Probability measure is a function that assigns a number between 0 and 1 to "nice" sets in the sample space, subject to the axioms of probability (WMS p.30 - below).

Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, $P(A)$, called the *probability* of A, so that the following axioms hold:

Axiom 1: $P(A) \geq 0$.
Axiom 2: $P(S) = 1$.
Axiom 3: If $A_1, A_2, A_3, \ldots$ form a sequence of pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \emptyset$ if $i \neq j$), then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

A random variable (rv) is a function from the sample space to the real line.

(Cumulative) distribution function (cdf) F of a rv X is defined as $F(x) = Pr(X \leq x)$ for every $x \in \mathbb{R}$.

Properties

- F is nondecreasing
- $F(x) \in [0, 1]$ for every valid x
- $F(\infty) = 1$.

Q: Why?

\infty

A rv X is called *discrete* iff there are at most *countably* many values $x_1, x_2, \ldots$ such that $\sum_{i=1}^{n} Pr(X = x_i) = 1$.

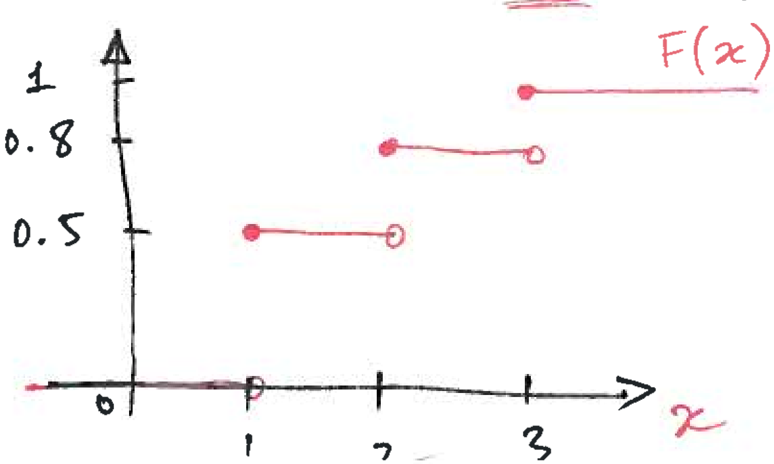
Aside: a set is *countable* if it can be enumerated ("counted") using positive integers.

$$F(x) = Pr(X \le x)$$
$$= Pr\Big(\bigcup_{i: x_i \le x} [X = x_i]\Big) = \sum_{i: x_i \le x} Pr(X = x$$

The function defined as $f(x_i) = Pr(X = x_i)$ is called the *probability mass function (pmf)* of the rv X.

x	$f(x)$
1	0.5
2	0.3
3	0.2

Q: How does the cdf of any discrete rv look like?



cdf $F(x)$ is made entirely of "jumps".

Some standard discrete distributions/rvs (WMS)

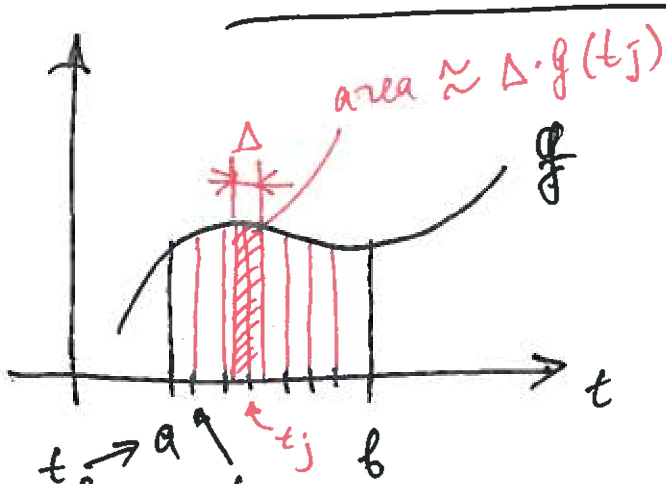
Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y};$ $y = 0, 1, \ldots, n$	np	$np(1-p)$	$[pe^t + (1-p)]^n$
Geometric	$p(y) = p(1-p)^{y-1};$ $y = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}};$ $y = 0, 1, \ldots, n$ if $n \leq r$, $y = 0, 1, \ldots, r$ if $n > r$	$\frac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!};$ $y = 0, 1, 2, \ldots$	λ	λ	$\exp[\lambda(e^t - 1)]$

mass

A rv X is called <u>*continuous*</u> if $Pr(X = x) = 0$ for every $x \in \mathbb{R}$.

Q: How does the cdf of any continuous rv look like?

A: continuous non-decreasing function.
$\Rightarrow$ "jumps are not allowed".



Numerical integration:

in R, use function "integrate".

want $\displaystyle\int_a^b g(t)\,dt \equiv$ area under the curve (AUC) from a to b

$t_0 = a$

$t_{j+1} = t_j + \Delta$

$\approx \displaystyle\sum_j g(t_j)\cdot\Delta$

If X is a continuous rv and has differentiable cdf F, then the <u>*probability density function (pdf)*</u> of X is defined as

$$f(x) = \frac{\partial}{\partial x}F(x). \quad \Rightarrow F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$\approx \frac{F(x+\Delta) - F(x)}{\Delta}$$

Some standard continuous distributions/rvs (WMS)

density

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Uniform	$f(y) = \frac{1}{\theta_2 - \theta_1}; \theta_1 \le y \le \theta_2$	$\frac{\theta_1 + \theta_2}{2}$	$\frac{(\theta_2 - \theta_1)^2}{12}$	$\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$
Normal	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$	μ	σ^2	$\exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta}; \quad \beta > 0$ $0 < y < \infty$	β	β^2	$(1 - \beta t)^{-1}$
Gamma	$f(y) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1} e^{-y/\beta};$ $0 < y < \infty$	$\alpha\beta$	$\alpha\beta^2$	$(1 - \beta t)^{-\alpha}$

Indicator function h of a set or event A is defined as $h(x) = \mathbb{I}(x \in A)$, where

$$\mathbb{I}(x \in A) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A. \end{cases}$$