Probability density function (p.d.f.) (2 conditions)

$$\begin{cases} p(x) \geq 0 \\ \int_{-\infty}^{+\infty} p(x)dx = 1 \end{cases}$$

Cumulative distribution function (c.d.f.) (definition)

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(y)dy$$

Properties of p.d.f. and c.d.f.

$$P(X = x) = 0 \qquad p(x) = F'(x) \qquad P(X \geq x) = 1 - P(X \leq x) = 1 - F(x)$$

$$\boxed{\boldsymbol{P(X \in [a; b])} = \int_{a}^{b} p(x)dx = F(x)|_a^b = \boldsymbol{F(b) - F(a)}}$$

Independent continuous random variables (definition)

$$P\big((X \leq x_i) \cap (Y \leq y_j)\big) = P(X \leq x_i) \cdot P\big(Y \leq y_j\big) \text{ for all } (x_i, y_j)$$

Expectation of a continuous random variable (definition)

$$\mathrm{E}X = \int_{-\infty}^{+\infty} xp(x)dx$$

Variance (definition)

$$\mathrm{Var}\, X = \int_{-\infty}^{+\infty} (x - \mathrm{E}X)^2 p(x)dx = \mathrm{E}(X^2) - (\mathrm{E}X)^2$$

Standard deviation (definition)

$$\sigma X = \sqrt{\mathrm{Var}\, X}$$

Clarification. Properties of characteristics of discrete random variables are hold for continuous ones.

Uniform distribution on $[a; b]$

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a; b] \\ 0, & x \notin [a; b] \end{cases} \qquad F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a; b] \\ 1, & x > b \end{cases}$$

Properties of the uniform distribution

$$\mathrm{E}X = \frac{a+b}{2} \qquad \mathrm{Var}\, X = \frac{(b-a)^2}{12}$$

$$X \sim \text{uniform on } [a; b];\ Y = cX + d \Rightarrow Y \sim \text{uniform on } [ca + d; cb + d]$$

Exponential distribution with parameter $\lambda > 0$

$$p(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases} \qquad F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

Properties of the exponential distribution

$$\mathrm{E}X = \frac{1}{\lambda} \qquad \mathrm{Var}\,X = \frac{1}{\lambda^2}$$

$P(X > t + s \mid X > t) = P(X > s)$ for any $t, s \geq 0$ ("absence of memory")

Standard normal distribution defined by its p.d.f. $\varphi(z)$

$$\boxed{\boldsymbol{\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}}}$$

Clarifications. z denotes a value of a standard normal random variable Z.
Notation: $Z \sim N(0, 1^2)$. Notation for c.d.f.: $\Phi(z)$.

Properties of standard normal variable

$$\varphi(z) = \varphi(-z) \qquad \mathrm{E}Z = 0 \qquad \mathrm{Var}\,Z = 1$$

General normal distribution of a general normal random variable $X = \sigma Z + \mu$

$$Z = \frac{X-\mu}{\sigma} \qquad z\text{-score of } x = \frac{x-\mu}{\sigma} \text{ (definition)}$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \boxed{\boldsymbol{F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)}} \qquad \mathrm{E}X = \mu \qquad \mathrm{Var}\,X = \sigma^2$$

Notation: $X \sim N(\mu, \sigma^2)$

Properties of general normal random variables

- $\begin{cases} X \sim N(\mu, \sigma^2) \\ Y = aX + b \end{cases} \Rightarrow Y \sim N(a\mu + b, (a\sigma)^2)$
- $X_i \sim N(\mu_i, {\sigma_i}^2)$ are independent $\Rightarrow \boldsymbol{\sum_i X_i \sim N(\sum_i \mu_i, \sum_i {\sigma_i}^2)}$
  Clarification. A **sum** of normal and independent random variables is again a normal random variable.
- $\begin{cases} X_i \sim N(\mu, \sigma^2) \\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \end{cases} \Rightarrow \boldsymbol{\bar{X}} \sim \frac{1}{n} N(n\mu, n\sigma^2) = \boldsymbol{N\left(\mu, \frac{\sigma^2}{n}\right)}$
  Clarification. X_i are independent and identically distributed (i.i.d.) normal random variables, and $\bar{X}$ is their **arithmetic average**.

One, two and three sigma rules

- $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95$
- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.997$

Assumptions for the general limit theorems

X_i are i.i.d. random variables $\mathrm{E}X_i = \mu < \infty$ $\mathrm{Var}\, X_i = \sigma^2 < \infty$ $S_n = X_1 + \ldots + X_n$

The Law of Large Numbers (general form)

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu$$

Meaning. As the number of trials increases unlimitedly, the mean of random outcomes approaches the population mean.

Justification. $\mathrm{E}(\bar{X}) = \mathrm{E}\left(\frac{S_n}{n}\right) = \frac{\mathrm{E}(S_n)}{n} = \frac{n\mathrm{E}(X_i)}{n} = \mathrm{E}(X_i) = \mu$

$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{\mathrm{Var}(S_n)}{n^2} = \frac{n\mathrm{Var}(X_i)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$

The Central Limit Theorem (general form)

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \Phi(x) \quad \text{or} \quad \frac{S_n - n\mu}{\sigma\sqrt{n}} \to Z \sim N(0, 1^2)$$

Meaning. As the number of trials increases unlimitedly, the normalized sum of random outcomes tends to the standard normal distribution.

Assumptions for the limit theorems for the binomial distribution

$X_i \sim Bernoulli(p)$ $\mathrm{E}X_i = p$ $\mathrm{Var}\, X_i = p(1-p)$ $S_n = X_1 + \ldots + X_n \sim Bin(n, p)$

The Law of Large Numbers for the binomial distribution

$$\lim_{n\to\infty} \frac{S_n}{n} = p$$

The Central Limit Theorem for the binomial distribution

$$\lim_{n\to\infty} P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \le x\right) = \Phi(x) \quad \text{or} \quad \frac{S_n - np}{\sqrt{np(1-p)}} \to Z \sim N(0, 1^2)$$

Normal approximation of the binomial distribution – consequence of the CLT

$$\boxed{\boldsymbol{P(S_n \le m) \approx \Phi\left(\frac{m - np}{\sqrt{np(1-p)}}\right)}}$$

Clarification. This formula is hold only for large n. It is considered large enough if $np \ge 5$ and $n(1-p) \ge 5$.

Student's t-distribution with n degrees of freedom

$$p(\tau) = c_n\left(1+\frac{\tau^2}{n}\right)^{-\frac{n+1}{2}}$$

Clarification. c_n is a constant depending on n. Notation: $T \sim t(n)$.

Properties of the t-distribution

$$p(\tau) = p(-\tau) \qquad \mathrm{E}T = 0,\ n > 1 \qquad \operatorname{Var} T = \frac{n}{n-2},\ n > 2$$

Properties of the t-statistic

- $\begin{cases} X_i \sim N(\mu, \sigma^2) \\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \\ s^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1} \end{cases} \Rightarrow \boxed{\boldsymbol{\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)}}$ (alternative definition of the t-distribution)

  Clarification. s^2 is sample variance.
- $\lim\limits_{n\to\infty} P\left(\frac{\bar{X}-\mu}{s/\sqrt{n}} \le y\right) = \Phi(y)$ or $\frac{\bar{X}-\mu}{s/\sqrt{n}} \to Z \sim N(0, 1^2)$

  Clarification. This property results in $t(n) \to N(0, 1^2)$.

Chi-square distribution with k degrees of freedom

$$Z_i \sim N(0, 1^2) \text{ are independent} \Rightarrow \boxed{\boldsymbol{X = \sum_{i=1}^{k} Z_i^{\,2} \sim \chi^2(k)}}$$

Properties of the χ^2-distribution

$$\mathrm{E}X = k \qquad \operatorname{Var} X = 2k$$

$$\begin{cases} X_1 \sim \chi^2(k_1) \\ X_2 \sim \chi^2(k_2) \end{cases} \Rightarrow (X_1 + X_2) \sim \chi^2(k_1 + k_2)$$