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

- 1 -

- EVENTS AND PROBABILITIES -

A random experiment $\rightsquigarrow$ 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 $\mathbb{P} : \mathcal{F} \to [0,1]$

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

The set of all possible outcomes of an experiment is called the <u>sample space</u> (usually Ω). Each member of Ω, $\omega \in \Omega$, is called an <u>elementary event</u>.

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

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

1st toss heads 2nd toss tails

The <u>event space</u> (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) $\overset{\text{empty set}}{\emptyset} \in \mathcal{F}$

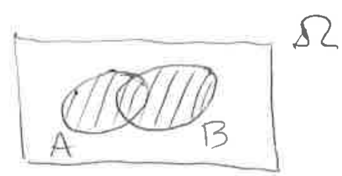
(ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ $\quad \left\{ \begin{array}{l} A^c = \Omega \setminus A \\ \text{the compliment of } A \end{array} \right.$

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

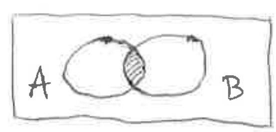
σ-field of events = "events which are interesting to us"

$\emptyset$ = impossible event, Ω = certain event

A, B two events :

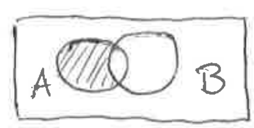


$A \cup B$ either A or B occurs



$A \cap B$ both A and B occurs

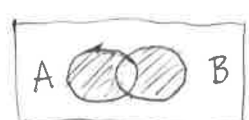
A^c A does not occur



$A \setminus B = A \cap B^c$

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

Proof · $(A \cap B)^c = \underbrace{A^c \cup B^c}_{\text{belong to } \mathcal{F} \text{ by (ii)}}$ ← belongs to $\mathcal{F}$ by (iii)

· $(A \cap B)^c \in \mathcal{F} \underset{(ii)}{\Longrightarrow} A \cap B \in \mathcal{F}$. $\square$

Fact · If $A, B \in \mathcal{F}$, then $A \setminus B,\ A \triangle B,\ B \setminus A \in \mathcal{F}$.

E.g. $\Omega = \{ (H,H),\ (H,T),\ (T,H),\ (T,T) \}$

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

· $\mathcal{F} = \{ \emptyset, \Omega, \underbrace{\{(H,H), (H,T)\}}_{1^{st} \text{ toss } H}, \underbrace{\{(T,H),(T,T)\}}_{1^{st} \text{ toss } T} \}$

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

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

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

(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

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

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

$\mathbb{P}(\{(T,H)\}) = \frac{1}{4}$, $\mathbb{P}(\{(T,T)\}) = \frac{1}{4}$.

Then $\mathbb{P}(1^{\text{st}} \text{ toss } H) = \mathbb{P}(\{(H,H),(H,T)\})$

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

$+ \mathbb{P}(\{(H,T)\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

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

Then $\mathbb{P}(\{\omega_i\}) = \frac{1}{N}$, that is

$$\mathbb{P}(A) = \frac{|A|}{N}, \quad A \subset \Omega$$

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

Def A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$

sample space, σ-field of events, prob. measure

E.g. Toss a fair coin until you get tails

$$\Omega = \{ T, HT, HHT, HHHT, \ldots \}$$

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

$$\mathbb{P}(\{ \underbrace{HH\ldots H}_{n}T \}) = \frac{1}{2^{n+1}}, \quad n \geq 0,$$ 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

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

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

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A), \quad \mathbb{P}(A) \leq \mathbb{P}(B)$$

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

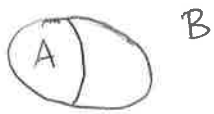
(4) $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

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

· if $A_1 \supset A_2 \supset \ldots$, then $\mathbb{P}(A_n) \xrightarrow[n\to\infty]{} \mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right)$.

<u>Proof</u> (1) define $A_{n+1} = \emptyset$, $A_{n+2} = \emptyset$, ... and use (iii)

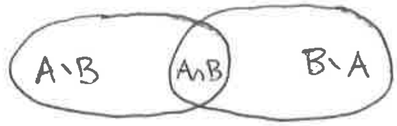
(2)



$$B = \underbrace{A \cup (B \setminus A)}_{\text{disjoint}}$$

$$\mathbb{P}(B) = \mathbb{P}(A) + \underbrace{\mathbb{P}(B \setminus A)}_{\geq 0} .$$

(3)



$$A = A \setminus B \cup A \cap B$$
$$B = B \setminus A \cup A \cap B$$

$$\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$$
$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \quad \Big/ +$$

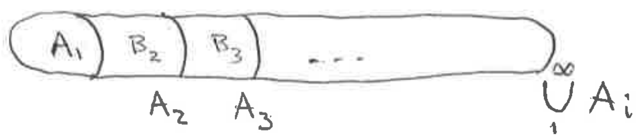
$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$$

disjoint

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

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

(5)



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

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) , \text{ so}$$

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

telescoping sum

$$= \lim_{n\to\infty} \mathbb{P}(A_n) . \quad \square$$

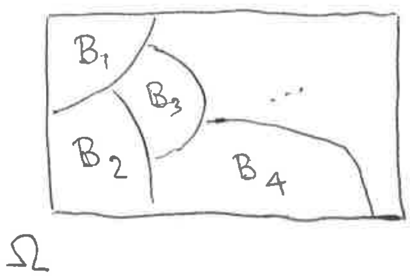
Discrete sample spaces

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a prob. space with Ω countable,

$\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$. Then any σ-algebra on Ω

is of the form :

- take a countable partition $\{B_i\}_{i=1}^{N}$ $(N \le \infty)$ of Ω,



that is $B_i \cap B_j = \emptyset$, $i \neq j$,

$$\bigcup_{i=1}^{N} B_i = \Omega$$

- take $\mathcal{F} = \left\{ \bigcup_{i \in I} B_i \,,\ I \subset \{1, 2, \dots, N\} \right\}$

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

$$A = \bigcup_{i \in I} B_i \quad \text{for some subset} \quad I \subset \{1, 2, \dots, N\}$$

- define $p_i = \mathbb{P}(B_i)$, $i = 1, \dots, N$ $\quad (p_i \geq 0, \ \sum p_i = 1)$

$$\mathbb{P}(A) = \sum_{i \in I} p_i$$

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

$$\mathbb{P}(A) = \sum_{i : \omega_i \in A} p_i \,.$$

E.g. Uniform measure : if $\Omega = \{1, \dots, N\}$, $B_i = \{i\}$, $\mathcal{F} = 2^\Omega$,

$$p_i = \mathbb{P}(\{i\}) = \frac{1}{N},$$

$$\mathbb{P}(A) = \sum_{i \in A} p_i = \sum_{i \in A} \frac{1}{N} = \frac{|A|}{N} = \frac{|A|}{|\Omega|}.$$

E.g. You are dealt 5 cards (from 52). What is the prob. of getting a full house (one pair + three of a kind)?

$$\Omega = \{ \text{subsets of size 5 of } \{1, \dots, 52\} \}$$

$$|\Omega| = \binom{52}{5}, \quad \mathbb{P} = \text{uniform}$$

$$A = \text{full house}, \quad |A| = \underbrace{\binom{13}{1}}_{\text{rank for a pair}} \underbrace{\binom{4}{2}}_{\text{suits to form the pair}} \cdot \underbrace{\binom{12}{1}}_{\text{rank for a three of a kind}} \underbrace{\binom{4}{3}}_{\text{suits}}$$

$$\mathbb{P}(A) = \frac{13 \cdot 12 \cdot \binom{4}{2}\binom{4}{3}}{\binom{52}{5}}.$$

E.g. What is the probability that each player in a game of bridge receives one ace?

$$\Omega = \{ \text{permutations of } \{1, 2, \dots, 52\} \}$$

$$|\Omega| = 52! = 52 \times 51 \times \dots \times 2 \times 1$$

$$\underbrace{\hspace{3em}}_{\text{Player } 1}^{\text{13 cards}} \; \underbrace{\hspace{3em}}_{2} \; \underbrace{\hspace{3em}}_{3} \; \underbrace{\hspace{3em}}_{4}$$

A = each player has one ace

$$|A| = \underbrace{13^4}_{\text{position of an ace in each of player's hand}} \cdot \underbrace{4!}_{\text{which ace to which player}} \cdot \underbrace{48!}_{\text{remaining cards}}$$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{13^4 \cdot 4! \cdot 48!}{52!} = 0.105\ldots$$

<u>E.g.</u> In how many ways can we put n oranges into k labeled boxes?

$$\circ\ \circ\ \circ\ \circ\ \circ\ \circ \qquad n=6$$

$$\underbrace{\lfloor \circ\circ \rfloor}_{1}\ \underbrace{\lfloor \circ\circ\circ \rfloor}_{2}\ \underbrace{\lfloor \circ \rfloor}_{3} \qquad k=3 \qquad \lfloor \circ\circ\circ\circ \rfloor\ \lfloor\quad\rfloor\ \lfloor \circ\circ \rfloor$$

encode

$$\circ\,\circ\ |\ \circ\,\circ\,\circ\ |\ \circ \qquad\qquad \circ\,\circ\,\circ\,\circ\ |\ |\ \circ\,\circ$$

the same as choosing positions for $k-1$ pencils

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

Answer: $\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$.

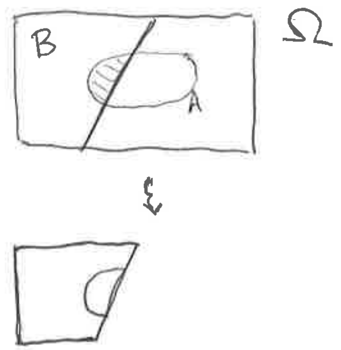
<u>E.g.</u> A fair die is thrown 3 times. What is the probability of getting at least two outcomes the same?

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

$$= 1 - \frac{6\cdot 5\cdot 4}{6\cdot 6\cdot 6} = \frac{16}{36} = \frac{4}{9}.$$

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.



Def If A, B are two events and $\mathbb{P}(B)>0$, then the conditional probability of A given B is
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Thm The conditional probability defines a prob. measure.

Proof WTS $\mathbb{Q}(B) = \mathbb{P}(\cdot \mid B) = \dfrac{\mathbb{P}(\cdot \cap B)}{\mathbb{P}(B)}$ is a prob. measure

- $\mathbb{P}(\Omega \mid B) = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$, so $\mathbb{Q}(\Omega) = 1$

- $A_1, A_2, \ldots$ disjoint events,

$$\mathbb{Q}(\cup A_i) = \mathbb{P}(\cup A_i \mid B) = \frac{\mathbb{P}((\cup A_i) \cap B)}{\mathbb{P}(B)}$$
$$= \frac{\mathbb{P}(\bigcup A_i \cap B)}{\mathbb{P}(B)} = \frac{\sum \mathbb{P}(A_i \cap B)}{\mathbb{P}(B)}$$
$$= \sum \mathbb{Q}(A_i). \quad \square$$

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

Let A = 1^{st} card an ace
B = 2^{nd} card an ace

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

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

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

$$\mathbb{P}(A) = \frac{4}{52}, \qquad \mathbb{P}(B) = \frac{\overset{1^{st}\text{ ace}}{4 \cdot 3} + \overset{1^{st}\text{ not ace}}{48 \cdot 4}}{52 \cdot 51} = \frac{4(3+48)}{52 \cdot 51} = \frac{4}{52}$$

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

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

<u>E.g.</u> (Cascade prob.) $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$,

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

(assuming $\mathbb{P}(B \cap C) > 0$ and $\mathbb{P}(C) > 0$).

n+ ○
1× ●

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

$$\mathbb{P}(A_3^c \cap A_2 \cap A_1) = \mathbb{P}(A_3^c \mid A_2 \cap A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_1)$$

$$= \frac{1}{n-2} \cdot \frac{n-2}{n-1} \cdot \frac{n-1}{n} = \frac{1}{n}.$$

Very useful for urn models

<u>Theorem</u> If $\{B_1, B_2, \dots\}$ is a partition of Ω s.t. $\mathbb{P}(B_i) > 0$ for every i, then for any event A,

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \mid B_i) \cdot \mathbb{P}(B_i)$$

(This is the so-called partition thm / the thm of total prob.)

<u>Proof</u> $\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}(A \cap (B_1 \cup B_2 \cup \dots))$

$$= \mathbb{P}(\underbrace{\bigcup A \cap B_i}_{\text{disjoint}}) = \sum_i \mathbb{P}(A \cap B_i)$$

$$= \sum_i \mathbb{P}(A \mid B_i)\mathbb{P}(B_i). \quad \square$$

<u>E.g.</u> A biased coin shows heads with prob $p \in (0,1)$. Let

$$u_n = \mathbb{P}(\underbrace{\text{in } n \text{ tosses no pair of heads occurs successively}}_{A_n})$$

Find u_n.

Let $B_i = \overbrace{H \dots H}^{i-1} T$ (first $i-1$ tosses H, then T), $i \geq 1$.

They form a partition,

$$u_n = \mathbb{P}(A_n) = \sum_{i=1}^{\infty} \mathbb{P}(A_n \mid B_i) \cdot \mathbb{P}(B_i)$$

$$= \mathbb{P}(A_n \mid B_1) \cdot \mathbb{P}(B_1) + \mathbb{P}(A_n \mid B_2) \cdot \mathbb{P}(B_2)$$

$$= \underbrace{u_{n-1} \cdot (1-p)}_{T\underbrace{\quad}_{n-1}} + \underbrace{u_{n-2} \cdot p(1-p)}_{HT\underbrace{\quad}_{n-2}}. \quad \square$$

Independence

If $\quad \mathbb{P}(A \mid B) = \mathbb{P}(A) \overset{\text{intuitively}}{\rightsquigarrow}$ B doesn't affect A

or A is independent of B

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

Def Events A, B are <u>independent</u> if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

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

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

E.g. Events A, B, C are independent if and only if

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

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

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

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

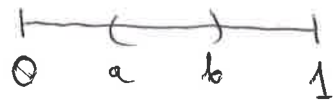
$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{1\}) = \frac{1}{4}$$
$$\mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}.$$

But they are pairwise indep., $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$, etc.

Subtleties

1) Why $\mathbb{P}$ countably additive ?

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



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

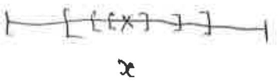
(Formally, $\Omega = [0,1]$

$\mathcal{F}$ = Borel σ-field of $[0,1]$ = smallest σ-field containg all intervals

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

Then

$$\mathbb{P}(\{x\}) = \mathbb{P}\left(\bigcap_{k\geq 1} \left[x-\tfrac{1}{k}, x+\tfrac{1}{k}\right]\right)$$



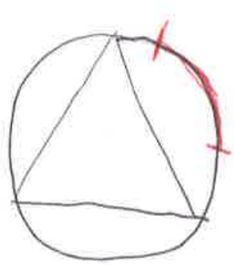
$$= \lim_{k\to\infty} \mathbb{P}\left(\left[x-\tfrac{1}{k}, x+\tfrac{1}{k}\right]\right) = \lim_{k\to\infty} \frac{2}{k} = 0.$$

If $\mathbb{P}$ was fully additive

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

In the def. we require $\mathbb{P}$ to be countably additive.

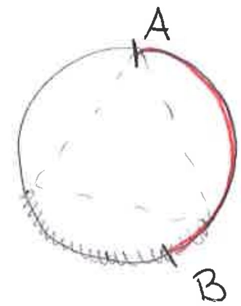
2) Bertrand's paradox



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 $\quad p = \frac{1}{3}$

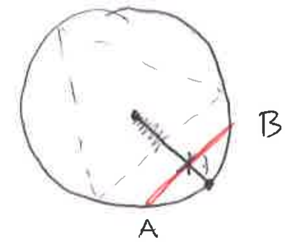
- choose A
- choose B $\rightsquigarrow$ chord $\widehat{AB}$



b) random radius model $\quad p = \frac{1}{2}$

- choose a radius
- choose a point on it

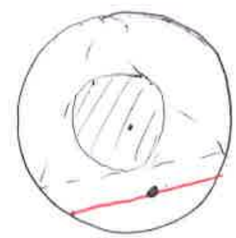
$\rightsquigarrow$ chord $\widehat{AB}$
$\quad ||$
$\perp$ radius, passing t. the point



c) random midpoint model

- choose a random point in the disc

$\rightsquigarrow$ chord $\widehat{AB}$
$\quad ||$
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 calulate probabilities within a certain model.