"In mathematics you don't understand things, you just get used to them."  
[ J. von Neumann ]

- 1 -

- EVENTS AND PROBABILITIES -

A random experiment $\Rightarrow$ Probability space

e.g. a fair die throw

- sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$
- event space $\mathcal{F} = \{\emptyset, \Omega, \{1, 3, 5\}, \{2, 4, 6\}\}$

- probability measure $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$

$\mathbb{P}(\{1, 3, 5\}) = \frac{1}{2} = \mathbb{P}(\{2, 4, 6\})$
$\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$

The set of all possible outcomes of an experiment is called the sample space (usually $\Omega$). Each member of $\Omega$, $w \in \Omega$, is called an elementary event.

E.g. if we toss a fair coin twice,

$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$

1st toss heads, 2nd toss tails

- 1 -
The **event space** (a.k.a. *σ*-field / *σ*-algebra of events) is a collection of subsets of Ω, \( \mathcal{F} = \{ A_i, i \in I \} \) which satisfies

(i) \( \emptyset \in \mathcal{F} \)

(ii) if \( A \in \mathcal{F} \), then \( A^c \in \mathcal{F} \)

(iii) if \( A_1, A_2, \ldots \in \mathcal{F} \), then \( \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \)

*σ*-field of events = "events which are interesting to us"

\( \emptyset \) = impossible event, \( \Omega \) = certain event

A, B two events:

\( A \cup B \) * either A or B occurs

\( A \cap B \) * both A and B occurs

\( A^c \) * A does not occur

\( A \setminus B = A \cap B^c \)

set minus * A occurs and B does not

\( A \triangle B = (A \setminus B) \cup (B \setminus A) \)

symmetric difference * exactly one of A, B occurs
Fact: If $A, B \in F$, then $A \cap B \in F$.

Proof: $(A \cap B)^c = A^c \cup B^c$ refers to $\in F$ by (iii) (by (ii))

$(A \cap B)^c \in F \Rightarrow A \cap B \in F$. \qed

Fact: If $A, B \in F$, then $A \setminus B$, $A \Delta B$, $B \setminus A \in F$.

Example: $\Omega = \{ (H, H), (H, T), (T, H), (T, T) \}$

$\mathcal{F} =$ all subsets of $\Omega = 2^\Omega$ (power set)

$\mathcal{F} = \{ \emptyset, \Omega, \{ (H, H), (H, T) \}, \{ (T, H), (T, T) \} \}

1^{st}$ toss $H$ \hspace{1cm} $1^{st}$ toss $T$

A function $P: \mathcal{F} \to \mathbb{R}$ is called a probability measure on $(\Omega, \mathcal{F})$ if

(i) $P(A) \geq 0 \quad \forall A \in \mathcal{F}$

(ii) $P(\Omega) = 1$ and $P(\emptyset) = 0$

(iii) if $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint (i.e. $A_i \cap A_j = \emptyset \ \forall i \neq j$), then
\[ P \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} P(A_i). \]

E.g. Let \( P(\{(H,H)\}) = \frac{1}{4}, \quad P(\{(H,T)\}) = \frac{1}{4}, \quad P(\{(T,H)\}) = \frac{1}{4}, \quad P(\{(T,T)\}) = \frac{1}{4}. \)

Then \( P(1^{st} \text{ toss } H) = P(\{(H,H), (H,T)\}) = P(\{(H,H)\}) + P(\{(H,T)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \)

E.g. Let \( \Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}, \quad \mathcal{F} = 2^\Omega \)

Then \( P(\{\omega_i\}) = \frac{1}{N}, \quad \text{that is} \)

\[ P(A) = \frac{|A|}{N}, \quad A \in \Omega \]

defines a probability measure (called the counting prob. measure on \( \Omega \) or the uniform prob. measure on \( \Omega \)).

\text{Def} \quad \text{A probability space is a triple } (\Omega, \mathcal{F}, P).
E.g. Toss a fair coin until you get tails

\[ \Omega = \{ T, HT, HHT, HHHT, \ldots \} \]

\[ \mathcal{F} = 2^{\Omega} \]

\[ P\left( \left\{ \underbrace{HH\ldots HT}_{n} \right\} \right) = \frac{1}{2^{n+1}}, \quad n \geq 0, \quad \text{check it sat. (i)-(iii)}! \]

**Properties**

Let \( A, B, A_1, A_2, \ldots \in \mathcal{F} \), then

1. if \( A_1, A_2, \ldots, A_n \) are disjoint, then

\[ P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i) \]

2. if \( A \subset B \), then

\[ P(B \setminus A) = P(B) - P(A), \quad P(A) \leq P(B) \]

3. \( P(A \cup B) = P(A) + P(B) - P(A \cap B) \)

4. \( P\left( \bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} P(A_i) \)

5. if \( A_1 \subset A_2 \subset \ldots \), then \( P(A_n) \xrightarrow{n \to \infty} P\left( \bigcup_{i=1}^{\infty} A_i \right) \)

if \( A_1 \supset A_2 \supset \ldots \), then \( P(A_n) \xrightarrow{n \to \infty} P\left( \bigcap_{i=1}^{\infty} A_i \right) \).
Proof  
(1) Define $A_{n+1} = \emptyset$, $A_{n+2} = \emptyset$, ... and use (iii)

(2) 
$B = A \cup (B \cap A)$

\[ P(B) = P(A) + P(B \cap A) \geq 0 \]

(3) 
$A = A \cup B \cup A \cap B$
$B = B \cap A \cup A \cap B$

\[ P(A) = P(A \cap B) + P(A \cap B) \]
\[ P(B) = P(B \cap A) + P(A \cap B) \]

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

\[ \text{disjoint} \]

$= P(A \cap B \cup A \cap B \cup B \cap A) + P(A \cap B)$

$= P(A \cup B) + P(A \cap B)$

(5) 
$B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_2$, ... disjoint

\[ P(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i), \; \text{so} \]

\[ P(\bigcup_{i=1}^{\infty} A_i) = \lim_{n \to \infty} \sum_{i=1}^{n} P(B_i) = \lim_{n \to \infty} \sum_{i=1}^{n} \left[ P(A_i) - P(A_{i-1}) \right]_{A_0 = \emptyset} \]

= $\lim_{n \to \infty} P(A_n)$  \[ \Box \]
Discrete sample spaces

Let \((\Omega, \mathcal{F}, P)\) be a prob. space with \(\Omega\) countable, 
\(\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}\). Then any \(\sigma\)-algebra on \(\Omega\) is of the form:

- take a countable partition

\[ \{B_i\}_{i=1}^{\infty} \quad (N \leq \infty) \]

of \(\Omega\), that is
\[ B_i \cap B_j = \emptyset, \quad i \neq j, \]
\[ \bigcup_{i=1}^{\infty} B_i = \Omega \]

- take \(\mathcal{F} = \left\{ \bigcup_{i \in I} B_i : I \subseteq \{1, 2, \ldots, N\} \right\}\)

that is, any \(A \in \mathcal{F}\) is of the form

\[ A = \bigcup_{i \in I} B_i \] for some subset \(I \subseteq \{1, 2, \ldots, N\}\)

- define \(p_i = P(B_i), \quad i = 1, \ldots, N\) \((p_i \geq 0, \sum p_i = 1)\)

\[ P(A) = \sum_{i \in I} p_i \]

Often: \(B_i = \{\omega_i\}\) (singletons), \(\mathcal{F} = 2^\Omega\),

\[ P(A) = \sum_{i : \omega_i \in A} p_i \]
E.g. Uniform measure: if \( \Omega = \{1, \ldots, N\} \), \( B_i = \{i\} \), \( \mathcal{F} = 2^\Omega \),
\[ p_i = \mathbb{P}(\{i\}) = \frac{1}{N}, \]
\[ \mathbb{P}(A) = \sum_{i \in A} \frac{1}{N} = \frac{|A|}{N} = \frac{|A|}{|\Omega|}. \]

E.g. You are dealt 5 cards (from 52). What is the prob.
of getting a full house (one pair + three of a kind)?
\( \Omega = \{ \text{3 subsets of size 5 of } \{1, \ldots, 52\} \} \)
\[ |\Omega| = \binom{52}{5}, \quad \mathbb{P} = \text{uniform} \]
\( A = \text{full house} \), \[ |A| = \binom{13}{1} \binom{4}{2} \cdot \binom{12}{1} \binom{4}{3} \]
\[ \mathbb{P}(A) = \frac{13 \cdot 12 \cdot \binom{4}{2} \binom{4}{3}}{\binom{52}{5}}. \]

E.g. What is the probability that each player in a game
of bridge receives one ace?
\( \Omega = \{ \text{permutations of } \{1, 2, \ldots, 52\} \} \)
\[ |\Omega| = 52! = 52 \times 51 \times \ldots \times 2 \times 1 \]

---

Player 1 2 3 4
A: each player has one ace

\[ |A| = \frac{13^4}{4!} \cdot 48! \]

which ace \( \uparrow \)
remaining cards \( \uparrow \)

\[ P(A) = \frac{|A|}{52!} = \frac{13^4 \cdot 4! \cdot 48!}{52!} = 0.105... \]

E.g. In how many ways can we put \( n \) oranges into \( k \) labeled boxes?

\[
\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}
\]

\( n = 6 \)

\( k = 3 \)

\( \binom{000}{100} \)

can encode

\( \binom{010}{000} \)

\( \binom{001}{000} \)

the same as choosing positions for \( k-1 \) pencils

(gaps between boxes) among \( n+k-1 \) positions

Answer: \( \binom{n+k-1}{k-1} = \binom{n+k-1}{n} \)

E.g. A fair die is thrown 3 times. What is the probability of getting at least two outcomes the same?

\[ P(\text{at least two outcomes the same}) = 1 - P(\text{all outcomes different}) \]

\[ = 1 - \frac{6 \cdot 5 \cdot 4}{6 \cdot 6 \cdot 6} = \frac{16}{36} = \frac{4}{9} \]
Conditional probabilities

If we throw a die and somebody tells us that an even number is showing, then this information affects all our calculations of probabilities.

\[
\Omega \\
\begin{array}{c}
\text{Def}\text{ If } A, B \text{ are two events and } P(B) > 0, \\
\text{then the conditional probability of } A \text{ given } B \\
is \quad P(A \mid B) = \frac{P(A \cap B)}{P(B)}.
\end{array}
\]

Thm. The conditional probability defines a prob. measure.

\text{Proof} \quad \text{WTS } Q(\Omega) = P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} \text{ is a prob. measure.}

\begin{align*}
\cdot & \quad P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1, \text{ so } Q(\Omega) = 1 \\
\cdot & \quad A_1, A_2, \ldots \text{ disjoint events,} \\
Q(\bigcup_{i=1}^{n} A_i) = & \quad P\left(\left\{ \bigcup_{i=1}^{n} A_i \right\} \mid B\right) = \frac{P\left(\left\{ \bigcup_{i=1}^{n} A_i \right\} \cap B\right)}{P(B)} \\
& \quad = \frac{P\left(\bigcup_{i=1}^{n} A_i \cap B\right)}{P(B)} = \frac{\sum P(A_i \cap B)}{P(B)} \\
& \quad = \sum Q(A_i). \quad \square
\end{align*}

E.g. Take two cards from a standard deck of cards. What's the probability that the 1st card was an ace if the 2nd card was an ace?
Let $A = 1^{st}$ card an ace
$B = 2^{nd}$ card an ace

Want $\Pr(A \mid B)$, Easy $\Pr(B \mid A) = \frac{3}{51},$

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A \cap B)}{\Pr(A)} \cdot \frac{\Pr(A)}{\Pr(B)}$$

$$= \Pr(B \mid A) \cdot \frac{\Pr(A)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)}$$

$\Pr(A) = \frac{4}{52}, \quad \Pr(B) = \frac{4 \cdot 3 + 48 \cdot 4}{52 \cdot 51} = \frac{4}{52} \cdot \frac{52}{52} = \Pr(A)$

$$\Pr(A \mid B) = \Pr(B \mid A) = \frac{3}{51}. \quad \square$$

E.g. (cascade prob.) $\Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B),$ $\Pr(B \cap C) = \Pr(A \mid B \cap C) \cdot \Pr(B \cap C)$

$$= \Pr(A \mid B \cap C) \cdot \Pr(B \cap C) \cdot \Pr(B \mid C) \cdot \Pr(C)$$

(assuming $\Pr(B \cap C) > 0$ and $\Pr(C) > 0$).

Suppose there are $n-1$ white and 1 black balls in a box. We draw one by one until we get the black one.

What is the prob. that we draw 3 times ?

$A_i$ = $i^{th}$ ball white

$$\Pr(A_3 \cap A_2 \cap A_1) = \Pr(A_3 \cap A_2 \cap A_1) \cdot \Pr(A_2 \mid A_1) \cdot \Pr(A_1)$$

$$= \frac{1}{n-2} \cdot \frac{n-2}{n-1} \cdot \frac{n-1}{n} = \frac{1}{n}.$$
Theorem 1: If \( \{ B_1, B_2, \ldots \} \) is a partition of \( \Omega \) s.t. \( P(B_i) > 0 \) for every \( i \), then for any event \( A \),
\[
P(A) = \sum_i P(A \mid B_i) \cdot P(B_i)
\]
(This is the so-called partition theorem of total prob.)

Proof:
\[
P(A) = P(A \cap \Omega) = P(A \cap (B_1 \cup B_2 \cup \ldots))
\]
\[
= P(\bigcup_i A \cap B_i) = \sum_i P(A \cap B_i)
\]
\[
= \sum_i P(A \mid B_i) P(B_i). \quad \Box
\]

E.g., A biased coin shows heads with prob \( p \in (0, 1) \). Let \( u_n = P(\text{in } n \text{ tosses no pair of heads occurs successively}) \).

Find \( u_n \).

Let \( B_i = \frac{i-1}{H \ldots HT} (\text{first } i-1 \text{ tosses } H, \text{then } T), i \geq 1. \)

They form a partition,
\[
u_n = P(A_n) = \sum_{i=1}^{\infty} P(A_n \mid B_i) \cdot P(B_i)
\]
\[
= P(A_n \mid B_1) \cdot P(B_1) + P(A_n \mid B_2) \cdot P(B_2)
\]
\[
= \frac{HT}{n-1} \cdot \frac{1}{n-1} + \frac{H}{n-2} \cdot \frac{p}{1-p}. \quad \Box
\]
Independence

If \( P(A \mid B) = P(A) \) \( \iff \) B doesn't affect A or A is independent of B

\[
\frac{P(A \cap B)}{P(B)} = P(A)
\]

Def Events A, B are independent if \( P(A \cap B) = P(A) \cdot P(B) \).

(Dependent o/w) A family of events \( \{A_i; i \in I\} \) is called independent if for all finite subsets \( J \) of \( I \)

\[
P\left( \bigcap_{j \in J} A_j \right) = \prod_{j \in J} P(A_j)
\]

Eg. Events A, B, C are independent \( \iff \) and only if

\[
P(A \cap B) = P(A) \cdot P(B), P(B \cap C) = P(B) \cdot P(C), P(C \cap A) = P(C) \cdot P(A),
P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)
\]

All the equations are needed! Eg. roll a fair 4-sided die,

\( \Omega = \{1, 2, 3, 4\} \), \( P(\{i\}) = \frac{1}{4} \),

\( A = \{1, 2\}, B = \{1, 3\}, C = \{1, 4\} \) are not indep.

\[
P(A \cap B \cap C) = P(\{1\}) = \frac{1}{4}
P(A) \cdot P(B) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}
\]

But they are pairwise indep., \( P(A \cap B) = P(A) \cdot P(B) \), etc.
Subtleties

1) Why $P$ countably additive?

Consider the unit interval $[0, 1]$ and choose a point in it uniformly at random.

$$P(\text{point } \in [a, b]) = \frac{b-a}{1}$$

(Formally,

$\Omega = [0, 1]$

$\mathcal{F} = \text{Borel } \sigma\text{-field of } [0, 1] = \text{smallest } \sigma\text{-field containing all intervals}$

$P = \text{Lebesgue measure on } [0, 1]$)

Then

$$P(\{x\}) = P(\bigcap_{k \geq 1} [x - \frac{1}{k}, x + \frac{1}{k}])$$

$$= \lim_{k \to \infty} P([x - \frac{1}{k}, x + \frac{1}{k}]) = \lim_{k \to \infty} \frac{2}{k} = 0.$$

If $P$ was fully additive

$$P(\Omega) = P\left(\bigcup_{x \in [0, 1]} \{x\}\right) = \sum_{x \in [0, 1]} P(\{x\}) = 0.$$

In the def, we require $P$ to be countably additive.

2) Bertrand's paradox

Consider an equilateral triangle inscribed in a circle. A chord of the circle is chosen at random. What is the prob. that the chord is longer than a side of the triangle?
a) random endpoints model \( p = \frac{1}{3} \)
   - choose A
   - choose B \( \Rightarrow \) chord \( \widehat{AB} \)

b) random radius model \( p = \frac{1}{2} \)
   - choose a radius \( \Rightarrow \) chord \( \widehat{AB} \)
   - choose a point on it \( \perp \) radius, passing through the point

c) random midpoint model
   - choose a random point in the disc \( \Rightarrow \) chord \( \widehat{AB} \)
   - unique chord whose midpoint is the point

\[
p = \frac{\pi \left( \frac{1}{2} \right)^2}{\pi \cdot 1^2} = \frac{1}{4}
\]

We got 3 different results because we chose 3 different models (prob. spaces). Prob. theory does not tell us which model to choose but how to calculate probabilities within a certain model.