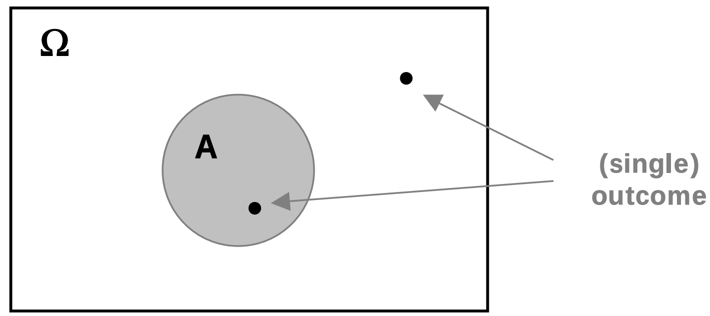
Short Tutorial on Probability and Random Variables

http://www.site.uottawa.ca/~nvlajic/ProbabilityTutorial.pdf

(January 2003, by N. Vlajic)

1. Probability

Let us envision an experiment, for which the result is unknown.



Def. 1 **Sample Space** (Ω) – collection of all possible outcomes.

Def.2 **Event** (A) – set of outcomes i.e. subset of the Sample Space.

Example: **experiment "measurement of voltage"**

1) measured voltage = 1 [V] - **outcome !**
2) measured voltage negative- **event !**

Def. 3 **Probability Space** - three-tuple (Ω, F, Pr), where

- Ω is a sample space
- F is a collection of events from the sample space (event space)
- Pr is a probability measure (law) that assigns a number to each event in F

Furthermore, Pr must satisfy, for $\forall$ A, B $\in F$ the following conditions

(1) $\Pr(A) \geq 0$ - **probability is a positive measure**

(2) $\Pr(\Omega) = 1$ - **probability is a finite measure**

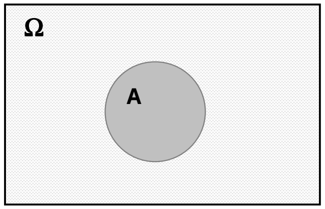
(3) A, B are disjoint events $\Rightarrow$ $\Pr(A+B) = \Pr(A) + \Pr(B)$ - **additive property**

(1), (2) and (3) are known as axioms of probability measure Pr.

From the above axioms, the following well-known properties of probability measure (Pr) are derived:

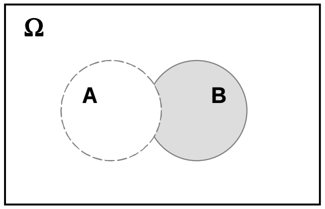
Property 1 $\Pr(A^C) = 1 - \Pr(A)$

Proof: $\Pr(\Omega) = \Pr(A+A^C) = \Pr(A^C) + \Pr(A) = 1$



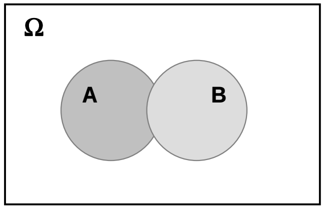
Property 2 $\Pr(B-A) = \Pr(B) - \Pr(AB)$

Proof: $\Pr(B) = \Pr(B-A+(AB)) = \Pr(B-A) + \Pr(AB)$



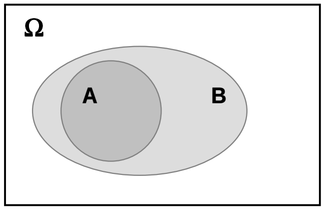
Property 3 $\Pr(A+B) = \Pr(A) + \Pr(B) - \Pr(A-B)$

Proof: $\Pr(A+B) = \Pr(A+(B-A)) = \Pr(A) + \Pr(B-A)$



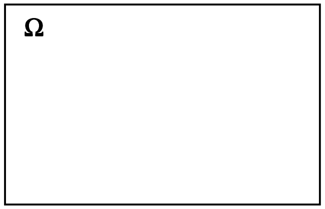
Property 4 if $A \subseteq B \Rightarrow \Pr(A) \leq \Pr(B)$

Proof: $\Pr(B) = \Pr((B-A)+A) = \Pr(B-A) + \Pr(A) \geq \Pr(A)$



Property 5 $\Pr(\varnothing) = 0$

Proof: $\Pr(\varnothing) = \Pr(\Omega^C) = 1 - \Pr(\Omega) \geq 0$



‘Relative Frequency’ Definition of Probability

Perform an experiment a number of times/trials (n), counting the occurrences of event A (n_A). Then the probability P(A) of event A can be found/defined as the limit:

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

Def. 4 **Conditional Probability** enables us to determine whether two events, A and B, are related in the sense that knowledge about the occurrence of one alters the likelihood of occurrence of the other.

probability of A given B has occurred:

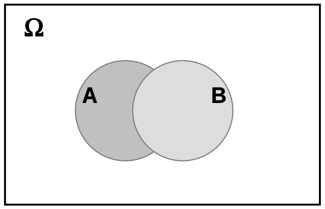
$$\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)}$$

A occurs, in reduced event space, only if A∩B occurs

reduced event space

consequently:

$$\Pr(AB) = \Pr(A|B) \cdot \Pr(B)$$



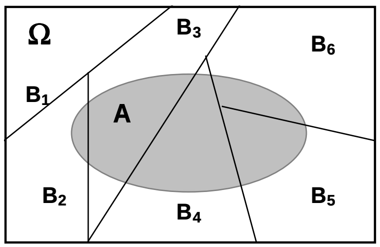
Def. 5 **Independent Events** - two events, A and B, are independent if

$$\Pr(AB) = \Pr(A) \cdot \Pr(B)$$

From Def. 4 and Def. $\Rightarrow$ A and B are independent if $\Pr(A|B)=\Pr(A)$.

Theorem 1 **Total Probability**

Let B_1, .., B_n be mutually exclusive events whose union equals the sample space Ω.

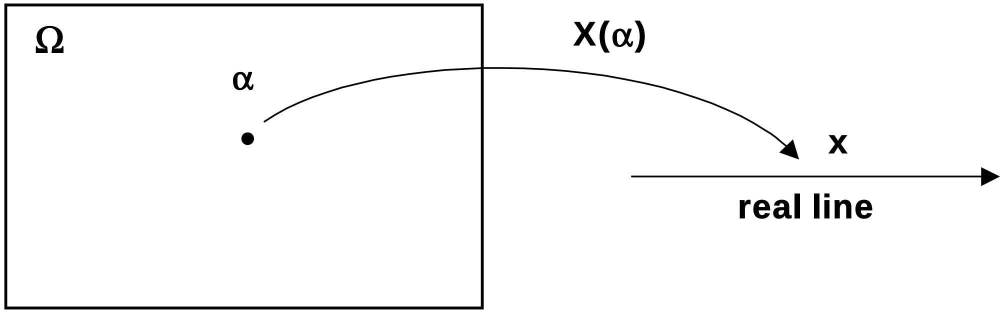


Then, the probability of any given event $A \subseteq \Omega$ can be expressed as

$$\Pr(A) = \Pr(A|B_1) \cdot \Pr(B_1) + \Pr(A|B_2) \cdot \Pr(B_2) + .. + \Pr(A|B_n) \cdot \Pr(B_n)$$

Proof: based on $A = A \cap \Omega = A(B_1 + B_2 + .. + B_n) = AB_1 + AB_2 + .. + AB_n$

Def. 6 **Random Variable** (X) is a function that assigns a real number ($X(\alpha)$) to each outcome α in the sample space.



Example: **experiment "measurement of voltage"**

α = (measured voltage = x [V]) - **outcome !**

$X(\alpha) = x$ - **r.v. associated with experiment !**

Def. 7 **Continuous Random Variable** - takes on an uncountablely infinite number of distinct values.

Def. 8 **Discrete Random Variable** - takes on a finite or countably infinite number of distinct values.

Def. 9 **Cumulative Distribution Function (cdf) $F_X(x)$** of a random variable X is defined as the probability of the event $\{X \leq x\}$.

$$F_X(x) = \Pr[X \leq x]$$

consequently:

$$\Pr[a < X \leq b] = F_X(b) - F_X(a)$$

$$\Pr[X > x] = 1 - F_X(x)$$

Def. 10 **Probability Density Function (pdf) $f_X(x)$**, if it exists, is defined as a derivative of $F_X(x)$.

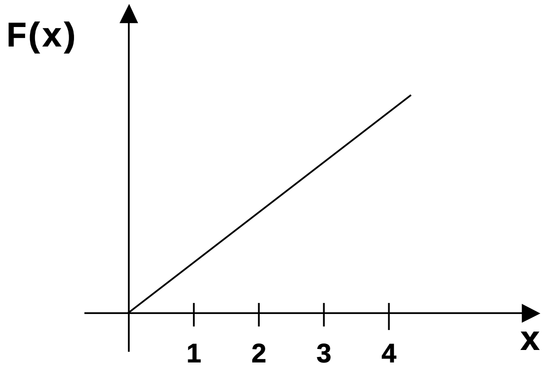
$$f_X(x) = \frac{d\,F_X(x)}{dx}$$

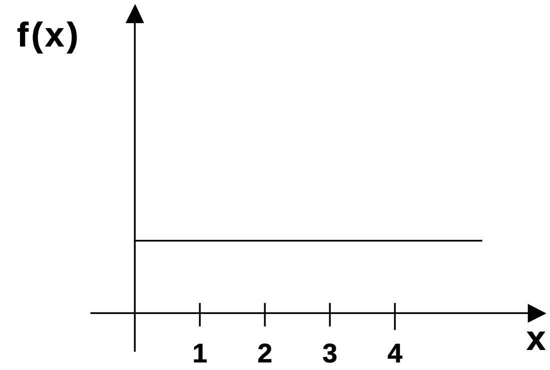
consequently:

$$\Pr[X \leq a] = F_X(a) = \int_{-\infty}^{a} f_X(x)dx$$

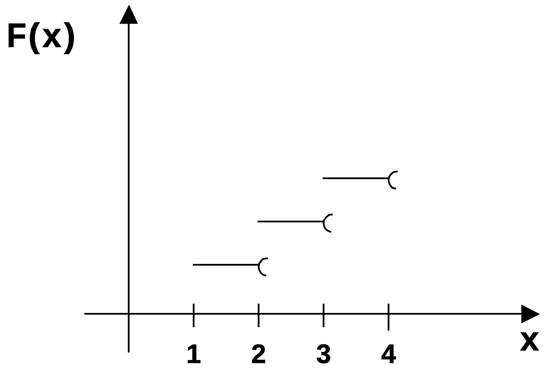
Note: $f_X(x)$ is called "density of probability" at point x, since the probability that X is in a small interval in the vicinity of x is approximately $f_X(x) \cdot x$.

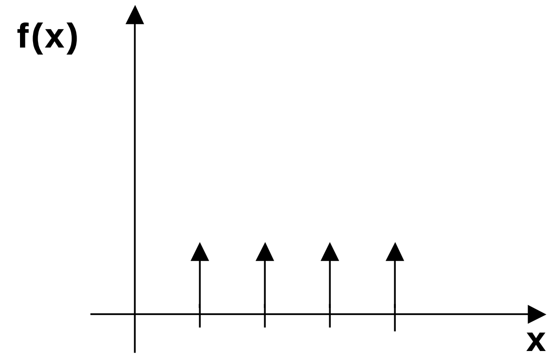
example: **cdf & pdf of a continuous r.v.**





example: **cdf & pmf (prob. mass func.) of a discrete r.v.**





We are often concerned with some characteristic of a random variable rather than the entire distribution:

<u>Def. 11</u> **mean of continuous r.v.**

$$E[X] = \mu_X = \int_{-\infty}^{\infty} x \cdot f_X(x)dx$$

<u>Def. 12</u> **mean of discrete r.v.**

$$E[X] = \sum_{\text{all } k} k \cdot \Pr(x = k)$$

important properties: $E[a \cdot X] = a \cdot E[X]$ **and** $E[X+Y] = E[X] + E[Y]$

<u>Def. 13</u> **variance (dispersion around mean)**

$$\mathrm{Var}[X] = E[(X-\mu_X)^2] = E[X^2] - \mu_X{}^2$$

<u>Def. 14</u> **standard deviation**

$$\sigma_x = \sqrt{\mathrm{Var}[X]}$$

important properties:

$$\mathrm{Var}[a \cdot X] = a^2 \cdot \mathrm{Var}[X]$$

Exponential Distribution - important in queueing theory because we often assume that the service/waiting time is exp. Distributed.

$$F_X(x) = 1 - e^{-\mu x}$$

$$f_X(x) = \mu e^{-\mu x}$$

$$E[X] = \mu$$

properties:

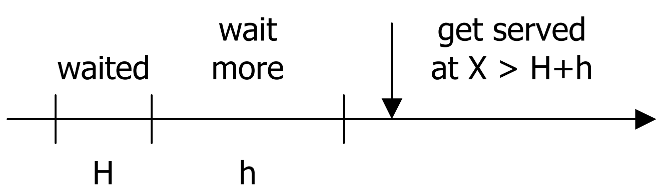
1) 'random distribution'

for service time with exp. distributed duration – each time at which service may finish is equally likely

2) memoriless property

for waiting time with exp distributed duration – the probability of waiting additional h [sec] is the same regardless of how long one has already been waiting

Proof: what is the prob. of waiting another h [sec], given H [sec] already waiting



$$\Pr[X \geq h + H \mid X \geq H] = \frac{\Pr[X \geq h + H \cap X \geq H]}{\Pr[X \geq H]} = \frac{\Pr[X \geq h + H]}{\Pr[X \geq H]} = \ldots = \Pr[X \geq h]$$

Poisson Process - important in queueing theory because we often assume that users/packets arrive according to a Poisson process.

$$\Pr[N(t)] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

(λ - expected # of arrivals per time unit)

$$E[N(t)] = \lambda t$$

properties: **1) exp. distributed inter arrival times**

interarrival times in a Poisson process form an iid sequence of <u>exponential random variables</u> with mean $1/\lambda$