Continuity property of probability P.

Consider a sequence $A_1, A_2, \ldots \in \mathcal{F}$ (sequence of events).

a) If $A_n \subset A_{n+1}$ for all n (sequence is increasing), then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

b) If $A_n \supset A_{n+1}$ for all n (decreasing sequence), then

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n} P(A_n).$$

Consequence of continuity. Consider $A_1, A_2, \ldots$

Then

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

(ii) $P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcap_{i=1}^{n} A_i\right)$

Why? $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty}\left(\bigcup_{i=1}^{n} A_i\right)$, and $\bigcup_{i=1}^{n} A_i$ is increasing sequence: $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_i\right)$.

Ex 1. Show that $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$

Answer. $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_i\right) \leq$

$$\leq \lim_{n} \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Ex 2. A coin with $P(H) = p$, $P(T) = q = 1-p$ is tossed repeatedly. Consider T_i = "T in the is tossed repeatedly. Consider T_k = "T in the k-th toss". Find

$$P\left(\bigcap_{n=1}^{\infty} T_n\right) = P(T_1 T_2 \ldots)$$

Answer. By continuity of prob. and independence,

$$P\left(\bigcap_{n=1}^{\infty} T_n\right) = \lim_{n\to\infty} P(T_1 \ldots T_n) = \lim_{n\to\infty} q^n = 0.$$

2.1 Random Variables (r.v.)

Often instead of an outcome itself, we are interested in a function of an outcome.

Def. A function $X: \Omega \longrightarrow \mathbb{R}$ is called r.v. if $\{X \le x\} = \{X \in (-\infty, x]\} \in \mathcal{F}$ for all $-\infty < x < \infty$.

Example 1. (Binomial r.v.) A coin with $P(H) = p$, $P(T) = q = 1-p$ is tossed n times.

$\Omega = \{$all "words" of length n made of H and $T\}$.

Consider X, the number of H's in n tosses: $X(\omega) = k$ if ω has exactly k H's and $n-k$ T's.

Range of $X = \{0, 1, \ldots, n\}$

Remark 1. Collection of probabilities $P(X=k)$, $k = 0, 1, \ldots, n$, allows to compute any probability related to X.

Computation of $P(X=k)$

(i) Event $\{X=k\}$ consists of all words with

k H's and $(n-k)$ T's.

(ii) Probability of every such word (assuming independence) is $p^k q^{n-k}$

(iii) There are $\binom{n}{k, n-k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$ distinct words with k H's and $n-k$ T's.

$$\boxed{\text{(iv)} \quad P(X=k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \dots, n}$$

X is called binomial(n,p)-r.v.

the function on $\mathbb{R}$, defined as

$$p(x) = P(X=x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, \dots, n \\ 0, & \text{otherwise} \end{cases}$$

is called probability mass function (pmf) of X.

Remark 2. a) if $p = q = \frac{1}{2}$, then $p^k q^{n-k} = \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} =$

$= \frac{1}{2^n} = 2^{-n}$, $\quad P(X=k) = \binom{n}{k} 2^{-n}$.

b) The simplest binomial $(n=1, p)$-r.v. is called Bernoulli(p)-r.v.:

$$X = \begin{cases} 1 & \text{if } H \\ 0 & \text{if } T \end{cases} \qquad \Big| \quad P(X=1) = p, \; P(X=0) = q = 1-p$$

Note. For an event A with $P(A) = p$,

$$X = 1_A = \begin{cases} 1 & \text{if } A \text{ happens} \\ 0 & \text{otherwise} \end{cases} \quad \text{is Bernoulli}(p)$$

$X = 1_A$ is called indicator r.v.

Example 2 (uniform r.v.) Let $\Omega = (0, 10]$, $\mathcal{F}$ be σ-field of subsets of Ω that have lengths,

$P(A) = \frac{|A|}{10}$, $A \in \mathcal{F}$, where $|A|$ is length of A.
Define $X(\omega) = \omega$, $\omega \in \Omega = (0,10)$: X is J's departure time. Range of $X = (0,10)$.

<u>Properties of X:</u>

(i) $P(X = x) = \frac{|\{x\}|}{10} = \frac{0}{10} = 0$, $0 < x < 10$, does not allow to compute $P(a < X \le b) = \frac{b-a}{10}$, $0 < a < b < 10$.

(ii) $P(X \in A) = P(A) = \frac{|A|}{10}$, $A \in \mathcal{F}$. Hence

$$\frac{P(X \in A)}{|A|} = \frac{P(A)}{|A|} = \frac{1}{10}$$ is <u>probability density</u>

(probability per unit length along possible values of X). Since it is <u>constant</u>, all values are <u>equally likely</u>.

(iii) For $0 < a < b < 10$,

$$P(a < X \le b) = \frac{b-a}{10} = \int_a^b \frac{1}{10}\,dx = \frac{1}{10}(b-a).$$

Also, $P(X \in A) = \frac{|A|}{10} = \int_A \frac{1}{10}\,dx = \frac{1}{10}|A|$, $A \in \mathcal{F}$.

(iv) the function on $\mathbb{R}$ defined as

$$\boxed{f(x) = \begin{cases} \frac{1}{10} & 0 < x < 10 \\ 0 & \text{otherwise} \end{cases}}$$

can be used to find all probabilities related to X:

$$P(X \in A) = \int_A f(x)\,dx,$$

$$P(a < X \le b) = \int_a^b f(x)\,dx, \quad -\infty < a \le b < \infty$$

$f(x)$ is called <u>probability density function</u> (pdf) of X.
X is called <u>uniform r.v.</u> in $(0,10)$

<u>Remark 3</u>. Compare uniform r.v. X to binomial(n,p), Y:

$$P(a < Y \le b) = \sum_{a < k \le b} P(Y = k) = \sum_{a < y \le b} P(Y = y)$$

<u>Note</u>: a) Y does not have pdf, role of <u>pdf</u> is played by <u>pmf</u>.
b) For uniform r.v. X, pmf is useless, role of pmf is played by pdf.

<u>Universal</u> function containing all info about probabilities related to a <u>general r.v.</u> X is <u>distribution function</u> (df or cdf) of X defined as

$$F(x) = P(X \le x), \quad -\infty < x < \infty.$$