MECA0010 – Reliability and stochastic modeling of engineered systems

Notations and review of background material

Maarten Arnst and Marco Lucio Cerquaglia

September 20, 2017

Outline

- System of notation.
- Fundamentals of probability.
  - Events and probability.
  - Mathematics of probability.
- Stochastic models of random phenomena.
  - Random variables and probability distributions.
  - Useful probability distributions.
  - Multiple random variables.
- Convergence of random variables.

System of notation

- A lowercase letter, for example, x, is a real deterministic variable.
- A boldface lowercase letter, for example, $\boldsymbol{x} = (x_1, \ldots, x_n)$, is a real deterministic vector.
- An uppercase letter, for example, X, is a real random variable. Exceptions: P (probability), Γ (gamma function), and E (expectation operator).
- A boldface uppercase letter, for example, $\boldsymbol{X} = (X_1, \ldots, X_n)$, is a real random vector.
- An uppercase letter between square brackets, for example, $[A]$, is a real deterministic matrix.
- A boldface uppercase letter between square brackets, for example, $[\boldsymbol{A}]$, is a real random matrix.

Fundamentals of probability

- Probability triple $(\mathcal{S}, \mathcal{E}, P)$

  $\mathcal{S}$ “sample space”

  $\mathcal{E}$ “event space”

  P “probability”

- Axioms of probability:

  (1) $P(\mathcal{A}) \geq 0$ for any event $\mathcal{A}$ in $\mathcal{E}$,

  (2) $P(\mathcal{S}) = 1$ for the “certain event” $\mathcal{S}$,

  (3) $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$ for any two mutually exclusive events $\mathcal{A}$ and $\mathcal{B}$ in $\mathcal{E}$.

- Addition rule:

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}).$$

Note that $P(\mathcal{A} \cap \mathcal{B}) = 0$ if $\mathcal{A}$ and $\mathcal{B}$ are mutually exclusive events.

- Complement rule:

$$P(\overline{\mathcal{A}}) = 1 - P(A).$$

- Conditional probability:

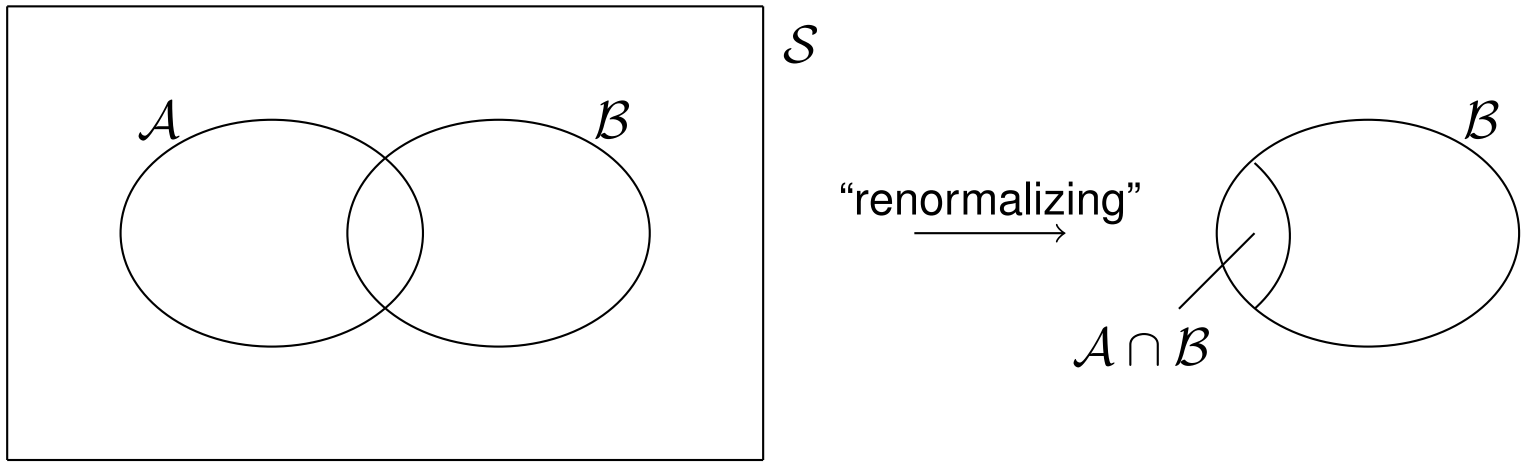
$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} \quad \text{if } P(\mathcal{B}) \neq 0.$$

- Multiplication rule:

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) = P(\mathcal{B}|\mathcal{A})P(\mathcal{A}).$$

- Conditional probability refers to the probability of an event given/dependent on another event:

$$\underset{\text{“given”}}{P(\mathcal{A}\underset{\downarrow}{|}\mathcal{B})} = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} \quad \text{if } P(\mathcal{B}) \neq 0.$$



$P(\mathcal{A}|\mathcal{B})$ is interpreted as the probability of a sample being in $\mathcal{A}$ given that it is in $\mathcal{B}$. Thus, the conditional probability pertains to the samples in $\mathcal{A}$ relative to those of $\mathcal{B}$, and it must thus be normalized with respect to $\mathcal{B}$.

- Two events $\mathcal{A}$ and $\mathcal{B}$ are statistically independent if the occurrence of $\mathcal{A}$ does not affect the probability of $\mathcal{B}$ occurring and vice versa. Thus, $\mathcal{A}$ and $\mathcal{B}$ are statistically indepdendent if

$$\begin{cases} P(\mathcal{A}|\mathcal{B}) = P(\mathcal{A}), \\ P(\mathcal{B}|\mathcal{A}) = P(\mathcal{B}), \end{cases} \quad \text{that is,} \quad P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B}).$$

- Multiplication rule:

$$P(\mathcal{A} \cap \mathcal{B}) = \begin{cases} P(\mathcal{A}|\mathcal{B})P(\mathcal{B}) = P(\mathcal{B}|\mathcal{A})P(\mathcal{A}) & \text{general case,} \\ P(\mathcal{A})P(\mathcal{B}) & \text{if } \mathcal{A} \text{ and } \mathcal{B} \text{ are statistically independent.} \end{cases}$$

Stochastic models of random phenomena

Single random variables

Random variables and probability distributions

- Definition of a (real) random variable:
  - Let us consider a (real) random variable denoted by X (uppercase letter).
  - Sample space $\mathcal{S} = \mathbb{R}$, that is, the sample space is the real line.
  - Event space $\mathcal{E}$ collects events, for example,

$$\begin{aligned}
\mathcal{A} &= \{a < X < b\}, \\
\mathcal{B} &= \{c < X < d\}, \\
\overline{\mathcal{A}} &= \{X \leq a\} \cup \{b \leq X\}, \\
\mathcal{C} &= \{X = a\}.
\end{aligned}$$

  - Probability P assigns probabilities to events. It satisfies the axioms of probability.

Random variables and probability distributions

- The probability P can be described by the cumulative distribution function

upper case

$$c_X(x) = P(X \leq x).$$

lower case

- For a (real) random variable X, the cumulative distribution function c_X is a function from $\mathbb{R}$ into $[0, 1]$, which possesses the following properties owing to the axioms of probabilty:
  - $c_X(x) \geq 0$,
  - c_X is monotonically increasing,
  - $c_X(-\infty) = 0$ and $c_X(+\infty) = 1$.

Random variables and probability distributions

- A discrete random variable can assume only a finite or listable infinite number of real values, e.g.,
  - rolling a dice: possible samples $\{1, 2, 3, 4, 5, 6\}$,
  - numer of microbes on a kitchen table: possible samples $\{0, 1, 2, 3, \ldots\} = \mathbb{N}$.
- The probability assigned to an elementary event can be nonzero for a discrete random variable, e.g.,
  - rolling a dice: $P(\{1\}) = 1/6$.
- A continuous random variable can assume a range of values, e.g.,
  - velocity of a car: possible samples $[0, +\infty[= \mathbb{R}^+$.
- The probability assigned to an elementary event is zero for a continuous random variable, e.g.,
  - velocity of a car: $P(\{75\}) = 0$ (probability that velocity is precisely 75 km/h).
- The probability assigned to an interval can be nonzero for a continuous random variable, e.g.,
  - velocity of a car: $P([25, 30]) = \ldots \geq 0$ (probability that velocity is between 25 and 30 km/h).

Random variables and probability distributions

DISCRETE

probability mass function (PMF)

$$P_X$$

$$P(X = x_i) = P_X(x_i)$$

relation to CDF:

$$\begin{aligned} c_X(x) &= P(X \leq x) \\ &= \sum_{x_i \leq x} P(X = x_i) \\ &= \sum_{x_i \leq x} P_X(x_i). \end{aligned}$$

properties:

$$0 \leq P_X(x_i) \leq 1,$$

$$\textstyle\sum_{x_i} P_X(x_i) = 1.$$

CONTINUOUS

probability density function (PDF)

$$\rho_X$$

$$P(a \leq X \leq b) = \textstyle\int_a^b \rho_X(x) dx$$

relation to CDF:

$$\begin{aligned} c_X(x) &= P(X \leq x) \\ &= P(-\infty \leq X \leq x) \\ &= \int_{-\infty}^{x} \rho_X(\xi) d\xi, \end{aligned}$$

$$\tfrac{dc_X}{dx}(x) = \rho_X(x).$$

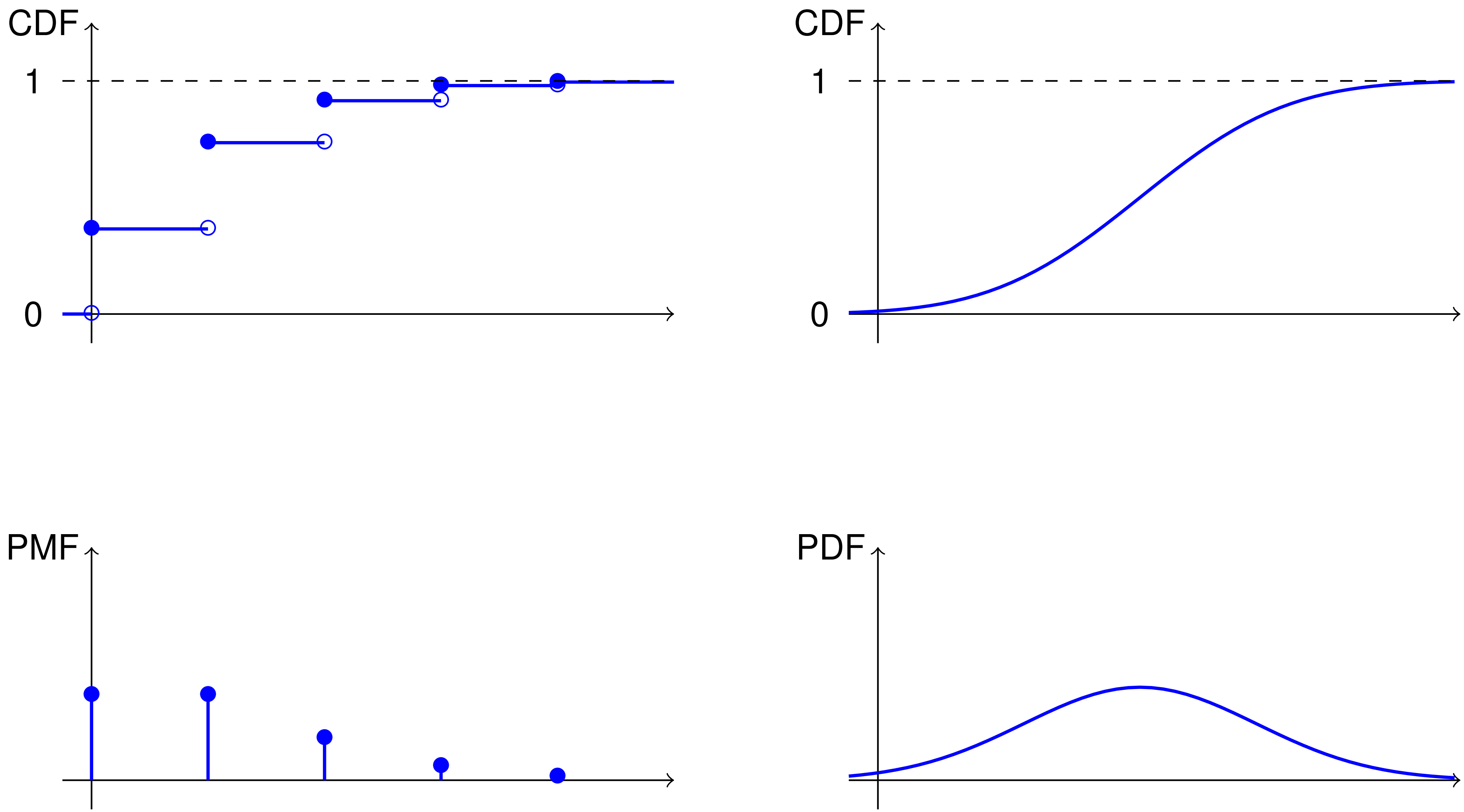
properties:

$$\rho_X(x) \geq 0,$$

$$\textstyle\int_{-\infty}^{+\infty} \rho_X(x) dx = 1.$$

Random variables and probability distributions

DISCRETE CONTINUOUS



Random variables and probability distributions

- Mean (expected value, average):

$$\overline{x} = m_X = E\{X\} = \begin{cases} \displaystyle\sum_{x_i} x_i P_X(x_i), & \text{(DISCRETE)}, \\ \displaystyle\int_{-\infty}^{+\infty} x \rho_X(x) dx, & \text{(CONTINUOUS)}. \end{cases}$$

- Variance (measure of dispersion):

$$\sigma_X^2 = E\{(X - m_X)^2\} = \begin{cases} \displaystyle\sum_{x_i} (x_i - m_X)^2 P_X(x_i), & \text{(DISCRETE)}, \\ \displaystyle\int_{-\infty}^{+\infty} (x - m_X)^2 \rho_X(x) dx, & \text{(CONTINUOUS)}. \end{cases}$$

- "E" is the expectation operator:

$$E\{g(X)\} = \begin{cases} \displaystyle\sum_{x_i} g(x_i) P_X(x_i), & \text{(DISCRETE)}, \\ \displaystyle\int_{-\infty}^{+\infty} g(x) \rho_X(x) dx, & \text{(CONTINUOUS)}. \end{cases}$$

Random variables and probability distributions

- Coefficient of variation:

$$\delta_X = \frac{\sigma_X}{m_X}.$$

- Coefficient of skewness

$$E\{(X - m_X)^3\}.$$

- Coefficient of kurtosis

$$E\{(X - m_X)^4\}.$$

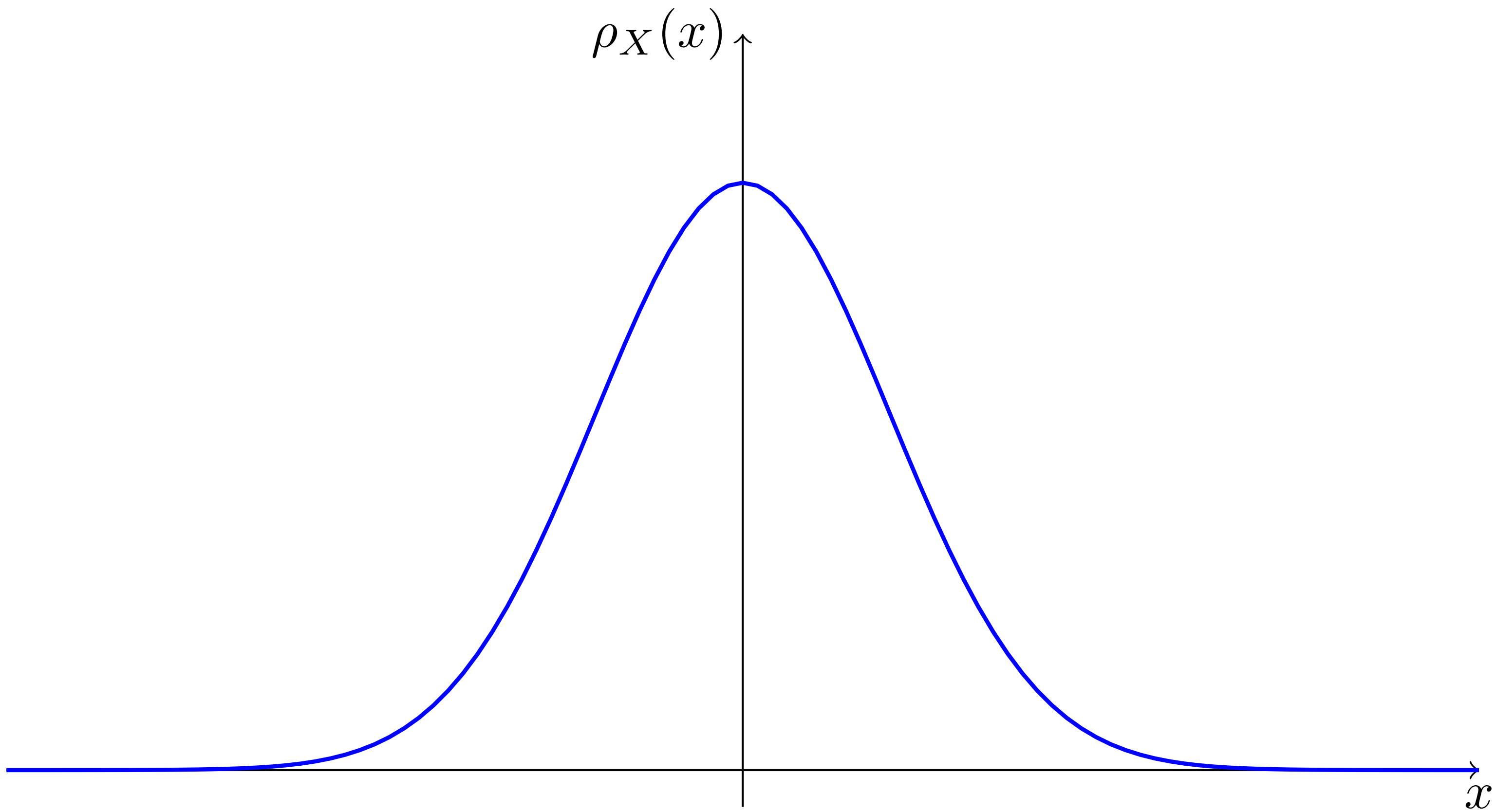
- n-th moment:

$$E\{X^n\}.$$

- n-th central moment:

$$E\{(X - m_X)^n\}.$$

- The Gaussian PDF is the PDF of a continuous random variable given by

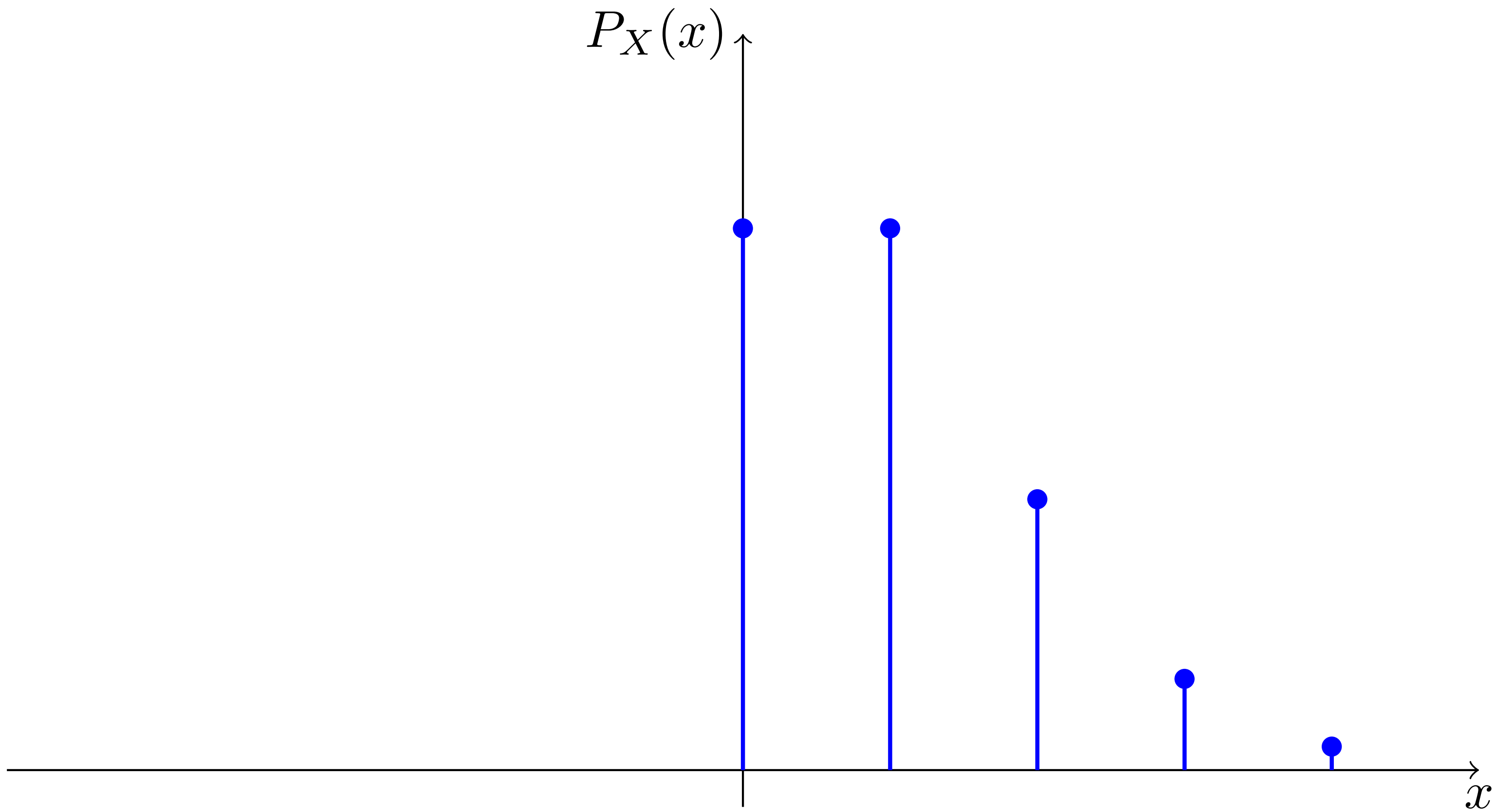


$$\rho_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right),$$

where m and σ are parameters of the PDF; in fact, m and σ are the mean and the standard deviation, respectively, that is, $E\{X\} = m$ and $E\{(X - E\{X\})^2\} = \sigma^2$.

Useful probability distributions

- The Poisson PMF is the PMF of a discrete random variable given by

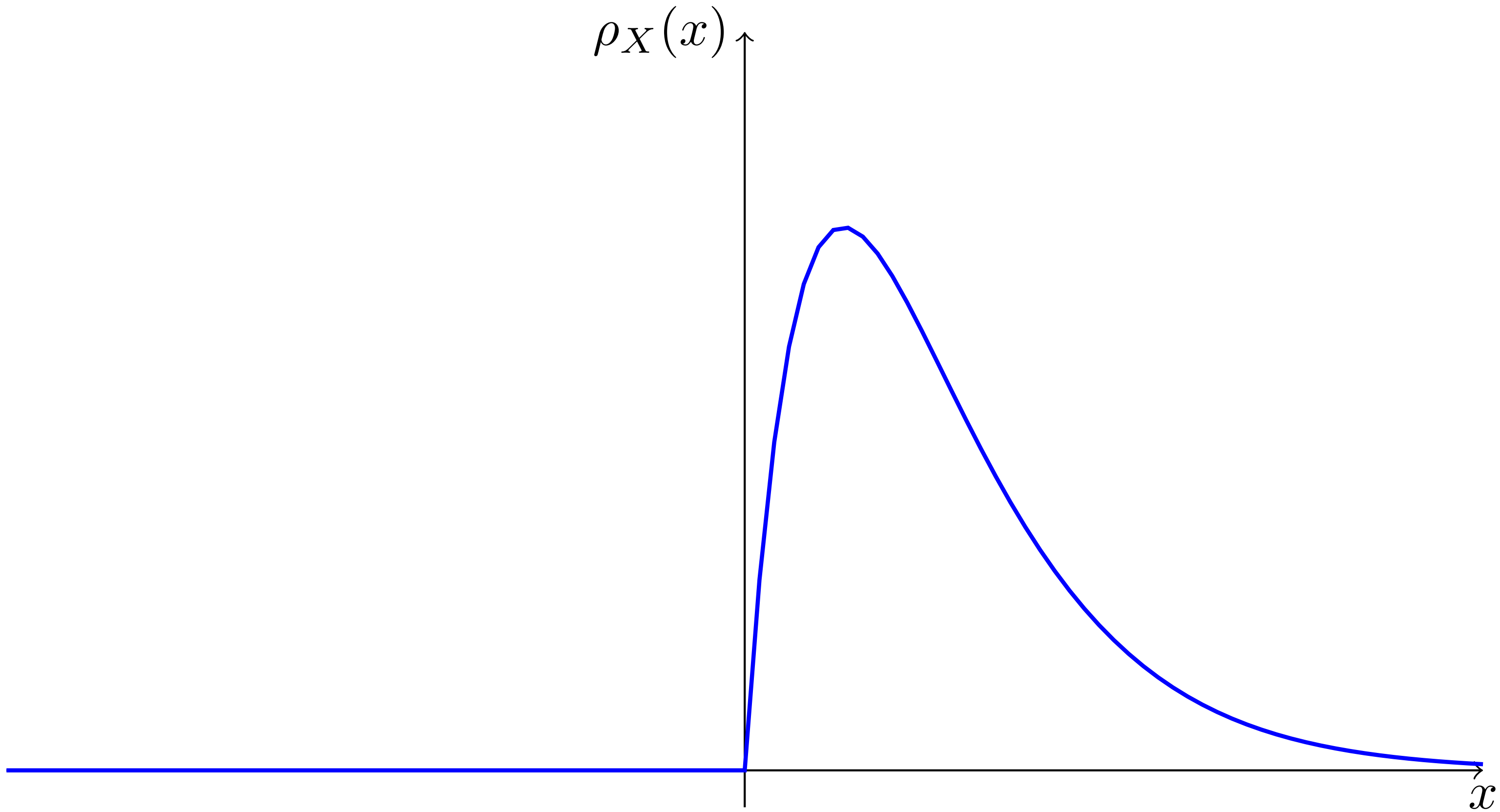


$$P_X(x) = \frac{\lambda^x}{x!}\exp(-\lambda), \quad x = 0, 1, 2, \ldots,$$

where λ is a parameter of the PMF; in fact, λ is equal to the mean and the variance, that is, $E\{X\} = E\{(X - E\{X\})^2\} = \lambda$.

Useful probability distributions

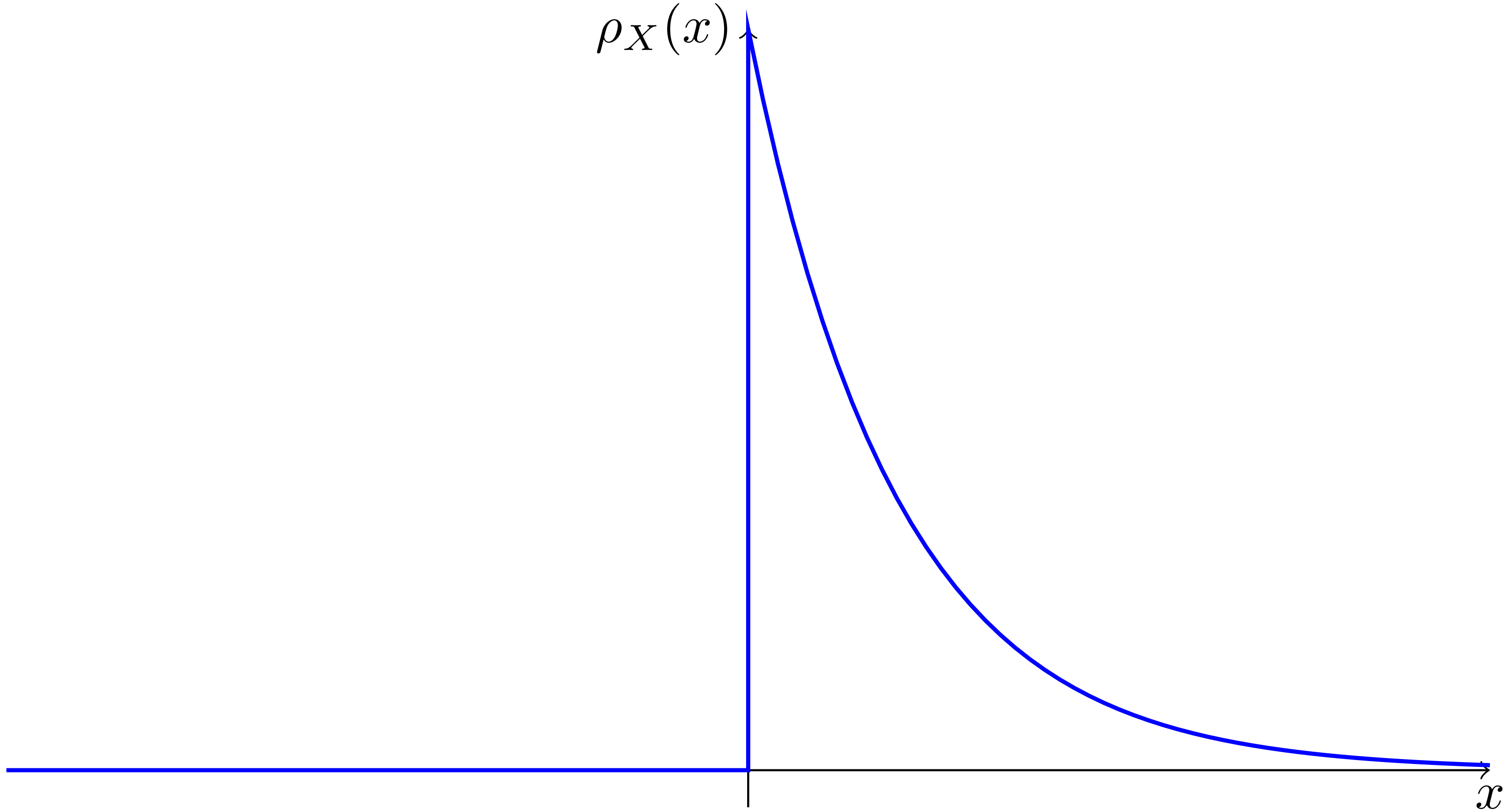
- The gamma PDF is the PDF of a continuous random variable given by



$$\rho_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\dfrac{x}{\beta}\right) & \text{if } x \geq 0, \\ 0 & \text{ortherwise,} \end{cases} \quad \text{with} \quad \Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} \exp(-x) dx,$$

where α and β are parameters of the PDF; in fact, α and β are related to the mean and the standard deviation as follows: $E\{X\} = \alpha\beta$ and $E\{(X - E\{X\})^2\} = \alpha\beta^2$.

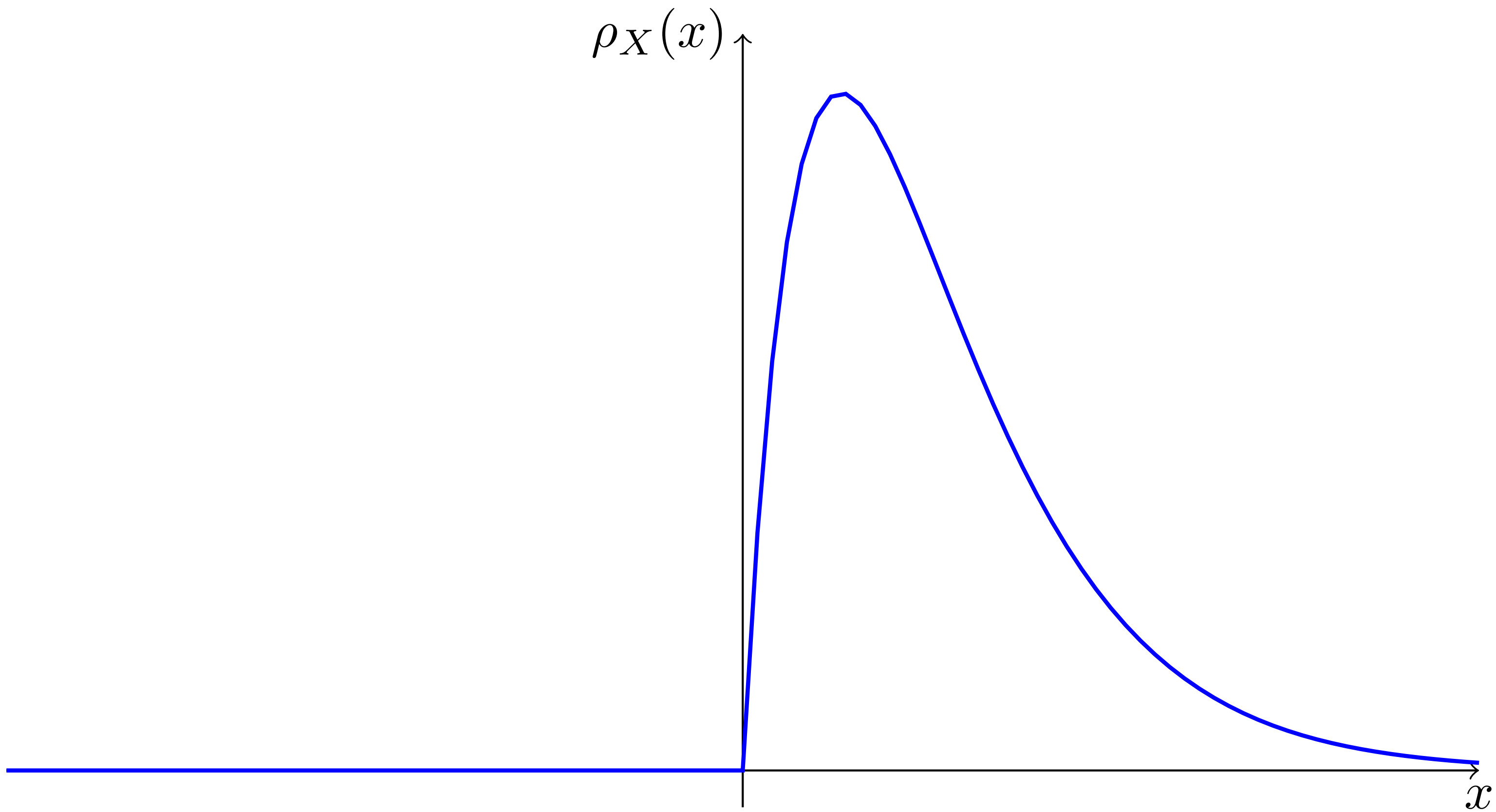
- The exponential PDF is the PDF of a continuous random variable given by



$$\rho_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \geq 0, \\ 0 & \text{ortherwise,} \end{cases}$$

where λ is a parameter of the PDF; in fact, λ is related to the mean and the standard deviation as follows: $E\{X\} = \lambda^{-1}$ and $E\{(X - E\{X\})^2\} = \lambda^{-2}$.

- The chi-squared PDF is the PDF of a continuous random variable given by



$$\rho_X(x) = \begin{cases} \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} x^{n/2-1} \exp\left(-\frac{x}{2}\right) & \text{if } x \geq 0, \\ 0 & \text{ortherwise,} \end{cases}$$

where n is a parameter of the PDF; in fact, n is related to the mean and the standard deviation as follows: $E\{X\} = n$ and $E\{(X - E\{X\})^2\} = 2n$.

Useful probability distributions

- There are many other useful probability distributions:
  - uniform PDF,
  - Weibull PDF,
  - ...
- One can establish various relationships among these probability distributions:
  - A gamma PDF with parameters $\alpha = 1$ and β is an exponential PDF with parameter $\lambda = \beta^{-1}$.
  - A gamma PDF with parameters $\alpha = n/2$ and $\beta = 2$ is a chi-squared PDF with parameter n.
  - The sum of n statistically independent Gaussian random variables with mean 0 and standard deviation 1 is a chi-squared random variable with parameter n.
  - ...

Multiple random variables

Multiple random variables

- For a pair of (real) random variables X and Y, the joint probability P can be described using the joint cumulative distribution function

$$c_{X,Y}(x,y) = P(X \leq x, Y \leq y).$$

- For a pair of (real) random variables X and Y, the cumulative distribution function $c_{X,Y}$ is a function from $\mathbb{R} \times \mathbb{R}$ into $[0,1]$, which possesses the following properties:
  - $c_{X,Y}(x,y) \geq 0$,
  - $c_{X,Y}$ is monotonically increasing,
  - $c_{X,Y}(-\infty,-\infty) = 0$, $c_{X,Y}(-\infty,y) = 0$, $c_{X,Y}(x,-\infty) = 0$, $c_{X,Y}(+\infty,y) = c_Y(y)$, $c_{X,Y}(x,+\infty) = c_X(x)$, and $c_{X,Y}(+\infty,+\infty) = 1$.

DISCRETE	CONTINUOUS
probability mass function (PMF)	probability density function (PDF)
$P_{X,Y}$	$\rho_{X,Y}$
$P(X=x_i, Y=y_j) = P_{X,Y}(x_i, y_j)$	$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d \rho_{X,Y}(x,y)dxdy$

DISCRETE

relation to CDF:

$$c_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} P_{X,Y}(x_i, y_j).$$

properties:

$$0 \leq P_{X,Y}(x_i, y_j) \leq 1,$$
$$\sum_{x_i} \sum_{y_j} P_{X,Y}(x_i, y_j) = 1.$$

CONTINUOUS

relation to CDF:

$$c_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} \rho_{X,Y}(\xi, \zeta) d\xi d\zeta,$$

$$\frac{\partial^2 c_{X,Y}}{\partial x \partial y}(x,y) = \rho_{X,Y}(x,y).$$

properties:

$$\rho_{X,Y}(x,y) \geq 0,$$
$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho_{X,Y}(x,y)dxdy = 1.$$

- For a pair of discrete random variables X and Y, we obtain conditional PMFs as

$$P_{X|Y}(x_i|y_j) = \frac{P_{X,Y}(x_i,y_j)}{P_Y(y_j)} \quad \text{if } P_Y(y_j) \neq 0,$$
$$P_{Y|X}(y_j|x_i) = \frac{P_{X,Y}(x_i,y_j)}{P_X(x_i)} \quad \text{if } P_X(x_i) \neq 0.$$

- For a pair of discrete random variables X and Y, we obtain the marginal PMFs as

$$P_X(x_i) = \sum_{y_j} P_{X,Y}(x_i,y_j),$$
$$P_Y(y_j) = \sum_{x_i} P_{X,Y}(x_i,y_j).$$

- If two discrete random variables X and Y are statistically independent, then we have

$$\begin{aligned} P_{X,Y}(x_i|y_j) &= P_X(x_i), \\ P_{Y|X}(y_j|x_i) &= P_Y(y_j), \end{aligned} \quad \text{hence} \quad P_{X,Y}(x_i,y_j) = P_X(x_i)P_Y(y_j).$$

- For a pair of continuous random variables X and Y, we obtain conditional PMFs as

$$\rho_{X|Y}(x|y) = \frac{\rho_{X,Y}(x,y)}{\rho_Y(y)} \quad \text{if } \rho_Y(y) \neq 0,$$
$$\rho_{Y|X}(y|x) = \frac{\rho_{X,Y}(x,y)}{\rho_X(x)} \quad \text{if } \rho_X(x) \neq 0.$$

- For a pair of continuous random variables X and Y, we obtain the marginal PMFs as

$$\rho_X(x) = \int_{-\infty}^{+\infty} \rho_{X,Y}(x,y)dy,$$
$$\rho_Y(y) = \int_{-\infty}^{+\infty} \rho_{X,Y}(x,y)dx.$$

- If two continuous random variables X and Y are statistically independent, then we have

$$\begin{aligned}\rho_{X,Y}(x|y) &= \rho_X(x),\\ \rho_{Y|X}(y|x) &= \rho_Y(y),\end{aligned} \quad \text{hence} \quad \rho_{X,Y}(x,y) = \rho_X(x)\rho_Y(y).$$

Multiple random variables

- The joint second moment of two random variables X and Y is defined as

$$E\{XY\} = \begin{cases} \sum_{x_i} \sum_{y_j} x_j y_j P_{X,Y}(x_i, y_j), & \text{(DISCRETE)}, \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy \rho_{X,Y}(x,y) dx dy, & \text{(CONTINUOUS)}. \end{cases}$$

- The joint second central moment of two random variables X and Y is defined as

$$E\{(X - m_X)(Y - m_Y)\} = \begin{cases} \sum_{x_i} \sum_{y_j} (x_j - m_X)(y_j - m_Y) P_{X,Y}(x_i, y_j), & \text{(DISCRETE)}, \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_X)(y - m_Y) \rho_{X,Y}(x,y) dx dy, & \text{(CONTINUOUS)}. \end{cases}$$

- The correlation coefficient of two random variables X and Y is defined as

$$\rho = \frac{E\{(X - m_X)(Y - m_Y)\}}{\sigma_X \sigma_Y} \quad \text{with} \quad -1 \leq \rho \leq 1.$$

Multiple random variables

- This can be extended from pairs of (real) random variables to n-uples of (real) random variables.
- Let us consider a random vector $\boldsymbol{X}$ with values in $\mathbb{R}^n$.
- The mean vector of $\boldsymbol{X} = (X_1, \ldots, X_n)$ is the vector $\boldsymbol{m_X}$ in $\mathbb{R}^n$ defined as

$$\boldsymbol{m_X} = E\{\boldsymbol{X}\} = \begin{bmatrix} E\{X_1\} \\ \vdots \\ E\{X_n\} \end{bmatrix}.$$

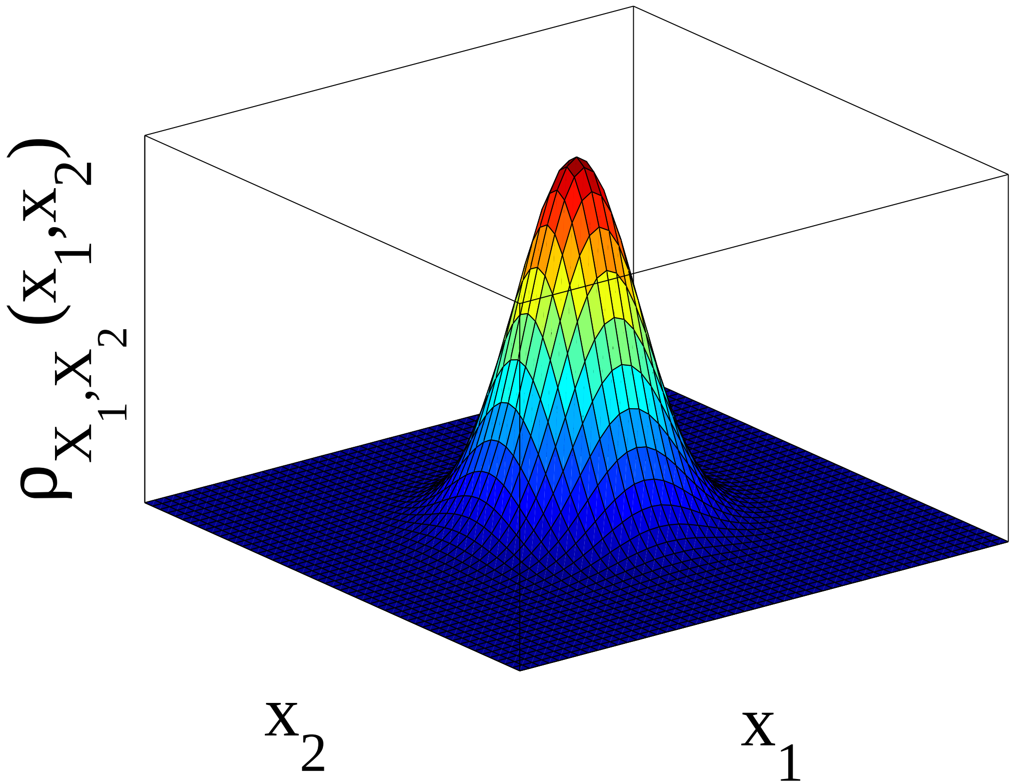
- The correlation matrix of $\boldsymbol{X}$ is the $(n \times n)$-dimensional real matrix $[R_{\boldsymbol{X}}]$ defined as

$$[R_{\boldsymbol{X}}] = E\{\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\} = \begin{bmatrix} E\{X_1 X_1\} & \ldots & E\{X_1 X_n\} \\ \vdots & & \vdots \\ E\{X_n X_1\} & \ldots & E\{X_n X_n\} \end{bmatrix}.$$

- The covariance matrix of $\boldsymbol{X}$ is the $(n \times n)$-dimensional real matrix $[C_{\boldsymbol{X}}]$ defined as

$$[C_{\boldsymbol{X}}] = E\{(\boldsymbol{X} - \boldsymbol{m_X})(\boldsymbol{X} - \boldsymbol{m_X})^{\mathrm{T}}\} = \begin{bmatrix} E\{(X_1 - m_{X_1})(X_1 - m_{X_1})\} & \ldots & E\{(X_1 - m_{X_1})(X_n - m_{X_n})\} \\ \vdots & & \vdots \\ E\{(X_n - m_{X_n})(X_1 - m_{X_1})\} & \ldots & E\{(X_n - m_{X_n})(X_n - -m_{X_n})\} \end{bmatrix}.$$

- The n-variate Gaussian PDF is the PDF of a continuous random vector given by



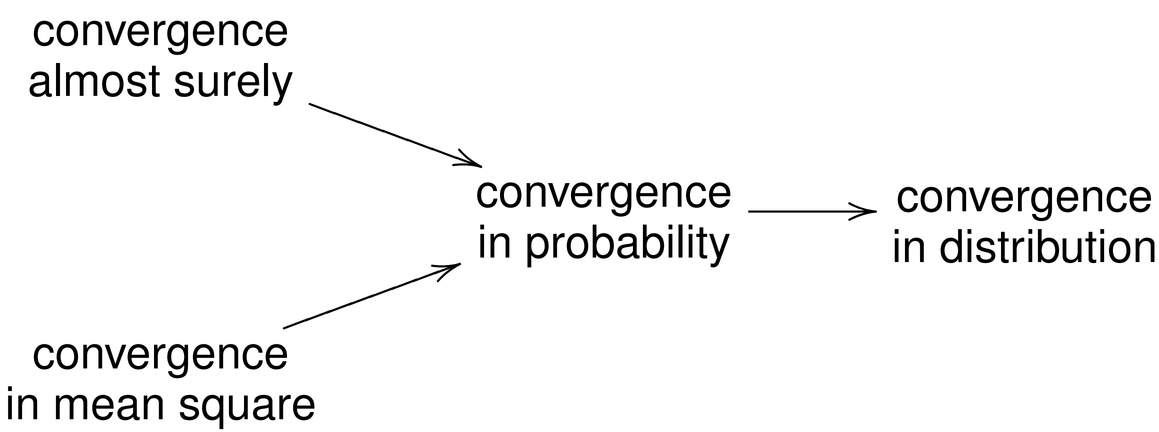
$$\rho_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \text{det}[C]}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^{\mathrm{T}}[C]^{-1}(\boldsymbol{x}-\boldsymbol{m})\right),$$

where the vector $\boldsymbol{m}$ in $\mathbb{R}^n$ and the $(n \times n)$-dimensional real matrix $[C]$ are parameters of the PDF; in fact, $\boldsymbol{m}$ and $[C]$ are the mean vector and the covariance matrix, respectively, that is, $E\{\boldsymbol{X}\} = \boldsymbol{m}$ and $E\{(\boldsymbol{X} - E\{\boldsymbol{X}\})(\boldsymbol{X} - E\{\boldsymbol{X}\})^{\mathrm{T}}\} = [C]$.

Convergence of random variables

Convergence of random variables

- Probability theory offers several ways in which a sequence of random variables can be considered to converge, namely, convergence almost surely, convergence in distribution, convergence in mean square, and convergence in probability among other ways.
- Let us consider a sequence of random vectors $\{\boldsymbol{X}_\nu\}_{\nu=0}^{+\infty}$ with values in $\mathbb{R}^n$ and a random vector $\boldsymbol{X}$ with values in $\mathbb{R}^n$. Then, we have that
  - $\lim_{\nu\to\infty} \boldsymbol{X}_\nu \overset{\text{a.s.}}{=} \boldsymbol{X}$ if and only if $P\big(\lim_{\nu\to\infty} \boldsymbol{X}_\nu = \boldsymbol{X}\}\big) = 1$.
  - $\lim_{\nu\to\infty} \boldsymbol{X}_\nu \overset{\text{distr.}}{=} \boldsymbol{X}$ if and only if $\lim_{\nu\to\infty} P_{\boldsymbol{X}_\nu} = P_{\boldsymbol{X}}$.
  - $\lim_{\nu\to\infty} \boldsymbol{X}_\nu \overset{\text{m.s.}}{=} \boldsymbol{X}$ if and only if $\lim_{\nu\to\infty} E\{\|\boldsymbol{X}_\nu - \boldsymbol{X}\|^2\} = 0$.
  - $\lim_{\nu\to\infty} \boldsymbol{X}_\nu \overset{\text{prob.}}{=} \boldsymbol{X}$ if and only if for every $\epsilon > 0$, $\lim_{\nu\to\infty} P\big(\|\boldsymbol{X}_\nu - \boldsymbol{X}\| \geq \epsilon\}\big) = 0$.
- These modes of convergence are related as follows:



References

Suggested reading material

- L. Wehenkel. Eléments du calcul des probabilités. ULg, 2013.

Additional references also consulted to prepare this lecture

- A. Ang and W. Tang. Probability concepts in engineering. John Wiley & Sons, 2007.
- C. Soize. The Fokker–Planck equation for stochastic dynamical systems and its explicit steady state solutions. World Scientific Publishing, 1994.