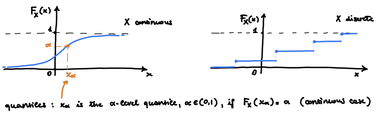
PREREQUISITES of PROBABILITY

- **RANDOM VARIABLE** X is a measurable function $X: \Omega \to \mathbb{R}$
 Ω is the **SAMPLE SPACE** : set of possible outcomes
 notation: uppercase for random variables (e.g. $X, Y, \dots$)
 lowercase for the realization (number) ($x, y, \dots$)
 $\Rightarrow$ e.g. $P(X=x)$ (X: r.v., x: value that assumes)
- **SUPPORT** of X (S_X): is the set of all the values X takes
 - if we can enumerate the values in $S_X \to$ X is DISCRETE
 - if X takes values in a real interval $\to$ X is CONTINUOUS

example: we roll two fair 6-sided dice. $\Omega = \{\underbrace{(1,1),(1,2),(1,3),\dots,(1,6),(2,1),\dots,(6,5),(6,6)}_{\text{all 36 combinations}}\}$
X = "sum of the outcomes of the two dice"
$S_X = \{2,3,4,5,\dots,10,11,12\}$ ($\to$ discrete)
$P(X=2) = 1/36$ $\quad P(X \le 12) = 1$
$P(X=1) = 0$

- the **CUMULATIVE DISTRIBUTION FUNCTION** (CDF) $F_X(x) = P(X \le x)$
 right-continuous | monotone increasing | $\lim_{x\to-\infty} F_X(x) = 0$; $\lim_{x\to+\infty} F_X(x) = 1$



- quantiles: x_α is the α-level quantile, $\alpha \in (0,1)$, if $F_X(x_\alpha) = \alpha$ (continuous case)

- discrete r.v.'s: the **PROBABILITY FUNCTION** $p_X(x) = P(X=x)$ for all $x \in S_X$
- continuous r.v.'s: **DENSITY FUNCTION** $f_X(x)$ $x \in \mathbb{R}$
 $f_X(x) \ge 0$ s.t. $F_X(x) = \int_{-\infty}^{x} f_X(u)\,du$

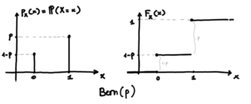
- **EXPECTED VALUE** of a r.v. $\mathbb{E}[X]$
 X discrete $\mathbb{E}[X] = \sum_{x \in S_X} x \cdot P(X=x)$
 X continuous $\mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, f_X(x)\,dx$
 X r.v., $a, b \in \mathbb{R}$: LINEARITY $\mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b$
- **VARIANCE** $\mathrm{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
 variance of a linear transformation $\mathrm{var}(aX+b) = a^2\,\mathrm{var}(X)$

IMPORTANT PROBABILITY DISTRIBUTIONS

→ DISCRETE

- **BERNOULLI**

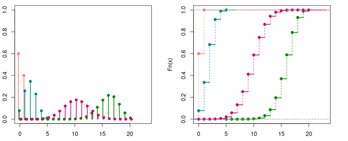
distribution of a binary variable. It models experiments with only two outcomes (e.g. toss of a coin).
support $S_X = \{0,1\}$
parameter $\pi \in [0,1]$ probability of success
$X \sim \mathrm{Bern}(\pi)$ $\quad p_X(x) = P(X=x) = \pi^x (1-\pi)^{1-x}$ if $x \in S_X$
$= \begin{cases} \pi & \text{if } x=1 \\ 1-\pi & \text{if } x=0 \end{cases}$
$\mathbb{E}[X] = \pi$ $\quad \mathrm{var}(X) = \pi(1-\pi)$



Bern(p)

- **BINOMIAL**

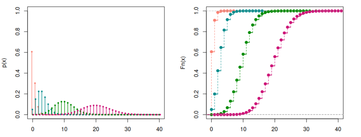
distribution of the number of successes in a sequence of n independent binary experiments (e.g. n tosses of a coin)
support $S_X = \{0,1,\dots,n-1,n\}$
parameters: $\pi \in (0,1)$ success probability
$n \in \{0,1,2,\dots\}$ number of trials
$X \sim \mathrm{Bi}(n,\pi)$
$p_X(x) = P(X=x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$ for $x \in S_X$
$\mathbb{E}[X] = n\pi$ $\quad \mathrm{var}(X) = n\pi(1-\pi)$



- The Bernoulli is a special case with $n=1$.
- Binomial from a sequence of n independent Bernoulli r.v.'s with the same success probability π:
 $X_i \sim \mathrm{Bern}(\pi)$ indep. for $i=1,\dots,n$ $\Rightarrow$ $X = \sum_{i=1}^{n} X_i$ $\quad X \sim \mathrm{Bi}(n,\pi)$

- **POISSON**

distribution to model counts
support $S_X = \{0,1,2,\dots\}$
parameter $\lambda \in (0,+\infty)$ rate
$X \sim \mathrm{Pois}(\lambda)$
$p_X(x) = P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for $x \in S_X$
$\mathbb{E}[X] = \mathrm{var}(X) = \lambda$

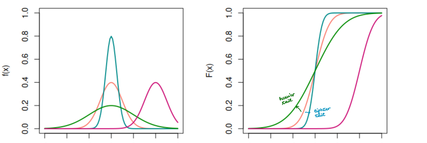


→ CONTINUOUS

- **GAUSSIAN / NORMAL**

support $S_X = \mathbb{R}$
parameters $\mu \in \mathbb{R}$ mean
$\sigma^2 \in (0,+\infty)$ variance
$X \sim N(\mu, \sigma^2)$
$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $\quad x \in \mathbb{R}$

- $\mathbb{E}[X] = \mu$ $\quad \mathrm{var}(X) = \sigma^2$
- unimodal, symmetric around μ
- closed under linear transformations: $X \sim N(\mu, \sigma^2)$; $a, b \in \mathbb{R}$
 $\Rightarrow$ $aX+b$ is normal $aX+b \sim N(a\mu+b,\ a^2\sigma^2)$



- **STANDARD NORMAL**

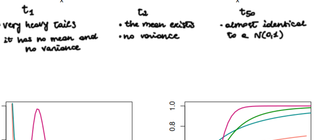
special case with $\mu = 0$ and $\sigma^2 = 1$
usually denoted with $Z \sim N(0,1)$
density $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ $\quad z \in \mathbb{R}$
CDF $\Phi(z) = P(Z \le z)$
$\mathbb{E}[Z] = 0$ $\quad \mathrm{var}(Z) = 1$

- "general" normal from a standard normal
 $X \sim N(\mu, \sigma^2) \iff X = \mu + \sigma Z$ with $Z \sim N(0,1)$
 indeed, $\mathbb{E}[X] = \mathbb{E}[\mu + \sigma Z] = \mu + \sigma\,\mathbb{E}[Z] = \mu$
 $\mathrm{var}(X) = \mathrm{var}(\mu + \sigma Z) = \sigma^2\,\mathrm{var}(Z) = \sigma^2$
 cdf of X $F(x) = P(X \le x) = P(\mu + \sigma Z \le x) = P\left(Z \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

- notable related distributions

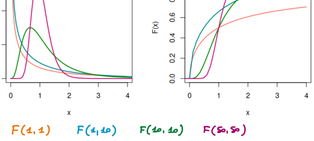
- **CHI SQUARED**

If $Z \sim N(0,1)$, then $V = Z^2$ $V \sim \chi^2_1$ chi-squared with 1 degree of freedom (d.o.f.)
If $Z_1, \dots, Z_k$ are independent standard normal r.v.'s, $V = \sum_{i=1}^{k} Z_i^2$ $V \sim \chi^2_k$ k dof.
support $S_V = (0,+\infty)$
parameter $k \in \{1,2,3,\dots\}$ degrees of freedom
$\mathbb{E}[V] = k$ $\quad \mathrm{var}(V) = 2k$



- **STUDENT'S T**

If $Z \sim N(0,1)$ and $V \sim \chi^2_k$ independent, then $T = \frac{Z}{\sqrt{V/k}}$ $\quad T \sim t_k$
t distribution with k degrees of freedom
support $S_T = \mathbb{R}$
parameter $k > 0$ degrees of freedom
$\mathbb{E}[T] = 0$ if $k > 1$ $\quad \mathrm{var}(T) = \begin{cases} \frac{k}{k-2} & \text{if } k > 2 \\ +\infty & \text{if } k = 1, 2 \end{cases}$



- **F DISTRIBUTION**

If $V_1 \sim \chi^2_{k_1}$ and $V_2 \sim \chi^2_{k_2}$ independent, then $Q = \frac{V_1/k_1}{V_2/k_2}$ $\quad Q \sim F_{k_1,k_2}$
F-distribution with k_1 and k_2 degrees of freedom.
support $S_Q = (0,+\infty)$
parameters $k_1, k_2 > 0$ degrees of freedom