Introduction to Probability

B39AX — Fall 2023

Heriot-Watt University

Random experiment

- Its output cannot be surely predicted in advance
- If we repeat it a large number of times, we observe a "regularity" in the average output

Examples:

- Toss a coin
- Play the lottery
- Predict the weather
- Detect a transmitted signal

Probability triple

$(\Omega, \mathcal{F}, \mathbb{P})$ - probability triple

Ω: **sample space**; set of all possible outcomes of an experiment

$\mathcal{F}$: **set of events**; an <u>event</u> is a subset of Ω, and is a "property" that holds or not after an experiment

$\mathbb{P}$: **probability measure**; it is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$; $\mathbb{P}(A)$ is the probability that event A will occur
(other common notation: $\mathbf{P}(A)$, $\mathbf{P}\{A\}$, $\mathbb{P}\{A\}$, $P(A)$, etc)

Examples

Countable Ω: $\left(\mathcal{F} = 2^{\Omega}\right)$

- Toss a coin once: $\Omega = \{H, T\}$, $\mathcal{F} = \big\{\emptyset, \{H\}, \{T\}, \{H, T\}\big\}$
- Toss a coin twice
- Select a student at random, out of 50

Uncountable Ω:

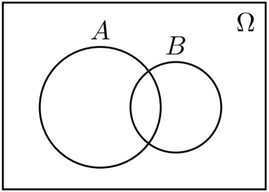
- Lifetime of a light bulb: $\Omega = [0, +\infty) = \mathbb{R}_+$
- Randomly select a number between 0 and 1: $\Omega = [0, 1]$

Uncountable case: 2^{Ω} too large to assign a probability to each event in a meaningful way. Requires *measure theory* (not on this course).

Events as sets

$A, B \in \mathcal{F}$ can be visualized as subsets of Ω:

- A^c: complement of A
- $A \cup B$: A or B
- $A \cap B$: A and B
- Ω: sure event
- $\emptyset$: impossible event
- $\omega \in \Omega$: elementary event



De Morgan laws: $\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c, \qquad \left(\bigcap_i A_i\right)^c = \bigcup_i A_i^c$

Example

Roll a die. $\Omega = \{1,2,3,4,5,6\}$

Consider the events

$$A = \{2,4,6\}$$

$$B = \{3,4,5\}$$

And compute

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

$$B^c = \{1,2,6\}$$

$$A \cap B = \{4\}$$

$$A \cup B = \{2,3,4,5,6\}$$

$$(A \cup B)^c = \{1\} = A^c \cap B^c$$

Axioms of probability

*advanced material

$(\Omega, \mathcal{F}, \mathbb{P})$

Assume $\mathcal{F}$ is a $\underline{\sigma\text{-algebra}}$, i.e.,

- $\Omega \in \mathcal{F}$
- $A \in \mathcal{F} \quad \Longrightarrow \quad A^c \in \mathcal{F}$
- $A_1,\, A_2,\, \ldots$: countable sequence of events in $\mathcal{F} \quad \Longrightarrow \quad \bigcup_i A_i \in \mathcal{F}$

Axioms of probability

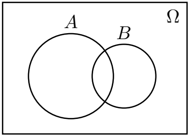
$(\Omega, \mathcal{F}, \mathbb{P})$

The probability measure $\mathbb{P} : \mathcal{F} \to [0,1]$ satisfies 3 axioms:

- **Nonnegativity**: $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$
- **Normalization**: $\mathbb{P}(\Omega) = 1$
- **Countable additivity**: if A_1, A_2, $\ldots$ is a countable sequence of pairwise disjoint events in $\mathcal{F}$ $\big(A_i \cap A_j = \emptyset$, for all $i \neq j\big)$, then

$$\mathbb{P}\Big(\bigcup_i A_i\Big) = \sum_i \mathbb{P}(A_i)$$

Consequences of the axioms



- $A \cap B = \emptyset \quad \Longrightarrow \quad \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
  (subcase of countable additivity)

- $\mathbb{P}(\emptyset) = 0$

- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

- $\mathbb{P}\big(A^c\big) = 1 - \mathbb{P}(A)$

- $\mathbb{P}\big(A \cap B^c\big) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$

- $A \subseteq B \quad \Longrightarrow \quad \mathbb{P}(A) \leq \mathbb{P}(B)$

- $\ldots$

Consequences of the axioms

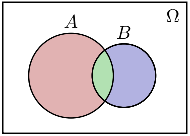
Proof of $\mathbb{P}(\emptyset) = 0$

Let A be any event. Then,

- $A \cup \emptyset = A \qquad \Longrightarrow \qquad \mathbb{P}(A \cup \emptyset) = \mathbb{P}(A)$
- $A \cap \emptyset = \emptyset$, which means that A and $\emptyset$ are disjoint. By the countable additivity, then $\mathbb{P}(A \cup \emptyset) = \mathbb{P}(A) + \mathbb{P}(\emptyset)$.
- From the two points above, $\mathbb{P}(A) = \mathbb{P}(A \cup \emptyset) = \mathbb{P}(A) + \mathbb{P}(\emptyset)$
- Cancelling terms gives $\mathbb{P}(\emptyset) = 0$ $\square$

Consequences of the axioms

Proof of $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$



A, B, and $A \cup B$ can be decomposed into disjoint sets as:

$$A = \big(A \cap B^c\big) \cup \big(A \cap B\big) \qquad B = \big(A \cap B\big) \cup \big(A^c \cap B\big)$$

$$A \cup B = \big(A \cap B^c\big) \cup \big(A \cap B\big) \cup \big(A^c \cap B\big)$$

By countable additivity, we have

$$\begin{aligned}
\mathbb{P}\big(A \cup B\big) &= \mathbb{P}\big(A \cap B^c\big) + \mathbb{P}\big(A \cap B\big) + \mathbb{P}\big(A^c \cap B\big) \\
&= \Big(\mathbb{P}(A) - \mathbb{P}\big(A \cap B\big)\Big) + \mathbb{P}\big(A \cap B\big) + \Big(\mathbb{P}(B) - \mathbb{P}\big(A \cap B\big)\Big) \\
&= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}\big(A \cap B\big) \qquad \square
\end{aligned}$$

How to construct a probability measure?

<u>Countable Ω</u>. Consider the toss of a coin.

$$\Omega = \{H, T\} \qquad \mathcal{F} = \Big\{\emptyset,\, \{H\}, \{T\},\, \{H, T\}\Big\} \qquad \mathbb{P} = ?$$

- Normalization implies $\mathbb{P}(\Omega) = \mathbb{P}\big(\{H, T\}\big) = \mathbb{P}\big(\{H\}\big) + \mathbb{P}\big(\{T\}\big) = 1$
- If the coin is fair, it is reasonable to assume $\mathbb{P}(H) = \mathbb{P}(T)$
- This gives $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$

For countable Ω: only need to specify probability of the elementary sets

Normalization + assumptions about the problem $\rightarrow$ $\mathbb{P}(\omega_i)\,,\quad \omega_i \in \Omega$

Exercise

You want to park your car in a parking lot with a given number of free spaces. You select your spot randomly such that the probability of selecting the ith free space is half the probability of selecting the $(i-1)$th free space, $i = 2, 3, \ldots$

- If there are 4 free spaces, what is the probability that you take the 3rd free space?
- What if there is an infinite number of free spaces?

Answers: $\frac{2}{15}$ and $\frac{2}{16}$

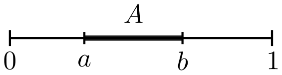
How to construct a probability measure?

*advanced concept

Uncountable (continuous) case. Select a number in $[0,\,1]$ at random.

$\Omega = [0,1]$ $\qquad$ $\mathcal{F} =$ Borel σ-algebra: $\sigma(\{(-\infty, a] \,:\, a \in \mathbb{Q}\})^\star$ $\qquad$ $\mathbb{P} = ?$

$\mathbb{P}(A)$ is the "volume" of A relative to Ω



$$\mathbb{P}\big([a,b]\big) = \mathbb{P}\big([a,b)\big) = \mathbb{P}\big((a,b]\big) = \mathbb{P}\big((a,b)\big) = \text{volume of } [a,b] = b-a$$

Note: $\mathbb{P}(\omega) = 0$ for all $\omega \in \Omega$

Example: $\Omega = [0,10]$, $\quad A =$ "Select a number between 2 and 5"

$$\mathbb{P}(A) = \tfrac{5-2}{10-0} = \tfrac{3}{10}$$

Frequentist interpretation of probability

Let $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ be a probability triple with a countable Ω.

Consider an event $A \in \mathcal{F}$, and repeat the experiment T times.

Relative frequency:

$$f_T(A) = \frac{\text{\# times A occurred}}{T}$$

Intuitively,

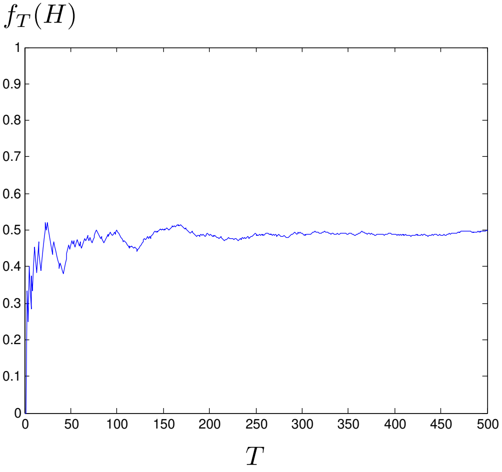
$$\mathbb{P}(A) = \lim_{T \to \infty} f_T(A)$$

This interpretation is justified by the law of large numbers

(not on this course).

Frequentist interpretation of probability

Example: Toss a fair coin T times. Count the number of heads H.



Frequentist interpretation of probability

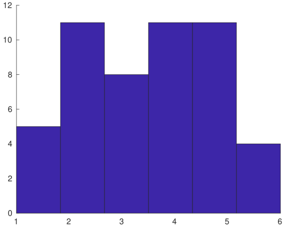
Matlab code to create the previous plot:

```
T = 500;                              % Number of experiments
successes = (randn(T,1)>0);  % Toss fair coin T times
% Plot relative frequency
op_matrix = tril(ones(T, T));
relative_frequency = (op_matrix*successes)./(1:T)';
plot(1:T, relative_frequency)
ylim([0,1])
```

Histograms and relative frequencies

Suppose we tossed a 6-faced die 50 times.

The **histogram** $\boldsymbol{h} = (h_1,\, h_2,\, h_3,\, h_4,\, h_5,\, h_6)$ counts the # of times each outcome was observed.



$$h = (5,\, 11,\, 8,\, 11,\, 11,\, 4)$$

The **relative frequencies** are in the normalized histogram: $\boldsymbol{f}_T = \frac{\boldsymbol{h}}{T}$

$$\boldsymbol{f}_T = \left(\frac{h_1}{50},\, \frac{h_2}{50},\, \frac{h_3}{50},\, \frac{h_4}{50},\, \frac{h_5}{50},\, \frac{h_6}{50}\right) = (0.10,\, 0.22,\, 0.16,\, 0.22,\, 0.22,\, 0.08)$$