Total
	D+	D-	
T+	30	15	45
T-	5	50	55
Total	35	65	100

$P(T+)?$

$P(D-)?$

Table in probability

New test	ELISA		Marginal probability
	D+	D-	
T+	.3000	.1500	.4500
T-	.0500	.5000	.5500
Marginal probability	.3500	.6500	1.000

Joint Probability

- Probability when the numerator is the <u>joint</u> count, i.e. for A & B, when both occurs
- Intersection between events
- Joint probability of A and B,

$$P(A \cap B) = P(A, B) = \frac{n_{A,B}}{N}$$

Table in count

New test	ELISA		Total
	D+	D-	
T+	30	15	45
T-	5	50	55
Total	35	65	100

$P(T+, D+)?$

$P(T-, D-)?$

Table in probability

New test	ELISA		Marginal probability
	D+	D-	
T+	.3000	.1500	.4500
T-	.0500	.5000	.5500
Marginal probability	.3500	.6500	1.000

Conditional Probability

- Probability calculated with a subset of the sample space as denominator
- Probability of A given B,

$$P(A|B)=\frac{P(A\cap B)}{P(B)}\text{, if } P(B)\neq 0$$

Table in count

New test	ELISA		Total
	D+	D-	
T+	30	15	45
T-	5	50	55
Total	35	65	100

$$P(D+|T+)?$$

$$P(D-|T-)?$$

Table in probability

New test	ELISA		Total
	D+	D-	
T+	.3000	.1500	.4500
T-	.0500	.5000	.5500
Total	.3500	.6500	1.000

$$P(D+|T+)$$

$$=\frac{P(D+\cap T+)}{P(T+)}$$

$$P(D-|T-)=\frac{P(T-\cap D-)}{P(D-)}$$

Probability Laws

Multiplication Law

- Calculate joint probability by,

$$P(A \cap B) = P(A|B)\,P(B)$$

Multiplication Law

New test	ELISA		Marginal probability
	D+	D-	
T+	.3000	.1500	.4500
T-	.0500	.5000	.5500
Marginal probability	.3500	.6500	1.000

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

$$P(D+ \cap T+) = P(D+ \mid T+)P(T+) = .6667 \times ?$$

Law of Total Probability

- For disjoint events B_1, B_2, ..., B_n with $P(B_i) > 0$ for all i, then

$$P(A)=\sum_{i=1}^{n} P(A\cap B_i)=\sum_{i=1}^{n} P(A|B_i)P(B_i)$$

Table in probability

New test	ELISA		Marginal probability
	D+	D-	
T+	.6667	.3333	.4500
T-	.0909	.9091	.5500
*	*	*	*

$$P(D+) = \sum P(D+|T_i)\,P(T_i) =$$
$$P(D+|T+)\,P(T+) + P(D+|T-)\,P(T-) = ?$$

Table in probability

New test	ELISA		Marginal probability
	D+	D-	
T+	.6667	.3333	.4500
T-	.0909	.9091	.5500
*	*	*	*

$$P(D\text{-}) = \sum P(D\text{-}|T_i)P(T_i) = ?$$

Bayes' Rule

- For disjoint events $B_1, B_2, \ldots, B_n$ with $P(B_i) > 0$ for all i, then,

$$P(B_j \mid A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{n} P(A \mid B_i)P(B_i)}$$

$$\text{Posterior probability} = \frac{\text{Prior probability} \times \text{Likelihood}}{\text{Marginal probability}}$$

Table in probability

New test	ELISA		*
	D+	D-	
T+	.8571	.2308	*
T-	.1429	.7692	*
Marginal probability	.3500	.6500	*

$$P(B_j \mid A) = \frac{P(A \cap B_j)}{P(A)}$$

Prior expert knowledge

$$P(D+) = .85$$
$$P(D-) = .15$$

$$P(D+ \mid T+) = \frac{P(T+ \cap D+)}{P(T+)}$$

Use multiplication law

Use law of total probability

Table in probability

New test	ELISA		*
	D+	D-	
T+	.8571	.2308	*
T-	.1429	.7692	*
Marginal probability	.3500	.6500	*

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)}$$

Prior expert knowledge

$$P(D+) = .85$$
$$P(D-) = .15$$

$$P(D+|T+) = \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+) + P(T+|D-)P(D-)} = ?$$

Table in probability

New test	ELISA		*
	D+	D-	
T+	.8571	.2308	*
T-	.1429	.7692	*
Marginal probability	.3500	.6500	*

$$P(B_j|A)=\frac{P(A\cap B_j)}{P(A)}$$

Prior expert knowledge

$$P(D+)=.85$$
$$P(D-)=.15$$

$$P(D+|T-)=\frac{P(T-\cap D+)}{P(T-)}=\frac{P(T-|D+)P(D+)}{P(T-|D+)P(D+)+P(T-|D-)P(D-)}=?$$