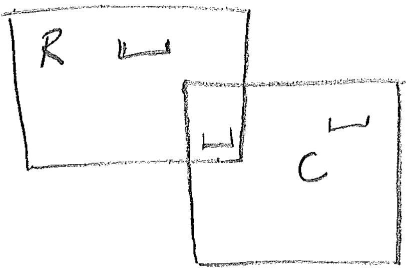
Ex.: A survey of ___ individuals here revealed that ___ liked RAP/HIP-HOP and ___ liked classical music and ___ liked both.



Let $|R|$ = # of indiv. liking rap/hip-hop

$|C|$ = # " " " classical

$$|R \cup C| = |R| + |C| - |R \cap C|$$

$$= \_\_\_ + \_\_\_ - \_\_\_$$

This idea in Property 2 generalizes to the "<u>inclusion-exclusion formula</u>".
Let $A_1, A_2, \cdots, A_n$ be any n events. Then,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \cdots$$

$$+ (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

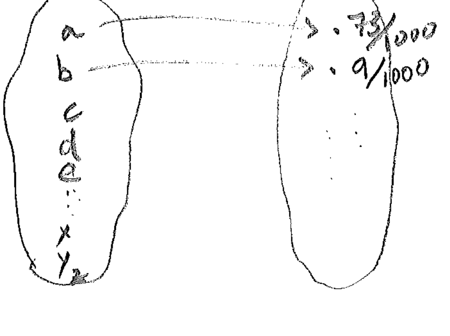
in words:

"we take all possible intersections of one, two, three, ..., n events and let the signs alternate".

<u>Qn</u>: Does the inclusion-exclusion formula agree with extended Axiom (ii): $A_1, A_2, \cdots, A_n$ are pair-wise disjoint event

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

<u>Example</u>: Frequency of letters in English (will do with SAGE)



$P(\text{vowel}) = P(\{a, e, i, o, u\})$

$= P(a) + P(e) + \quad + P(u).$

$\approx 0.378.$

SO Probability is a function from the set of events to [0,1]. And events are certain subsets of the sample space Ω. And the function satisfies four axioms: (i)

(ii)

(iii)

(iv)

and has two properties: (1)

(2).

But, what exactly is stipulated by the axioms with regard to the domain of the Probability (function).

Domain should be a Sigma Field or Sigma Algebra denoted $\mathcal{F}(\Omega)$. or $\mathcal{F}$

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup A_i \in \mathcal{F}$

$\boxed{(\Omega, \mathcal{F}(\Omega), P)}$ is called probability space or probability triple

Ex:

Let $\Omega = \{H,T\}$, $\mathcal{F}(\Omega) \overset{?}{=} \{\{H,T\}, \phi, \{H\}, \{T\}\}$. ← finest σ-field.

a smaller $\mathcal{F}'(\Omega) \overset{?}{=} \{\{H,T\}, \phi\}$ ← trivial σ-field.

Ex:

$\Omega = \{\omega_1, \dots, \omega_n\}$, $\mathcal{F}(\Omega) = 2^{\Omega}$ = The set of all subsets of Ω

$|2^{\Omega}| = 2^n$

Ex:

$\Omega = \{\omega_1, \omega_2, \dots\}$ ↑ countably ∞

Ex: $\Omega = \{\omega_1, \omega_2, \dots\}$ ↑ uncountably ∞

You need a course in measure theory 400 level or Prob. Theory to fully appreciate such Ω's.

Independence

Two events A and B are independent if

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

Intuitively, A and B are independent if the occurrence of A has no influence on the probability of occurrence of B.

Ex. Flip two fair coins A = "The first coin is 'H' "
B = "The second coin is 'H' "

$$P(A) = \tfrac{1}{2}, \quad P(B) = \tfrac{1}{2}.$$

Flips are independent

$$P(A \cap B) = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}.$$

Ex. roll a fair dice twice! (here the faces of the dice are enumerated 1,2,3,4,5,6)

A = 'first throw is 5'

B = 'second throw is 1'

$$P(A) = \tfrac{1}{6}, \quad P(B) = \tfrac{1}{6} \qquad P(A \cap B) = \tfrac{1}{36}.$$

if the dice were thrown independently

Ex Flip a fair coin n times.

$$P(\underbrace{H\,T\,T\cdots H}_{\text{any sequence } \omega\in\Omega}) = \left(\tfrac{1}{2}\right)^n = \tfrac{1}{2^n}.$$

$\Omega = \boxed{\{H,T\}^n}$ ← The set of all sequences of H & T of length n.

Qn ① $P(TT\cdots T) \overset{?}{=}$

Qn ② $P(HH\cdots H) \overset{?}{=}$

Qn ③ Show that $P(\Omega) = 1 = \sum_{\omega\in\Omega} P(\omega) = 2^n \frac{1}{2^n} = 1$

Ex. Flip a biased coin with $P(H) = \theta$ n times, $\theta \in [0,1]$

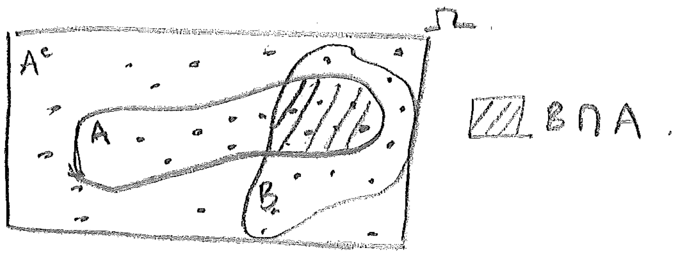
$$P(HHH\cdots H) = \theta^n, \quad P(HTHH\cdots H) = \theta(1-\theta)\theta\,\theta\cdots\theta = \theta^{n-1}(1-\theta), \ \ldots$$

Conditional Probability

Suppose we are told that the event A with $P(A) > 0$ occurs. Then, the sample space is shrunk from Ω to A and the probability that another event B will occur given that A has occurred is defined by:

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

We understand this definition of conditional probability by noting that (i) only the outcomes in B that also belong to A can possibly occur, and (ii) since the new sample space is A, we have to divide by $P(A)$ to make $P(A|A) = \frac{P(A \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$.



Ex1: Suppose two events A & B are independent

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

"in words"

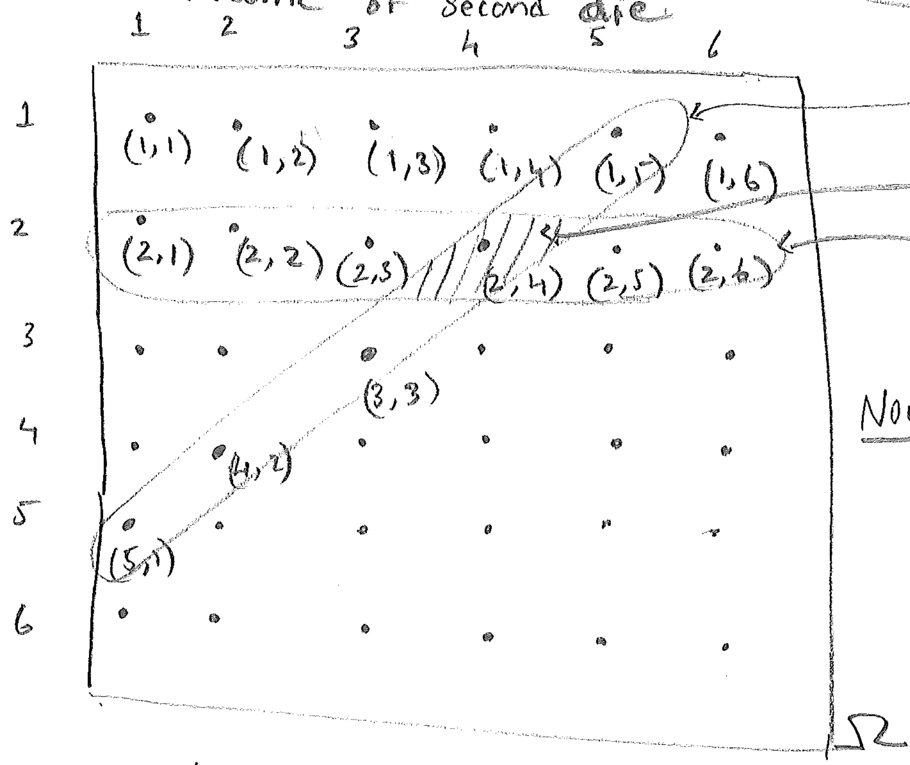
"The occurrence of event A has no influence on the probability of occurrence of B".

Ex 2: Suppose we roll two dice and

A = 'the sum is 6'

B = 'The first die is 2'

outcome of second die



outcome of first dice

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

$A \cap B = \{(2,4)\}$.

$B = \{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\}$

Now:

$$P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5}$$

Makes sense?

Let C = 'The first die is 6'

Then $P(C|A) \overset{?}{=} \frac{P(C\cap A)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0$ makes sense?

Proposition Conditional Probability is a probability.

Suppose $(\Omega, \mathcal{F}(\Omega), P)$ is a probability space, and $A \in \mathcal{F}(\Omega)$ with $P(A) > 0$.

Then $(\Omega, \mathcal{F}(\Omega), P(\cdot|A))$ is a probability space.

i.e., $P(\cdot|A) : \mathcal{F}(\Omega) \to [0,1]$ is a probability

check that

(i) for event $B \in \mathcal{F}(\Omega)$, $0 \le P(B|A) \le 1$ ✓ o.k., since $0 \le P(B\cap A) \le P(A)$.

(ii) $P(\Omega|A) = \frac{P(\Omega\cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$ ✓ ok.

(iii) & (iv) is also ✓ o.k. since if $B_1, B_2, \ldots$ are disjoint then $A\cap B_1, A\cap B_2, \ldots$ are disjoint and $\left(\bigcup_{i=1}^{\infty} B_i\right) \cap A = \bigcup_{i=1}^{\infty}(B_i \cap A)$, so using the dfn. of conditional probability and (iii) & (iv) of the dfn. of probability, we have:

$$P\left(\bigcup_{i=1}^{\infty} B_i \mid A\right) = \frac{P\left(\bigcup_{i=1}^{\infty} (B_i \cap A)\right)}{P(A)} = \frac{\sum_{i=1}^{\infty} P(B_i \cap A)}{P(A)} = \sum_i P(B_i \mid A)$$

Finally, conditional probabilities have the same properties 1 & 2 that ordinary probabilities do.

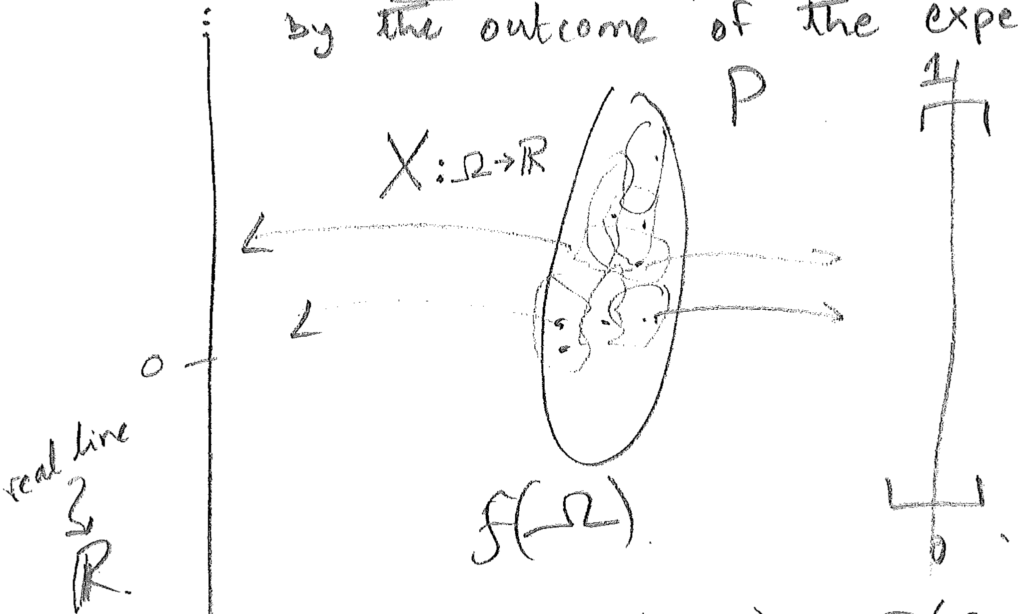
1) $P(B^c \mid A) = 1 - P(B \mid A)$

2) for any two events B_1, B_2

$$P(B_1 \cup B_2 \mid A) = P(B_1 \mid A) + P(B_2 \mid A) - P(B_1 \cap B_2 \mid A).$$

Random Variable (RV)

A random variable is a numerical value determined by the outcome of the experiment.



$$P(X=x) := P(\{\omega : X(\omega)=x\}) = P(X^{[-1]}(x))$$

Ex 1: Roll two dice and let X = the sum of two numbers that appear.

$\Omega = \{(1,1),(1,2),\dots,(1,6),(2,1),\dots(6,6)\}$ $X:\Omega\to\mathbb{R}$, e.g. $X(3,2)=5$, $X(6,6)=12$, $X^{[-1]}(3)=\{(1,2),(2,1)\}$

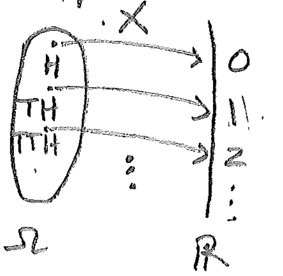
Ex 2. Pull a fruit out of our bowl and give it numerical values. $\Omega = \{\text{apple}, \text{orange}, \text{lemon}\}$, $X(\text{apple})=0$, $X(\text{orange}) = X(\text{lemon}) = 1$

Ex 3 Toss a fair coin until a H appears and let X = # of tosses before H.

$\Omega = \{H, TH, TTH, TTTH, \dots\}$

$X(H)=0$, $X(TH)=1$, $X(TTH)=2$, ...

X

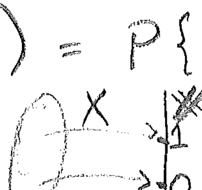


When there is a finite (as in Ex 1 & Ex 2) or countable sequence of values taken by a R.V. it is said to be discrete

Probability Mass Function of a RV X (discrete)

$$f(X=x) = P(X=x) = P\{\omega : X(\omega)=x\} = P(X^{[-1]}(x))$$

Ex :- Bernoulli(θ) R.V.



$f(X=1) = \theta$
$f(X=0) = 1-\theta$

Expectation : $E(g(X)) = \sum_{x\in\mathbb{X}} g(x) f(x)$

Mean $E(X)$
Variance $E(X-E(X))^2$