

LESSON 4

Probability: Sample Space, Events, Classical And Axiomatic Definitions of Probability, Properties of Probability Measures

What is probability?

It is a mathematical theory of "probabilistic" uncertainty.

Remark: There are other types of uncertainty that

Cannot be modeled using probability.

Example: Make a random number generator select $x \in [0,1]$:

★

What is the probability of selecting ★?

Is it probable to select ☆ ?

Is it possible to select ☆ ?

probable $\longrightarrow$ possible

possible $\not\longrightarrow$ probable

because $\mathbb{P}$(Something possible) may be 0.

Possibility is modeled by

It is NOT AN EASY THEORY.

Example: You have three decks of 52 cards

blue	blue	red
blue	red	red

You pick a card and observe that one side of the card is blue. What is the probability that the other side is also blue?

$=1/2$

$<1/2$

$>1/2$

Classical Definition of probability:

Ratio of favorable outcomes and the total number of possible outcomes, if all outcomes are equally likely

$$\mathbb{P}(A) = \frac{n_A}{N}$$

Example: we have Five balls, one is red, the rest are blue.

$\mathbb{P}(\text{blue}) = $ ______

Relative Frequency Definition:

$$\mathbb{P}(A) = \lim_{N \to \infty} \frac{n_A}{N}$$

Example: We ~~fdsp~~ flip a coin 10,000,000 times.

$\mathbb{P}(\text{Heads}) =$ ______

Axiomatic Definition:

A countably additive function defined on the set of events with a range $[0,1]$.

Sample Space:

Set of all possible outcomes of an experiment is called the sample Space, Ω

Example: Coin toss

(a) Interested in knowing whether the toss produces a head or a tail $\Omega =$

(b) Interested in the number of tumbles before the coin hits the ground $\Omega =$

- Interested in the speed with which the coin hits the ground $\Omega =$

Thus,
For the same experiment, Ω can be different, depending on what the observer is interested in.

Example: (a) Roll of a die

$\Omega =$

(b) Tossing two coins

$\Omega = \{ \quad \} \times \{ \quad \} =$

Example: What is Ω when a coin is tossed until heads appear?

$\Omega = \{ \qquad \}$

Example: Life expectancy of a random person

$\Omega =$

Def (Informal): An event is a subset of the sample space Ω, to which probabilities will be assigned

Def (Informal): An event is said to have occured if the outcome of the experiment (ω) belongs to it.

Example: Coin toss $\Omega = \{H, T\}$

$E =$

Roll of a die: $\Omega =$

$E = \{2, 4, 6\} \doteq \boxed{}$

Life expectancy $\Omega = [0, 120]$

$E = [50, 120] \doteq \boxed{}$

Assignment of probability to events:

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

$\mathbb{P}(H) =$ $\mathbb{P}(T) =$

$\mathbb{P}(\Omega) =$

Roll of a die

$\mathbb{P}(A) = \qquad \forall A \subseteq \Omega$

Question: Can we always assign probabilities to all subsets of Ω?

- Yes,

- No,

Relation between set theory and probability

We use the algebra of sets to model events

A or B occured

A and B occured

A occured, but B didn't occur

To model events, we need some structure that includes all possible events

σ- field of events

Def: A subset of 2^{Ω}, $\mathcal{F} \subseteq 2^{\Omega}$

is called a σ- field (or σ-algebra)

iff

(1)

(2)

(3)

This structure allows us to

model events like

$A^c \cup (B^c \cap C^c)$

Example: The smallest σ-field for any Ω

$\mathcal{F} =$

Example: $A \subseteq \Omega$ $\mathcal{F} =$

Example: is 2^{Ω} a σ-field?

Probability Measure (Kolmogorov)

Assume that Ω is a sample Space and $\mathcal{F}$ is a σ-field of subsets of Ω, i.e. $\mathcal{F} \subseteq 2^{\Omega}$.

A probability measure on $(\Omega, \mathcal{F})$ is a function

$$\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$$

s.t.

(1) Probability of Ω

(2) Probability of countable unions of disjoint events (Countable additivity property)

Alternatively $\mathbb{P}: \mathcal{F} \to \mathbb{R}$

(1) Probability of $\emptyset$

(2) Probability of Ω

(3) Countable additivity

We call $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space or a probability model

Aside: $\mathbb{P}$ is a special case of an additive measure:

$$\mu : \mathcal{F} \longrightarrow [0, +\infty)$$

$$\mu(\emptyset) = 0$$

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \boxed{}$$

when A_i's are mutually exclusive

Examples:

Coin toss: $\Omega =$

$\mathcal{F} =$

$\mathbb{P}(\emptyset) =$ $\mathbb{P}(\{H\}) =$

$\mathbb{P}(\{T\}) =$ $\mathbb{P}(\Omega) =$

Roll of a die

$\Omega =$

$\mathcal{F} =$

$\mathbb{P}(\{i\}) = p_i$, $\sum_{i=1}^{6} p_i = 1$

$\mathbb{P}(A) =$

$\forall A \subseteq \Omega$

Lemma

(1) Additivity for a finite number of events

$A_1, \dots, A_n \in \mathcal{F}$, $A_i \cap A_j = \emptyset$
$\forall\, i \neq j$

$\Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \boxed{}$

(2) Probability of complement

$A \in \mathcal{F} \Rightarrow \mathbb{P}(A^c) =$

(3) Preservation of order

$A \subseteq B \Rightarrow \mathbb{P}(A) \quad \mathbb{P}(B)$

(4) Boole's Inequality (Union Bound)

$A_1, A_2, \cdots \in \mathcal{F}$

$$\Rightarrow \quad \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \qquad \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

(5) Inclusion-Exclusion

$A_1, A_2 \in \mathcal{F}$

$\mathbb{P}(A_1 \cup A_2) =$

(6) Inclusion-Exclusion

(General form)

$A_1, A_2, \cdots, A_n \in \mathcal{F}$

$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) =$

for $n=3$

$\mathbb{P}(A_1 \cup A_2 \cup A_3) =$

for $n=4$

$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup A_4) =$

Proof.

(1) Finite sequence of disjoint events $A_1, \ldots, A_n \in \mathcal{F}$

To use axioms of probability, we create an infinite sequence of disjoint events

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

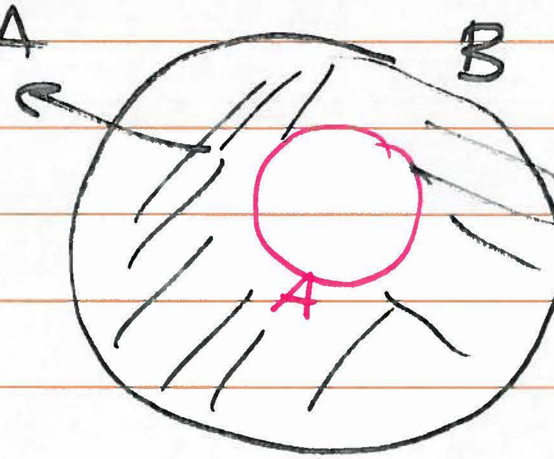
Note that

$A \cup A^c =$ $A \cap A^c =$

$\mathbb{P}(A \cup A^c) =$

(3) $A \subseteq B$ $\mathbb{P}(A) \leq \mathbb{P}(B)$

$B \setminus A$



$B =$

Important Corollary:

$$A \subseteq B \Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$$

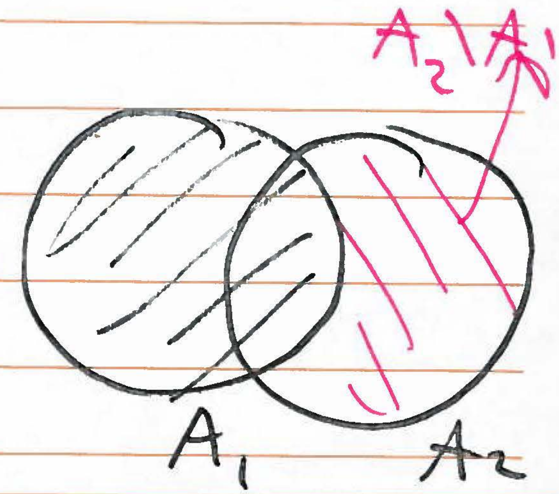
(4) Boole's Inequality

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

(5) Inclusion Exclusion

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

$A_1 \cup A_2 =$ $\cup$



How can we prove the general case?

By mathematical induction

Let us review the concept of mathematical induction

Principle of Mathematical Induction

Theorem: Let $p(n)$ be a predicate, $n \in \mathbb{N}$. Assume that:

(a) $p(m_0)$ is true for some $m_0 \in \mathbb{N}$

(b) If $p(k)$ is true for some $k \geq m_0$ then $p(k+1)$ is also true

Then $p(n)$ is true for all $n \geq m_0$, $n \in \mathbb{N}$

Example: Prove that

$$S(n): 1+3+5+\cdots+(2n-1)=n^2 \quad \forall n \in \mathbb{N}$$

Induction on n

$n=1 \Rightarrow 1=1$ true ✓

Assume for $n = k$

$S(k): 1 + 3 + 5 + \cdots + (2k-1) = k^2$. is true

To use induction, we must show that $S(k+1)$ is also true, i.e.

$S(k+1)$:

Back to proving inclusion-exclusion

$n=2$ $\quad \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$

$- \mathbb{P}(A_1 \cap A_2)$

Assume that

$n=\ell$: $\mathbb{P}(A_1 \cup A_2 \cup \dots \cup A_\ell)$

$$= \sum_{i=1}^{\ell} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j)$$

$$+ \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots$$

$$+ (-1)^{\ell+1} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_\ell)$$

We must prove that for

$n = \ell+1$:

$n = l+1 \quad \mathbb{P}\left(\bigcup_{i=1}^{l+1} A_i\right) =$

(It is a good exercise in applying probability properties to solving problems)

Continuity of Probability Measure

Lemma: Assume that

$\forall i \in \mathbb{N},\ A_i \in \mathcal{F}$ (A_i is an event, or is $\mathcal{F}$-measurable) then

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{n \to \infty} \underbrace{\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)}_{p_n}$$

This lemma defines a sequence of numbers $p_n = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)$ and states that the limit of the sequence p_n is equivalent to the probability $\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right)$.

Aside: Real Sequences and their limits

A sequence $\{a_n\}_{n=1}^{\infty}$ is a function from $\mathbb{N}$ to another set, therefore a real sequence is $a_n : \mathbb{N} \rightarrow \mathbb{R}$

Def: limit of a sequence

We say a_n converges to L as $n \rightarrow \infty$ and write

$$\lim_{n \rightarrow \infty} a_n = L$$

iff:

Example: $\lim_{n \to \infty} 1/n = 0$

Back To Continuity of Probability:

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)$$

Proof:

Corollary :

If $A_i{}'^{\,i \in \mathbb{N}}$ is a sequence of nested events:

(a). If A_i is increasing, i.e.

$A_i \subseteq A_{i+1} \quad \forall i \in \mathbb{N}$, then

$$\boxed{\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{n \to \infty} \mathbb{P}(A_n)}$$

(b) If A_i is decreasing, i.e.

$A_i \supseteq A_{i+1} \quad \forall i \in \mathbb{N}$, then

$$\boxed{\mathbb{P}\left(\bigcap_{i \in \mathbb{N}} A_i\right) = \lim_{n \to \infty} \mathbb{P}(A_n)}$$

Proof

(a) From continuity of probability

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right)$$

$A_i \subseteq A_{i+1}$ $\quad \bigcup_{i=1}^{n} A_i = A_n$

$$\Rightarrow \mathbb{P}\Big(\underbrace{\bigcup_{i\in\mathbb{N}} A_i}_{\lim_{i\to\infty} A_i}\Big) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

$$\mathbb{P}\left(\lim_{i\to\infty} A_i\right) = \lim_{i\to\infty} \mathbb{P}(A_i)$$

$$\lim_{x\to a} f(x) = f\Big(\overbrace{\lim_{x\to a} x}^{a}\Big)$$ continuous function

(b) We use (a)

$A_i \supseteq A_{i+1} \Rightarrow \underbrace{A_i^c \subseteq A_{i+1}^c}$

by (a)

$$\mathbb{P}\Big(\underbrace{\bigcup_{i\in\mathbb{N}} A_i^c}_{\Updownarrow}\Big) = \lim_{n\to\infty} \underbrace{\mathbb{P}(A_n^c)}_{\Downarrow}$$

$$= \mathbb{P}\Big(\underbrace{\Big[\bigcap_{i\in\mathbb{N}} A_i\Big]^c}\Big) = \lim_{n\to\infty} \{1 - \mathbb{P}(A_n)\}$$

$$= 1 - \mathbb{P}\Big(\bigcap_{i\in\mathbb{N}} A_i\Big) = 1 - \lim_{n\to\infty} \mathbb{P}(A_n)$$

$$\Rightarrow \quad \mathbb{P}\Big(\bigcap_{i\in\mathbb{N}} A_i\Big) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

Why is it called continuity?

$$\lim_{i\to\infty} \mathbb{P}(A_i) = \mathbb{P}\Big(\lim_{i\to\infty} A_i\Big)$$

The lim operator passes probability, like continuous functions:

$$\lim_{x\to x_0} f(x) = f\Big(\lim_{x\to x_0} x\Big) = f(x_0)$$

Remark (Zero Probability and Almost Sure Events)

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

If $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$,

Can we conclude that $A = \emptyset$?

No. Example: Selecting $\{\frac{1}{2}\}$ from $[0,1]$.

A is called a zero probability event

If $A \in \mathcal{F}$, $P(A) = 1$,

Can we conclude that $A = \Omega$? No, Example $[0,1] \setminus \{1/2\}$

A is called an event that occurs almost surely (a.s.)

Almost sure event

$[0,1] \setminus \{x \in \mathbb{Q} \mid x \in [0,1]\}$

Conditional Probability

Def: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and the event $B \in \mathcal{F}$ be such that $\mathbb{P}(B) > 0$. For any event

$A \in \mathcal{F}$, the conditional probability that A occurs given B is denoted as $\mathbb{P}(A|B)$ and is defined by:

$$\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}$$

$$\Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A|B)\,\mathbb{P}(B)$$

Example: Two fair dice are thrown. Given that the first die shows 3, what is the probability that the sum exceeds 6?

Intuition: The second die must be $\boxed{4, 5, 6}$. Answer: $\boxed{3/6}$

Principled answer:

$\Omega = \{1, 2, -6\} \times \{1, 2, -, 6\}$

$\mathcal{F} = 2^{\Omega}$ $|\mathcal{F}| = |2^{\Omega}| = 2^{|\Omega|} = 2^{36}$

$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \quad \forall A \in \mathcal{F}$

So we defined the probability space $(\Omega, \mathcal{F}, \mathbb{P})$

Assume

A = the event that the total number exceeds 6

$A = \{(a,b) \mid a,b \in \{1,2,\dots,6\},\ a+b>6\}$

B = the event that the first die shows 3

$B = \{(3,1),(3,2),(3,3),\dots,(3,6)\}$

Find $\mathbb{P}(A|B)$: $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

$A \cap B = \{(3,4),(3,5),(3,6)\}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/\cancel{36}}{6/\cancel{36}} = 3/6 = 1/2$$

Remark: $\mathbb{P}(A|B)$ is left undefined when $\mathbb{P}(B)=0$, i.e. when B is a zero probability event. In other words, we cannot condition on zero probability events.

Theorem (Conditional Probability is a Probability Measure Itself)

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $B \in \mathcal{F}$ and $\mathbb{P}(B) > 0$.

$\forall A \in \mathcal{F}$: $\mathbb{P}_B(A) = \mathbb{P}(A|B)$

is a probability measure, i.e. it satisfies Kolmogorov's axioms.

Proof:

(1) $0 \leq \mathbb{P}_B(A) \leq 1$ $\qquad A \cap B \subseteq B$

$0 \leq \mathbb{P}(A \cap B) \leq \mathbb{P}(B) \Rightarrow 0 \leq \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \leq 1$

$\Rightarrow 0 \leq \mathbb{P}_B(A) \leq 1$

(2) $\mathbb{P}_B(\Omega) = 1$

$$\mathbb{P}_B(\Omega) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$$

(3) Countable Additivity

For disjoint events $A_1, A_2, \dots$

$$\mathbb{P}_B\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}_B(A_i)$$

$$\mathbb{P}_B\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \frac{\mathbb{P}\left(\left(\bigcup_{i \in \mathbb{N}} A_i\right) \cap B\right)}{\mathbb{P}(B)}$$

$$= \frac{\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} (A_i \cap B)\right)}{\mathbb{P}(B)}$$

Claim: $i \neq j$ $\underbrace{(A_i \cap B)}_{C_i} \cap \underbrace{(A_j \cap B)}_{C_j} \neq \emptyset$

Because

$$(A_i \cap B) \cap (A_j \cap B) = \underbrace{(A_i \cap A_j)}_{\emptyset} \cap B = \emptyset$$

C_i's $= A_i \cap B$ are also disjoint

$$T = \frac{\sum_{i=1}^{\infty} \mathbb{P}(A_i \cap B)}{\mathbb{P}(B)}$$

$$= \sum_{i=1}^{\infty} \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_{i=1}^{\infty} \mathbb{P}_B(A_i)$$

Exercise: Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $B \in \mathcal{F}$, $\mathbb{P}(B) > 0$.

(a) Show that the "reduced σ-field" $\mathcal{F} \cap B = \{ B \cap A \mid A \in \mathcal{F} \}$ is a σ-field itself.

(b) Show that $(\underbrace{B \cap \Omega}_{B}, \underbrace{B \cap \mathcal{F}}_{\mathcal{F}_B}, \underbrace{\mathbb{P}_B(A)}_{\mathbb{P}_B})$ is also a probability space.

Remark: This exercise presents the "reduced sample space" view of conditional probability.

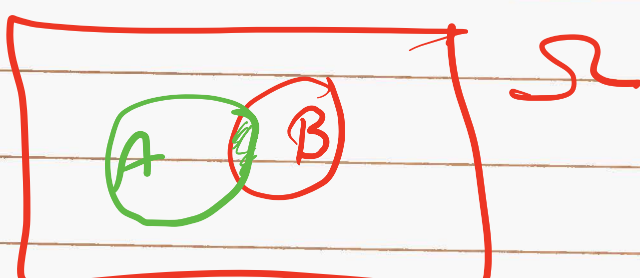
In this viewpoint, conditioning on B is equivalent to selecting B as the new sample space.

Exercise: Make sure you know how to do this.

Hint: $\mathcal{F} \cap B$ is a sigma field

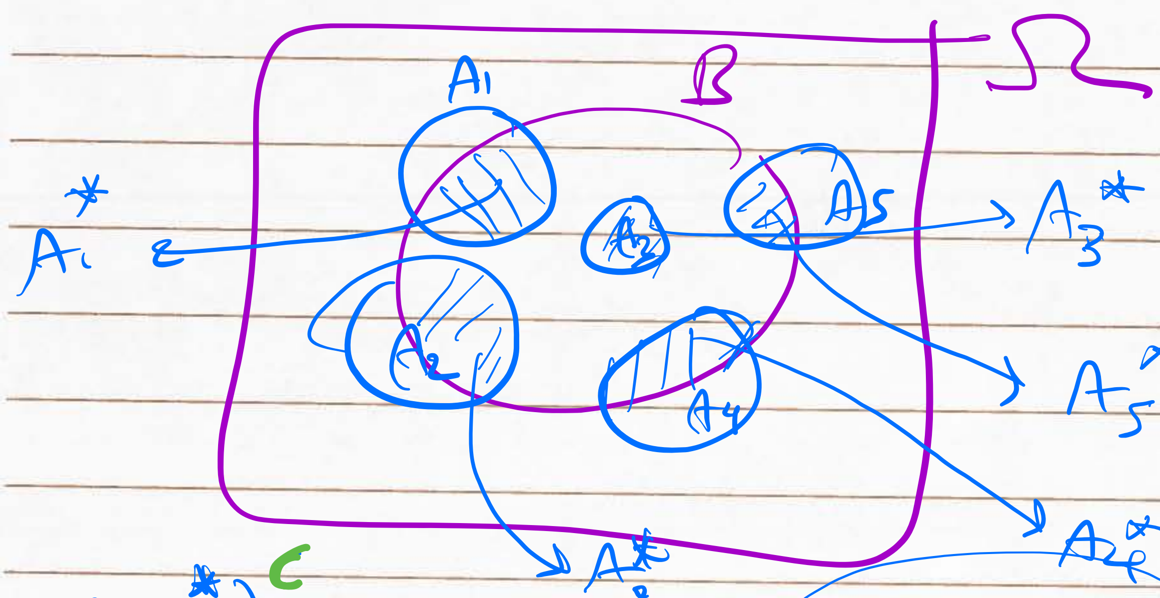
$C \in \mathcal{F} \cap B$ we want to show that $C^c \in \mathcal{F} \cap B$.

$A \cap B$, $A \in \mathcal{F}$



Complement of C with respect to the new sample space is

$$B\setminus C = \underbrace{B\cap(\underbrace{A\cap B}_{\in\mathcal{F}})^c}_{\in\mathcal{F}\cap B}$$



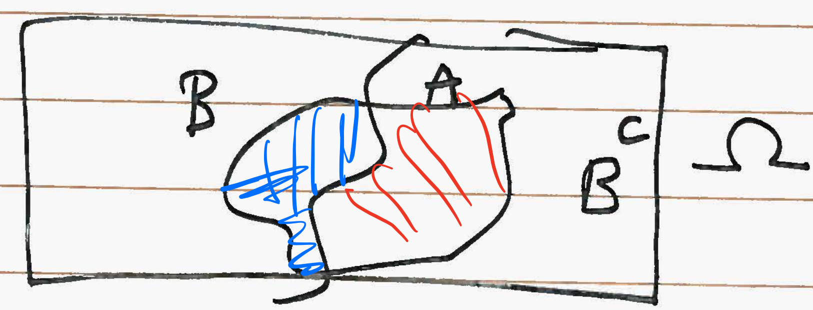
$$(A_i^*)^c = B\setminus A_i^* = \underbrace{B\setminus(\underbrace{A_i\cap B}_{\in\mathcal{F}})^c}_{\in\mathcal{F}\cap B}$$

$(F \cap B, \cap, \cup, e)$

Theorem (The Law of Total Probability):

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\exists\, A, B \in \mathcal{F}$ s.t. $0 < \mathbb{P}(B) < 1$.

Then: $\mathbb{P}(A) = \mathbb{P}(A|B)\,\mathbb{P}(B) + \mathbb{P}(A|B^c)\,\mathbb{P}(B^c)$



Proof:

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

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

mutually exclusive

$$(A \cap B) \cap (A \cap B^c) = \emptyset$$

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

$$= P(A|B)P(B) + P(A|B^c)P(B^c)$$

Note that we conditioned on B and B^c and because $0 < P(B) < 1$, $P(B) \neq 0$ and $P(B^c) \neq 0$

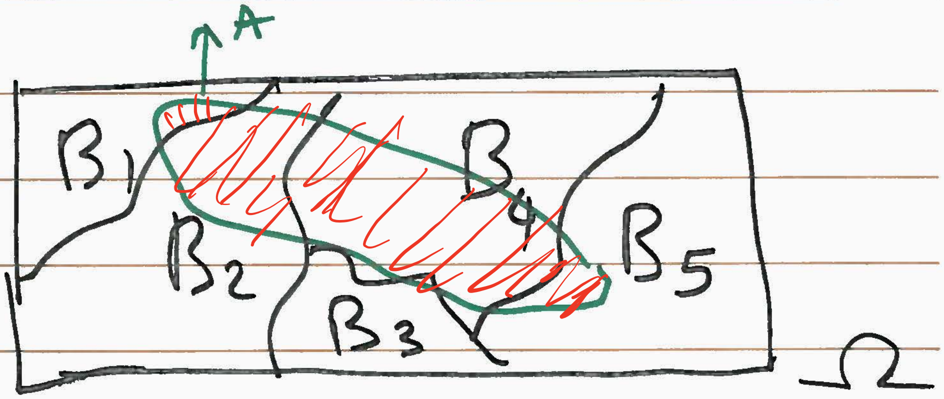
Theorem (General Total Probability)

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $B_1, B_2, \ldots$ is a partition of Ω such that $\mathbb{P}(B_i) > 0$, then:

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)$$

Remark: For a finite partition $B_1, B_2, \ldots, B_n$ the theorem is still true, i.e.

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



Proof: $A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{n} B_i\right)$

$$= \bigcup_{i=1}^{n} \underbrace{(A \cap B_i)}$$

$$\underbrace{(A \cap B_i)}_{C_i} \cap \underbrace{(A \cap B_j)}_{C_j} = \emptyset$$

$$P(A) = P\left(\bigcup_{i=1}^{n} \{A \cap B_i\}\right)$$

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

Infinite Case

$$A = A \cap \Omega = A \cap \left(\bigcup_{i \in \mathbb{N}} B_i \right)$$

$$= \bigcup_{i \in \mathbb{N}} (A \cap B_i) \longrightarrow \text{Mutually exclusive}$$

$$P(A) = P\left(\bigcup_{i \in \mathbb{N}} \{A \cap B_i\} \right) \quad \text{By continuity of prob}$$

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

$$= \lim_{n \to \infty} \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$

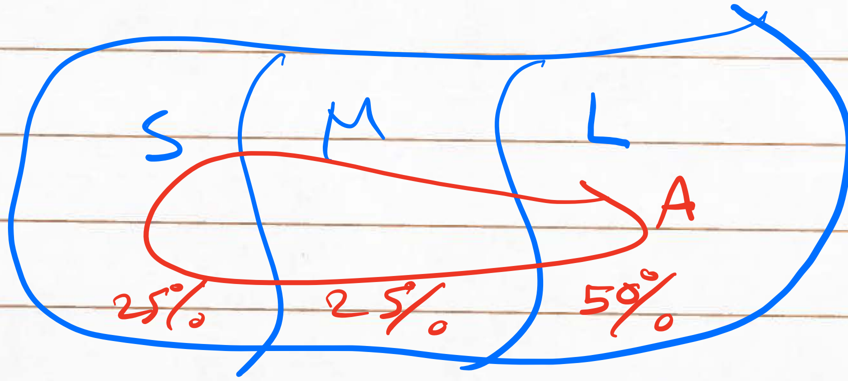
$$= \sum_{i=1}^{\infty} P(A \mid B_i) P(B_i)$$

Example: A USC student is participating in a charity fundraiser event. CEO's of small sized, medium sized, and large companies are participating in the event as well. 50% of the CEO's are from large companies, 25% are from medium companies, and 25% are from small companies.

Probability of raising funds from the CEO of a large company is 20%, while it is 70% for the CEO of a medium company and 30% for the CEO of a small company.

What is the probability of raising funds from a randomly selected CEO?

Solution:



B_1 : small sized

B_2 : Medium sized

B_3 : Large companies

A = raising funds

$$P(A) = \underset{.3}{P(A|B_1)}\,\underbrace{P(B_1)}_{.25} + \underset{.7}{P(A|B_2)}\,\underbrace{P(B_2)}_{.25} + \underset{.2}{P(A|B_3)}\,\underbrace{P(B_3)}_{.5} =$$

Theorem (The Baye's Rule)

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $A, B \in \mathcal{F}$ and $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$.

Then:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B)\,\mathbb{P}(A|B)}{\mathbb{P}(A)}$$

$$= \frac{\mathbb{P}(B)\,\mathbb{P}(A|B)}{\mathbb{P}(A|B)\,\mathbb{P}(B) + \mathbb{P}(A|B^c)\,\mathbb{P}(B^c)}$$

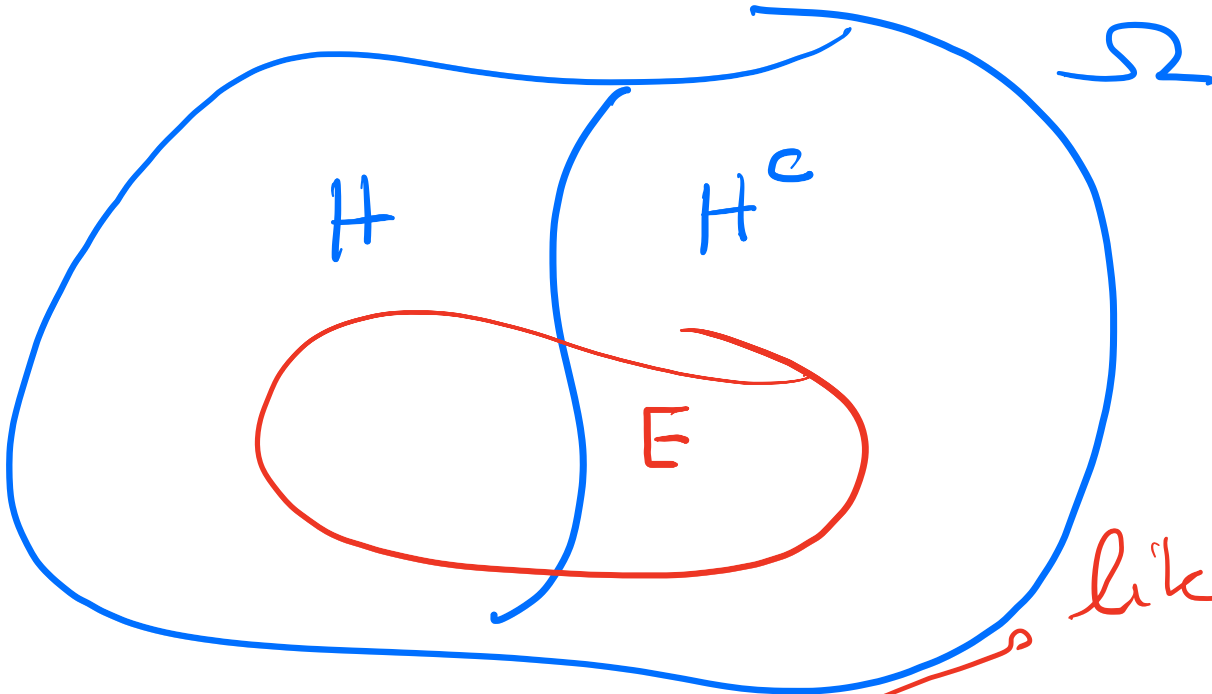
Total prob.

Proof:

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

Total prob

$$= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$



$$\underbrace{P(H \mid E)}_{\text{posterior}} = \frac{\overbrace{P(E \mid H)}^{\text{likelihood}}\ \overbrace{P(H)}^{\text{prior}}}{P(E)}$$

$$P(E) = P(E \mid H)\,P(H) + P(E \mid H^c)\,P(H^c)$$

$$P(H^c \mid E) = \frac{P(E \mid H^c)\,P(H^c)}{P(E \mid H^c)\,P(H^c) + P(E \mid H)\,P(H)}$$

$$P(H \mid E) \propto P(E \mid H)\,P(H)$$

$$P(H^c \mid E) \propto P(E \mid H^c)\,P(H^c)$$

Theorem (General Baye's Rule)

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$, and $A, B_i \in \mathcal{F}$ $\forall i \in \mathbb{N}$. Furthermore, B_i's form a partition and $\forall i \in \mathbb{N}$, $\mathbb{P}(B_i) > 0$. Then:

$$\mathbb{P}(B_i | A) = \frac{\mathbb{P}(A | B_i)\, \mathbb{P}(B_i)}{\mathbb{P}(A)}$$

$$= \frac{\mathbb{P}(A | B_i)\, \mathbb{P}(B_i)}{\sum_{j=1}^{\infty} \mathbb{P}(A | B_j)\, \mathbb{P}(B_j)}$$

Proof: Exercise. Write the definition of conditional prob and use the law of total prob in the denum.

$$\underbrace{P(H_i \mid E)}_{\text{posterior}} = \frac{\underbrace{P(E \mid H_i)}_{\text{likelihood}}\ \underbrace{P(H_i)}_{\text{prior}}}{\sum_{j=1}^{n} P(E \mid H_j)\, P(H_j)}$$

Example : (Charity Fundraiser)

Assume that the USC student raised some funds from a CEO in the event. What is the probability that it was from the CEO of a large company?

$\mathbb{P}(B_3) = .5$

$$\mathbb{P}(B_3 \mid A) = \frac{\mathbb{P}(A \mid B_3)\,\mathbb{P}(B_3)}{\mathbb{P}(A \mid B_1)\mathbb{P}(B_1) + \mathbb{P}(A \mid B_2)\mathbb{P}(B_2) + \mathbb{P}(A \mid B_3)\mathbb{P}(B_3)}$$

Plug in probabilities

Remark: Total probability and Baye's Rule deal with conditional probabilities of events given partitions of a sample Space:

When The (Unconditional) Probability of An Event Is Needed — USE TOTAL PROB

When The (Conditional) Probability of One of The Partitions Is Needed — USE BAYE'S

(You need evidence)

Remark: In Baye's Rule, A
E
is an event that we observe,
H_i
while B_i's are events that we do not observe, but would like to make some inference about.

Before any observation we know the prior probabilities $\mathbb{P}(B_i)$, and also know the
$P(H_i)$ $P(E|H_i)$
conditional probabilities $\mathbb{P}(A|B_i)$.

After observing A, we compute
E

the posterior probabilities

$P(B_i|A)$. We actually
$P(H_i|E)$

"update our belief" about

the occurence of B_i (H_i), given that A

was observed.

Theorem: (The Multiplication Rule)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A_1, A_2, \ldots, A_n$ be events, i.e. $A_i \in \mathcal{F}$ $\forall i \in \{1, 2, \ldots, n\}$.

Then:

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\ldots\mathbb{P}(A_n|A_1 \cap A_2 \cap A_3 \ldots \cap A_{n-1}) = \boxed{\mathbb{P}(A_1)\prod_{i=2}^{n}\mathbb{P}(A_i|A_1 \cap A_2 \cap \ldots \cap A_{i-1})}$$

Provided that the conditional probabilities are defined, i.e. $\forall \ell \in \{1, 2, \ldots, n\}$ $\mathbb{P}\left(\bigcap_{j=1}^{\ell} A_j\right) \neq 0$.

(First order) Markovian Property

$$\underline{\mathbb{P}(A_i \mid A_1 \cap A_2 \cap \cdots \cap A_{i-1})}$$

$$= \underline{\mathbb{P}(A_i \mid A_{i-1})}$$

Proof (by induction)

$n=1$: $\quad \mathbb{P}(A_1) = \mathbb{P}(A_1)$

$n=2$: $\quad \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 | A_1)$

Next, assume $\mathbb{P}\left(\bigcap_{i=1}^{k} A_i\right) = \mathbb{P}(A_1) \times$

$\times\, \mathbb{P}(A_2|A_1) \ldots \mathbb{P}(A_k | A_1 \cap A_2 \cap \ldots \cap A_{k-1})$.

One must show that

$$\mathbb{P}\left(\bigcap_{i=1}^{k+1} A_i\right) = \mathbb{P}(A_1) \prod_{i=2}^{k+1} \mathbb{P}(A_i | A_1 \cap A_2 \cap \ldots \cap A_{i-1})$$

$$\mathbb{P}\left(\bigcap_{i=1}^{k+1} A_i\right) = \mathbb{P}\left(\bigcap_{i=1}^{k} A_i \cap A_{k+1}\right)$$

$$= \underbrace{\mathbb{P}\left(\bigcap_{i=1}^{k} A_i\right)}\, \underbrace{\mathbb{P}\left(A_{k+1} \,\Big|\, \bigcap_{i=1}^{k} A_i\right)}$$

$$= \mathbb{P}(A_1)\,\mathbb{P}(A_2|A_1) \text{———} \mathbb{P}(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1})$$

$$\times\ \mathbb{P}\left(A_{k+1} \,\middle|\, \bigcap_{i=1}^{k} A_i\right)$$

$$= \mathbb{P}(A_1) \prod_{i=2}^{k+1} \mathbb{P}(A_i | A_1 \cap A_2 \cap \cdots \cap A_{i-1})$$

Theorem: (General Multiplication Rule):

The multiplication rule holds for an infinite number of events $A_1, A_2, \ldots$, where $A_i \in \mathcal{F}\ \forall i \in \mathbb{N}$

i.e.

$\bigcap_{i \in \mathbb{N}} A_i$

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \mathbb{P}(A_1) \prod_{i=2}^{\infty} \mathbb{P}(A_i \mid A_1 \cap \cdots \cap A_{i-1})$$

provided that all conditional probabilities are defined, i.e.

$$\mathbb{P}\left(\bigcap_{j=1}^{l} A_j\right) \neq 0 \quad \forall l \in \mathbb{N}$$

Proof: Take a $\lim_{n\to\infty}$ from the multiplication rule and use continuity of probability.

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\prod_{i=2}^{n}\mathbb{P}(A_i \mid A_1\cap A_2\cap\cdots\cap A_{i-1})$$

$$\lim_{n\to\infty}\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\lim_{n\to\infty}\prod_{i=2}^{n}\mathbb{P}(A_i \mid A_1\cap A_2\cap\cdots\cap A_{i-1})$$

By continuity of prob.:

$$\underbrace{\mathbb{P}\left(\lim_{n\to\infty}\bigcap_{i=1}^{n} A_i\right)}_{\mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right)} = \mathbb{P}(A_1)\prod_{i=2}^{\infty}\mathbb{P}(A_i \mid A_1\cap A_2\cap\cdots\cap A_{i-1})$$

- Independence

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $A, B \in \mathcal{F}$. A, B are called independent iff

$$\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)}$$

and if $\mathbb{P}(B) > 0$, it means:

$$\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\cancel{\mathbb{P}(B)}}{\cancel{\mathbb{P}(B)}} = \mathbb{P}(A)}$$

Question: Can an event be independent from itself?

$$\mathbb{P}(A \cap A) = \mathbb{P}(A)\,\mathbb{P}(A)$$

$$\Rightarrow \quad \mathbb{P}(A) = [\mathbb{P}(A)]^2$$

$$\mathbb{P}(A) = 0 \qquad \mathbb{P}(A) = 1$$

Almost sure and zero prob. events

Example: $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $A, B \in \mathcal{F}$, and $0 < \mathbb{P}(B) < 1$. If A and B are independent, show that the probability of A given B is equal to the probability of A given B^c.

In other words, the probability of A doesn't change whether or not B occurs.

$$\mathbb{P}(A|B) = \mathbb{P}(A|B^c)$$

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

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

$$= \underbrace{\mathbb{P}(A|B)}_{\mathbb{P}(A)} \mathbb{P}(B) + \mathbb{P}(A|B^c) \underbrace{\mathbb{P}(B^c)}_{1-\mathbb{P}(B)}$$

$$\mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A|B^c)(1-\mathbb{P}(B))$$

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

$$\mathbb{P}(A)\cancel{(1-\mathbb{P}(B))} = \mathbb{P}(A|B^c)\cancel{(1-\mathbb{P}(B))}$$

$$\frac{\mathbb{P}(A|B)}{\mathbb{P}(A)} = \mathbb{P}(A|B^c)$$

Exercise:

Also show that A^c is indep from both B & B^c.

More rigorously, we say the σ-field generated by A is independent from the σ-field generated by B.

$$\mathcal{F}_A = \{A, A^c, \emptyset, \Omega\}$$

$$\mathcal{F}_B = \{B, B^c, \emptyset, \Omega\}$$

Because every member of $\mathcal{F}_A$ is indep from every member of $\mathcal{F}_B$.

Remark: The concept of mutually exclusive events should not be confused with that of independent events.

Mutually exclusive sets (events) are sets whose overlap is empty. We do not need a probability measure to assess whether they are mutually exclusive or not. On the other hand, independent events are meaningful within the concept of probability. Without having a probability space equipped with a probability measure, one cannot define independence.

Is there an event that is mutually exclusive with itself

$A \cap A = \emptyset \Rightarrow A = \emptyset$

When are two mutually exclusive events indep?

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

$$\Rightarrow \mathbb{P}(A)\,\mathbb{P}(B) = 0$$

Only if at least one of them is zero probability

Conditional Independence

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and C is $\mathcal{F}$-measurable, i.e. $C \in \mathcal{F}$. Also assume that $\mathbb{P}(C) > 0$.

$A, B \in \mathcal{F}$ are called conditionally independent given C iff:

$$\boxed{\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C)}$$

Use: Bayesian Networks / Graphical Models / Naive Bayes classifier

Example: Assume that A, B are independent events. Are they conditionally independent given any $C \in \mathcal{F}$, for which $\mathbb{P}(C) > 0$?

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

we would like to know if

$$\mathbb{P}(A \cap B \mid C) \overset{?}{=} \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C)$$

$$\frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} \overset{?}{=} \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)}$$

Does not hold for all C

(Provide a counterexample)

Counter-example (to disprove):

Roll of a die:

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

$P(A) = \frac{2}{6} = \frac{1}{3}$ $P(B) = \frac{3}{6} = \frac{1}{2}$

$P(A \cap B) = P(\{2\}) = \frac{1}{6} = P(A)P(B)$

A, B are indep (but not mutually exclusive)

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{1/6}{2/6} = \frac{1}{2}$$

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{1/6}{2/6} = \frac{1}{2}$$

$$P(A \cap B|C) = \frac{P(A \cap B \cap C)}{P(C)} = \frac{1/6}{2/6} = 1/2$$

$\neq P(A|C)\,P(B|C)$

Independence for More Than Two Events

For an indexed family of events $\{A_i \in \mathcal{F} \mid i \in I\}$, A_i's are independent if for all J, finite subsets of I,

$$\mathbb{P}\Big(\bigcap_{i\in J} A_i\Big) = \prod_{i\in J} \mathbb{P}(A_i) \quad \forall J \subseteq I, \ |J| < \infty$$

Therefore, if A, B, C are independent

then: $\mathbb{P}(A\cap B\cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$

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

$\mathbb{P}(B\cap C) = \mathbb{P}(B)\mathbb{P}(C)$

$\mathbb{P}(A\cap C) = \mathbb{P}(A)\mathbb{P}(C)$

· Note: I can be infinite

Example: We have a pack of well-shuffled 52 playing cards. Is the suit of a card independent from its rank?

$\text{suit} = \{\diamondsuit, \clubsuit, \spadesuit, \heartsuit\}$

$\text{rank} = \{2, 3, 4, \ldots, 10, J, Q, K, A\}$

WLOG

Assume $\text{suit} = \heartsuit$

$\text{rank} = Q$

$P(\text{suit} = \heartsuit) = \frac{13}{52} = 1/4$

$P(\text{rank} = Q) = \frac{4}{52} = 1/13$

$P(\text{rank} = Q, \text{suit} = \heartsuit) = 1/52$

$$= \underbrace{P(\text{rank}=Q)}_{1/13}\ \underbrace{P(\text{suite}=\heartsuit)}_{1/4}$$

The above calculation would be true for any rank & suite.

Pairwise Independence.

$\{A_i \in \mathcal{F} \mid i \in I\}$ is called pairwise independent if

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j) \quad i \neq j$$

Question: which of the following implications is true?

Not true

pairwise independence $\cancel{\Rightarrow}$ Independence

Independence $\Rightarrow$ Pairwise Independence

This is because if $\{A_i | i \in I\}$ are independent, $\{A_i, A_j | i \neq j\}$, $\{A_i, A_j, A_k \mid i, j, k \text{ are distinct}\}$, etc must be independent but the reverse is not true.

Exercise: Consider mutually exclusive events B_1, B_2, B_3, and C on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{P}(B_1) = \mathbb{P}(B_2) = \mathbb{P}(B_3) = p$ and $\mathbb{P}(C) = q$, where $3p+q\leq 1$. Assume that $p=-q+\sqrt{q}$, and show that the events $B_1 \cup C$, $B_2 \cup C$, and $B_3 \cup C$ are pairwise independent. Also, is there any $p>0$ and $q>0$ such that these three events are independent.

Pairwise indep if for $B_i \cup C$ and $B_j \cup C$ $\quad i \neq j$

$$P((B_i \cup C) \cap (B_j \cup C)) = P(B_i \cup C)\, P(B_j \cup C)$$

$$P((B_i \cup C) \cap (B_j \cup C)) = P((\underbrace{B_i \cap B_j}_{\emptyset}) \cup C)$$

$$= P(C) = q$$

$$P(B_i \cup C) = P(B_i) + P(C) = p + q$$

$$= \cancel{-q} + \sqrt{q} + \cancel{q} = \sqrt{q}$$

$B_i \cup C$ and $B_j \cup C$ are pairwise indep

because

$$P((B_i \cup C) \cap (B_j \cup C)) = q$$

$$= \underbrace{P(B_i \cup C)}_{\sqrt{q}} \underbrace{P(B_j \cup C)}_{\sqrt{q}}$$

For independence

$$P([B_1 \cup C] \cap [B_2 \cup C] \cap [B_3 \cup C])$$

$$= \underbrace{P(B_1 \cup C)}_{\sqrt{q}} \underbrace{P(B_2 \cup C)}_{\sqrt{q}} \underbrace{P(B_3 \cup C)}_{\sqrt{q}}$$

$$\Rightarrow \mathbb{P}\big(\overbrace{(B_1 \cap B_2 \cap B_3)}^{\emptyset} \cup C\big) = (\sqrt{q})^3$$

$$\Rightarrow \underbrace{\mathbb{P}(C)}_{q} = (\sqrt{q})^3$$

$$\Rightarrow \begin{cases} q = 0 \longrightarrow \text{not acceptable} \\ q = 1 \end{cases}$$

$$p = -q + \sqrt{q} \Rightarrow p = 0 \quad \text{not acceptable}$$

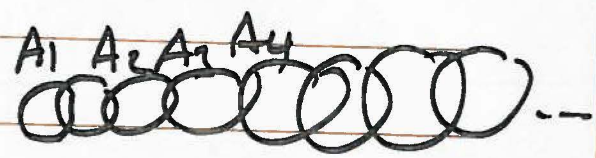
This serves as a general counterexample for events that are pairwise indep, but not indep.

The Borel - Cantelli Lemmas

An event that occurs infinitely often:

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $A_i \in \mathcal{F}$ $\forall i \in \mathbb{N}$.

Define:



$$B_1 = \bigcup_{i=1}^{\infty} A_i$$

$$B_2 = \bigcup_{i=2}^{\infty} A_i$$

$$\vdots$$

$$B_n = \bigcup_{i=n}^{\infty} A_i$$

If $A_{i.o.}$ is an event that occurs <u>infinitely often</u>, then $\therefore$

$x \in A_{i.o.}$ must be in all B_n's,

i.e. $\forall n,\ x \in B_n \Rightarrow x \in \bigcap_{n \geq 1}^{\infty} B_n$

$\Rightarrow A_{i.o.} \subseteq \bigcap_{n\geq 1}^{\infty} B_n$ $\qquad \bigcap_{n\geq 1}^{\infty} \bigcup_{i\geq n}^{\infty} A_i$

On the other hand, if

$x \in B_k \ \forall k \ \forall k \in \mathbb{N}$, we can show that x is a member of an infinite number of A_i's, i.e. $x \in A_{i.o.}$ $\therefore$

$x \in B_1 \Rightarrow x \in \bigcup_{i=1}^{\infty} A_i \Rightarrow \exists\, i > 0$ s.t. $x \in A_i$

$\vdots$

$x \in B_2 \Rightarrow x \in \bigcup_{i=2}^{\infty} A_i \Rightarrow \exists\, i > 1$ s.t. $x \in A_i$

$\vdots$

$x \in B_{l+1} \Rightarrow x \in \bigcup_{i=l+1}^{\infty} A_i \Rightarrow \exists\, i > l$ s.t $x \in A_i$

$\vdots$

$\forall k \;\; x \in B_k \Rightarrow x \in \bigcap_{k=1}^{\infty} B_k$

Therefore, there is an infinite sequence of integers $\{j_l\}$ s.t.

$x \in A_{j_l} \;\; \forall\, l \in \mathbb{N}. \Rightarrow$

$x \in A_{i.o.} \Rightarrow \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i \subseteq A_{i.o}$

$A_{i.o} = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$

Lemma (Borel–Cantelli):

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $A_1, A_2, \ldots,$ are events, i.e. $A_i \in \mathcal{F}\ \forall i \in \mathbb{N}$.

Define $A_{i.o.} = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$. Then:

(1) $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty \Rightarrow \mathbb{P}(A_{i.o.}) = 0$

(2) $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$ and A_i's are independent $\Rightarrow \mathbb{P}(A_{i.o.}) = 1$.

Proof: $A_{i.o.} = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{i=n}^{\infty} A_i}_{B_n}$

$A_{i.o.} \subseteq B_n \Rightarrow P(A_{i.o.}) \leq P(B_n)$

$P(B_n) = P\left(\bigcup_{i=n}^{\infty} A_i\right) \leq \sum_{i=n}^{\infty} P(A_i)$

$\Rightarrow P(A_{i.o.}) \leq \sum_{i=n}^{\infty} P(A_i)$

$\underbrace{\lim_{n\to\infty} P(A_{i.o.})}_{P(A} \leq \underbrace{\lim_{n\to\infty} \sum_{i=n}^{\infty} P(A_i)}_{0}$

Show that

Hint $\sum_{i=n}^{\infty} P(A_i) = \lim_{m\to\infty} S_m$

$S_m = \sum_{i=n}^{m} P(A_i)$

Example (Erdös – Rényi):

Consider an experiment in which a coin is tossed independently an infinite number of times. Let

$A_i = \{H \text{ occurs in the } i^{th} \text{ coin toss}\}$

A_5 O O O O H — — —

(1) Assume that $\mathbb{P}(A_i) = 1/2$.

Then $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} 1/2 = \infty$

By the second Borel - Cantelli lemma, it follows that infinitely many heads will occur with probability 1 (almost surely)

(2) Next assume that $\mathbb{P}(A_i) = \left(\frac{1}{2}\right)^i$.

Then $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1/2}{1 - 1/2} = 1 < \infty$

Hence, by the first Borel-Cantelli lemma, the probability that infinitely many heads occur is 0.

In other words, almost surely, only finitely many heads will occur!

It might appear surprising that in the first example, where probability of getting heads is $1/i$, for any i we select (no matter how large), there occurs a heads beyond i almost surely.

The reason is that the decay rate $1/i$ of the probability of observing heads is not "adequately fast."

On the other hand, in the second example, the rate of decay $(1/2)^i$ is adequately fast that after a finite i, there will almost surely be no heads.

Exercise: (Infinite Monkey):

A monkey sits in front of a computer and starts pressing keys randomly on the keyboard.

Using the Borel-Cantelli Lemmas,

show that the monkey will type any text of your choice (e.g. ABRACADABRA or the whole declaration of independence), with probability 1.

You can assume that the monkey selects the keys uniformly at random, and that it selects the successive keys independently.