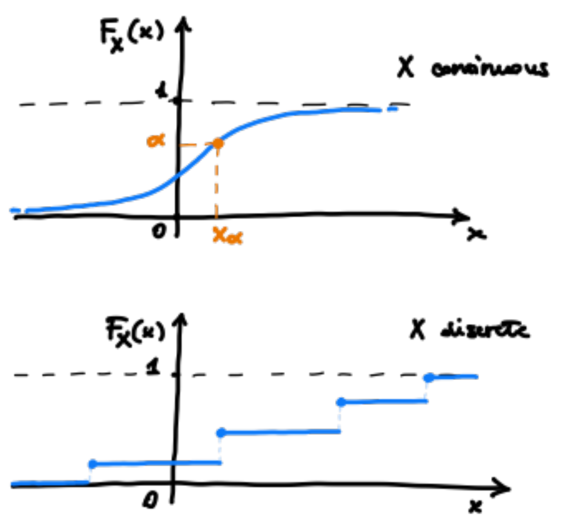
1 oct 2024 - LEC 1

PREREQUISITES of PROBABILITY

- random variable $X : \Omega \to \mathbb{R}$ Ω "sample space"

notation: uppercase for random variables (e.g. $X, Y, \dots$)
lowercase for the realisation (number) $(x, y, \dots)$ } ⇒ e.g. $\mathbb{P}(X = x)$ (X: r.v.; x: value that assumes)



- the cumulative distribution function (CDF) $F_X(x) = \mathbb{P}(X \le x)$
  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 $p_X(x) = \mathbb{P}(X = x)$
- continuous r.v.'s : density function $f_X(x)$
  $f_X(x) \ge 0$ s.t. $F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$
- expected value of a r.v. $\mathbb{E}[X]$
  X discrete $\mathbb{E}[X] = \sum_{x \in S_X} x \cdot \mathbb{P}(X = x)$
  X continuous $\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\, dx$
  X r.v., a, b constants: LINEARITY $\mathbb{E}[aX + b] = a\,\mathbb{E}[X] + b$
- variance $var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
  variance of a linear transformation $var(aX + b) = a^2\, var(X)$

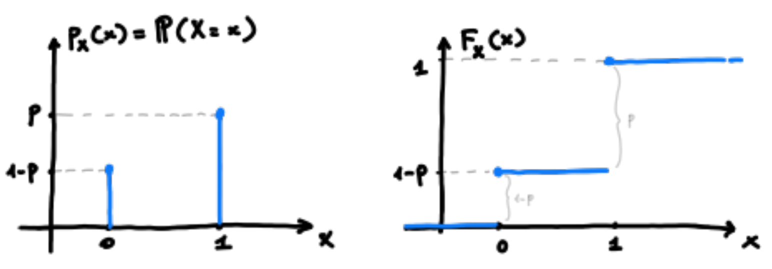
IMPORTANT PROBABILITY DISTRIBUTIONS

→ DISCRETE

- **BERNOULLI**

distribution of a single binary variable (e.g. toss of a coin)
support $S_X = \{0, 1\}$
parameter $\pi \in [0,1]$ probability of success
$X \sim Bern(\pi)$ $p_X(x) = \mathbb{P}(X = x) = \pi^x (1-\pi)^{1-x}$ if $x \in S_X$
$= \begin{cases} \pi & \text{if } x = 1 \\ 1-\pi & \text{if } x = 0 \end{cases}$

$\mathbb{E}[X] = \pi$ $var(X) = \pi(1-\pi)$

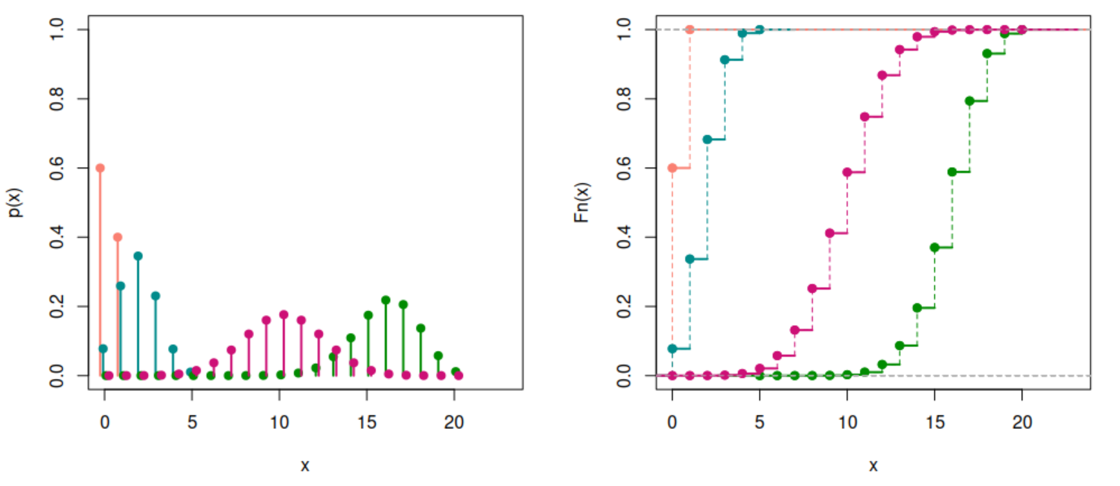


Ber(p)

- **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

$X \sim Bi(n, \pi)$
$p_X(x) = \mathbb{P}(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$ for $x \in S_X$
$\mathbb{E}[X] = n\pi$ $var(X) = n\pi(1-\pi)$



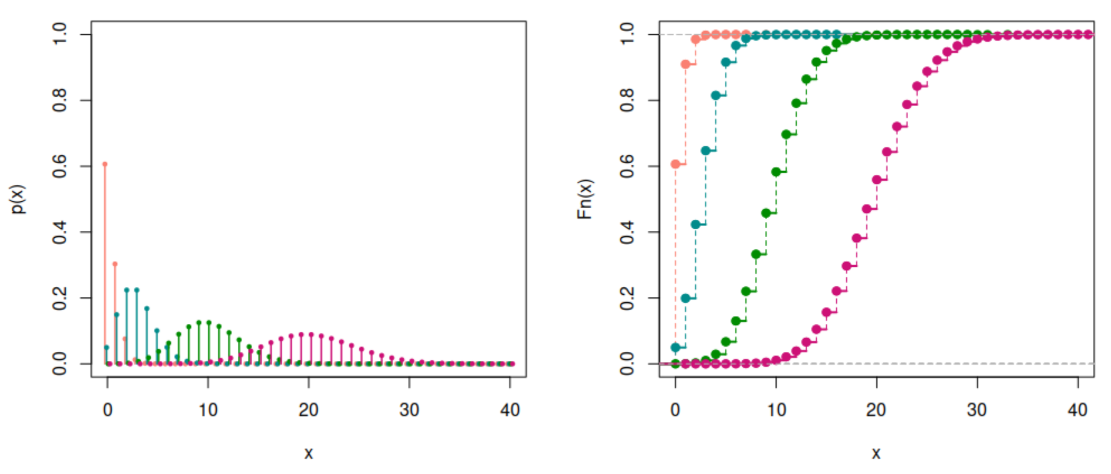
Bi(1, 0.4) = Bern(0.4)   Bi(5, 0.4)   Bi(20, 0.5)   Bi(20, 0.8)

Bi(20, 0.5) → symmetric; Bi(20, 0.5), Bi(20, 0.8): same support $S_X = \{0, 1, \dots, 20\}$

- 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 :
  $X_i \sim Bern(\pi)$ indep. for $i = 1, \dots, n$ ⇒ $X = \sum_{i=1}^{n} X_i$ $X \sim Bi(n, \pi)$

- **POISSON**

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



Pois(0.5)   Pois(3)   Pois(10)   Pois(20)

↓ they all have the same support $S_X = \{0, 1, 2, \dots\}$
↳ the larger the rate, the more symmetric the distribution

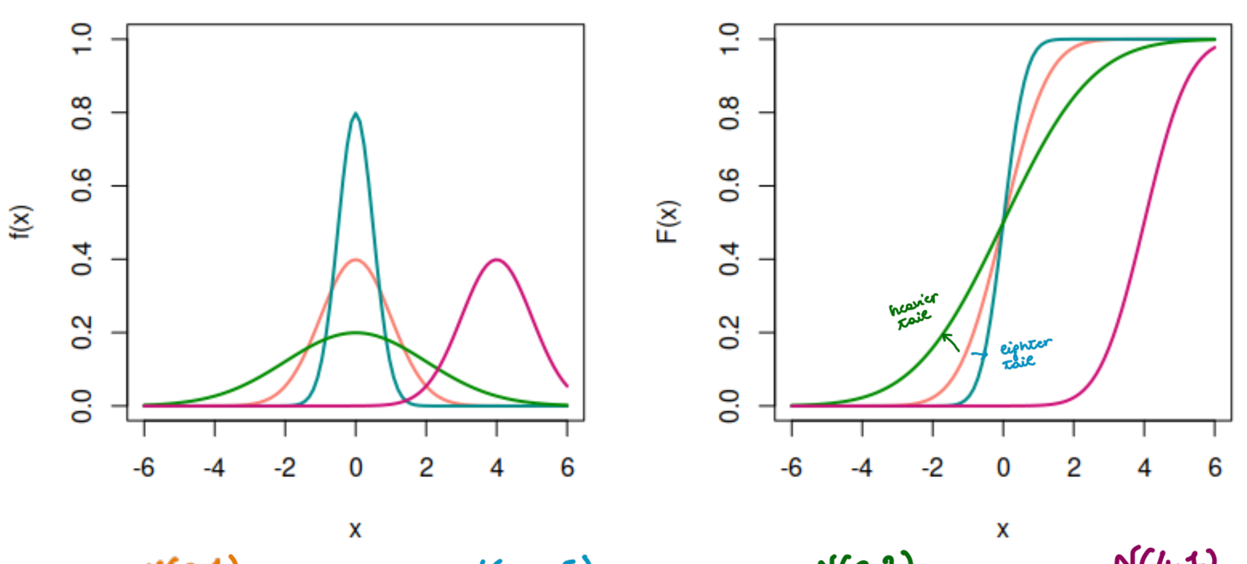
→ CONTINUOUS

- **GAUSSIAN / NORMAL**

support $S_X = \mathbb{R}$
parameters $\mu \in \mathbb{R}$ mean
$\sigma^2 \in (0, +\infty)$ variance

$X \sim N(\mu, \sigma^2)$
$f_X(x) = \phi_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $x \in \mathbb{R}$

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



N(0,1) ↓
- symmetric around 0
- has most of its mass between -3 and 3

N(0, 0.5) ↓
- symmetric around 0
- more concentrated

N(0, 2) ↓
- symm. around 0
- more spread

N(4, 1) ↓
- symmetric around 4

- **STANDARD NORMAL**

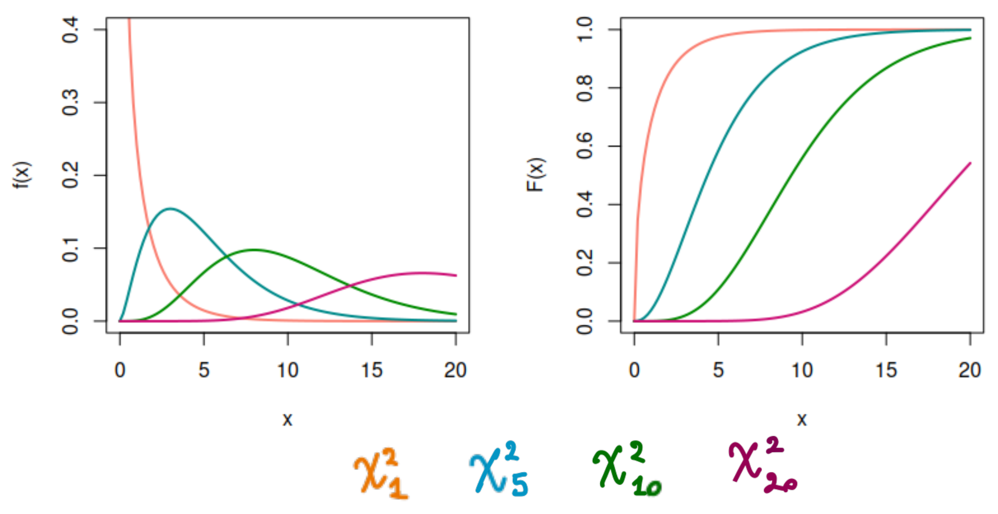
special case with $\mu = 0$ and $\sigma^2 = 1$
usually denoted with $Z \sim N(0,1)$
density $\phi_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ $z \in \mathbb{R}$
CDF $\Phi_Z(z) = \mathbb{P}(Z \le z)$
$\mathbb{E}[Z] = 0$ $var(Z) = 1$

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

- 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 d.o.f.
support $S_V = (0, +\infty)$
parameter $k \in \{1, 2, 3, \dots\}$ degrees of freedom
$\mathbb{E}[V] = k$ $var(V) = 2k$

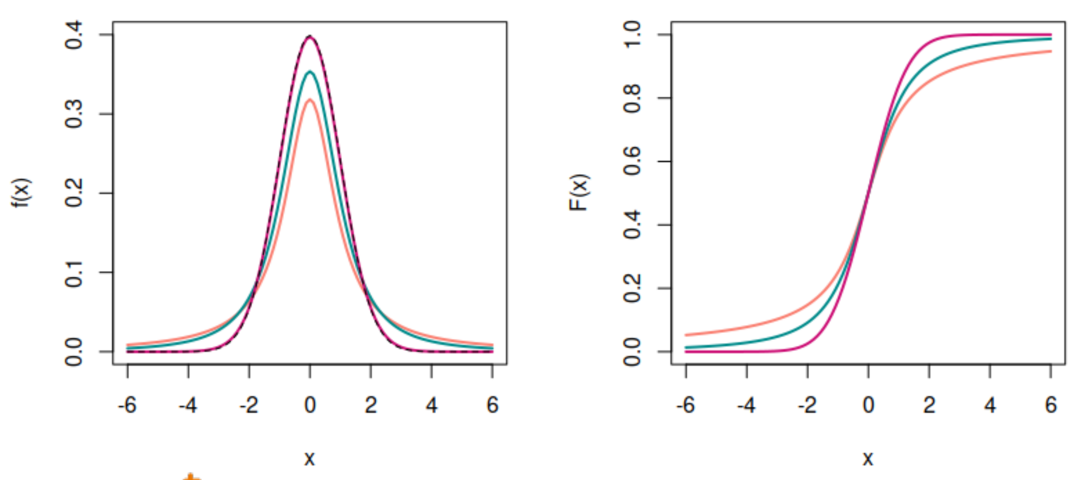


χ^2_1 χ^2_5 χ^2_{10} χ^2_{20}

increasing the d.o.f. I get larger mean and larger variance

- **STUDENT'S T**

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



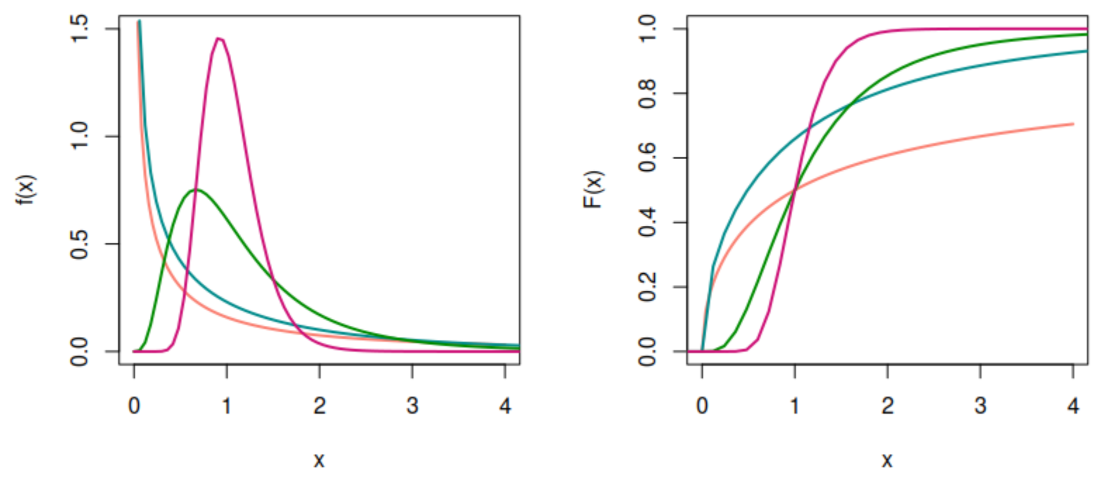
t_1
- very heavy tails
- it has no mean and no variance

t_2
- the mean exists
- no variance

t_{50}
- almost identical to a N(0,1)

- **F DISTRIBUTION**

If $V_1 \sim \chi^2_{k_1}$ and $V_2 \sim \chi^2_{k_2}$ independent, then $Q = \frac{V_1 / k_1}{V_2 / k_2}$ $Q \sim F_{k_1, k_2}$
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



F(1, 1)   F(1, 10)   F(10, 10)   F(50, 50)