6 Probability, Random Variables, and Random Processes

6.1 Introduction and Definitions

Deterministic Signals:

- can be expressed in form of a mathematical equation
- can be reproduced exactly with repeated measurements

Random Signals:

- cannot be repeated in a predictable manner
- cannot be described by mathematical equations or tables

Examples for Random Signals:

- Number of car accidents in Lubbock during 1 year
- Quantization noise in an A/D converter
- Background noise in speech transmission
- Speckle noise in synthetic aperture radar images
- Tossing a coin
- Life expectance of an engine
- Outcome of a roulette game
- Data transmission
- Bio-medical signals
- Rolling a dice

Random Experiment: Experiment with random outcome that is repeatable and for which the outcome of each experiment is not predictable.

Sample Space S**:** Set of all possible outcomes if a random experiment.

Examples:

1.	Tossing a coin	$S_1 = \{\text{head}, \text{tail}\}$
2.	Rolling a die	$S_2 = \{1, 2, 3, 4, 5, 6\}$
3.	Rolling two dice	$S_3 = S_2 \times S_2 =$\{(1,1),(1,2),. . . ,(6,5), (6,6)\}
4.	No. of phone calls per day	$S_4 = 0, 1, 2, 3, \ldots$
5.	Life expectance of an engine	$S_5 = [0, \infty)$
6.	Angular position of a pointer	$S_6 = [0, 2\pi)$
7.	Quantization error in an ADC	$S_7 = [0, Q)$ or $S_7 = [-Q/2, Q/2)$

Event A**:** Subset of the sample space S.

$A = S$:	certain event
$A = \emptyset$:	impossible event
A consists of a single element of S:	elementary event

Examples:

- Tossing a coin: $S = \{\text{head}, \text{tail}\}$, $A_1 = \{\text{head}\}$, $A_2 = \{\text{tail}\}$, $A_3 = \{\text{head, tail}\}$
- Rolling a die: $S = \{1, 2, 3, 4, 5, 6\}$, $A_1 = \{3\}$, $A_2 = \{2, 4, 6\}$, $A_3 = \{\# \leq 5\}$
- Quantization error in an ADC: $S = [-Q/2, Q/2)$, $A_1 = [0, Q/2)$, $A_2 = Q/4$, $A_3 = (-Q/4, Q/4)$

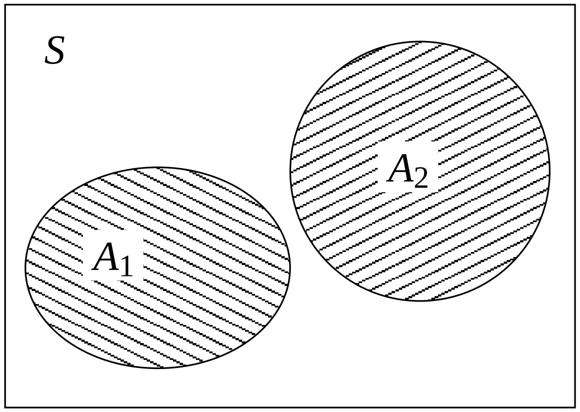
6.2 Probability

Definition: In order for an operator P that assigns a real number, called the probability, to each event A in the sample space S, to be a valid probability assignment, it has to satisfy the following 3 axioms:

- $0 \leq P(A) \leq 1$
- $P(S) = 1$
- if A_1 and A_2 are mutually exclusive (or disjoint) events, i.e. the occurrence of one outcome precludes the occurrence of the other, then:

  $P(A_1 \text{ or } A_2) = P(A_1 \cup A_2) = P(A_1 + A_2) = P(A_1) + P(A_2)$

Venn Diagram:



Example: Tossing a Coin
$S = \{\text{head}, \text{tail}\}$, $A_1 = \{\text{head}\}$, $A_2 = \{\text{tail}\}$, $A_3 = \{\text{head, tail}\}$

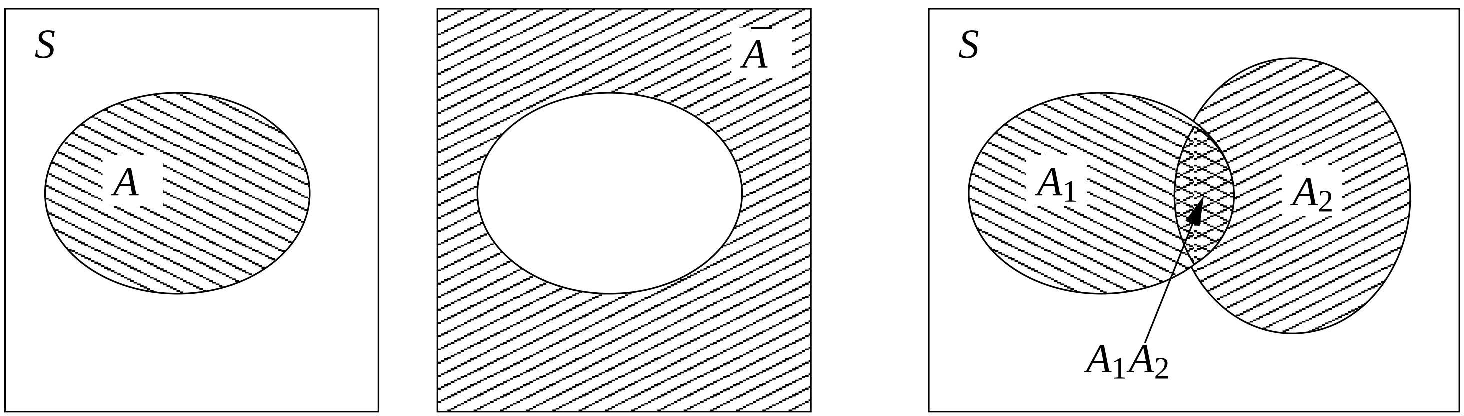
$P(\text{head}) = P(A_1) = 0.5$,

$P(\text{tail}) = P(A_2) = 0.5$,

$P(\text{head or tail}) = P(A_3) = P(A_1) + P(A_2) = P(S) = 1$

Properties of Probabilities:

- $P(\bar{A}) = 1 - P(A), \quad$ with $\bar{A} = S - A$: complementary event to A
- $P(\emptyset) = 0$
- $P(A_1 + A_2) = P(A_1) + P(A_2) - P(A_1 A_2)$
  with $P(A_1 A_2)$: probability of the joint occurrence of A_1 and A_2



Example: Rolling a Die

Sample space: $S = \{1, 2, 3, 4, 5, 6\}$

Certain event:

$$P(S) = P(1 + 2 + 3 + 4 + 5 + 6) = \sum_{k=1}^{6} P(k) = 1$$

For a fair die all elementary events are equally likely:

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

Probability of an odd number:

$$P(1 + 3 + 5) = P(1) + P(3) + P(5) = \frac{3}{6}$$

Probability of a number smaller 5:

$$\begin{aligned} P(\# < 5) = P(1 + 2 + 3 + 4) &= P(1) + P(2) + P(3) + P(4) \\ &= 1 - P(5) + P(6) = \frac{4}{6} \end{aligned}$$

Example: Number of Phone Calls to a Given Number During a Certain Time Period

Sample space: $S = \mathbb{N}_0 = 0, 1, 2, 3, \ldots$

Elementary event A_i: exactly i calls come in.

$P(A_i)$ cannot be equally likely since:

$$\sum_{i=0}^{\infty} P(A_i) = 1$$

Estimation of Probability $P(A)$:

An estimate $\hat{P}(A)$ for the probability $P(A)$ that the event A occurs is given by the relative frequency of occurrence of A:

$$\hat{P}(A) = \frac{N_A}{N}$$

with

N:	number of times the experiment is repeated
N_A:	number of times the event A occurred

The estimate is the more accurate the larger N is. For $N \to \infty$ the deviation from the probability $P(A)$ generally tends towards zero.

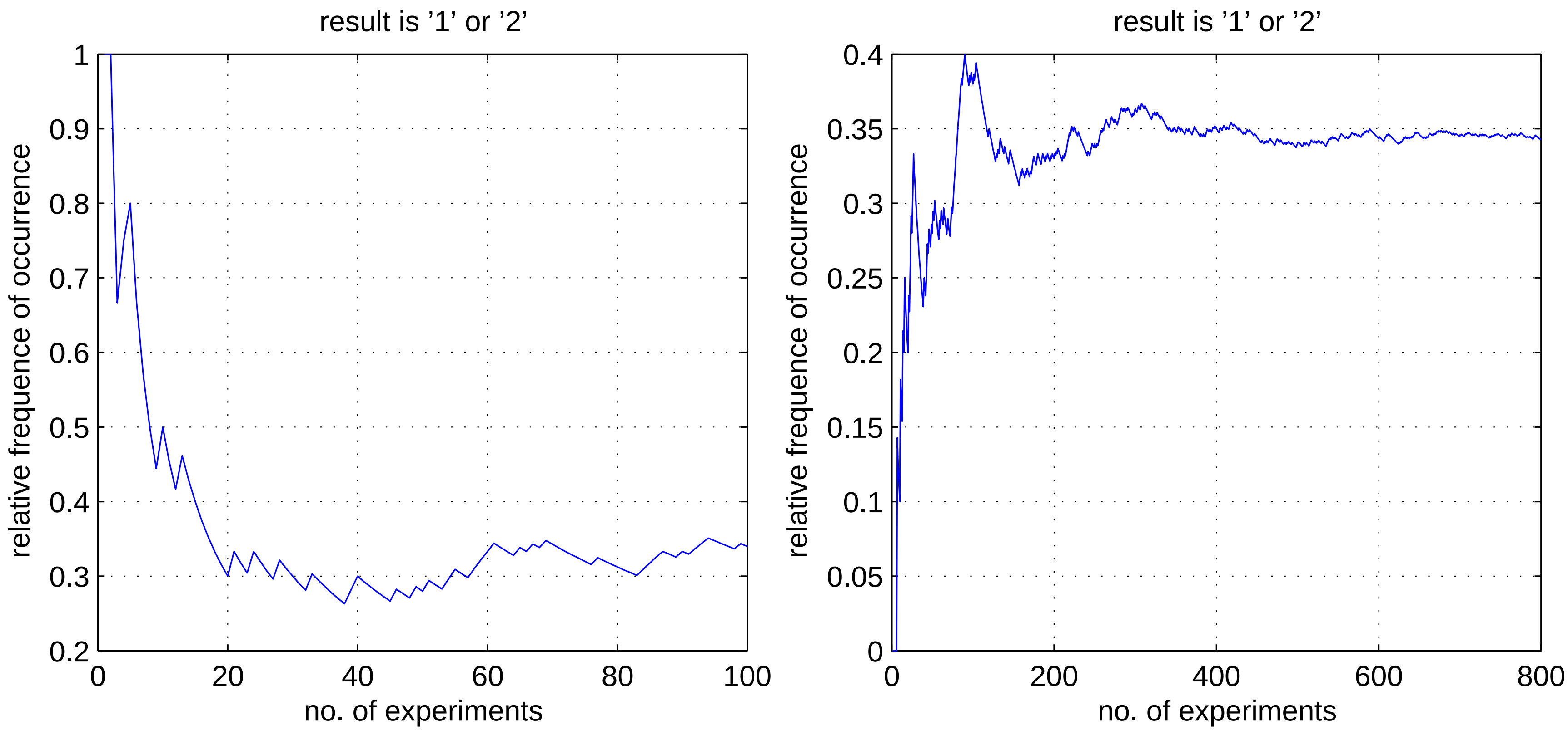
Example: Tossing a Coin

A fair coin is tossed 100 times, head occurs 51 times, tail occurs 49 times.

$$\hat{P}(\text{head}) = \frac{51}{100}, \qquad \hat{P}(\text{tail}) = \frac{49}{100}$$

Example: Rolling a Fair Die

The following plots show the relative frequency of occurrence of the event that the number is 1 or 2 when rolling the die 100 times and 800 times.



Compare to:

$$P(1+2) = P(1) + P(2) = \frac{2}{6} = 0.33$$

6.3 Conditional Probability and Statistical Independence

Conditional Probability $P(B|A)$**:**

Probability of the event B knowing that event A has occurred.

$$P(AB) = P(A) \cdot P(B|A)$$

Conditional Probability $P(A|B)$**:**

Probability of the event A knowing that event B has occurred.

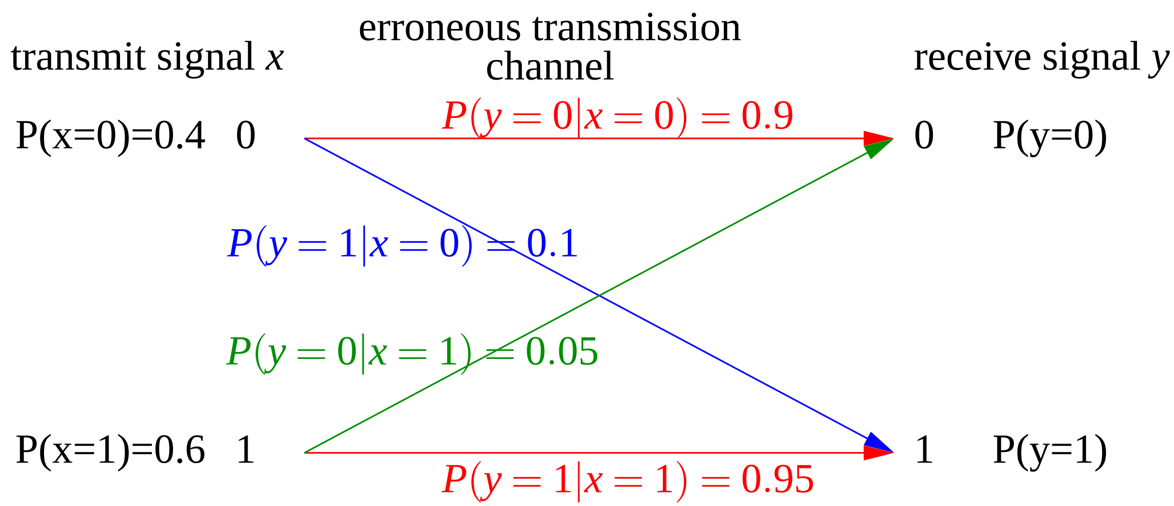
$$P(BA) = P(B) \cdot P(A|B)$$

With $P(AB) = P(BA)$ we obtain **Bayes' Theorem**:

$$P(A) \cdot P(B|A) = P(B) \cdot P(A|B) \quad \text{or} \quad P(B|A) = \frac{P(B)P(A|B)}{P(A)}$$

Example: Binary Data Transmission

A binary data source sends a '0' with probability 0.4 and a '1' with probability 0.6. The data is transmitted through an erroneous channel. The probability for a '0' to be received correctly is 0.9, the probability for a '1' to be received correctly is 0.95. In case of an error, a '1' is received instead of a '0' and vice versa.



Conditional Probabilities (knowing what happens at the transmitter):

- Probability to receive a '0' if a '0' was sent: $P(y=0|x=0)=0.9$
- Probability to receive a '1' if a '0' was sent: $P(y=1|x=0)=0.1$
- Probability to receive a '1' if a '1' was sent: $P(y=1|x=1)=0.95$
- Probability to receive a '0' if a '1' was sent: $P(y=0|x=1)=0.05$

Joint Probabilities:

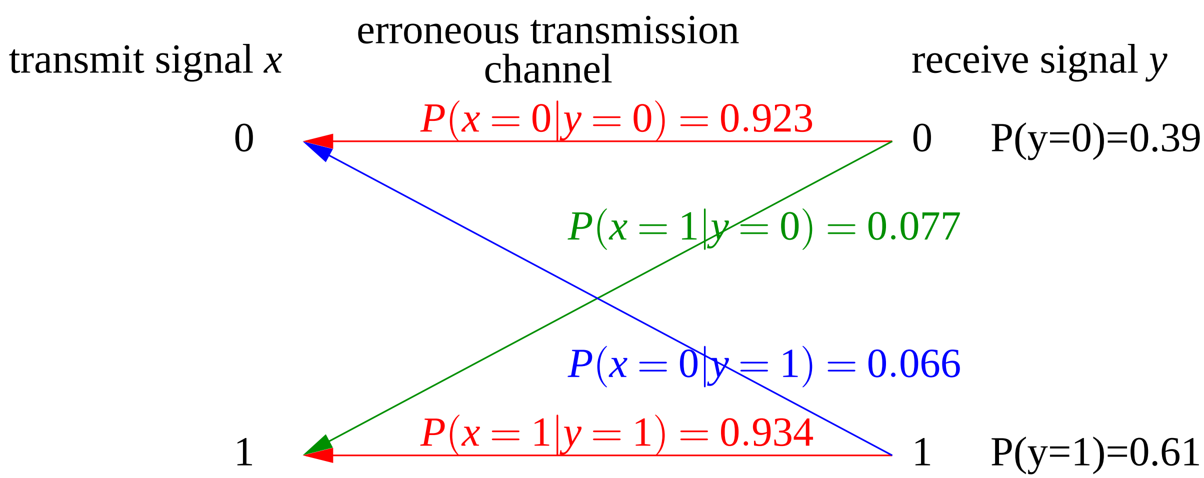
- Probability to send a '0' and receive a '0':
  $P(x=0\,y=0) = P(x=0)\cdot P(y=0|x=0) = 0.4\cdot 0.9 = 0.36$
- Probability to send a '0' and receive a '1':
  $P(x=0\,y=1) = P(x=0)\cdot P(y=1|x=0) = 0.4\cdot 0.1 = 0.04$
- Probability to send a '1' and receive a '1':
  $P(x=1\,y=1) = P(x=1)\cdot P(y=1|x=1) = 0.6\cdot 0.95 = 0.57$
- Probability to send a '1' and receive a '0':
  $P(x=1\,y=0) = P(x=1)\cdot P(y=0|x=1) = 0.6\cdot 0.05 = 0.03$

Probabilities to receive a '0' or '1':

- Probability to receive a '0':
  $P(y=0) = P(x=0\,y=0) + P(x=1\,y=0) = 0.36 + 0.03 = 0.39$
- Probability to receive a '1':
  $P(y=1) = P(x=0\,y=1) + P(x=1\,y=1) = 0.04 + 0.57 = 0.61$

Conditional Probabilities (knowing what happens at the receiver):

- Probability that a '0' was sent if a '0' is received:
$$P(x=0|y=0) = \frac{P(x=0\;y=0)}{P(y=0)} = \frac{P(x=0)P(y=0|x=0)}{P(y=0)} = 0.923$$
- Probability that a '1' was sent if a '0' is received:
$$P(x=1|y=0) = \frac{P(x=1\;y=0)}{P(y=0)} = \frac{P(x=1)P(y=0|x=1)}{P(y=0)} = 0.077$$
- Probability that a '1' was sent if a '1' is received:
$$P(x=1|y=1) = \frac{P(x=1\;y=1)}{P(y=1)} = \frac{P(x=1)P(y=1|x=1)}{P(y=1)} = 0.934$$
- Probability that a '0' was sent if a '1' is received:
$$P(x=0|y=1) = \frac{P(x=0\;y=1)}{P(y=1)} = \frac{P(x=0)P(y=1|x=0)}{P(y=1)} = 0.066$$



Statistical Independence:

Two events A and B are called statistically independent, if $P(B|A) = P(B)$ and thus:

$$P(AB) = P(A) \cdot P(B|A) = P(A) \cdot P(B)$$

Example: Rolling Three Dice

Determine the probability that all three dice show a '6'.

$$\begin{aligned} P(\text{die } 1 = 6 \text{ die } 2 = 6 \text{ die } 3 = 6) &= P(\text{die } 1 = 6) \cdot P(\text{die } 2 = 6) \cdot P(\text{die } 3 = 6) \\ &= \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216} \end{aligned}$$

Example: Binary Data Transmission

Check for the binary data transmission example, whether the events x and y are statistically independent.

In case of statical independence we have:

$$P(x = i\, y = j) = P(x = i) \cdot P(y = j), \qquad i, j = 0, 1$$

- $P(x = 0) \cdot P(y = 0) = 0.4 \cdot 0.39 = 0.156 \neq P(x = 0\, y = 0) = 0.36$
- $P(x = 0) \cdot P(y = 1) = 0.4 \cdot 0.61 = 0.244 \neq P(x = 0\, y = 1) = 0.04$
- $P(x = 1) \cdot P(y = 1) = 0.6 \cdot 0.61 = 0.366 \neq P(x = 1\, y = 1) = 0.57$
- $P(x = 1) \cdot P(y = 0) = 0.6 \cdot 0.39 = 0.234 \neq P(x = 1\, y = 1) = 0.03$

6.4 Random Variables

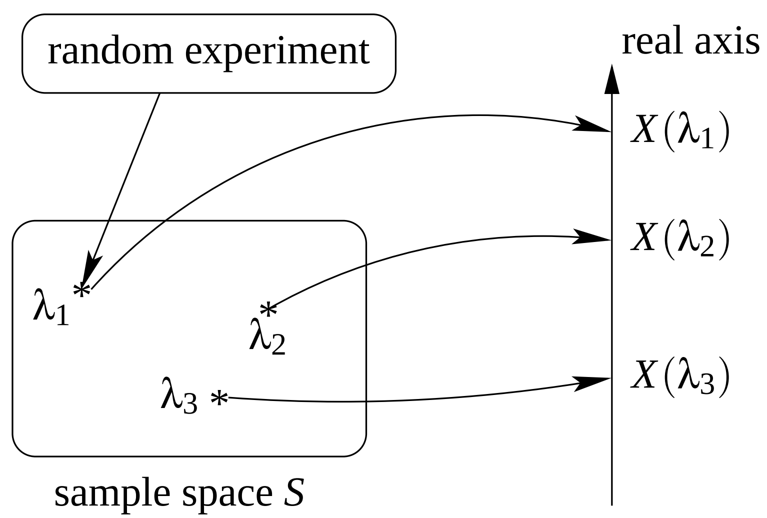
Motivation:

So far, the outcomes of a random experiment were described verbally, e.g. 'head' and 'tail' for tossing a coin. The description of an event A can become long and cumbersome.

Random Variable:

A random variable X maps each of the possible outcomes of a random experiment to a real number.

$$X : S \rightarrow \mathbb{R}$$



Examples:

- Tossing a coin: $S = \{\text{head, tail}\}$

$$X(\text{head}) = 0, \qquad X(\text{tail}) = 1$$

- Rolling a die: $S = \{\text{'1', '2', '3', '4', '5', '6'}\}$

$$X(\text{'1'}) = 1,\ X(\text{'2'}) = 2,\ X(\text{'3'}) = 3,\ X(\text{'4'}) = 4,\ X(\text{'5'}) = 5,\ X(\text{'6'}) = 6$$

  Another valid but less intuitive mapping is:

$$X(\text{'1'}) = -5.6,\ X(\text{'2'}) = \pi,\ X(\text{'3'}) = 0.5,$$
$$X(\text{'4'}) = -34.6,\ X(\text{'5'}) = -99,\ X(\text{'6'}) = 56.7$$

- Rolling two dice: $S = \{(\lambda_1, \lambda_2) :\ \lambda_1, \lambda_2 \in \{\text{'1', '2', '3', '4', '5', '6'}\}\}$
  Event A: sum of numbers is larger than 8

$$A = \{(\lambda_1, \lambda_2) : \lambda_1 + \lambda_2 > 8\}$$

  Random variable:

$$X((\lambda_1, \lambda_2)) = \lambda_1 + \lambda_2$$

  Probability:

$$P(A) = P(X > 8)$$

- Angular position of a pointer on a rotating wheel: $S = [0, 2\pi)$
  Event A: angle lies in the interval $[\pi/4, \pi/2)$
  $$A = \{\lambda : \lambda \in [\pi/4, \pi/2)\}$$
  Random variable:
  $$X(\lambda) = \lambda/\text{rad}$$
  Probability:
  $$P(A) = P(\pi/4 \leq X(\lambda) < \pi/2)$$

6.5 Cumulative Distribution Function and Probability Density Function

The **cumulative distribution function** $F_X(x)$ associated with a random variable X is defined as:

$$F_X(x) = P(X(\lambda) \leq x)$$

The **probability density function (pdf)**, denoted by $p(x)$ or $p_X(x)$, of a random variable X with cumulative distribution function $F_X(x)$ is given by:

$$p_X(x) = \frac{d}{dx}F_X(x) \quad \text{or} \quad F_X(x) = \int_{-\infty}^{x} p_X(t)\,dt$$

Properties of the Cumulative Distribution Function $F_X(x)$:

- $0 \leq F_X(x) \leq 1$ for all $x \in I\!R$
- $F_X(x_1) \leq F_X(x_2)$ for all $x_1, x_2 \in I\!R$ with $x_1 < x_2$
- $\lim\limits_{x \to \infty} F_X(x) = 1, \qquad \lim\limits_{x \to -\infty} F_X(x) = 0$
- $\lim\limits_{x \to x_0 + 0} F_X(x) = F_X(x_0)$, i.e. $F_X(x)$ is continuous to the right
- For a discrete random variable $X(\lambda_i)$ with associated probabilities $P(\lambda_i)$, the cumulative distribution function can be expressed as:

$$F_X(x) = \sum_i P(\lambda_i) u(x - X(\lambda_i))$$

Properties of the Probability Density Function $p(x)$:

- $p(x) \geq 0$ for all $x \in I\!R$
- $\int\limits_{-\infty}^{\infty} p(x)\, dx = 1$
- For a discrete random variable $X(\lambda_i)$ with associated probabilities $P(\lambda_i)$, the probability density function can be expressed as:

$$p(x) = \sum_{i=0}^{\infty} P(\lambda_i) \delta(x - X(\lambda_i))$$

Example: Tossing a Fair Coin

$$S = \{\text{head, tail}\}, \qquad P(\text{head}) = P(\text{tail}) = 0.5$$

Random Variable:

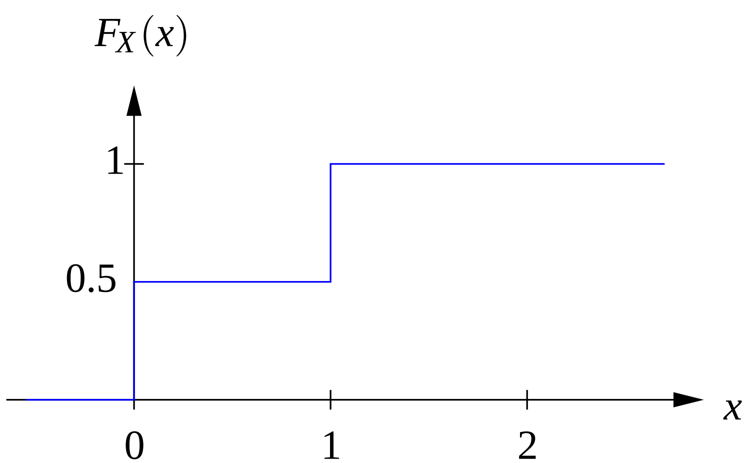
$$X(\text{head}) = 0, \quad X(\text{tail}) = 1$$

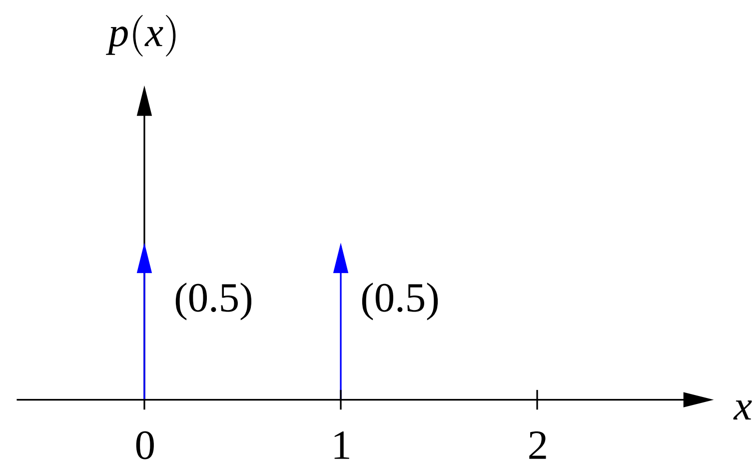
Cumulative distribution function:

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ 0.5 & \text{for } 0 \leq x < 1 \\ 1 & \text{for } x \geq 1 \end{cases}$$

Probability density function:

$$p(x) = 0.5\,\delta(x) + 0.5\,\delta(x-1)$$





Example: Angular Position of a Pointer on a Rotating Wheel

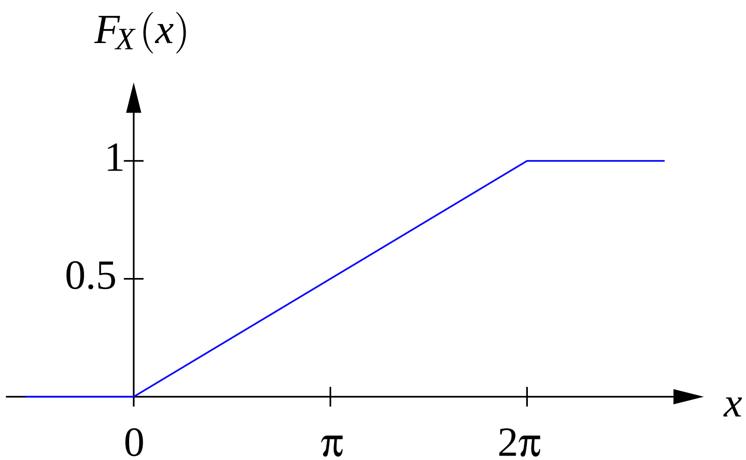
Sample Space: $S = [0, 2\pi)$

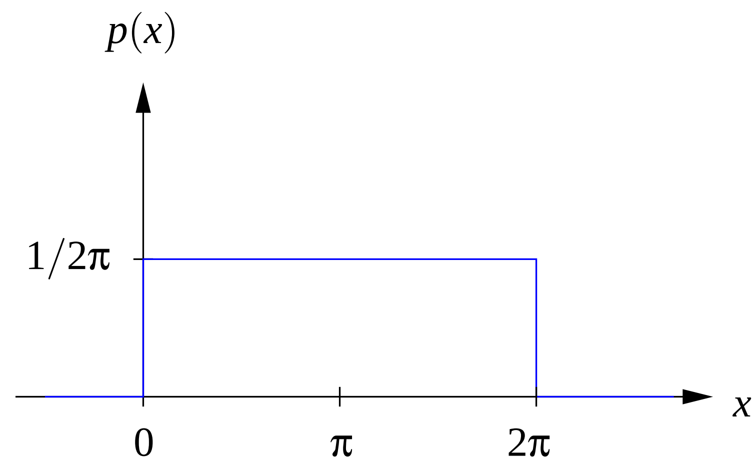
All angles in the range from 0 to 2π are equally likely

$$p(x) = \begin{cases} 0 & \text{for } x < 0 \\ K & \text{for } 0 \le x < 2\pi \\ 0 & \text{for } x \ge 2\pi \end{cases} \qquad \int_{-\infty}^{\infty} p(x)dx = \int_{0}^{2\pi} p(x)dx = 2\pi K = 1$$

Probability density function and cumulative distribution function:

$$p(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1/2\pi & \text{for } 0 \le x < 2\pi \\ 0 & \text{for } x \ge 2\pi \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ x/2\pi & \text{for } 0 \le x < 2\pi \\ 1 & \text{for } x \ge 2\pi \end{cases}$$





Probability that the angle lies between $\pi/2$ and $3\pi/2$:

$$P(\frac{\pi}{2} < X \leq \frac{3\pi}{2}) = F_X(\frac{3\pi}{2}) - F_X(\frac{\pi}{2}) = \int_{\pi/2}^{3\pi/2} p(t)\, dt = \frac{1}{2}$$

Probability that the angle is exactly π:

$$P(X = \pi) = \int_{\pi}^{\pi} p(x)\, dx = 0$$

6.6 The Histogram

Definition:

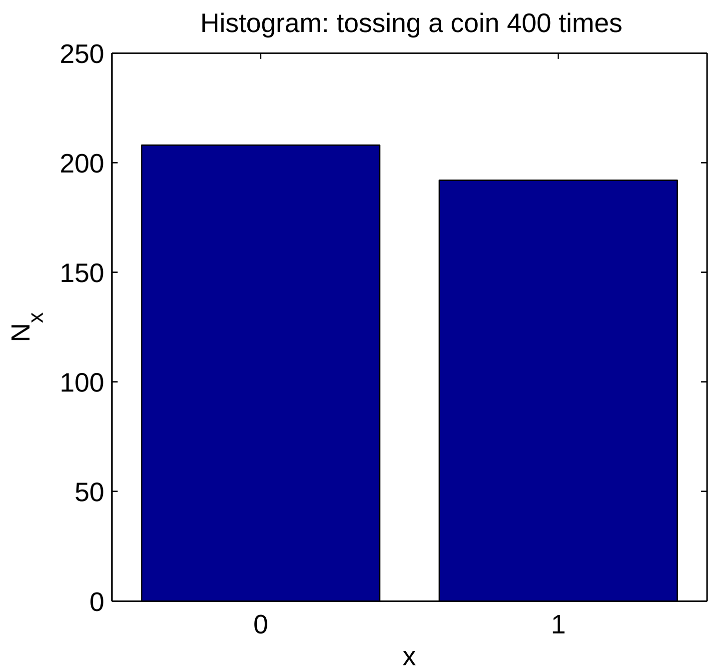
Tabulation of the frequency of occurrence of a random variable within preset ranges. The total number of events is equal to the number of sets (also called cells).

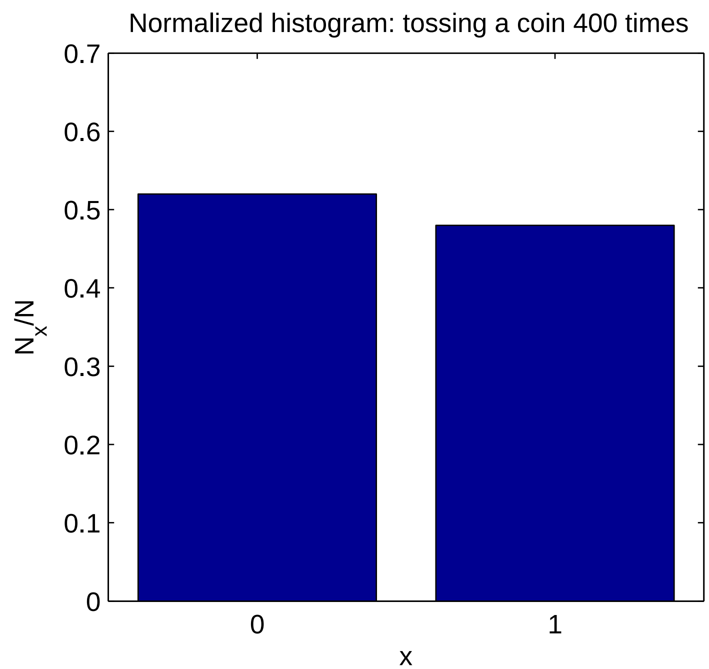
The normalized histogram is obtained by dividing the height of each cell by the total number of events.

The normalized histogram provides an estimate for the probability density function $p(x)$ of a random variable.

Example: Tossing a Fair Coin

$$S = \{\text{head, tail}\}, \qquad X(\text{head}) = 0, \quad X(\text{tail}) = 1, \qquad P(0) = P(1) = 0.5$$





Estimated probability for $X = 0$ and $X = 1$:

$$\hat{P}(0) = \frac{N_0}{N} = \frac{208}{400} = 0.52, \qquad \hat{P}(1) = \frac{N_1}{N} = \frac{192}{400} = 0.48$$

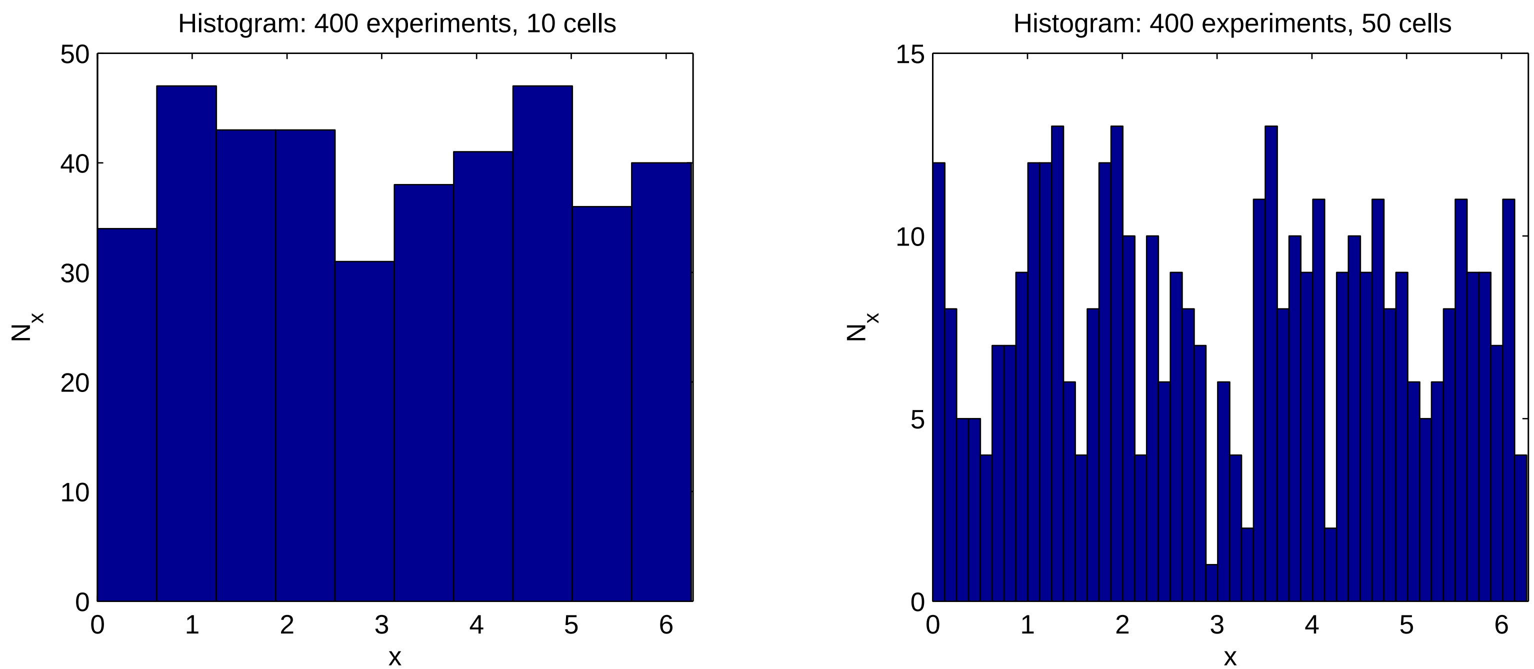
Example: Angular Position of a Pointer on a Rotating Wheel

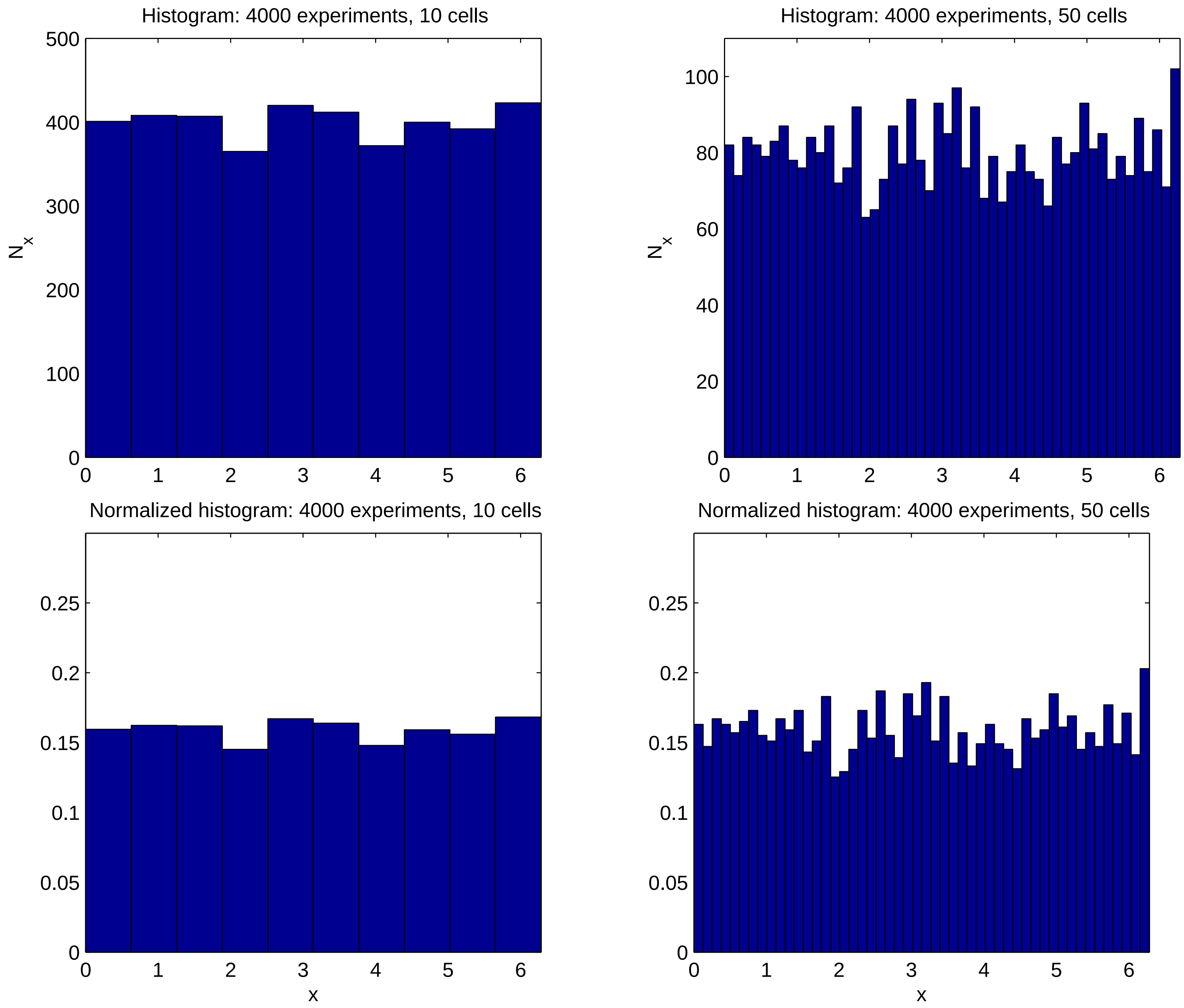
$$S = [0, 2\pi), \qquad X(\text{angle}) = \text{angle}/\text{rad}, \qquad p(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1/2\pi & \text{for } 0 \le x < 2\pi \\ 0 & \text{for } x \ge 2\pi \end{cases}$$

Normalization of the histogram for a continuous random variable:

$$\hat{P}(x_i) = \frac{N_{x_i}}{N} \cdot \Delta x_i$$

where x_i denotes a cell of width Δx_i.





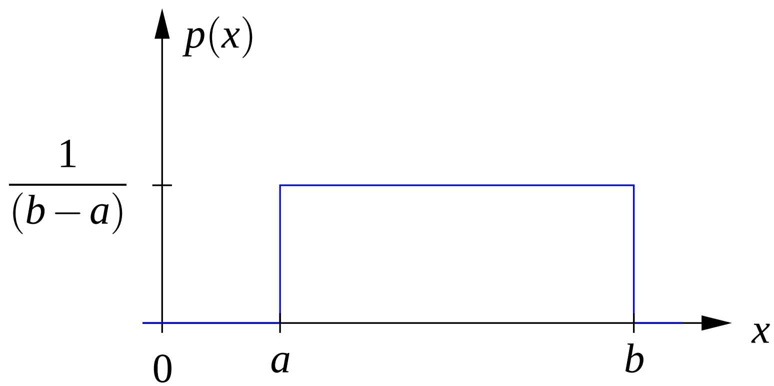
Histogram: 4000 experiments, 10 cells
Histogram: 4000 experiments, 50 cells
Normalized histogram: 4000 experiments, 10 cells
Normalized histogram: 4000 experiments, 50 cells
N_x
x

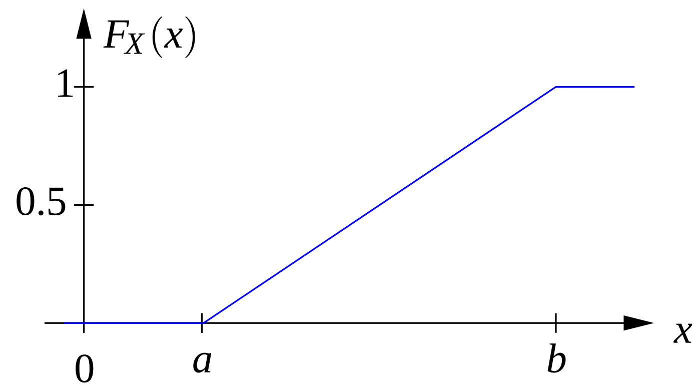
6.7 Some Important Probability Distributions

6.7.1 The Uniform Distribution

Random variable X is equally likely within the range $[a, b]$.

$$p(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } \ a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{for } \ x < a \\ \dfrac{x-a}{b-a} & \text{for } \ a \leq x \leq b \\ 1 & \text{for } \ b < x \end{cases}$$





Examples: quantization noise of an ADC, angular position of a pointer on a rotating wheel

6.7.2 The Binomial Distribution

Let's assume that a random experiment has two mutual exclusive events A and B with probabilities $P(A) = p$ and $P(B) = 1 - p = q$. The experiment is repeated n times. The random variable X is assigned the value i if event A occurs exactly i times out of the n repetitions. The probability for $X = i$, $i = 0, \ldots, n$ is given by:

$$P(X = i) = \binom{n}{i} p^i q^{n-i}, \qquad \text{binomial coefficient: } \binom{n}{i} = \frac{n!}{i!(n-i)!}$$

Number of combinations to place i occurrences of event A on n possible positions:

$$n(n-1)(n-2)\cdots(n-i+1) = \frac{n!}{(n-i)!}$$

Out of these there are $i!$ combinations that we cannot distinguish between.

Probability density function and cumulative distribution function:

$$p(x) = \sum_{i=0}^{n} P(X = i)\delta(x - i) = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \delta(x - i)$$

$$F_X(x) = \sum_{i=0}^{n} P(X = i)u(x - i) = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} u(x - i)$$

Example: Rolling a die n times

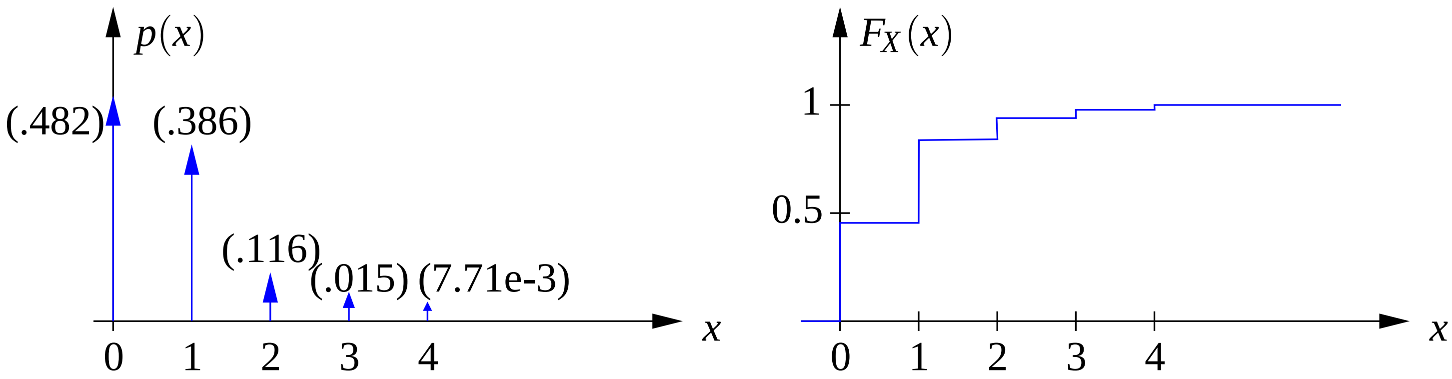
A die is rolled 4 times. Sketch $p(x)$ and $F_X(x)$ for the event that a '3' occurs on top of die $i = 0, \ldots, 4$ times.

Event A: number on top of die is '3', $P(A) = p = 1/6$.
Event B: number on top of die is not '3', $P(B) = q = 1 - p = 5/6$.

Probability density function:

$$p(x) = \sum_{i=0}^{4} P(X = i)\delta(x - i) = \sum_{i=0}^{4} \binom{4}{i} (1/6)^i (5/6)^{4-i} \delta(x - i)$$

i	$P(X = i)$
0	$\binom{4}{0}(1/6)^0(5/6)^4 = \frac{4!}{4! \cdot 0!}(5/6)^4 = (5/6)^4 = 0.482$
1	$\binom{4}{1}(1/6)^1(5/6)^3 = \frac{4!}{3! \cdot 1!}(1/6)^1(5/6)^3 = 4 \cdot (1/6)^1(5/6)^3 = 0.386$
2	$\binom{4}{2}(1/6)^2(5/6)^2 = \frac{4!}{2! \cdot 2!}(1/6)^2(5/6)^2 = 6 \cdot (1/6)^2(5/6)^2 = 0.116$
3	$\binom{4}{3}(1/6)^3(5/6)^1 = \frac{4!}{1! \cdot 3!}(1/6)^3(5/6)^1 = 4 \cdot (1/6)^3(5/6)^1 = 0.015$
4	$\binom{4}{4}(1/6)^4(5/6)^0 = \frac{4!}{4! \cdot 0!}(1/6)^4 = (1/6)^4 = 7.71 \cdot 10^{-3}$



6.7.3 The Poisson Distribution

The exponential distribution depends on a parameter t. It approximates the binomial distribution for large n.

$$P(X = i) = \exp(-t)\,\frac{t^i}{i!}$$

Probability density function and cumulative distribution function:

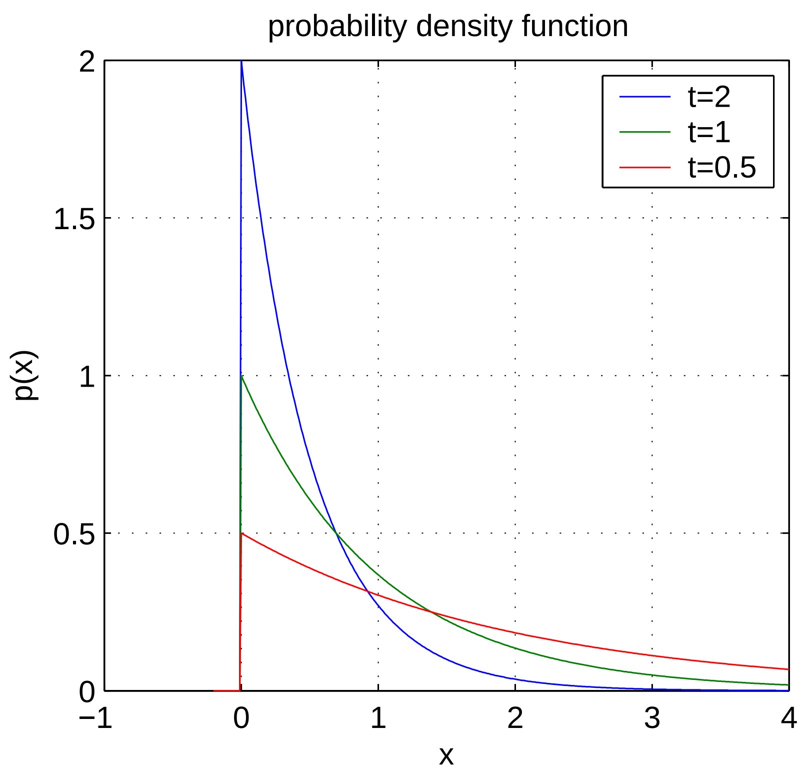
$$p(x) = \sum_i \exp(-t)\,\frac{t^i}{i!}\,\delta(x-i), \quad F_X(x) = \sum_i \exp(-t)\,\frac{t^i}{i!}\,u(x-i)$$

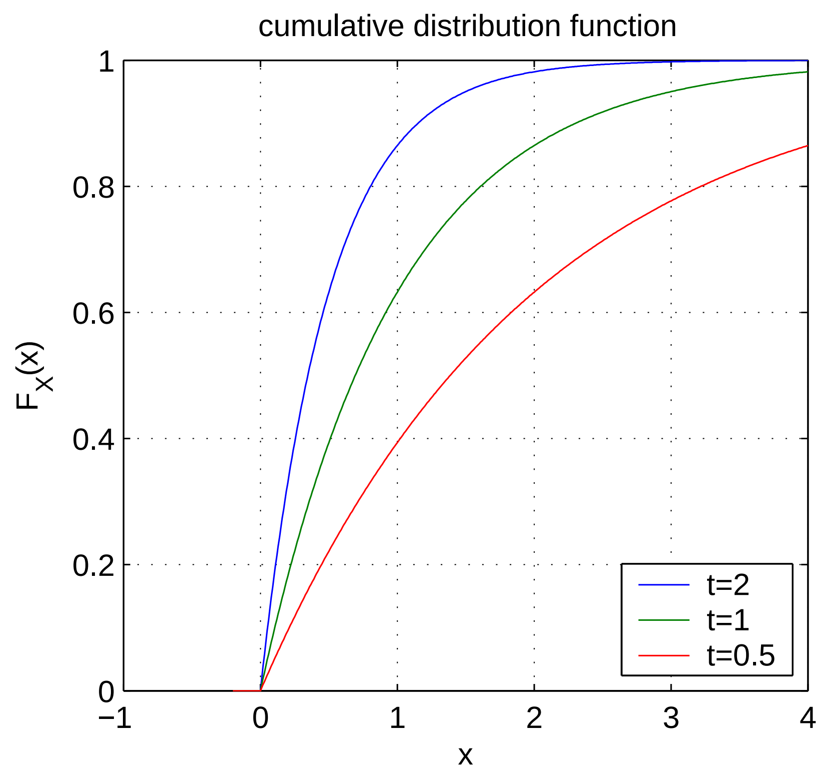
Examples: number of particles emitted over a time duration in radioactive decay, number of telephone calls arriving at a switching venter during a certain time period, binary data transmission with low error rates

6.7.4 The Exponential Distribution

The exponential distribution depends on a parameter t.

$$p(x) = \begin{cases} 0 & \text{for } x < 0 \\ t \cdot \exp(-xt) & \text{for } x \geq 0 \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - \exp(-xt) & \text{for } x \geq 0 \end{cases}$$





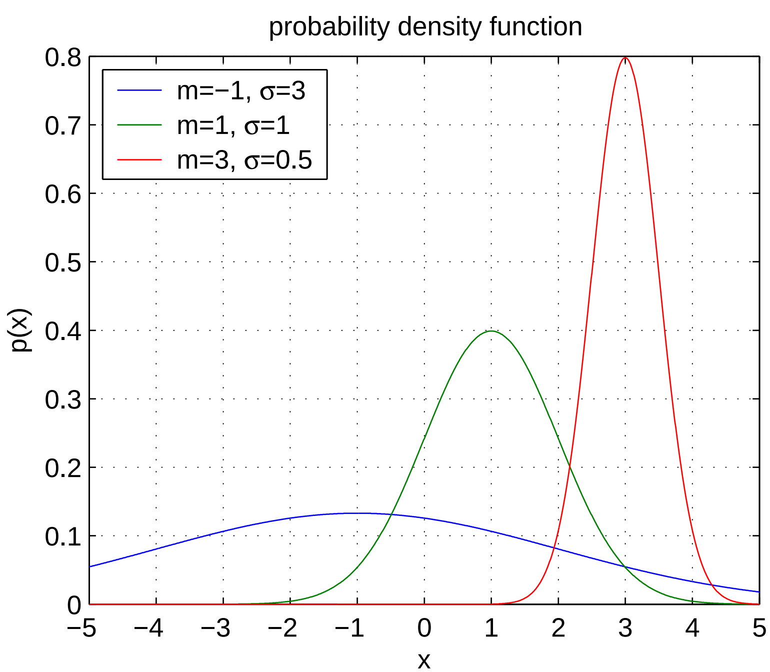
Examples: characterization of waiting times for processes without aging, e.g. time period between two calls at a calling center, time to repair a defect component of a system

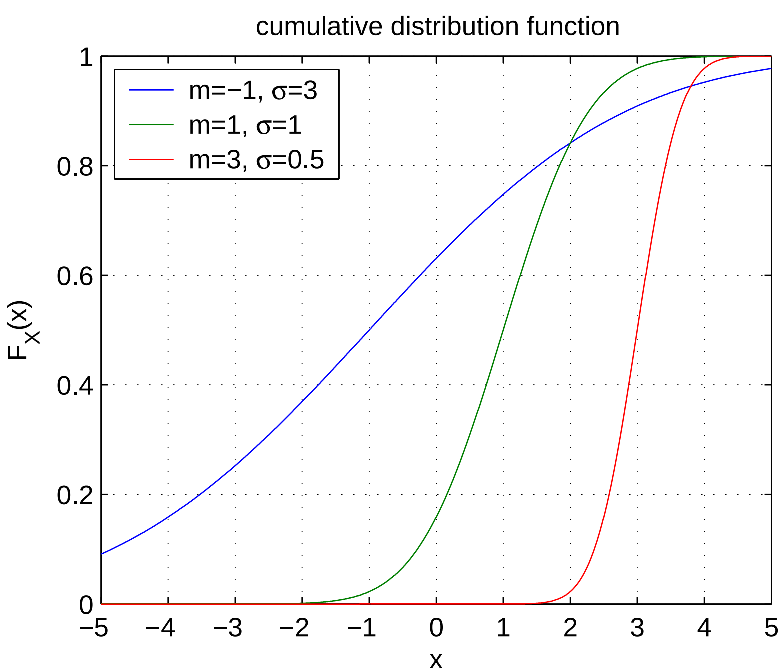
6.7.5 The Gaussian (Normal) Distribution

The Gaussian distribution depends on parameters $m \in \mathbb{R}$ and $\sigma > 0$. It is also denoted as $N(m, \sigma^2)$ distribution.

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

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right) dy$$





The integral for the cumulative distribution function $F_X(x)$ cannot be evaluated in closed form and requires numerical evaluation.

Normalized Gaussian Distribution:

$$P(X \leq (m + k\sigma)) = \int_{-\infty}^{m+k\sigma} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right) dy$$

Substitute $z = (y - m)/\sigma$:

$$P(X \leq (m + k\sigma)) = \int_{-\infty}^{k} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) dz$$

Using the Q-Function (tabulated in appendix I of the textbook)

$$Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right) dz$$

we obtain

$$P(X \leq (m + k\sigma)) = 1 - Q(k)$$

A random variable has a Gaussian distribution if it depends on many independent random variables with identical distribution.

Examples: Electrical energy consumption in a metropolitan area, noise in a transmission channel, arbitrary measurement errors

6.8 Statistical Averages

A random variable is completely determined by its probability density function $p(x)$ or its cumulative distribution function $F_X(x)$, respectively.

In many cases $p(x)$ and $F_X(x)$ are difficult to estimate and knowledge about the average outcome of a random experiment is sufficient (e.g. average life expectance of an engine, average time spent in a waiting queue, average gain/loss of a slot machine).

Expected Value $E\{X\}$:

For a random variable X with probability density function $p(x)$, the expected value $E\{X\} = m_X$ is given by:

$$E\{X\} = \int_{-\infty}^{\infty} x \cdot p(x) dx = m_X$$

If X is a discrete random variable that takes the value x_i, $i = 1, 2 \ldots N < \infty$, with probability $P(X = x_i)$, then

$$p(x) = \sum_{i=1}^{N} P(X = x_i) \cdot \delta(x - x_i), \qquad E\{X\} = \sum_{i=1}^{N} x_i \cdot P(X = x_i)$$

The expected value is also called mean value, average value, or first moment.

nth Moment of a Random Variable X:

$$E\{X^n\} = \int_{-\infty}^{\infty} x^n \cdot p(x)dx$$

A knowledge of all the moments of a random variable, provided they are finite, specifies a random variable as completely as a knowledge of the probability density function.

The second moment is called the **mean-square** value of X and its square root is called the **root mean square (rms)** value of X.

nth Central Moment of a Random Variable X:

$$E\{(X - m_X)^n\} = \int_{-\infty}^{\infty} (x - m_X)^n \cdot p(x)dx$$

The second order central moment $E\{(X - m_X)^2\}$ is called the **variance** of a random variable X. It measures the average quadratic deviation of X from its average value m_X.

$$\sigma_X^2 = E\{|(X - m_X)|^2\} = \int_{-\infty}^{\infty} (x - m_X)^2 \cdot p(x)dx$$

σ_X is called the **standard deviation** of X and is an indicator for the effective width of the probability density function (the larger σ_X the more spread out is the pdf of X).

Chebyshev's Inequality:

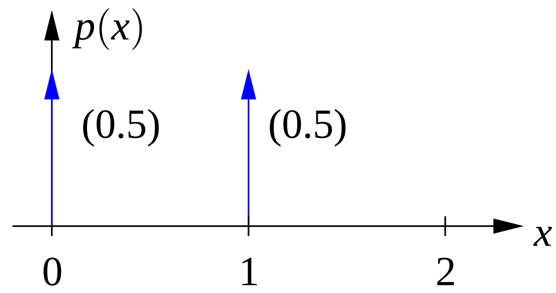
$$P(|X - m_X| > k \cdot \sigma_X) < 1/k^2$$

Example: Tossing a Fair Coin

$$S = \{\text{head, tail}\}, \qquad X(\text{head}) = 0, \quad X(\text{tail}) = 1, \qquad P(0) = P(1) = 0.5$$

Probability density function:

$$p(x) = 0.5\,\delta(x) + 0.5\,\delta(x-1)$$



Average value m_X:

$$E\{X\} = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

$$= 0.5 \int_{-\infty}^{\infty} x\big(\delta(x) + \delta(x-1)\big)\,dx$$

$$= 0.5(0+1) = 0.5$$

Variance σ_X^2:

$$E\{(X - m_X)^2\} = \int_{-\infty}^{\infty} (x-0.5)^2 p(x)\,dx$$

$$= 0.5 \int_{-\infty}^{\infty} (x-0.5)^2\big(\delta(x) + \delta(x-1)\big)\,dx$$

$$= 0.5\big((0-0.5)^2 + (1-0.5)^2\big) = 0.25$$

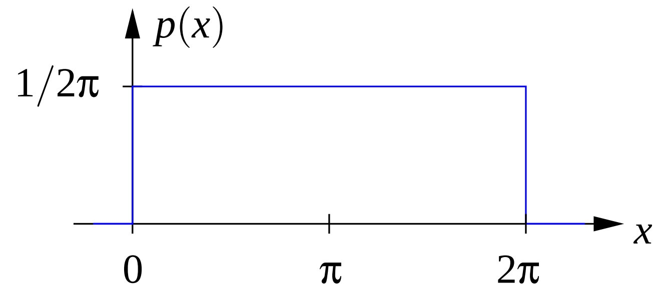
Standard deviation: $\sigma_X = 0.5$

Example: Angular Position of a Pointer on a Rotating Wheel

$$S = [0, 2\pi), \qquad X(\text{angle}) = \text{angle/rad}$$

Probability density function:

$$p(x) = \begin{cases} 0 & \text{for } \; x < 0 \\ 1/2\pi & \text{for } \; 0 \le x < 2\pi \\ 0 & \text{for } \; x \ge 2\pi \end{cases}$$



Average value m_X:

$$E\{X\} = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

$$= \frac{1}{2\pi} \int_0^{2\pi} x \, dx = \pi$$

Variance σ_X^2:

$$E\{(X - m_X)^2\} = \int_0^{2\pi} (x - \pi)^2 p(x) \, dx$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (x - \pi)^2 dx = \frac{1}{3} \pi^2$$

Standard deviation: $\sigma_X = \frac{1}{3} \pi$

Properties of the Mathematical Expectance Operator $E\{.\}$:

- Linearity:

$$X = aX_1 + bX_2, \qquad E\{X\} = E\{aX_1 + bX_2\} = a \cdot E\{X_1\} + b \cdot E\{X_2\}$$

  with X, X_1, X_2: random variables and a, b: real constants

- Function of a random variable:

$$Y = g(X), \qquad E\{Y\} = E\{g(X)\} = \int_{-\infty}^{\infty} g(x)p(x)dx$$

  with X, Y: random variables

- Repeatedly taking the mathematical expectance:

$$E\{E\{X\}\} = E\{X\}$$

- Properties of the variance:

$$\sigma_X^2 = E\{(X - m_X)^2\} = E\{X^2\} - m_X^2$$

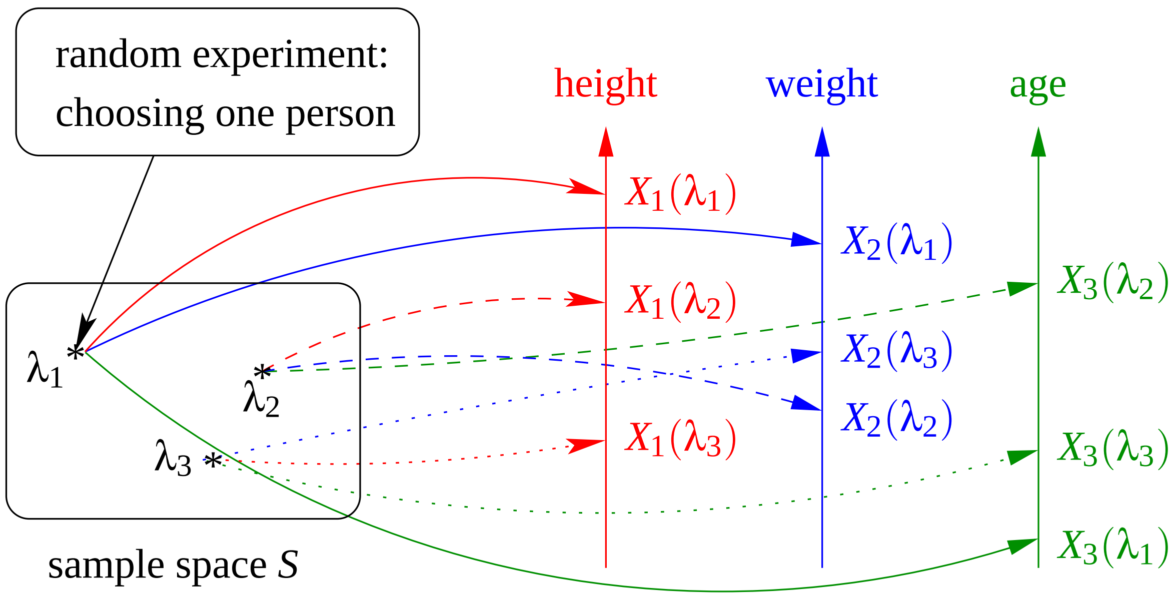
$$Y = aX + b, \qquad \sigma_Y^2 = E\{(Y - m_Y)^2\} = a^2 \cdot \sigma_X^2$$

  with X, Y: random variables and a, b: real constants

6.9 Joint and Conditional Density Functions

In many cases a random experiment depends on several random variables.

Example: Determine how the likelihood to be overweight ed depends on the age of a person.



Joint Cumulative Function:

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

where X and Y are two random variables defined on the same sample space S.

Joint Probability Function:

$$p_{X,Y}(x,y) = \frac{\partial^2}{\partial x\,\partial y}\,F_{X,Y}(x,y)$$

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} p_{X,Y}(t,s)\,ds\,dt$$

where X and Y are two random variables defined on the same sample space S.

Marginal Probability Function:

$$p_X(x) = \int\limits_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy, \qquad p_Y(y) = \int\limits_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx$$

Marginal Cumulative Distribution Function:

$$F_X(x) = P(X \le x, Y \le \infty) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{\infty} p_{X,Y}(t,s)\,ds\,dt = \int\limits_{-\infty}^{x} p_X(t)\,dt$$

$$F_Y(y) = P(X \le \infty, Y \le y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{y} p_{X,Y}(t,s)\,ds\,dt = \int\limits_{-\infty}^{y} p_Y(s)\,ds$$

Conditional Density Function:

$$p_X(x|y_1) = \frac{p_{X,Y}(x, y_1)}{p_Y(y_1)}, \qquad p_Y(y|x_1) = \frac{p_{X,Y}(x_1, y)}{p_X(x_1)}$$

$$p_{X,Y}(x, y) = p_X(x|y) \cdot p_Y(y) = p_Y(y|x) \cdot p_X(x)$$

Statistical Independence:

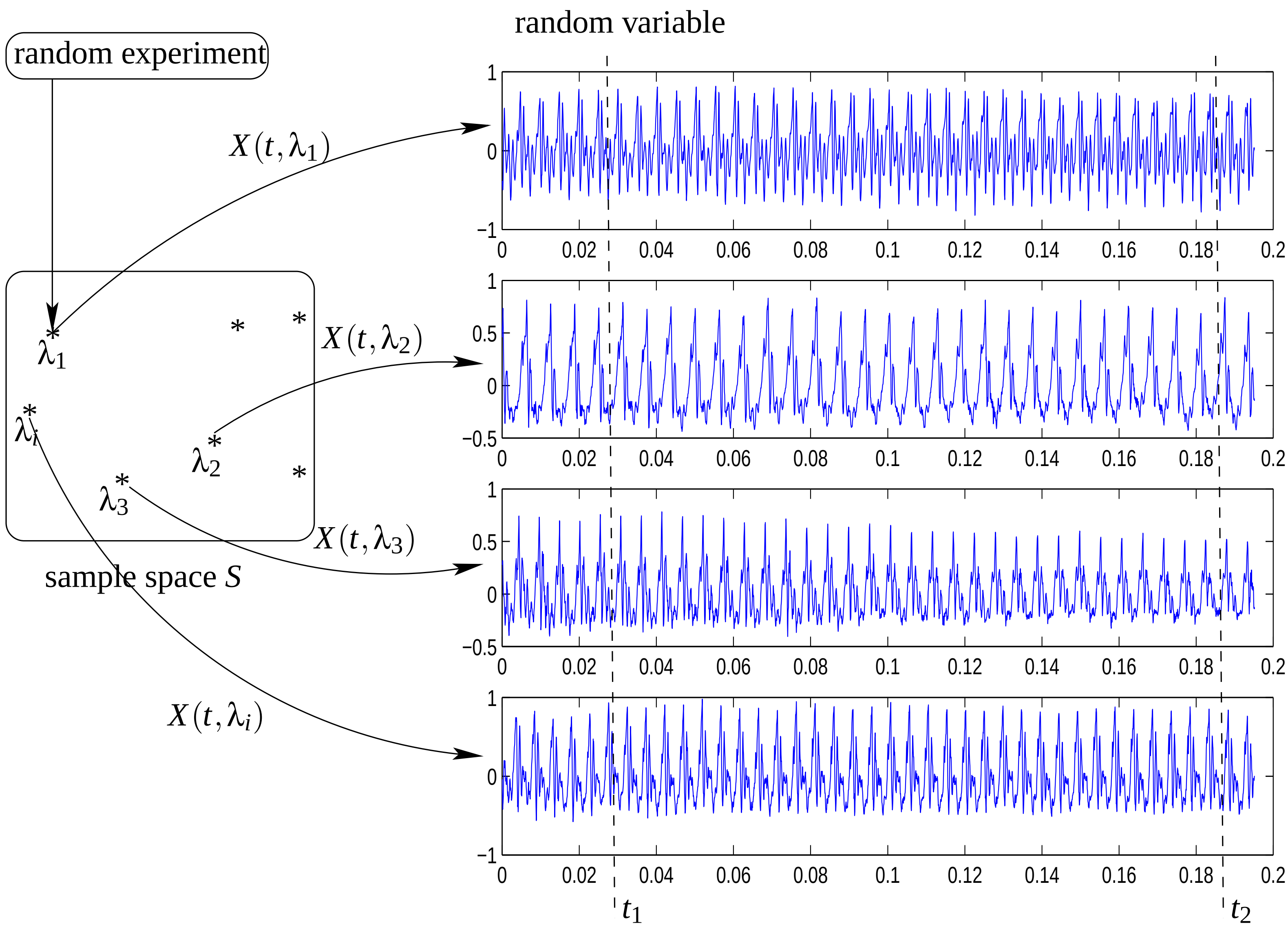
Two random variables X and Y are called statistically independent if their joint probability density function satisfies

$$p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y),$$

i.e. the value of one random variable is independent of the value of the other.

6.10 Random Processes

6.10.1 Definition and Examples



Random Process:

A random process $X(t)$ describes the mapping of a random experiment with sample space S onto an ensemble of sample functions $X(t, \lambda_i)$. For each point in time t_1, $X(t_1)$ describes a random variable.

Example: Rolling a Die

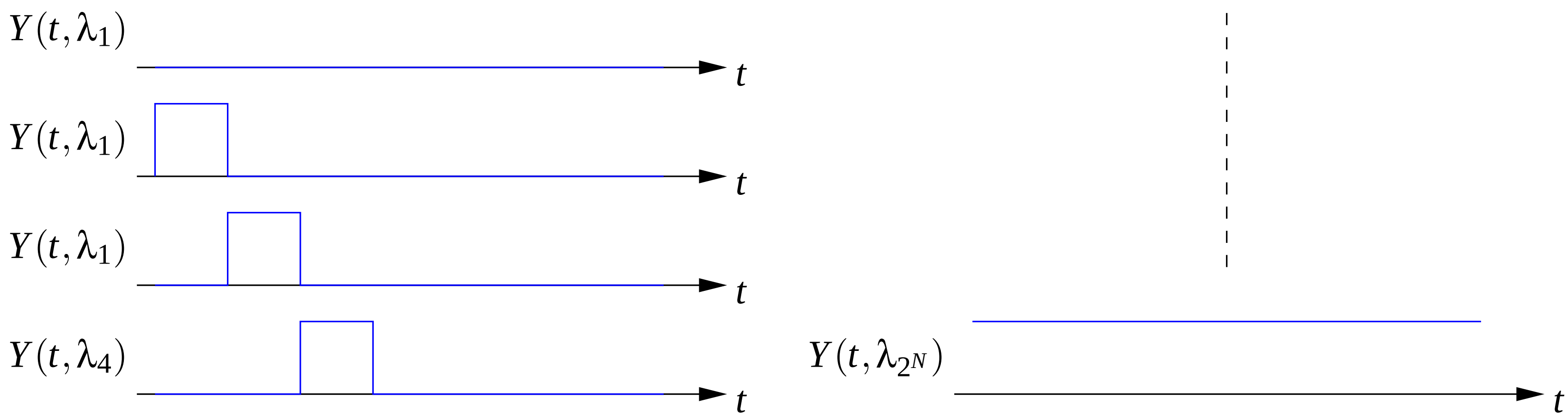
Random variable: $X = i$ if number i is on top of the die

Random process: $Y(t) = X \cdot \cos(\omega_0 t)$.

Example: Tossing a Coin N Times

Random variable: $X_n = 0$ if the nth result is head, $X_n = 1$ if the nth result is tail

Random process: $Y(t) = \sum_{n=1}^{N} X_n \cdot \text{rect}(t - n + 0.5)$.



Example: Filtering a Random Process

$$Y(t) = h(t) * X(t) = \int_{-\infty}^{\infty} h(\tau)X(t-\tau)d\tau$$

$X(t)$, $Y(t)$: random processes, $h(t)$: filter impulse response

6.10.2 Ensemble Averages and Stationarity

For each time instance of a random process, the average value, variance etc. can be calculated from all sample functions $X(t, \lambda_i)$.

Expected Value $E\{X(t)\}$:

For a random process $X(t)$ with probability density function $p_{X(t)}(x)$, the expected value $E\{X(t)\} = m_X(t)$ is given by:

$$E\{X(t)\} = \int\limits_{-\infty}^{\infty} x \cdot p_{X(t)}(x)dx = m_X(t)$$

Variance $\sigma_X(t)$:

$$\sigma_X^2(t) = E\{|X(t) - m_X(t)|^2\} = \int\limits_{-\infty}^{\infty} |x - m_X(t)|^2 \cdot p_{X(t)}(x)dx$$

For a **stationary** random process the probability density function is independent of time t, thus the expected value and the variance are also a constant over time.

$$p_{X(t)}(x) = p_{X(t+t_0)}(x), \qquad \forall \quad t, t_0$$

$$m_X(t) = m_X(t+t_0) = m_X, \qquad \sigma_X(t) = \sigma_X(t+t_0) = \sigma_X$$

6.10.3 Time Averages and Ergodicity

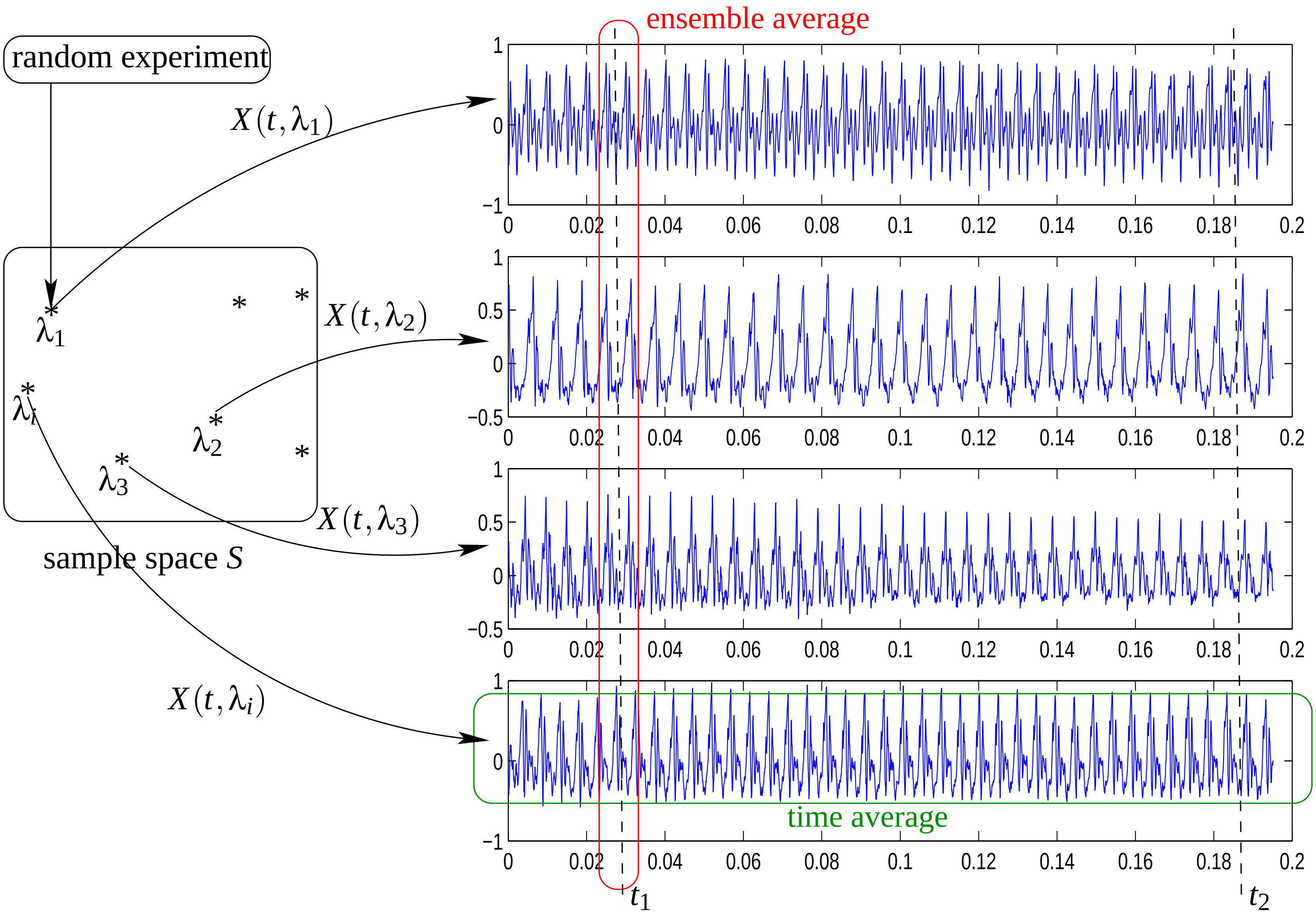
So far, the average value and the variance of a random process $X(t)$ were calculated based on the probability density function $p_{X(t)}$. However, in practical experiments the probability density function of a random process is often unknown. Also, in many cases, there is only one sample function $X(t, \lambda_i)$ available. Therefore, it is favorable to average over time instead of taking the ensemble average.

Average Value $m_{X(\lambda_i)}$**:**

$$m_{X(\lambda_i)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t, \lambda_i)dt$$

Variance $\sigma^2_{X(\lambda_i)}$**:**

$$\sigma^2_{X(\lambda_i)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} (X(t, \lambda_i) - m_{X_{\lambda_i}})^2 dt$$



ensemble average
random experiment
$X(t,\lambda_1)$
$X(t,\lambda_2)$
$X(t,\lambda_3)$
$X(t,\lambda_i)$
λ_1
λ_2
λ_3
λ_i
sample space S
time average
t_1
t_2

Ergodicity:

A stationary random process $X(t)$ is called **ergodic**, if the time averages of each sample function $X(t, \lambda_i)$ converge towards the corresponding ensemble average with probability one.

In practical applications ergodicity is often assumed since just one sample function is available and therefore the ensemble averages cannot be taken.

Example 1:

Random process: $X(t) = A\cos(\omega_0 t)$
A: discrete random variable with $P(A = 1) = P(A = 2) = 0.5$
ω_0: constant

Ensemble average:

$$m_X(t) = E\{X(t)\} = E\{A\}\cos(\omega_0 t) = 1.5\cos(\omega_0 t)$$

For $\omega_0 \neq 0$ the random process is not stationary and we are not allowed to take the time average.

Example 2:

Random process: $X(t) = A$

A: discrete random variable with $P(A = 1) = P(A = 2) = 0.5$

Ensemble average:

$$m_X(t) = E\{X(t)\} = E\{A\} = 1.5$$

$\Rightarrow$ the ensemble average is independent of time.

Time averages:

$$m_{X(1)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t,1)dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} 1\,dt = 1$$

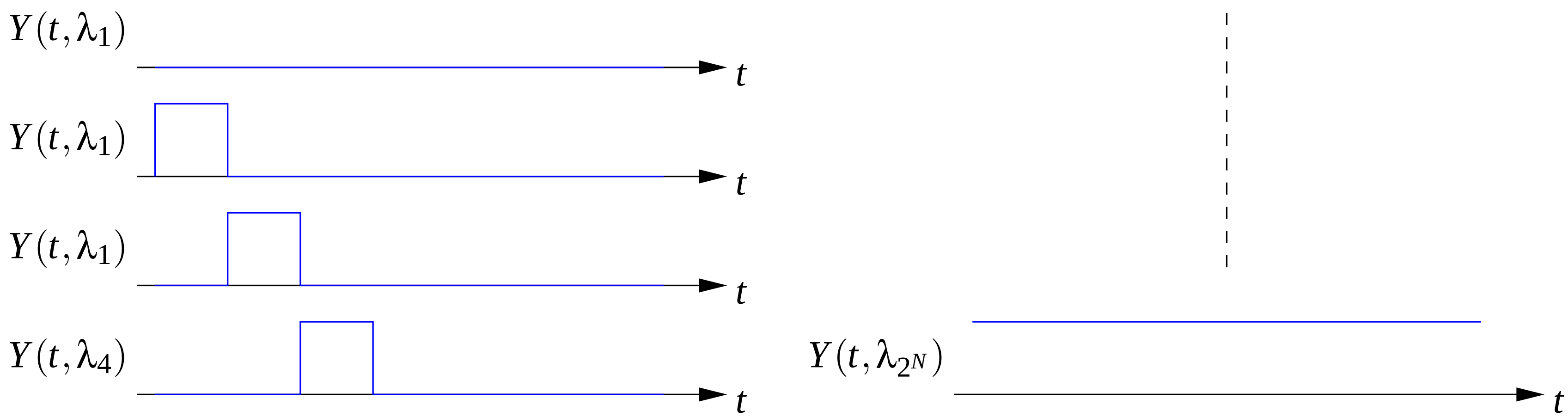
$$m_{X(2)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t,2)dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} 2\,dt = 2$$

$\Rightarrow$ time averages taken for different sample functions are not identical to the ensemble average, the random process is thus not ergodic.

Example 3: Tossing a Coin N Times

Random variable: $X_n = 0$ if the nth result is head, $X_n = 1$ if the nth result is tail.

Random process: $Y(t) = \sum_{n=1}^{N} X_n \cdot \text{rect}(t - n + 0.5)$



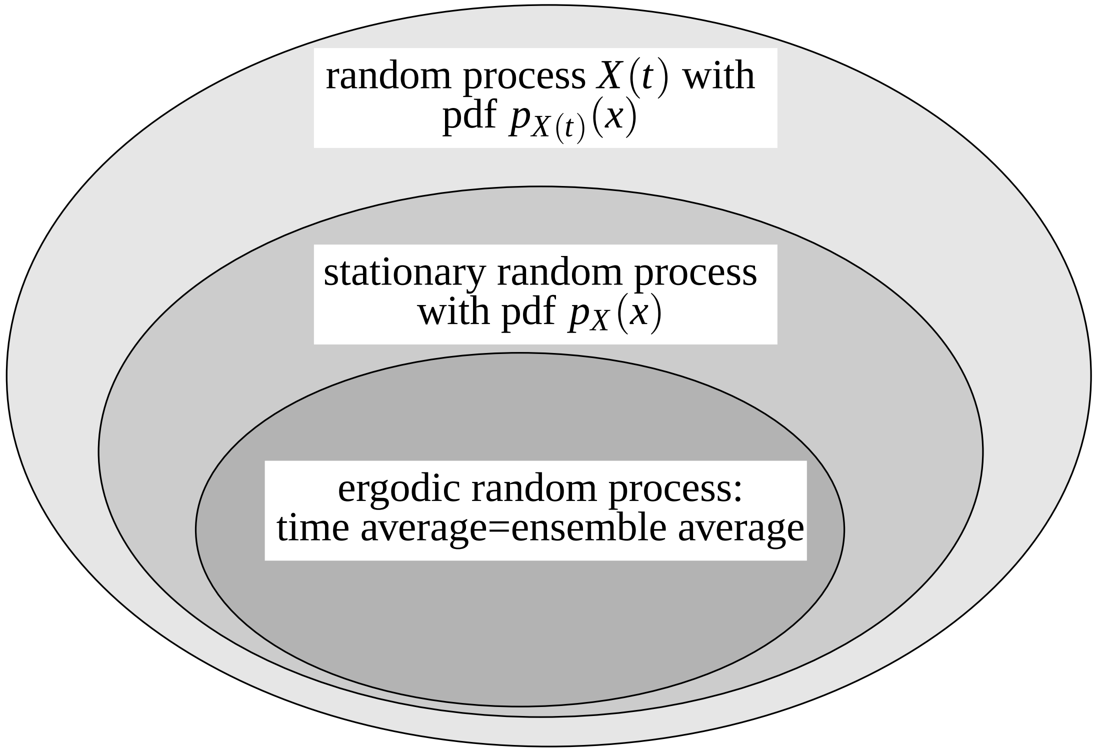
Ensemble average:

$$m_Y(t) = E\{Y(t)\} = 0.5$$

Time average: Sample function $Y(t, \lambda_i)$: Coin is tossed N times and we observe n_1 times head and n_2 times tail.

$$m_{Y(\lambda_i)} = \frac{1}{N} \int_0^N Y(t, \lambda_i) dt = \frac{1}{N}(n_1 \cdot 0 + n_2 \cdot 1) = \frac{n_2}{N}$$

$\Rightarrow$ for $N \to \infty$, n_1 and n_2 converge towards $N/2$ and $m_{Y(t,\lambda_i)} = m_Y(t) = m_Y$. The random process is thus ergodic.


random process $X(t)$ with
pdf $p_{X(t)}(x)$
stationary random process
with pdf $p_X(x)$
ergodic random process:
time average=ensemble average

6.11 Autocorrelation and Power Spectra

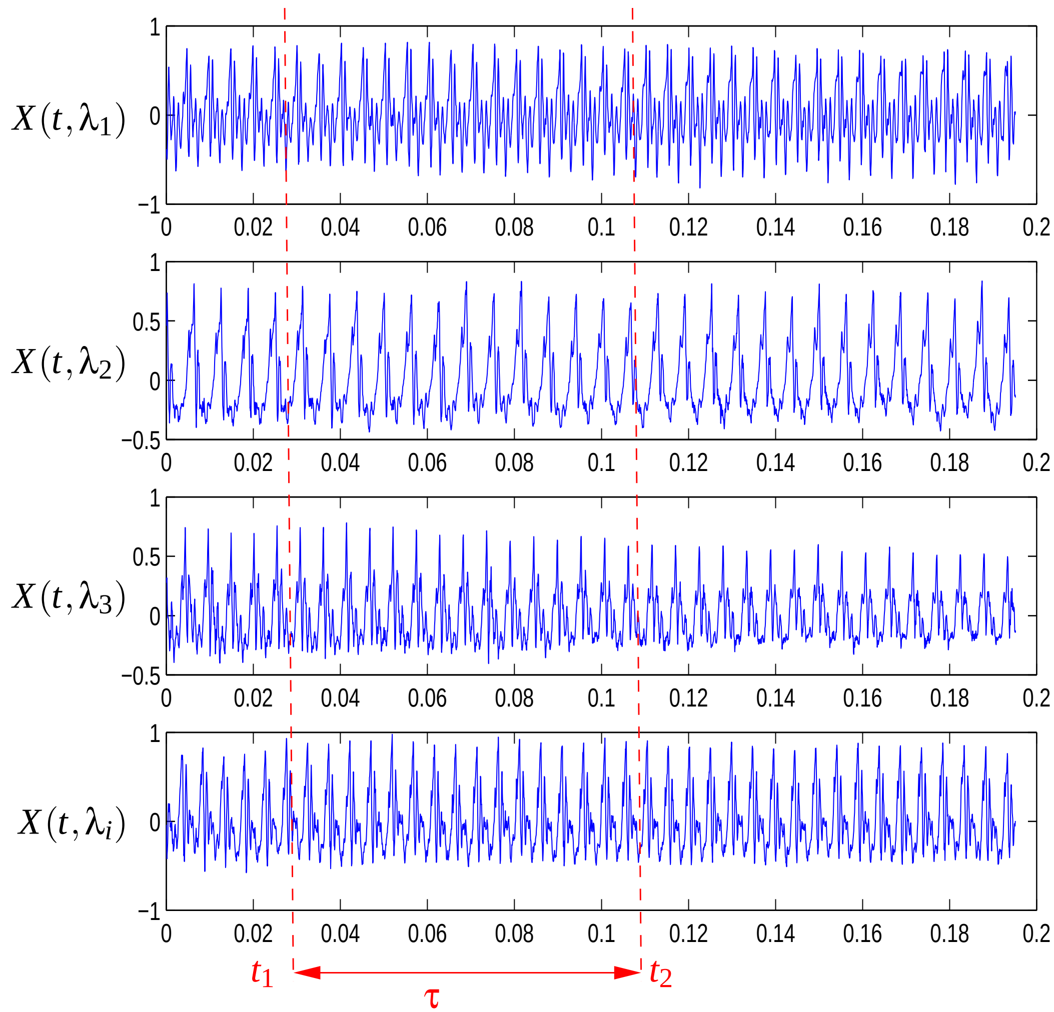
6.11.1 Autocorrelation and Autocovariance

We are interested in how the value of a random process $X(t)$ evaluated at t_2 depends on its value at time t_1.

At t_1 and t_2 the random process is characterized by random variables X_1 and X_2, respectively.

The relationship between X_1 and X_2 is given by the joint probability density function

$$p_{X_1 X_2}(x_1, x_2)$$



Autocorrelation Function:

$$R_{XX}(t_1, t_2) = E\{X_1\, X_2\}$$

$$= E\{X(t_1)\, X(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p_{X_1 X_2}(x_1, x_2) dx_1 dx_2$$

Autocovariance Function:

$$C_{XX}(t_1,t_2) = E\{(X(t_1) - m_X(t_1))\,(X(t_2) - m_X(t_2))\}$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x_1 - m_X(t_1))(x_2 - m_X(t_2))\, p_{X(t_1)X(t_2)}(x_1,x_2)dx_1dx_2$$

$$= R_{XX}(t_1,t_2) - m_X(t_1)m_X(t_2)$$

$C_{XX}(t,t)$ describes the variance $\sigma_X^2(t)$ of a random process.

Autocorrelation and Autocovariance Function of a Stationary Random Process:

The joint probability density function of a stationary process does not change if a constant value t is added to both t_1 and t_2.

$$p_{X_1\,X_2}(x_1,x_2) = p_{X(t_1)X(t_2)}(x_1,x_2) = p_{X(t_1+t)X(t_2+t)}(x_1,x_2)$$

The autocorrelation function then only depends on the difference τ between t_1 and t_2

$$R_{XX}(t_1,t_2) = E\{X(t_1)\,X(t_2)\} = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x_1x_2p_{X(t_1)X(t_2)}(x_1,x_2)dx_1dx_2$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} x_1x_2p_{X(0)X(t_2-t_1)}(x_1,x_2)dx_1dx_2 = R_{X,X}(0,t_2-t_1) = R_{X,X}(\tau)$$

Since the average value is a constant, the autocovariance function is given by:

$$
\begin{aligned}
C_{XX}(t_1, t_2) &= E\{(X(t_1) - m_X)\,(X(t_2) - m_X)\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - m_X)(x_2 - m_X)\, p_{X(t_1)X(t_2)}(x_1, x_2)dx_1dx_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - m_X)(x_2 - m_X)\, p_{X(0)X(t_2-t_1)}(x_1, x_2)dx_1dx_2 \\
&= C_{XX}(0, t_2 - t_1) = C_{XX}(\tau)
\end{aligned}
$$

Properties of the Autocorrelation Function of a Stationary Random Process:

- Symmetry: $R_{XX}(\tau) = R_{XX}(-\tau)$
- Mean Square Average: $R_{XX}(0) = E\{X(t)^2\} \geq 0$
- Maximum: $R_{XX}(0) \geq |R_{XX}(\tau)|$
- Periodicity: if $R_{XX}(0) = R_{XX}(t_0)$, then $R_{XX}(\tau)$ is periodic with period t_0.

Wide Sense Stationary (WSS) Random Process:

A random process $X(t)$ is called **WSS** if the following three properties are satisfied:

- The average value of the random process is a constant: $m_X(t) = m_X$
- The autocorrelation and autocovariance function only depend on the time difference $\tau = t_1 - t_2$:

$$R_{XX}(t_1, t_2) = R_{XX}(t_2 - t_1) = R_{XX}(\tau)$$

$$C_{XX}(t_1, t_2) = C_{XX}(t_2 - t_1) = C_{XX}(\tau)$$

- The variance is a constant and finite: $\sigma_X^2 = C_{XX}(0) = R_{XX}(0) - m_X^2 < \infty$

Autocorrelation and Autocovariance Function of an Ergodic Random Process:

$$R_{XX}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X_T(t, \lambda_i) X_T(t+\tau, \lambda_i) dt$$

$$C_{XX}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} (X_T(t, \lambda_i) - m_X)(X_T(t+\tau, \lambda_i) - m_X) dt$$

$X_T(t, \lambda_i)$: sample function of random process $X(t)$ windowed to be of length T (starting at $-T/2$ ending at $T/2$).

6.11.2 Power Spectral Density

Motivation:

- Description of random processes in the frequency domain
- Calculation of the Fourier Transform of a sample function is not useful
- We assume in the following that the random process considered is at least WSS if not stationary.

The power spectral density (psd) of a WSS random process $X(t)$ is defined as the Fourier Transform of the autocorrelation function $R_{XX}(\tau)$:

$$S_{XX}(\omega) = \mathcal{F}\{R_{XX}(\tau)\} = \int\limits_{-\infty}^{\infty} R_{XX}(\tau)\exp(-j\omega\tau)d\tau$$

Inverse transform:

$$R_{XX}(\tau) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_{XX}(\omega)\exp(j\omega\tau)\,d\omega$$

Properties of the Power Spectral Density:

$$S_{XX}(\omega) = S_{XX}(-\omega), \qquad S_{XX}(\omega) \geq 0, \qquad \mathrm{Im}\{S_{XX}(\omega)\} = 0$$

Ergodic Random Process $x(t)$:

Autocorrelation Function:

$$R_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x_T(t,\lambda_i)x_T(t+\tau,\lambda_i)dt = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x_T(t)x_T(t+\tau)dt$$

Power Spectral Density:

$$\begin{aligned}
S_{xx}(\omega) &= \int_{-\infty}^{\infty} R_{xx}(\tau)\exp(-j\omega\tau)d\tau \\
&= \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_T(t)x_T(t+\tau)dt\ \exp(-j\omega\tau)d\tau \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_T(t) \underbrace{\int_{-\infty}^{\infty} x_T(t+\tau)\exp(-j\omega\tau)d\tau}_{X_T(\omega)\exp(j\omega t)}\ dt \\
&= X_T(\omega) \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_T(t)\exp(j\omega t)\,dt \\
&= \lim_{T\to\infty} \frac{X_T(\omega)X_T^*(\omega)}{T} = \lim_{T\to\infty} \frac{|X_T(\omega)|^2}{T}
\end{aligned}$$

6.11.3 Deterministic Power and Energy Signals

Power Signal:

The autocorrelation function and the power spectral density can also be calculated for a deterministic power signal $f(t)$. In this case, the signal simply replaces the random process.

With $f_T(t) = f(t) \cdot \mathrm{rect}(t/T)$ we obtain for the **autocorrelation function**:

$$R_{ff}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f_T(t) f_T(t+\tau) dt$$

and for the **power spectral density**:

$$S_{ff}(\omega) = \int_{-\infty}^{\infty} R_{ff}(\tau) \exp(-j\omega\tau) d\tau = \lim_{T\to\infty} \frac{F_T(\omega) F_T^*(\omega)}{T} = \lim_{T\to\infty} \frac{|F_T(\omega)|^2}{T}$$

Note that we obtain the power of $f(t)$ as

$$P = R_{ff}(0) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f_T^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{ff}(\omega) d\omega$$

Energy Signals:

For an energy signal $f(t)$ an **energy autocorrelation function** can be defined as

$$R_{ff}^{E}(\tau) = \int\limits_{-\infty}^{\infty} f(t)f(t+\tau)dt = f(\tau) * f(-\tau)$$

Applying the Fourier Transform to the energy autocorrelation function, we obtain the **energy spectral density** as:

$$S_{ff}^{E}(\omega) = \int\limits_{-\infty}^{\infty} R_{ff}^{E}(\tau)\exp(-j\omega\tau)d\tau = F(\omega)F^{*}(\omega) = |F(\omega)|^{2}$$

We obtain the energy as

$$E = R_{ff}^{E}(0) = \int\limits_{-\infty}^{\infty} f(t)^{2}dt = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_{ff}^{E}(\omega)d\omega = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} |F(\omega)|^{2}$$

which restates Parseval's Theorem.

6.11.4 Examples of Autocorrelation Functions and Power Spectral Densities

Example 1: Sinusoid with Random Phase Angle

Random process:

$$X(t) = A\,\sin(\omega_0 t + \phi)$$

A, ω_0: constant values,

ϕ: random variable with probability density function $p_\phi(x)$:

$$p_\phi(x) = \begin{cases} 1/2\pi & \text{for } -\pi \leq x < \pi \\ 0 & \text{otherwise} \end{cases}$$

Average value:

$$\begin{aligned} m_X(t) = E\{X(t)\} &= E\{A\,\sin(\omega_0 t + \phi)\} \\ &= \int_{-\infty}^{\infty} A\,\sin(\omega_0 t + x) p_\phi(x)\,dx \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} A\,\sin(\omega_0 t + x)\,dx = 0 \end{aligned}$$

The average value is a constant and independent of t. Since $m_X = 0$ the autocorrelation and autocovariance functions are identical.

Autocorrelation function:

$$R_{XX}(t_1, t_2) = E\{X(t_1)X(t_2)\} = E\{A\ \sin(\omega_0 t_1 + \phi)\, A\ \sin(\omega_0 t_2 + \phi)\}$$

Applying: $\quad 2\sin(A)\sin(B) = \cos(A - B) - \cos(A + B)$

$$R_{XX}(t_1, t_2) = E\{0.5A^2 \cos(\omega_0(t_2 - t_1)) - 0.5A^2 \cos(\omega_0(t_2 + t_1) + 2\phi)\}$$

$$= 0.5A^2 E\{\cos(\omega_0(t_2 - t_1))\} - 0.5A^2 E\{\cos(\omega_0(t_2 + t_1) + 2\phi)\}$$

$$= 0.5A^2 \cos(\omega_0(t_2 - t_1)) - 0 = R_{XX}(\tau)$$

The autocorrelation function only depends on $\tau = t_2 - t_1$ but not on the absolute values of t_1 and t_2.

Power spectral density:

$$S_{XX}(\omega) = \mathcal{F}\{R_{XX}(\tau)\} = \mathcal{F}\{0.5A^2 \cos(\omega_0 \tau)\} = 0.5A^2 \mathcal{F}\{\cos(\omega_0 \tau)\}$$

$$= 0.5A^2 \pi\left(\delta(\omega - \omega_0) + \delta(\omega + \omega_0)\right)$$

Ergodicity:

If the random process is ergodic, we obtain the same results for the average value and the autocorrelation function by taking the time averages over one sample function.

Let us assume the sample function $X(t, \lambda_i)$ has a phase angle ϕ_i:

$$X(t, \lambda_i) = A\ \sin(\omega_0 t + \phi_i)$$

Average value (time average):

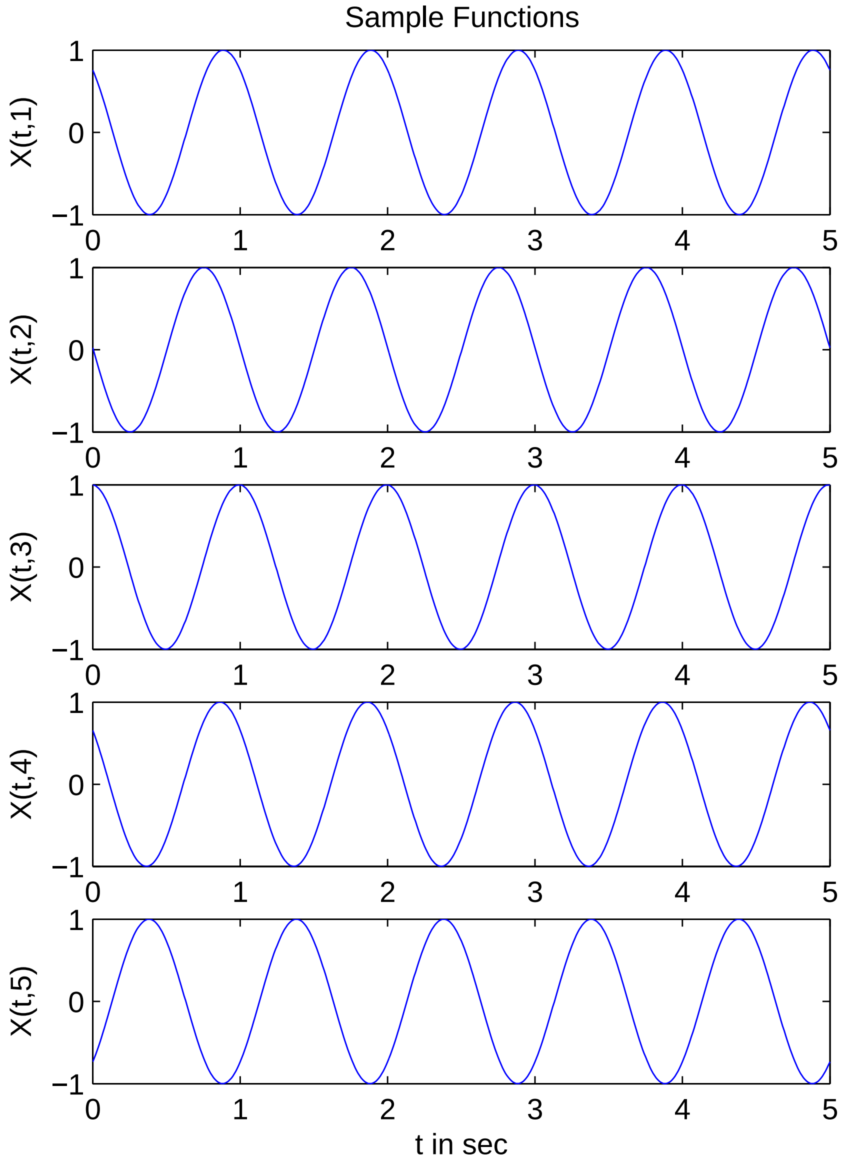
$$\begin{aligned} m_{X(\lambda_i)} &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t, \lambda_i) dt \\ &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} A\ \sin(\omega_0 t + \phi_i) dt = 0 = m_X \end{aligned}$$

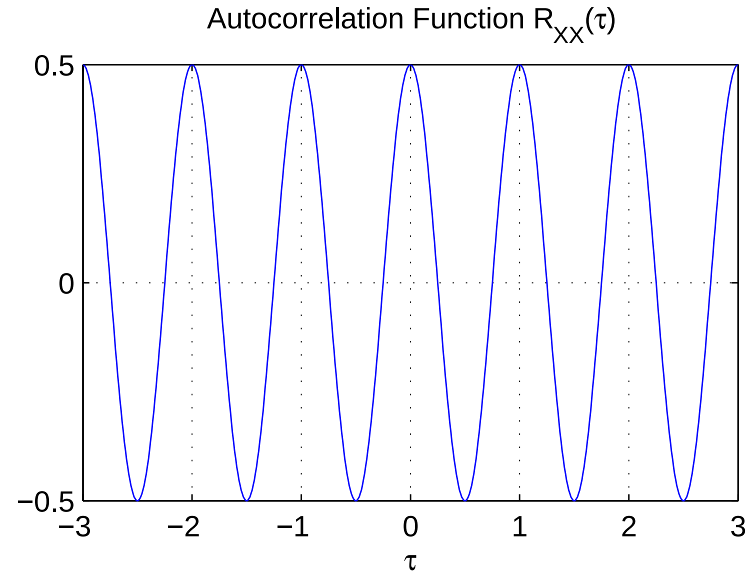
Autocorrelation function (time average):

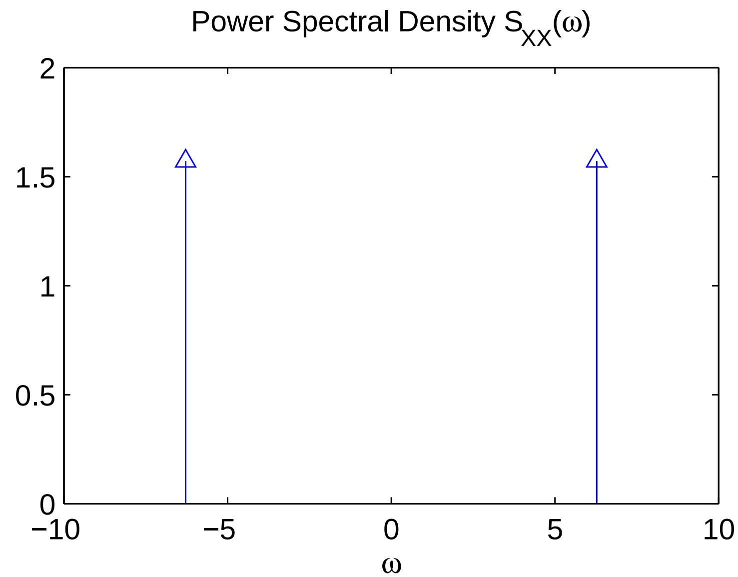
$$\begin{aligned} R_{X(\lambda_i)X(\lambda_i)}(\tau) &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t, \lambda_i) X(t+\tau, \lambda_i) dt \\ &= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} A^2\ \sin(\omega_0 t + \phi_i) \sin(\omega_0(t+\tau) + \phi_i) dt \\ &= \frac{A^2}{2} \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \cos(\omega_0 \tau) + \cos(\omega_0(2t+\tau) + 2\phi_i) dt \\ &= \frac{A^2}{2} \cos(\omega_0 \tau) = R_{XX}(\tau) \end{aligned}$$

Time averages of one sample function and ensemble averages are identical
$\Rightarrow$ the random process is ergodic.

$A = 1,\ \omega_0 = 2\pi$







Example 2: Binary Data Transmission

A binary sequence is transmitted by rectangular pulses of width T_b. The amplitude of the pulse is determined by each bit, i.e. it is one if the bit is one and zero if the bit is zero. We assume that ones and zeros are equally likely and that each bit is statistically independent of all others. Using ergodicity, we obtain the following results from a sample function $X(t,1)$:

Average value (sample function of length N bits with n_1 ones):

$$m_X = E\{X(t,1)\} \lim_{T\to\infty} \frac{1}{T}\int_0^T X(t,1)dt = \lim_{N\to\infty} \frac{1}{NT_b}\int_0^{NT_b} X(t,1)dt$$

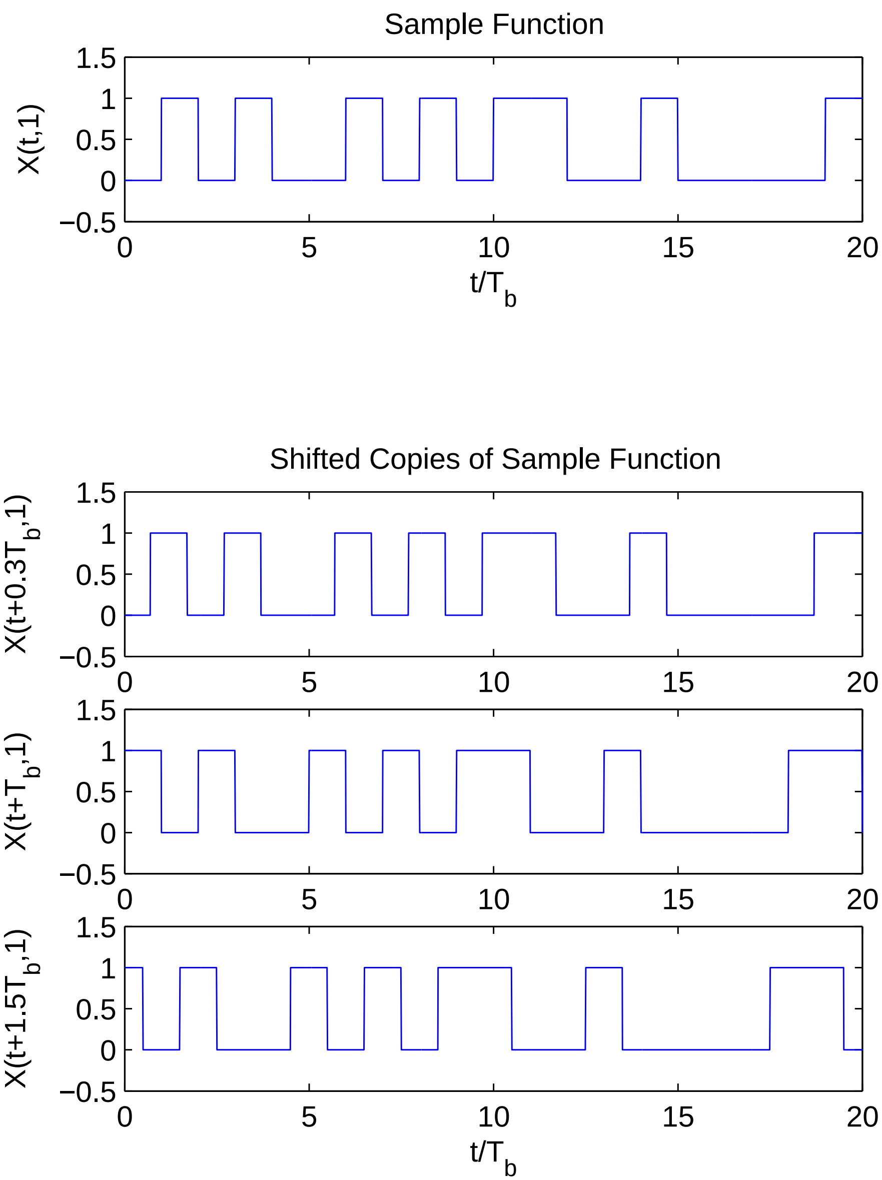
$$= \lim_{N\to\infty} \frac{1}{NT_b}[n_1 \cdot 1 \cdot T_b + (N-n_1)\cdot 0 \cdot T_b] = \lim_{N\to\infty} \frac{n_1}{N} = 0.5$$

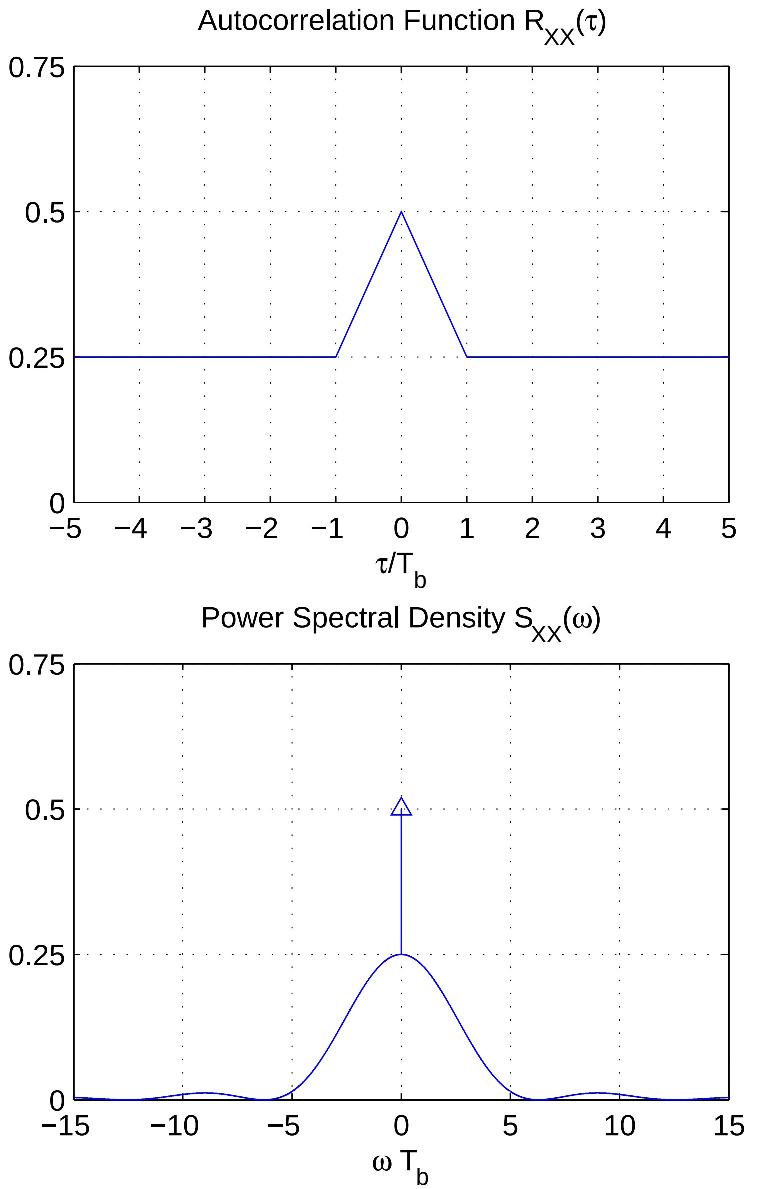
Autocorrelation function:

$$R_{XX}(\tau) = E\{X(t,1)X(t+\tau,1)\} = \lim_{T\to\infty} \frac{1}{T}\int_0^T X(t,1)X(t+\tau,1)dt$$
$$= 0.25(\delta(t) + \Lambda(\tau/T_b))$$

Power spectral density:

$$S_{XX}(\omega) = \mathcal{F}\{R_{XX}(\tau)\} = 0.25(1 + \mathsf{Sa}^2(\omega T_b/2))$$


Sample Function
X(t,1)
t/T_b
Shifted Copies of Sample Function
X(t+0.3T_b,1)
X(t+T_b,1)
X(t+1.5T_b,1)
t/T_b


Autocorrelation Function R_XX(τ)
τ/T_b
Power Spectral Density S_XX(ω)
ω T_b

Example 3: Deterministic Energy Signal

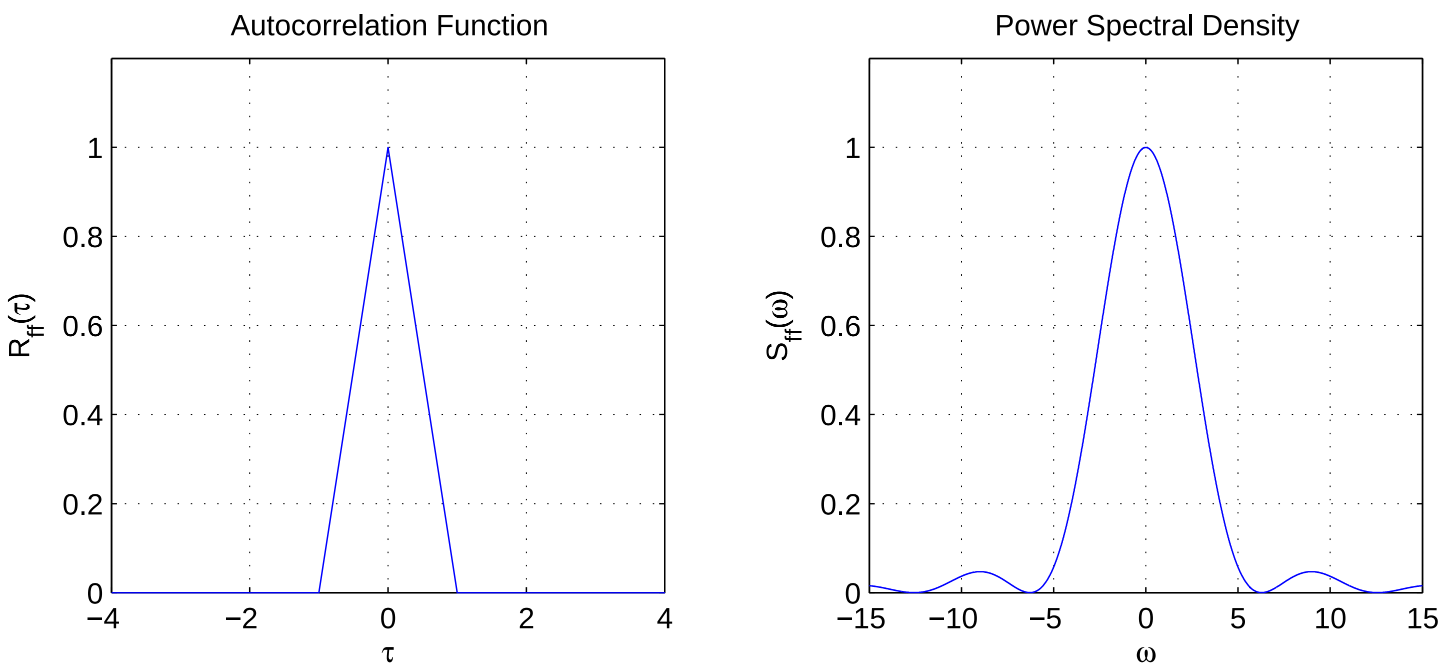
$$f(t) = \text{rect}(t - 0.5)$$

Energy autocorrelation function:

$$R_{ff}^{E}(\tau) = \int_{-\infty}^{\infty} f(t) f(t+\tau) dt = \Lambda(\tau)$$

Energy spectral density:

$$S_{ff}^{E}(\omega) = \mathcal{F}\{R_{ff}(\tau)\} = \text{Sa}^2(\omega/2)$$



Example 4: White Noise

A random process $n(t)$ is called white noise, if it has a constant power spectral density of $\eta/2$ watts per Hz measured over positive frequencies. If in addition the random process has zero mean ($m_n = 0$), the power spectral density is given by:

$$S_{nn}(\omega) = \eta/2 \qquad \text{for all } \omega$$

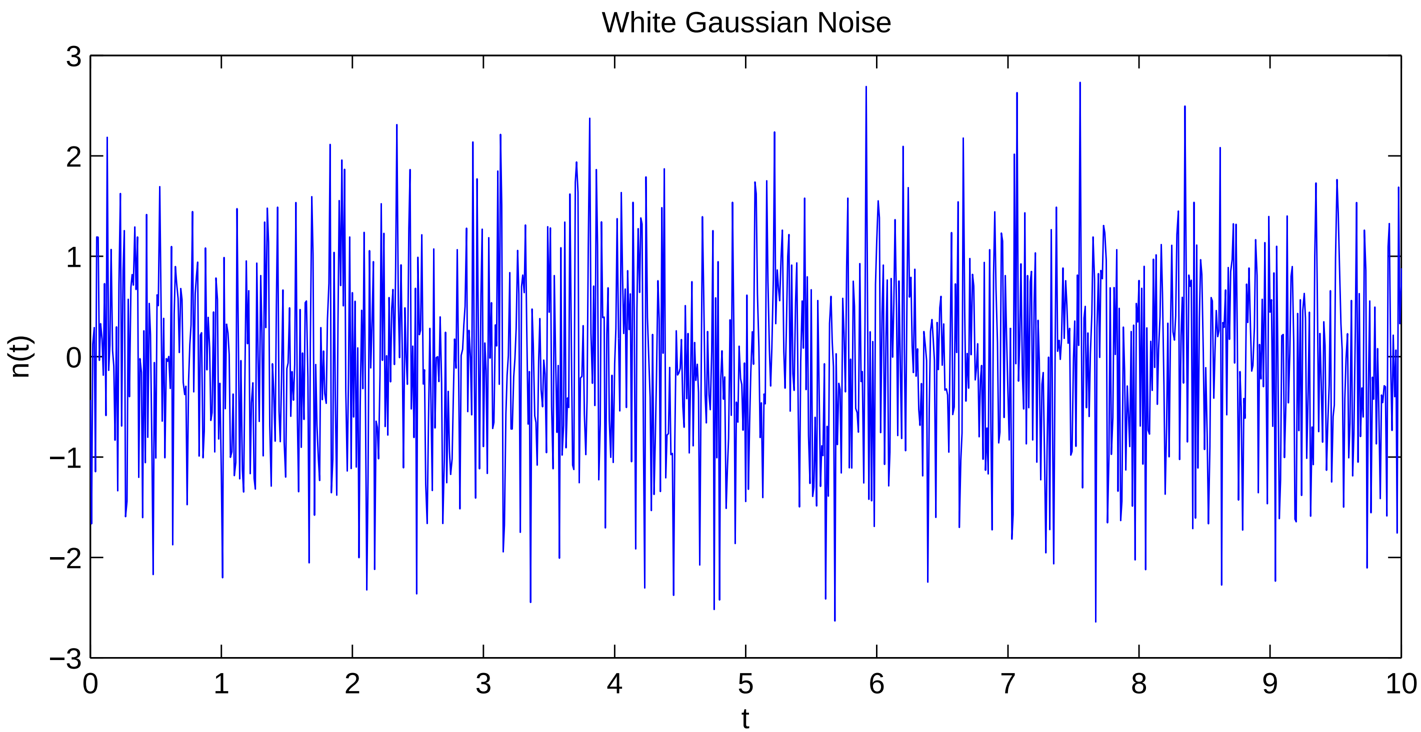
Autocorrelation function:

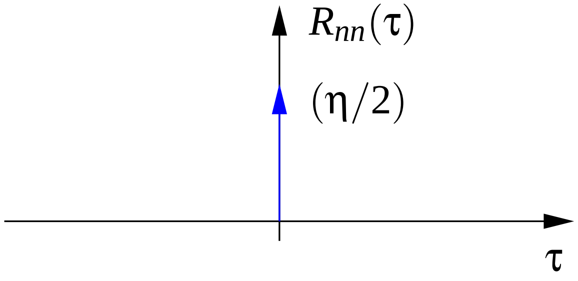
$$R_{nn}(\tau) = \mathcal{F}^{-1}\{S_{nn}(\omega)\} = \frac{\eta}{2}\,\delta(t)$$

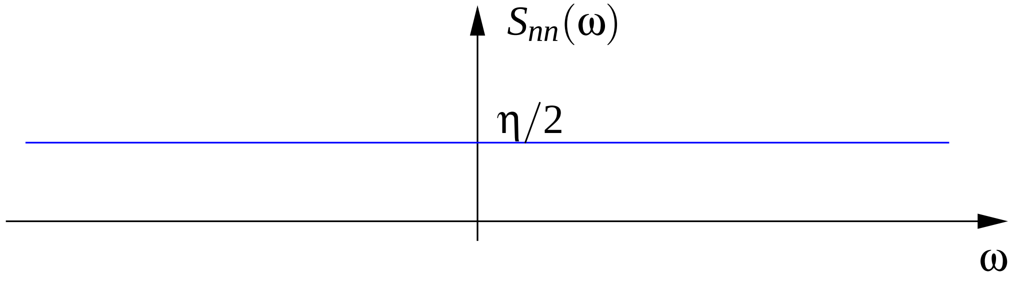
Since only the first and second moment of the process are known, the probability density function cannot be uniquely determined.

In the case of a Gaussian probability density function, the process is called **white Gaussian noise**.

If the white Gaussian noise is added to the signal, we denote it as **additive white Gaussian noise (AWGN)**.

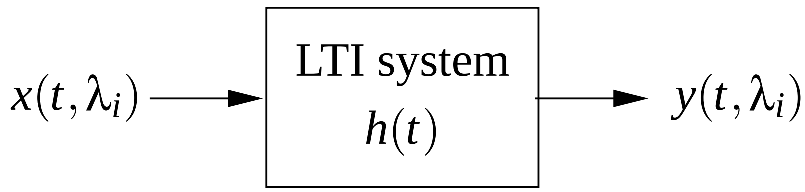

White Gaussian Noise
n(t)
t
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
10


$R_{nn}(\tau)$
$(\eta/2)$
τ


$S_{nn}(\omega)$
$\eta/2$
ω

6.12 Excitation of LTI Systems with Stationary Random Processes

Excitation of an LTI system with sample function $x(t, \lambda_i)$ of a stationary random process $x(t)$.



Sample function $y(t, \lambda_i)$ of the output random process $y(t)$:

$$y(t, \lambda_i) = h(t) * x(t, \lambda_i)$$

6.12.1 Expected Value of the Output Random Process

$$m_y = E\{y(t)\} = E\{x(t) * h(t)\} = E\left\{\int_{-\infty}^{\infty} x(t-\nu)h(\lambda)d\nu\right\}$$

$$= \int_{-\infty}^{\infty} E\left\{x(t-\nu)\right\} h(\nu)d\nu = \int_{-\infty}^{\infty} m_x\, h(\nu) \underbrace{\exp(j0\nu)}_{1} d\nu = m_x H(0)$$

6.12.2 Autocorrelation Function of the Output Random Process

$$
\begin{aligned}
R_{yy}(\tau) &= E\{y(t)y(t+\tau)\} = E\{(x(t)*h(t))(x(t+\tau)*h(t+\tau))\} \\
&= E\left\{\int_{-\infty}^{\infty} h(\nu)x(t-\nu)d\nu \int_{-\infty}^{\infty} h(\mu)x(t+\tau-\mu)d\mu\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\nu)h(\mu)E\left\{x(t-\nu)x(t+\tau-\mu)\right\} d\nu d\mu \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\nu)h(\mu)R_{xx}(\tau-\mu+\nu)d\nu d\mu \\
&\overset{\lambda=\mu-\nu}{=} \int_{-\infty}^{\infty}\underbrace{\int_{-\infty}^{\infty} h(\nu)h(\nu+\lambda)d\nu}_{R_{hh}^{E}(\lambda)}\, R_{xx}(\tau-\lambda)d\lambda \\
&= \int_{-\infty}^{\infty} R_{hh}^{E}(\lambda)R_{xx}(\tau-\lambda)d\lambda = R_{hh}^{E}(\tau)*R_{xx}(\tau)
\end{aligned}
$$

$R_{hh}^{E}(\tau)$: Energy autocorrelation function of the system impulse response

6.12.3 Power Spectral Density of the Output Random Process

$$S_{yy}(\omega) = \mathcal{F}\{R_{yy}(\tau)\} = \mathcal{F}\{R^E_{hh}(\tau) * R_{xx}(\tau)\}$$

$$= \mathcal{F}\{R^E_{hh}(\tau)\} \cdot \mathcal{F}\{R_{xx}(\tau)\} = \mathcal{F}\{R^E_{hh}(\tau)\} \cdot S_{xx}(\omega)$$

with

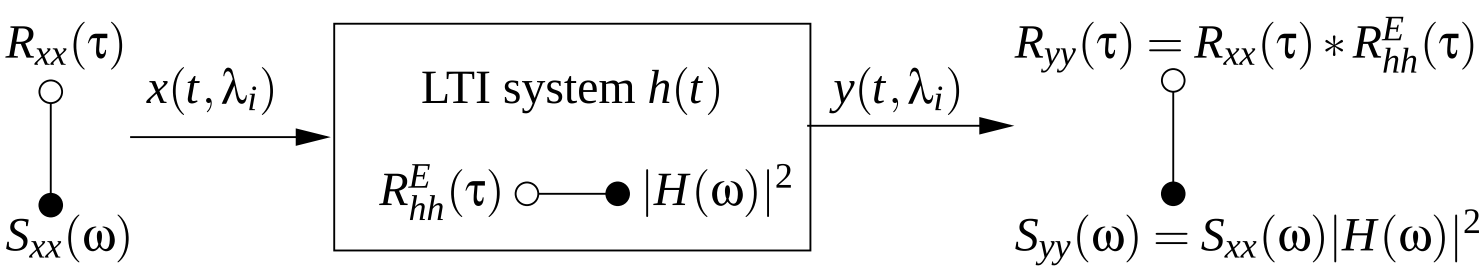
$$R^E_{hh}(\tau) = h(\tau) * h(-\tau)$$

$$h(\tau) \circ\!\!-\!\!\bullet\ H(\omega), \qquad h(-\tau) \circ\!\!-\!\!\bullet\ H^*(\omega)$$

$$R^E_{hh}(\tau) \circ\!\!-\!\!\bullet\ H(\omega)H^*(\omega) = |H(\omega)|^2$$

and thus

$$S_{yy}(\omega) = S_{xx}(\omega) \cdot |H(\omega)|^2$$



Example: Ideal Lowpass Filtering of White Noise

Input random process $n_i(t)$:

$$S_{n_in_i}(\omega) = 0.5\eta \bullet\!\!-\!\!\circ\; 0.5\eta\,\delta(\tau) = R_{n_in_i}(\tau)$$

Ideal lowpass filter:

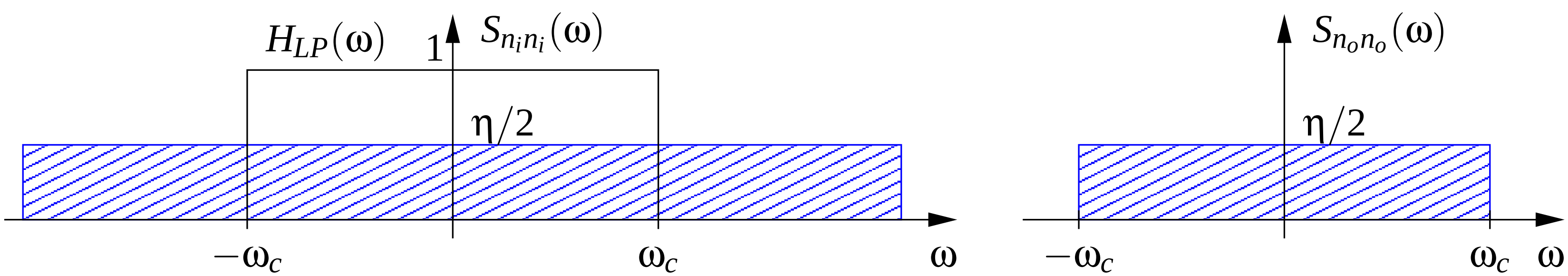
$$H_{LP}(\omega) = \begin{cases} 1 & \text{for } |\omega| < \omega_c \\ 0 & \text{otherwise} \end{cases}$$

Output random process $n_o(t)$:

$$S_{n_on_o}(\omega) = S_{n_in_i}(\omega)|H_{LP}(\omega)|^2 = \begin{cases} 0.5\eta & \text{for } |\omega| < \omega_c \\ 0 & \text{otherwise} \end{cases}$$

Power of output random process:

$$P_{n_o} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{n_on_o}(\omega)d\omega = \frac{1}{2\pi}\int_{\omega_c}^{\omega_c} 0.5\eta d\omega = \frac{\eta\omega_c}{2\pi} = \eta f_c$$



6.12.4 Cross-Correlation between Input and Output Random Process

The autocorrelation function describes the statistical properties of two random variables X_1 and X_2 taken from the same random process at times t_1 and t_2, respectively, $X_1 = X(t_1)$ and $X_2 = X(t_2)$.

The **cross-correlation function** describes the statistical properties of two random variables X_1 and Y_2 taken from two different random processes $X(t)$ and $Y(t)$ (here input and output of an LTI system) at times t_1 and $t-2$, respectively, such that $X_1 = X(t_1)$ and $Y_2 = Y(t_2)$. It is defined as:

$$R_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)\}$$

For stationary processes, it simplifies to:

$$R_{XY}(\tau) = E\{X(t)\,Y(t+\tau)\}$$

Two random processes $X(t)$ and $Y(t)$ are called **uncorrelated** if

$$R_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)\} = E\{X(t_1)\} \cdot E\{Y(t_2)\} = m_X(t_1) \cdot m_Y(t_2)$$

They are called **orthogonal** if

$$R_{XY}(t_1, t_2) = 0, \qquad \text{for all } t_1, t_2.$$

Here:

$$\begin{aligned}
R_{xy}(\tau) &= E\{x(t)\, y(t+\tau)\} = E\{x(t)(x(t+\tau) * h(t+\tau))\} \\
&= E\left\{ x(t) \int_{-\infty}^{\infty} h(\mu) x(t+\tau-\mu) d\mu \right\} \\
&= \int_{-\infty}^{\infty} h(\mu) E\left\{x(t)x(t+\tau-\mu)\right\} d\mu \\
&= \int_{-\infty}^{\infty} h(\mu) R_{xx}(\tau-\mu) d\mu = h(\tau) * R_{xx}(\tau)
\end{aligned}$$

Example: System Identification

An LTI system with unknown impulse response is excited with a white noise random process $n_i(t)$ with power spectral density $S_{n_i n_i} = \eta/2$. The output noise process is $n_o(t)$. The cross-correlation between input and output noise process is given by:

$$R_{n_i n_o}(\tau) = h(\tau) * R_{n_i,n_i}(\tau) = h(\tau) * \frac{\eta}{2}\,\delta(\tau) = \frac{\eta}{2}\, h(\tau)$$