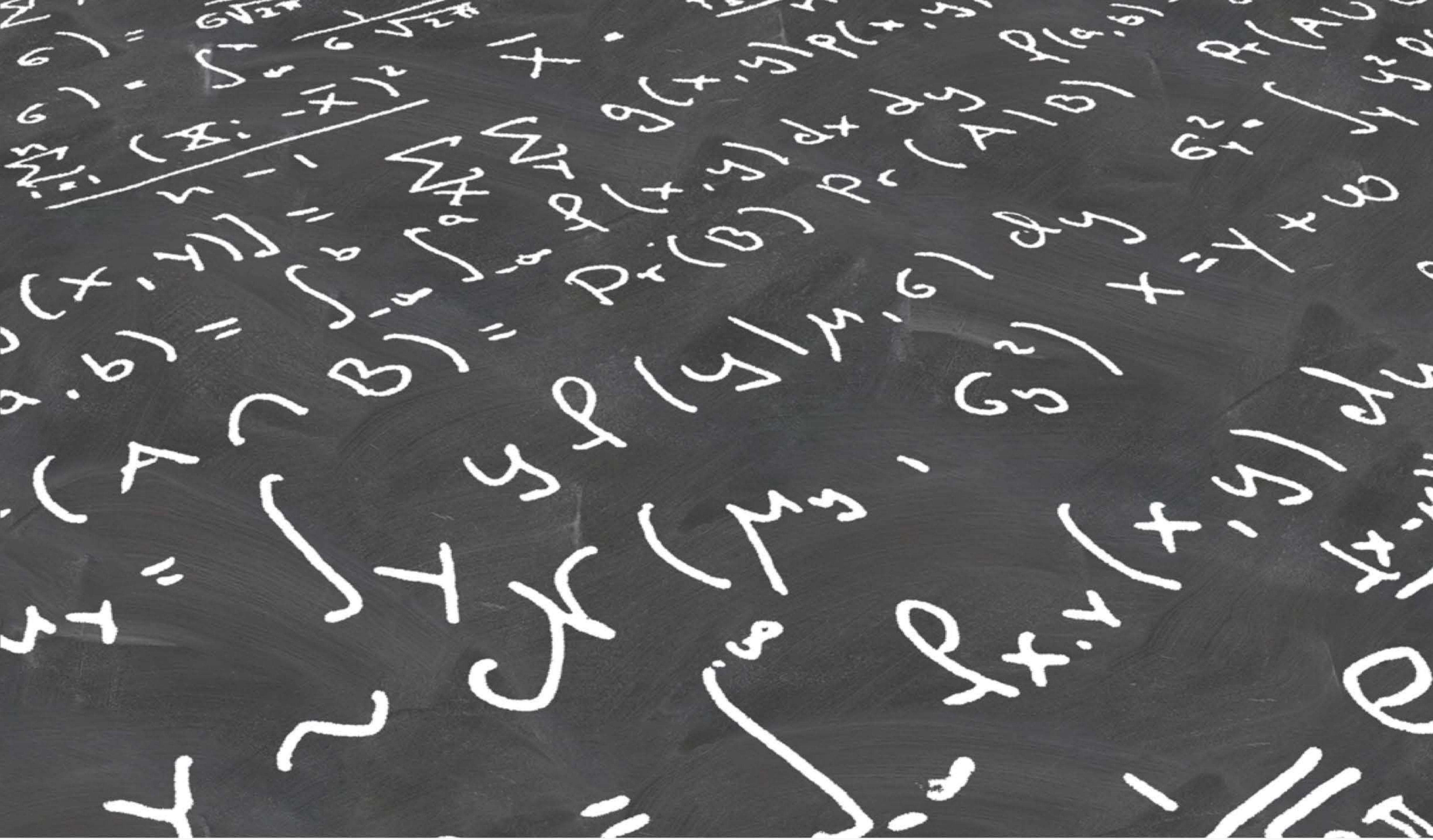


# 6. Introduction to Stochastic Processes

Gunvor Elisabeth Kirkelund
Lars Mandrup

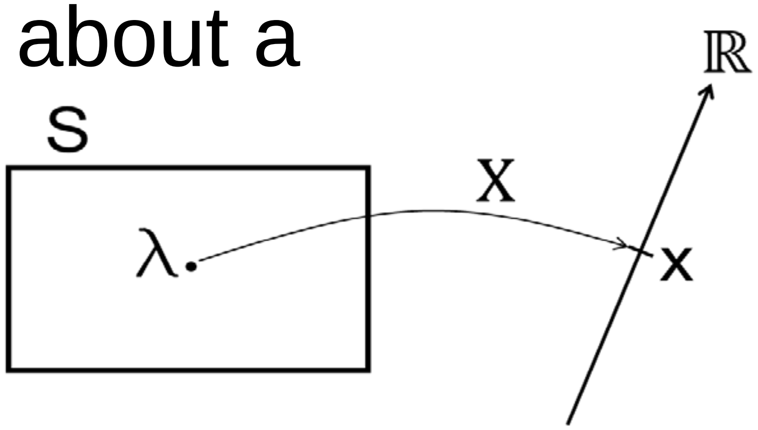
# Agenda for Today

- Repetition from last time
  - Random Variables
  - The Central Limit Theorem
- Stochastic Processes
  - Stationarity (WSS, SSS)
  - Ergodic Processes

Also just called a random variables

# Stochastic Random Variables

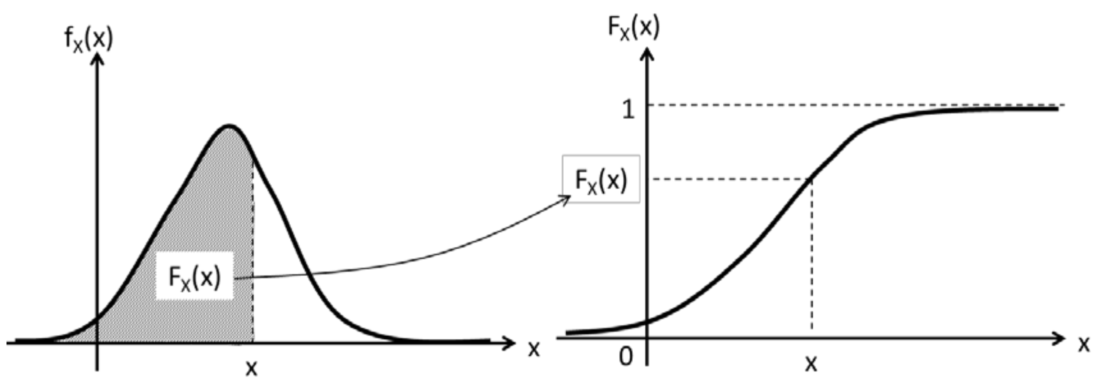
- A random variable tells something important about a stochastic experiment.
- Can be discrete or continous




- Probability density function (pdf):

$$Pr(a \leq X \leq b) = \int_a^b f_X(x)\, dx \qquad f_X(x) \geq 0 \qquad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

- Cumulative distribution function (cdf):




$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du = Pr(X \leq x)$$

$$0 \leq F_X(x) \leq 1$$

$$\lim_{x \to -\infty} F_X(x) = 0 \qquad \lim_{x \to \infty} F_X(x) = 1$$

# Two Random Variables X, Y

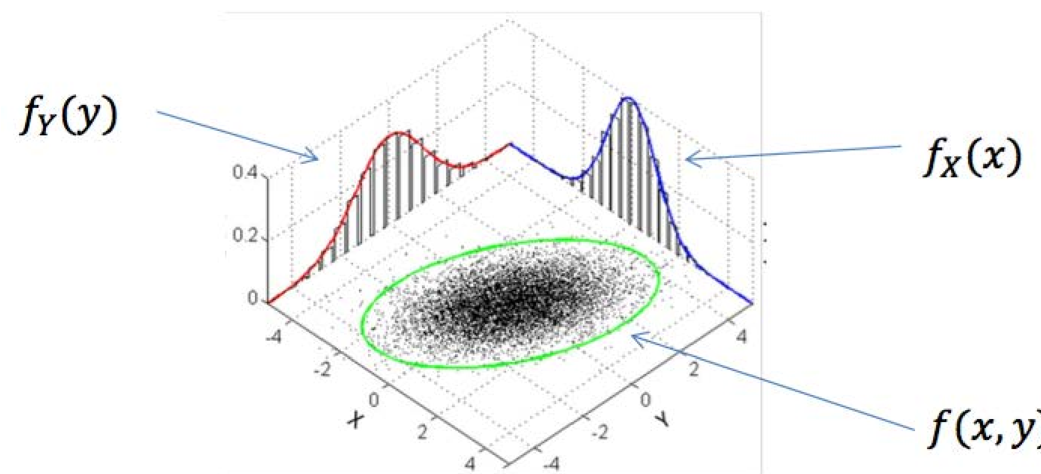
**Joint (Simultaneous) pdf:** $f_{X,Y}(x,y) \geq 0 \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$

$$Pr\big((a \leq X \leq b) \cap (c \leq Y \leq d)\big) = \int_c^d \int_a^b f_{X,Y}(x,y)dxdy$$

**Marginals:**

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx$$



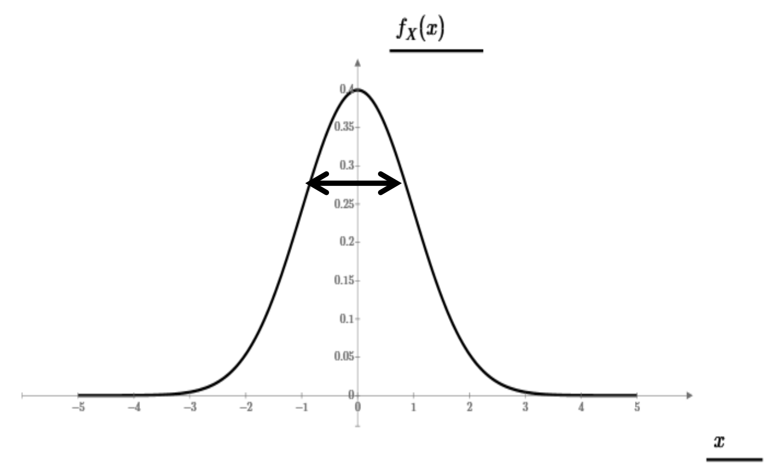

**Cumulative Distribution Function cdf:**

*cdf* $\quad F_{X,Y}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(x,y)dxdy = Pr(X \leq x \wedge Y \leq y)$

*pdf* $\quad f_{X,Y}(x,y) = \dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$

# Expectations

- Mean value: $E[X] = \bar{X} = \mu_X = \int_{-\infty}^{\infty} x \cdot f_X\,(x)dx \quad (\sum_{i=1}^{n} x_i\, f_X(x_i))$
- Variance: $Var(\mathrm{X}) = \sigma_X{}^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 \cdot f_X\,(x)dx = E\left[X^2\right] - E[X]^2$
- Standard deviation: $\sigma_X = \sqrt{Var(\mathrm{X})}$




- A function: $E[g(X)] = \overline{g(X)} = \int_{-\infty}^{\infty} g(x) \cdot f_X\,(x)dx \quad (\sum_{i=1}^{n} g(x_i)\, f_X(x_i))$

$$Var\big(g(\mathrm{X})\big) = \int_{-\infty}^{\infty} \Big(g(x) - \overline{g(x)}\Big)^2 \cdot f_X\,(x)dx = E\left[g(X)^2\right] - E[g(X)]^2$$

- Linear function: $E[aX + b\,] = a \cdot E[X] + b$

$$Var[aX + b\,] = a^2\big(E[X^2] - E[X]^2\big) = a^2 \cdot Var(X)$$

# Correlation, Covariance and summation

Two random variables: $X$ and $Y$

- Correlation: $corr(X, Y) = E[XY]$
- Covariance: $cov(X, Y) = E[XY] - E[X]E[Y]$
- Correlation coefficient: $\rho = \frac{E[XY]-E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$ $\quad -1 \leq \rho \leq 1$

# Sampling From Any Distribution

For test or simulation you need testdata (”measurements”) randomly sampled from a given distribution:

- Find the cdf of the distribution: $F_X(x)$
- Find the inverse of the cdf: $y = F_X(x) \;\Rightarrow\; x = F_X^{-1}(y)$
- Draw a ramdom sample: $y \sim \mathcal{U}[0;1]$
- Insert into the inverse cdf: $x = F_X^{-1}(y)$
- The samples $X = x$ is distributed according to: $F_X(x)$

# Transformation of Variable X to Y

- Given:
  - Pdf: $f_X(x)$
  - Function/Transformation: $\mathrm{Y} = g(X)$
  - Limits: $\mathrm{a} \leq X \leq b$

- Find new pdf: $f_Y(y)$:
  1. Inverse: $x = g^{-1}(y)$
  2. Differentiate: $\frac{dg^{-1}(y)}{dy} = \frac{dx(y)}{dy} = \frac{1}{\frac{dg(x)}{dx}}$
  3. Limits: Find $g(a) = a_Y \leq Y \leq b_Y = g(b)$ based on $\mathrm{a} \leq X \leq b$
  4. New pdf: $f_Y(y) = \sum \left|\frac{dx(y)}{dy}\right| f_X\left( g^{-1}(y) \right) = \sum \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$

# Distribution of the Sum of Two Random Variables

- Two random variables $X$ and $Y$ have density functions $f_X(x)$ and $f_Y(y)$.
- If we define a new random variable $Z = X + Y$, and $Z$ have density function $f_Z(z)$.

Convolution of Two functions

- Then $f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$
- Expectation: $E[Z] = E[X] + E[Y]$
- Variance: $var(Z) = var(X) + var(Y) + 2cov(X, Y)$

Very important!

# i.i.d.: Independent and Identically distributed

- We define that for series of random variables that is taken from the same distribution (identically distributed), and are sampled independent of each other, that they are i.i.d.

$$\boxed{\text{i.i.d. = Independent and Identically distributed}}$$

- i.i.d. is a very important characteristic in stochastic variable processing and statistics

# Central Limit Theorem

- Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables with mean $\mu$ and variance $\sigma^2$
- Let $\bar{X}$ be the random variable (average):

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

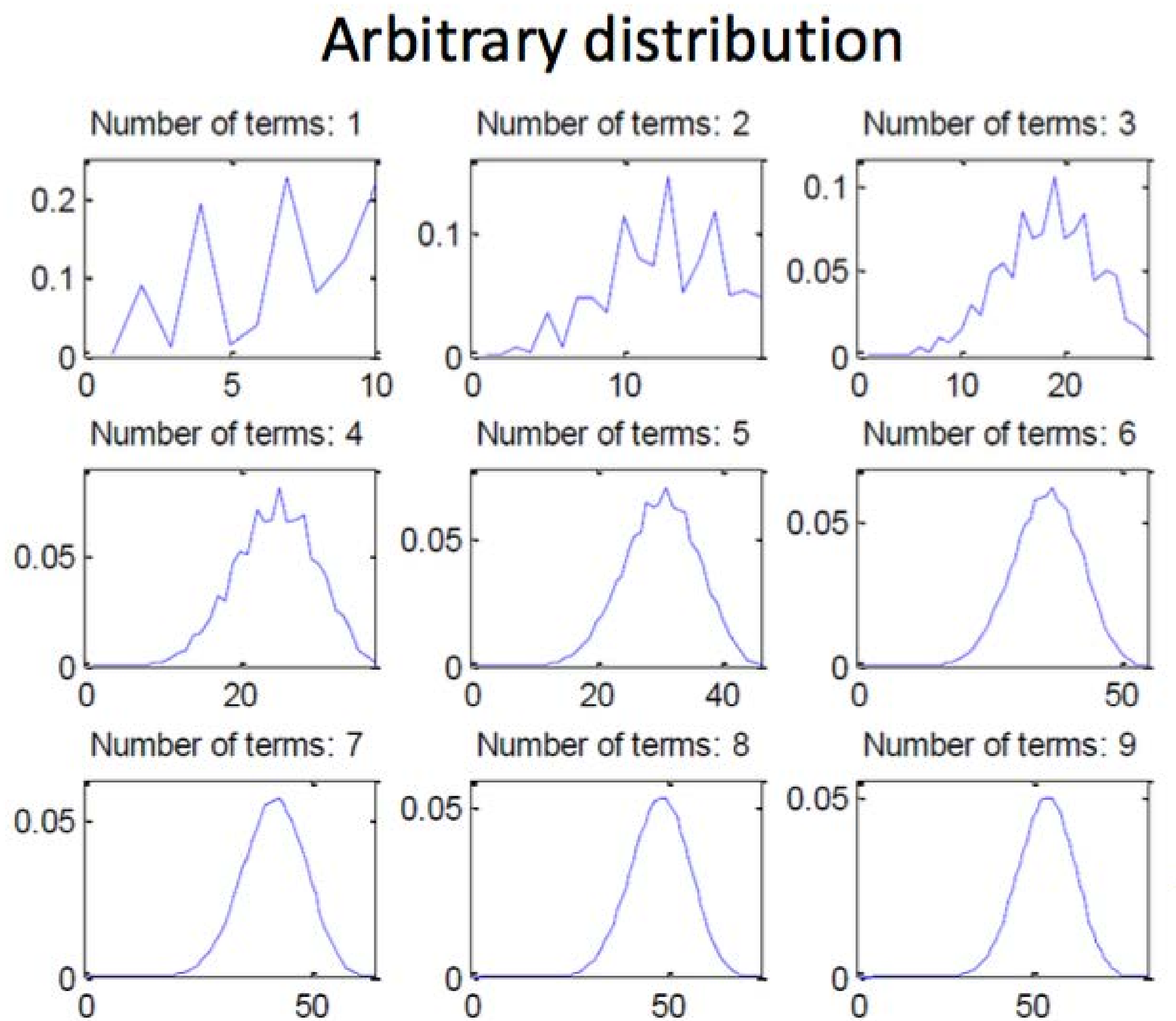
- Then in the limit: $n \to \infty$ we have that: $\underline{\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})}$

  i.e. in the limit $\bar{X}$ will be normally distributed with mean = $\mu$ and variance = $\frac{\sigma^2}{n}$.

The variance is reduced with a factor 1/n

# Sum of Random Variables

- The random variables are i.i.d and taken from the same distribution.

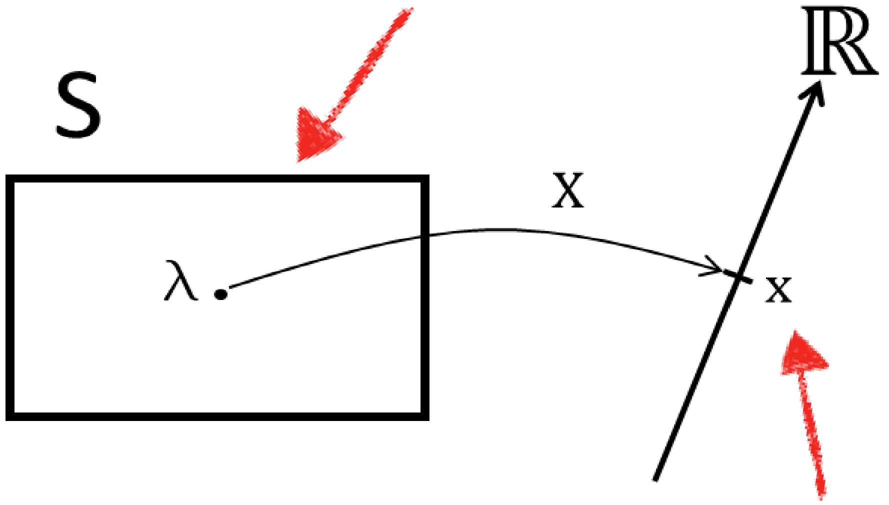


# Stochastic Processes

## Stochastic Variables

- Sample space for stochastic experiment

Sample space for
stochastic experiment



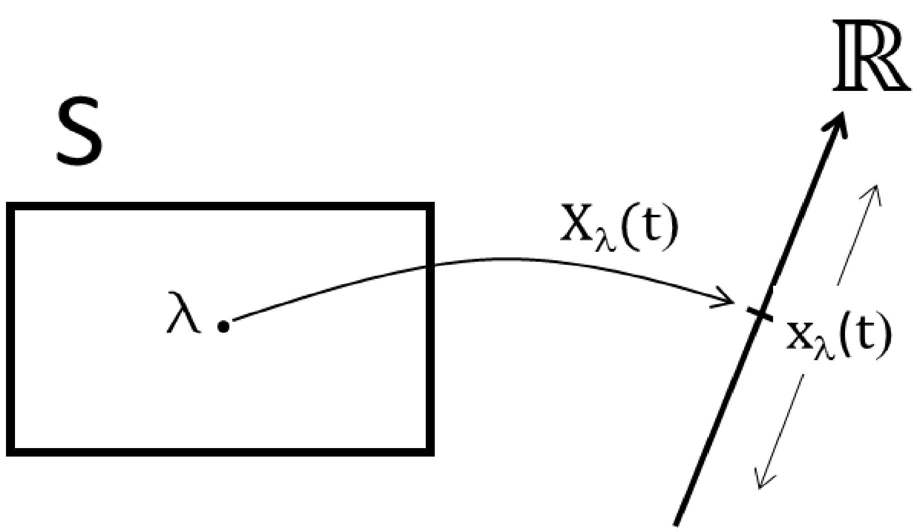

Sample space for
stochastic variable

Time dependent

## Stochastic Processes (signals)

- Sample space for stochastic experiment
- Random events that develops in time

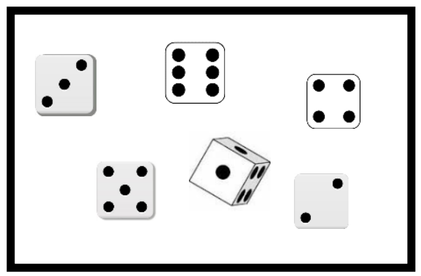
Sample space for
stochastic experiment



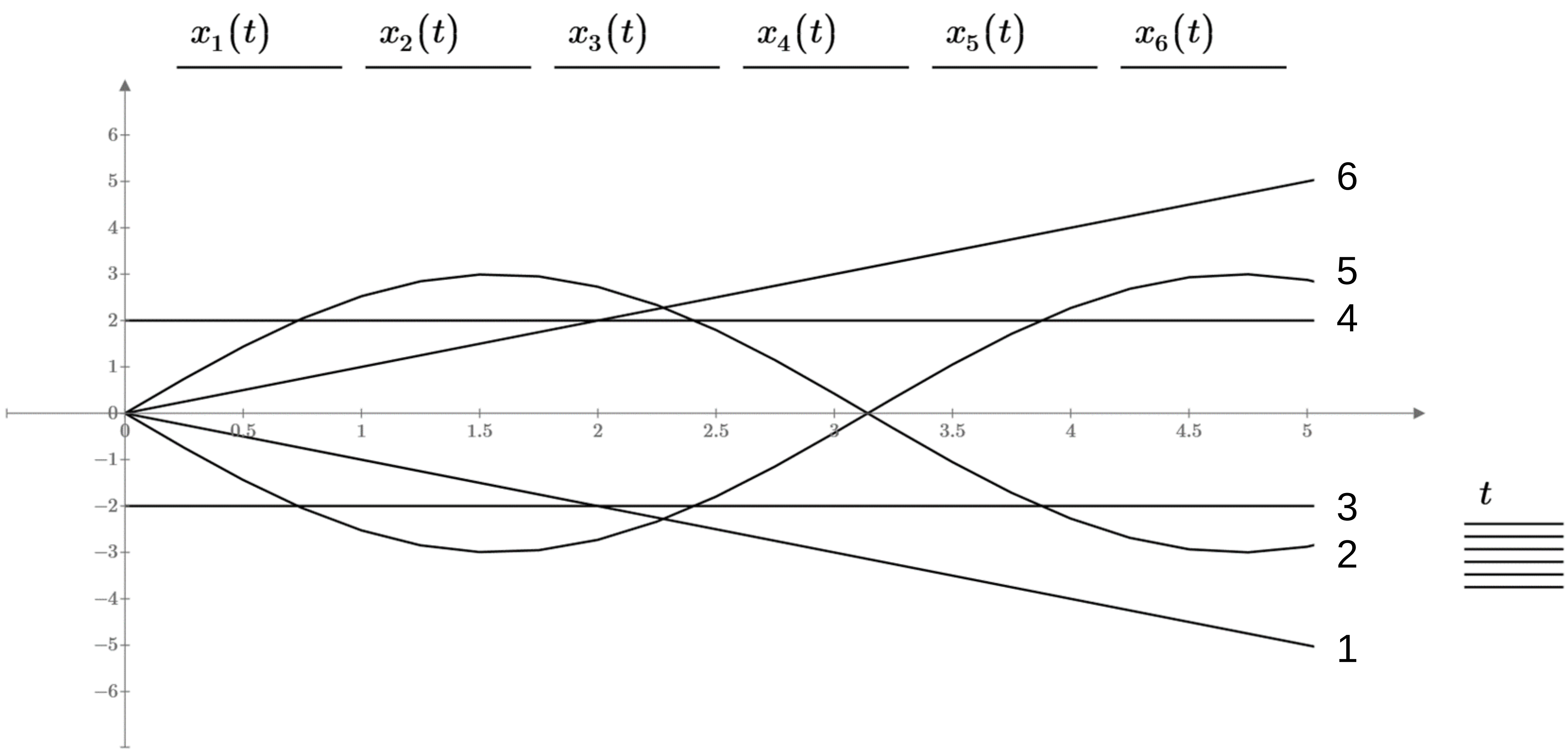

Sample space for
stochastic proces

# Stochastic Processes – Example

S



$$X_n(t): \quad x_1(t) = -t \qquad x_2(t) = 3\sin(t)$$
$$x_3(t) = -2 \qquad x_4(t) = 2$$
$$x_5(t) = -3\sin(t) \qquad x_6(t) = t$$

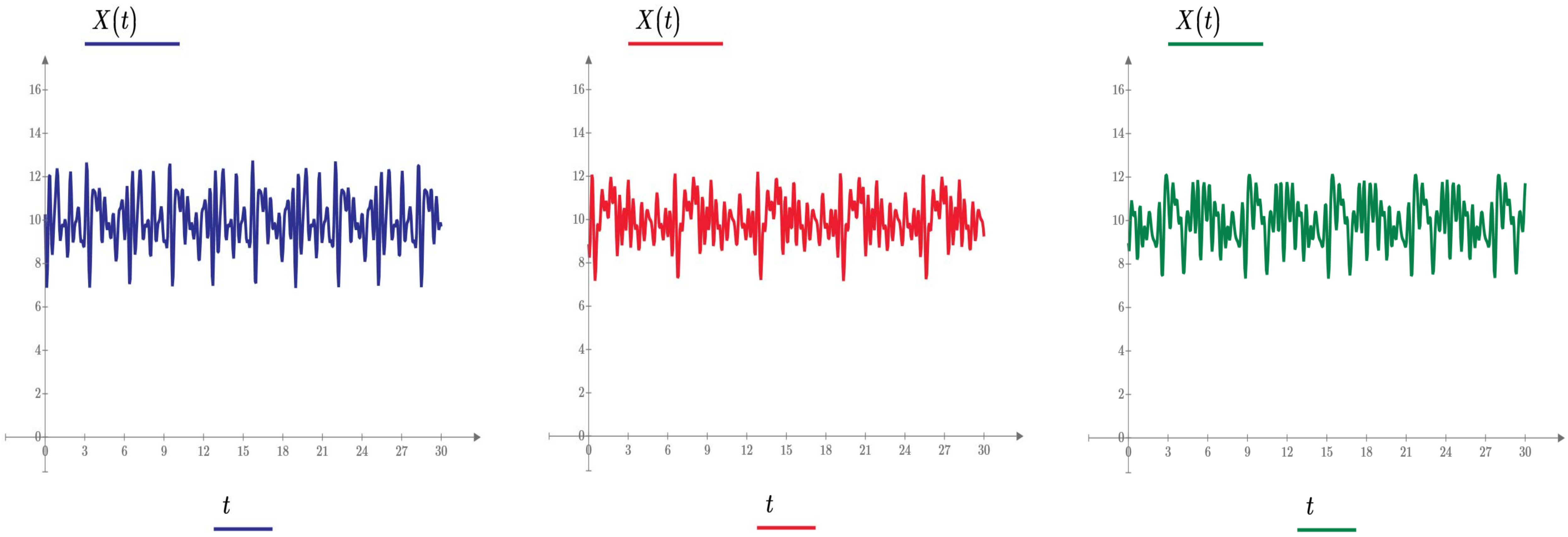


# Stochastic Processes – Signals

## Additive Noisemodel

$$observed\ signal = signal\ +\ noise$$

Three Realizations of the Stochastic Process



# Stochastic Processes

## Definitions:

- A stochastic process is a time dependent stochastic variable:

$$X(t)$$

- A discrete stochastic process is given by:

time

$$X[n] = X(nT)$$

where $n$ is an integer.

## Notice:

- When we sample a signal from a stochastic process, we observe only one realization of the process

# Sample Functions

**Definition:**

- A sample function $x(t)$ is a realization of a stochastic process $X$

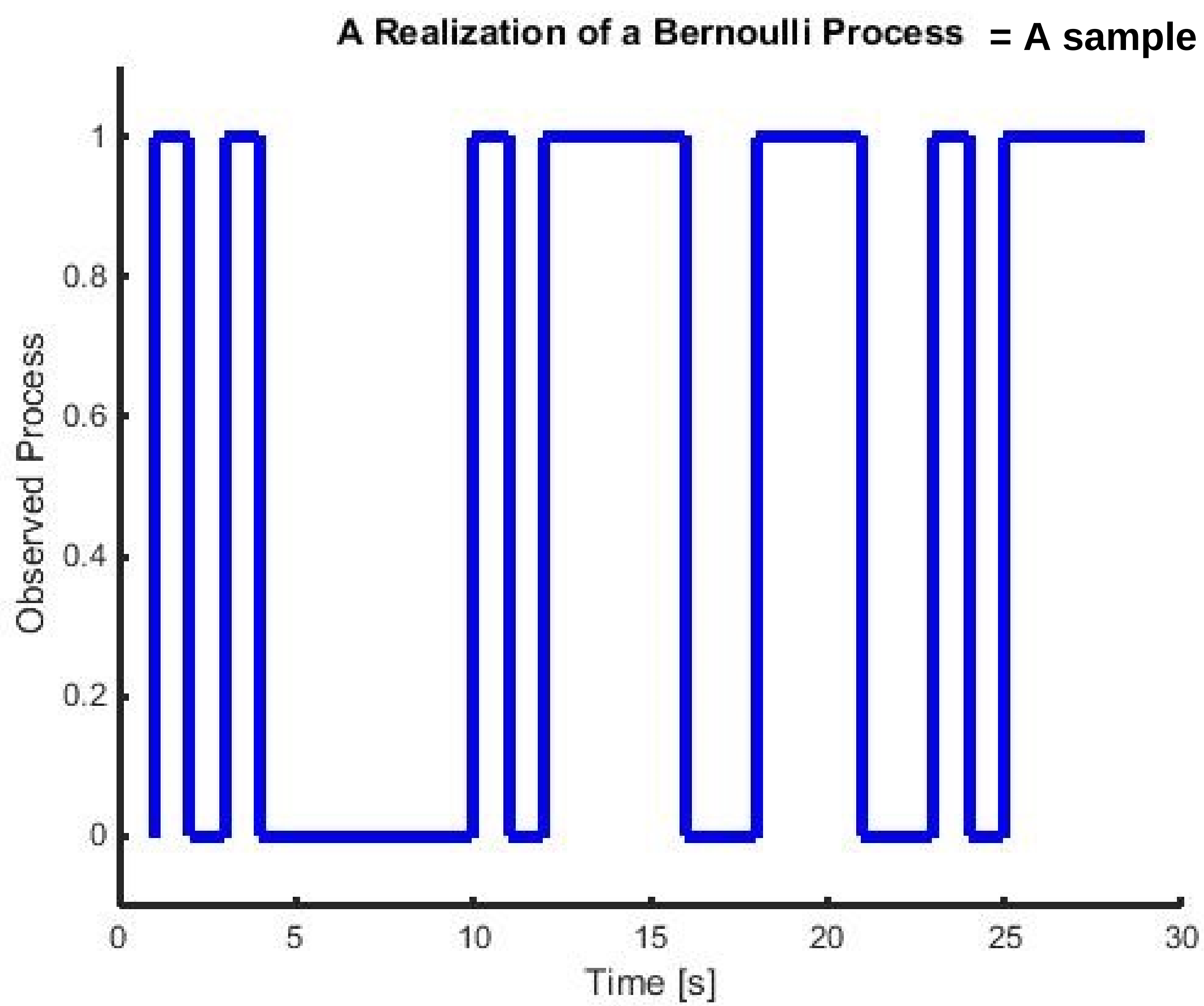
**Example:**

- A coin is thrown every minute: $H$ = head, $T$ = tail
- One realization of the stochastic signal is:

$$HTHT$$

# Example – Random Binary (digital) Signal

- Bernoulli process.
- A sequence of $1$ and $0$s.
- Is a sequence of i.i.d of Bernoulli trials.




= A sample function x(t)

# Ensemble

**Definition:**

- The Ensemble of the Stochastic Process is the collection of all possible realizations $x(t)$ of the Stochastic Process $X$

**Example:**

- A coin is thrown every minute: $H$ = head, $T$ = tail
- The Ensemble of the stochastic signals is:

$$HTHT, HHTT, TTHH, THTH, THHT, TTHT, HHHH\ldots$$

# Time Dependent Probability Functions

- Probability density function (pdf):

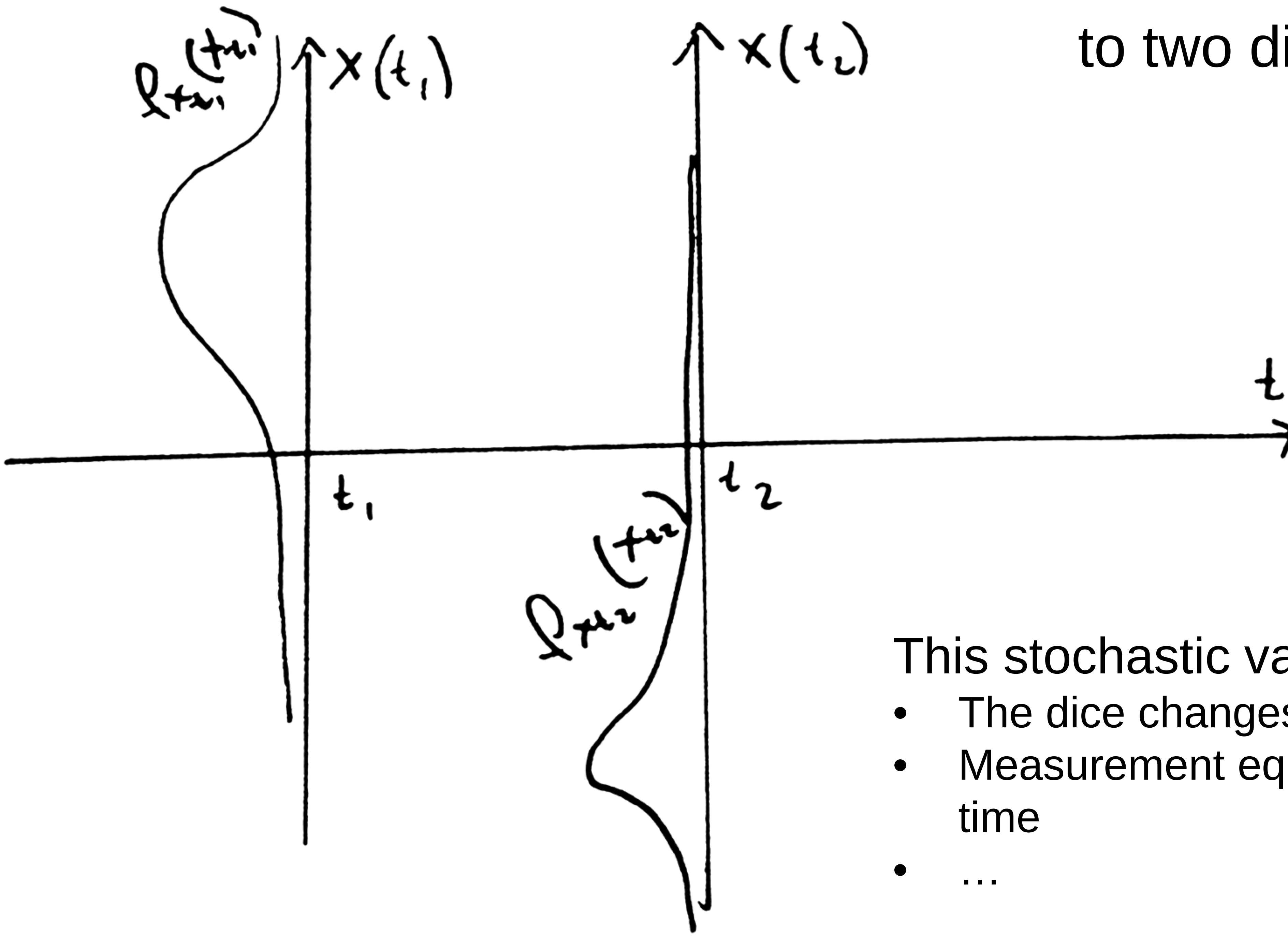
$$f_{X(t)}(x(t))$$

- Cumulative distribution function (cdf):

$$F_{X(t)}\big(x(t)\big) = \int_{-\infty}^{x(t)} f_{X(t)}(x(t))\ dx(t)$$

# Time Dependent Stochastic Process

The same stochastic variable
to two different times

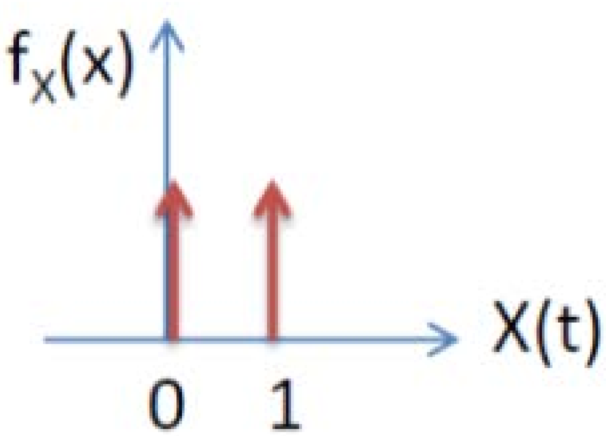



This stochastic variable is not i.i.d.:

- The dice changes it's properties (wears)
- Measurement equipment changes with time
- …

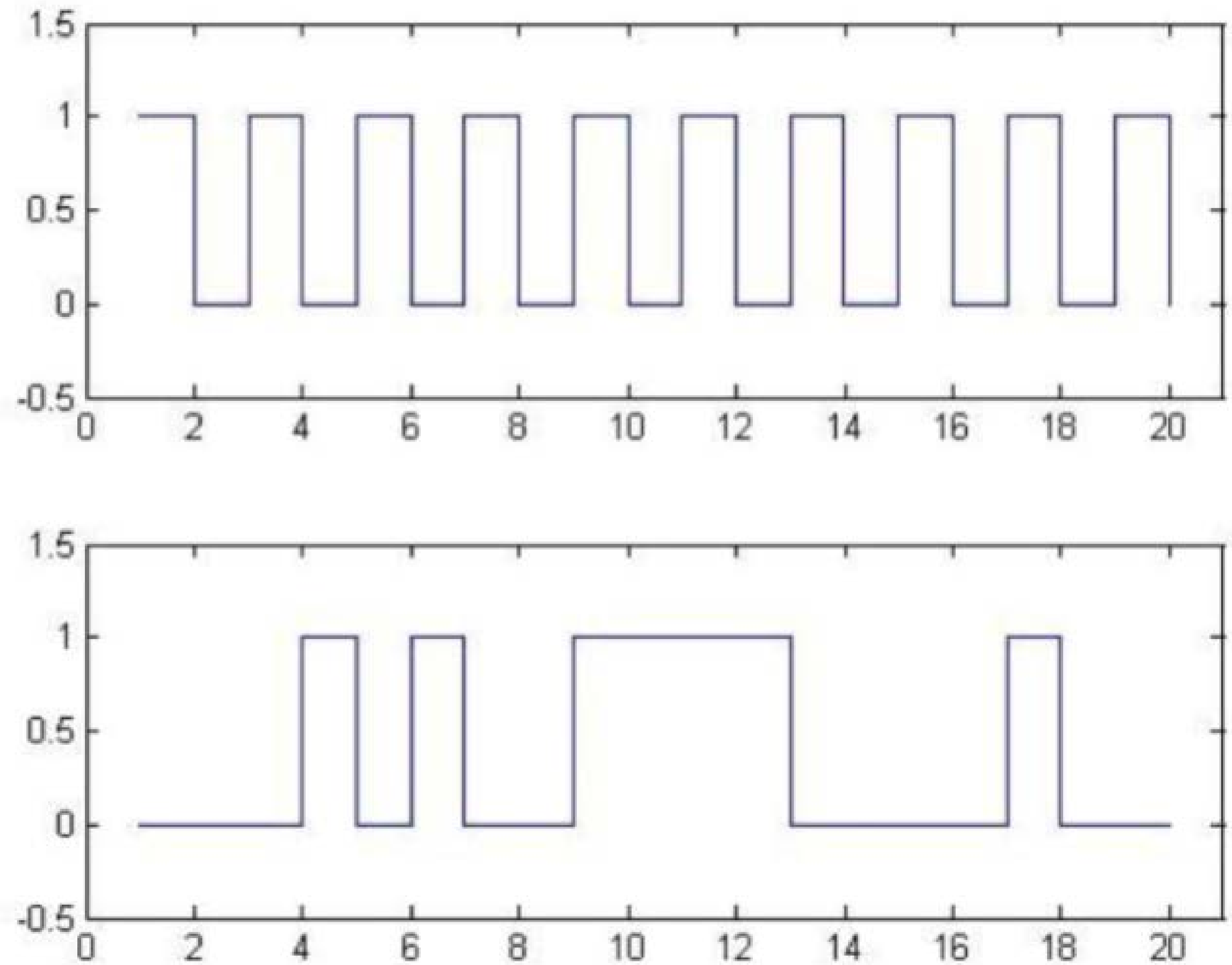
# Deterministic Functions

- We find a sample function from a stochastic process.
- The two samples have the same pmf.




pmf

Deterministic



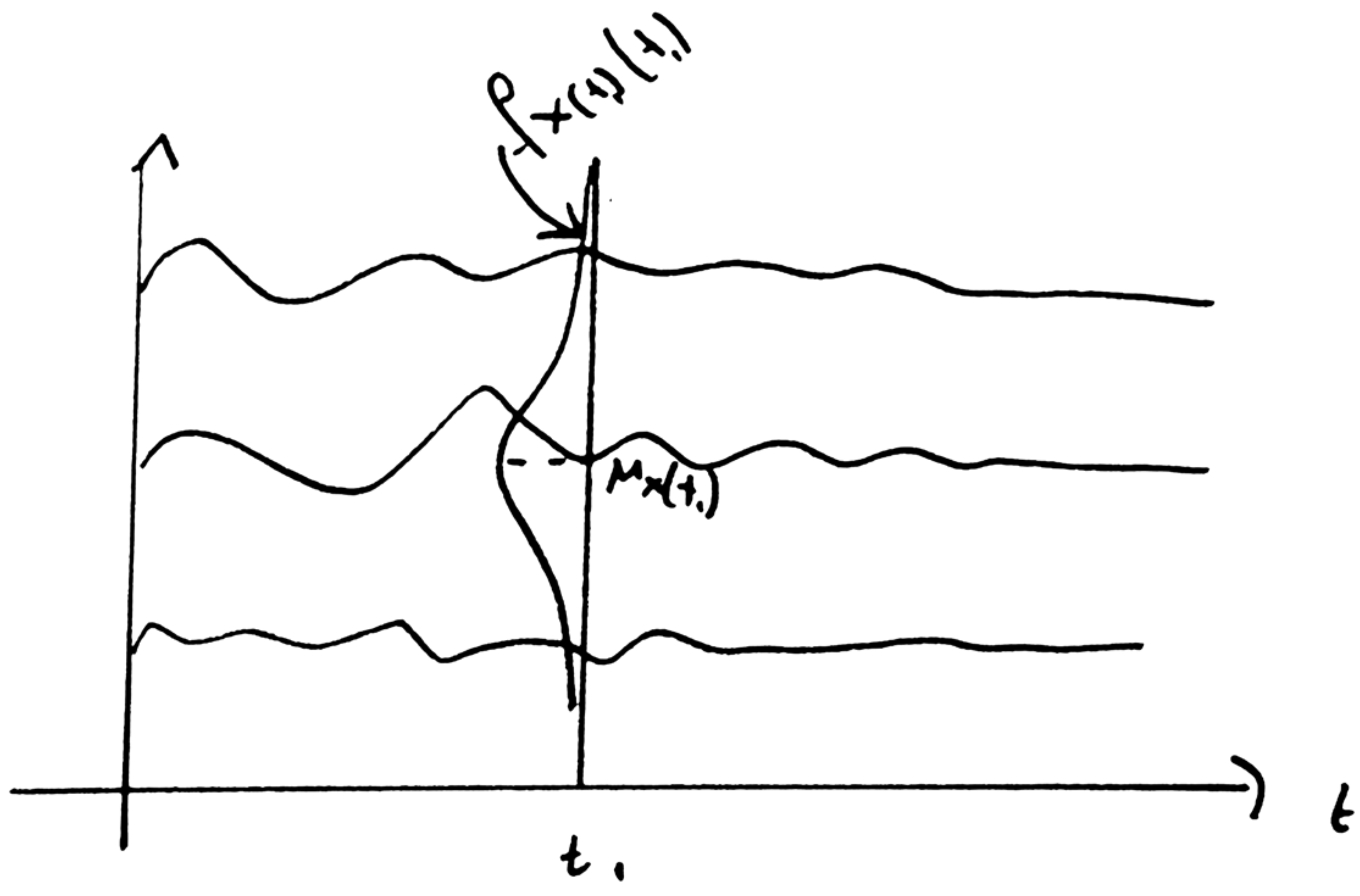

Random (non-deterministic)

# Ensemble mean

- The mean value function:

$$\mu_{X(t)}(t) = E[X(t)] = \int_{-\infty}^{\infty} x(t) f_{X(t)}(x(t))\, dx(t)$$

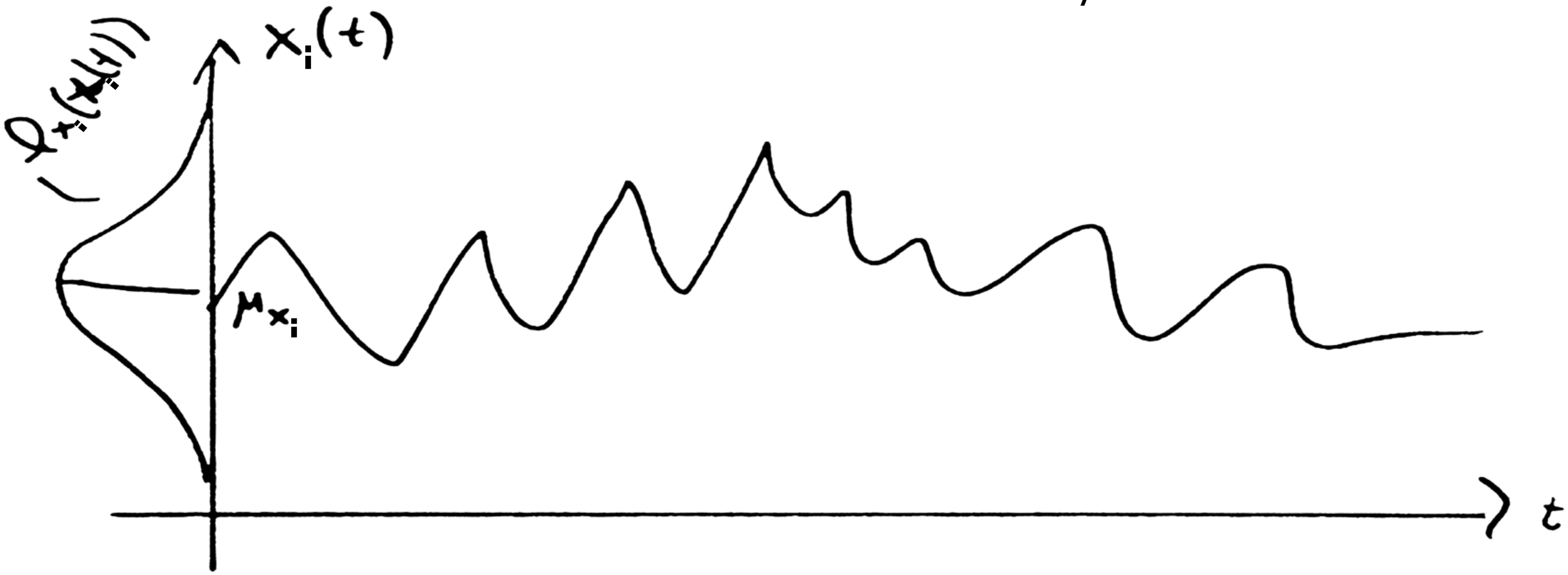
- The mean of <u>all possible realizations</u> to time t



# Temporal Mean

- The time average for <u>one realization</u> of the stochastic process
- The temporal mean can differ from the ensemble mean

$$\hat{\mu}_{X_i} = \langle X_i \rangle_T = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)\, dt$$

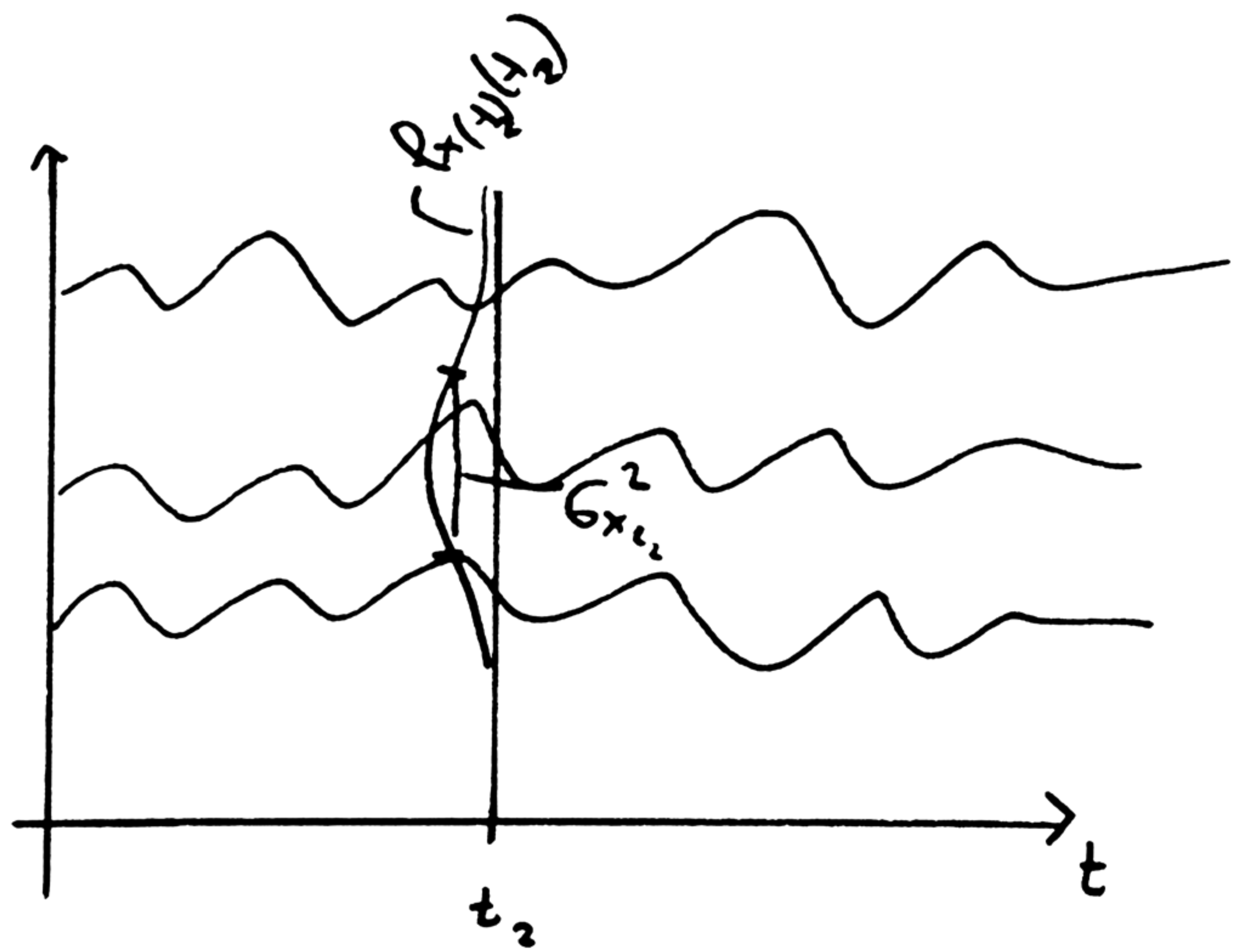


# Ensemble Variance

- The variance function:

$$var\big(X(t)\big) = \sigma^2_{X(t)}(t) = E[\Big(X(t) - \mu_{X(t)}(t)\Big)^2]$$

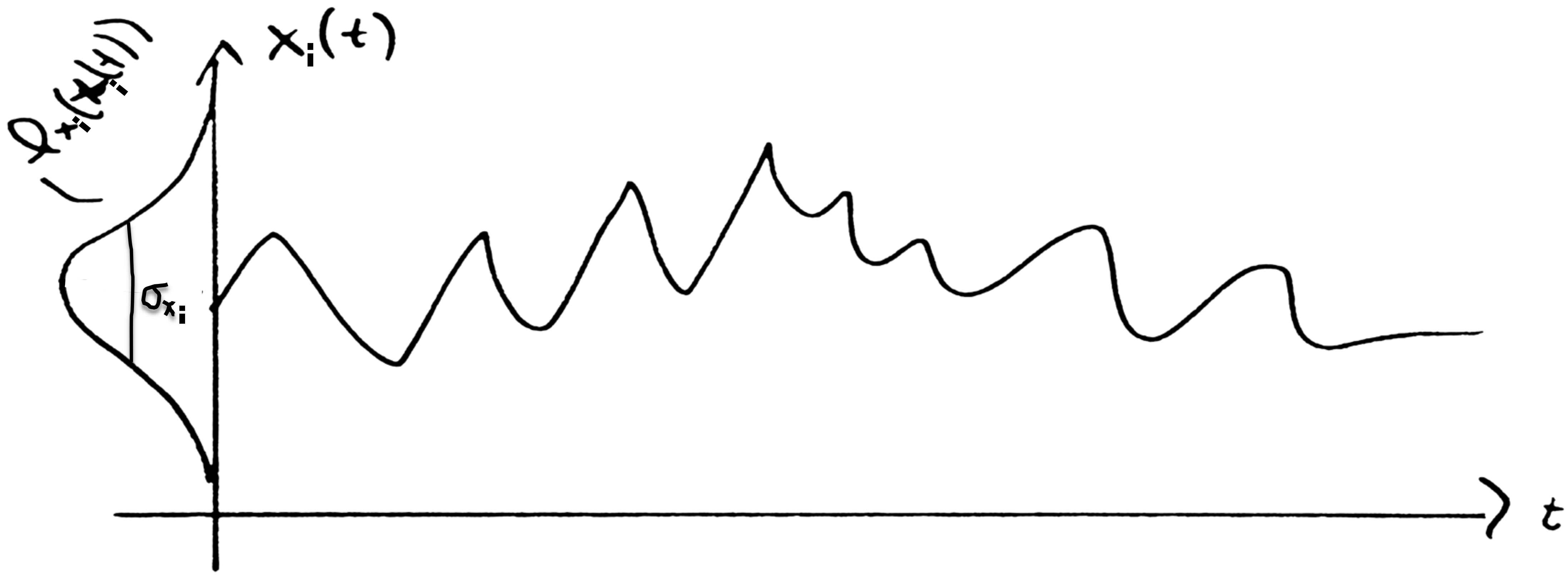
- The variance of <u>all possible realizations</u> to time t



# Temporal Variance

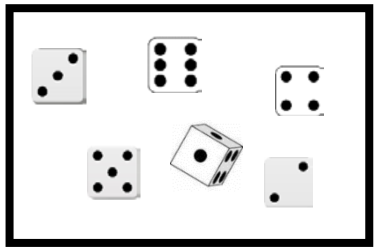
- The variance over time for one realization of the stochastic process
- The temporal variance can differ from the ensemble variance

$$\hat{\sigma}_{X_i}^2 = \left\langle X_i^2 \right\rangle_T - \langle X_i \rangle_T{}^2 = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left( x_i(t)^2 - \hat{\mu}_{X_i}{}^2 \right) dt = Var(X_i)$$

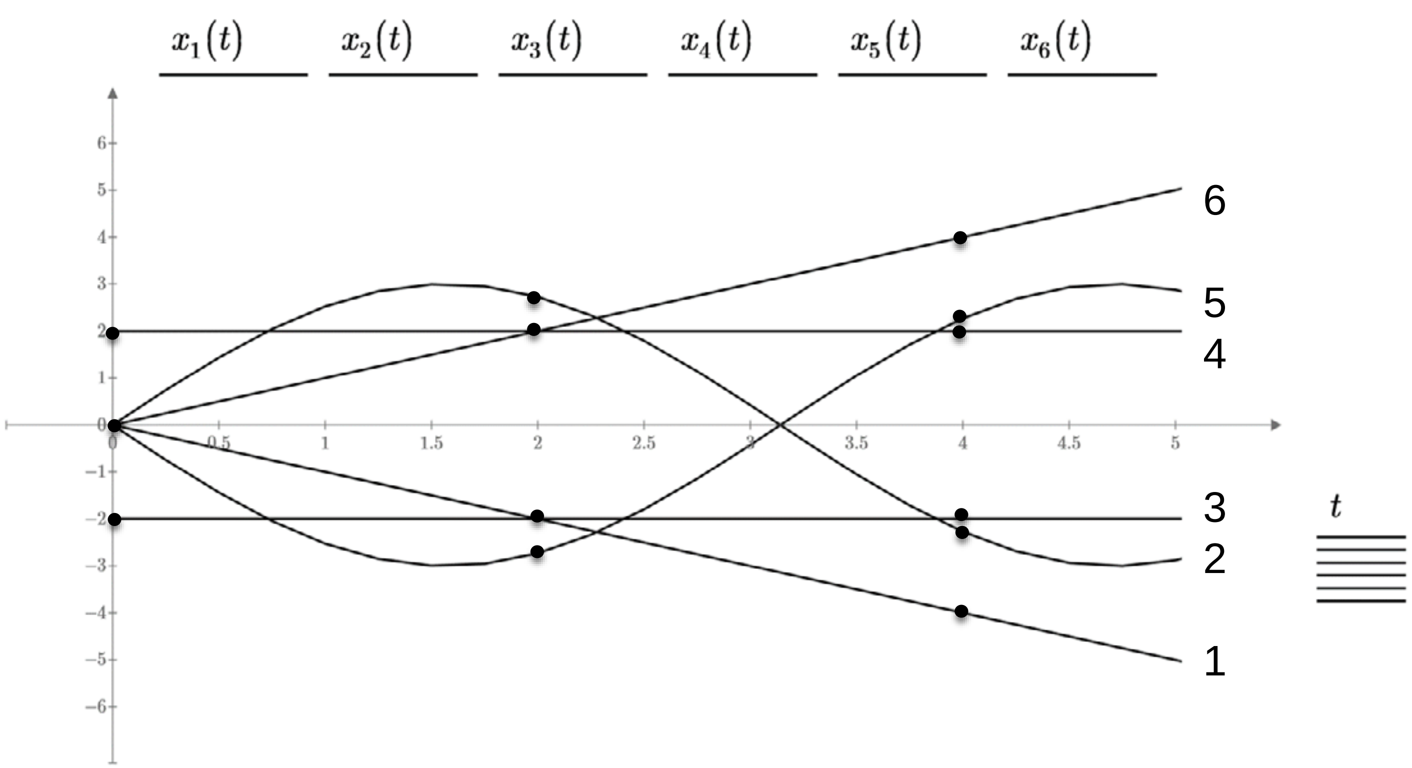


# Stochastic Processes – Example

S



$$X_n(t): \quad x_1(t) = -t \qquad x_2(t) = 3\sin(t)$$
$$x_3(t) = -2 \qquad x_4(t) = 2$$
$$x_5(t) = -3\sin(t) \qquad x_6(t) = t$$




$$X(0) = \{-2, 0, 2\}$$
$$X(2) = \{-2.7, -2, 2, 2.7\}$$
$$X(4) = \{-4, -2.3, -2, 2, 2.3, 4\}$$

$$\Pr(X(0) = 0) = 2/3$$
$$\Pr(X(2) = 2) = 1/3$$
$$\Pr(X(4) = -4) = 1/6$$

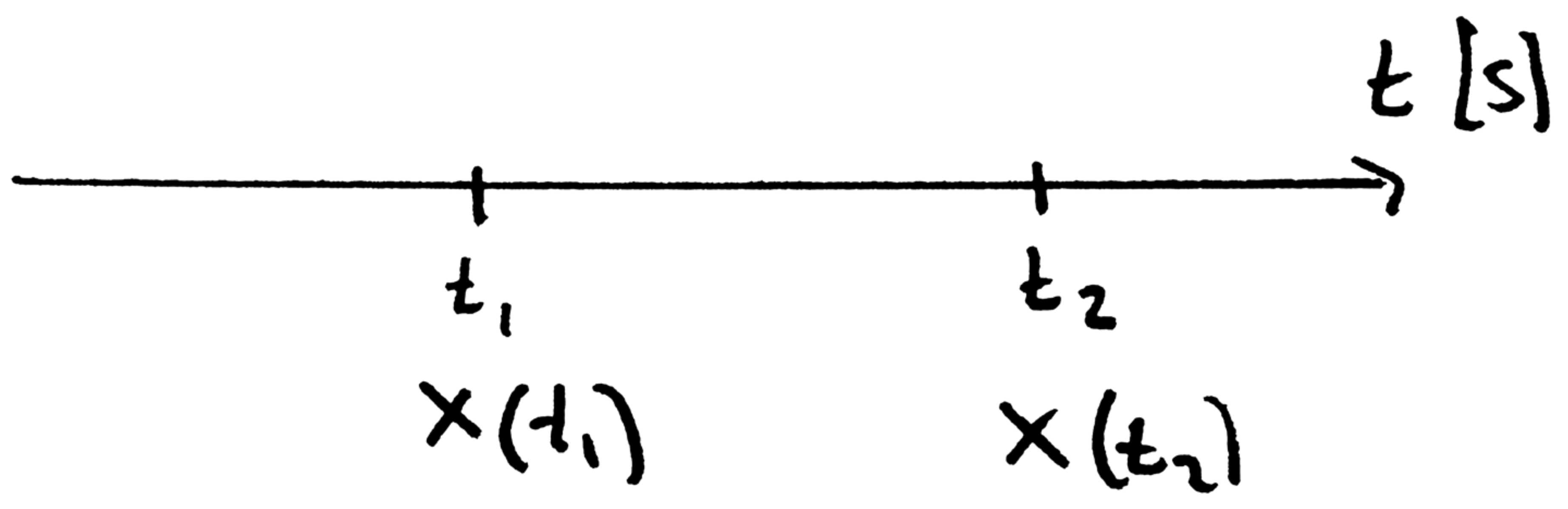
Ensemble: $\mu_{X(t)}(t) = E[X(t)] = 0$ $\quad var\big(X(t)\big) = \sigma^2_{X(t)}(t) = \frac{1}{3}(t^2 + 9\sin(t) + 4)$

Temporale: $\hat{\mu}_{X_2} = 0$ $\quad \hat{\mu}_{X_3} = -2$ $\quad \hat{\sigma}^2_{X_2} = 4.5$ $\quad \hat{\sigma}^2_{X_3} = 0$

# Correlations

Comparing realizations

- Autocorrelation Correlation of a realization with itself
- Cross-correlations Correlation of two different realizations
- We compare the processes at two different times



# Autocorrelations

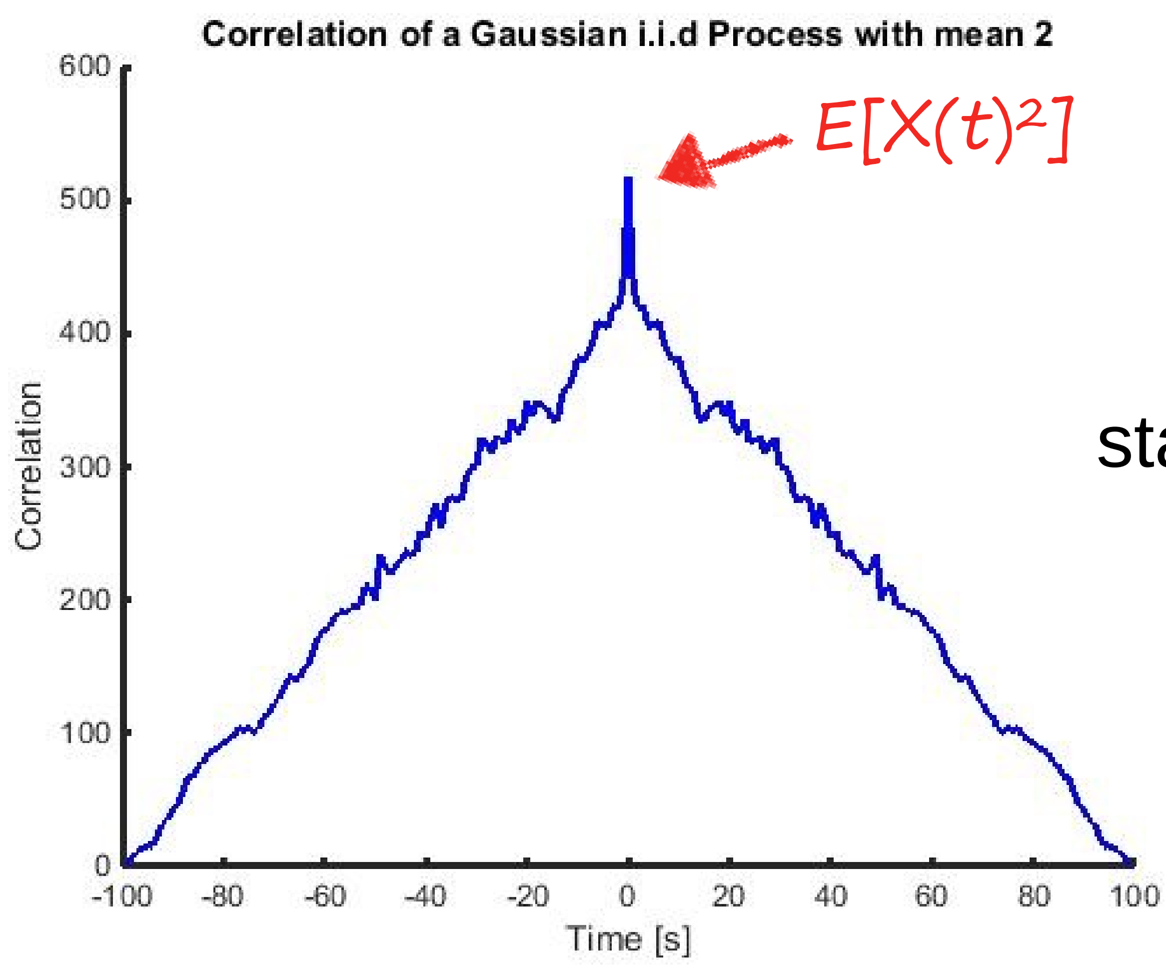
Tells about the connection at two different times

- Autocorrelation function:

Complex conjugated

$$R_{XX}(t_1, t_2) = E[X(t_1)X(t_2)^*]$$

$$= \iint_{-\infty}^{\infty} x(t_1)\, x(t_2)^* f_{X(t_1),X(t_2)}\big(x(t_1), x(t_2)\big) dx(t_1) dx(t_2)$$




Autocorrelation of a stationary process at time $t_1$ as a function of $\tau = t_1$-$t_2$

# Autocovariances

Tells about how much we can predict the future

- Autocovariance function:

$$C_{XX}(t_1,t_2) = E\left[\Big(X(t_1) - \mu_X(t_1)\Big)\big(X(t_2) - \mu_X(t_2)\big)^*\right]$$
$$= R_{XX}(t_1,t_2) - \mu_X(t_1)\mu_X(t_2)$$

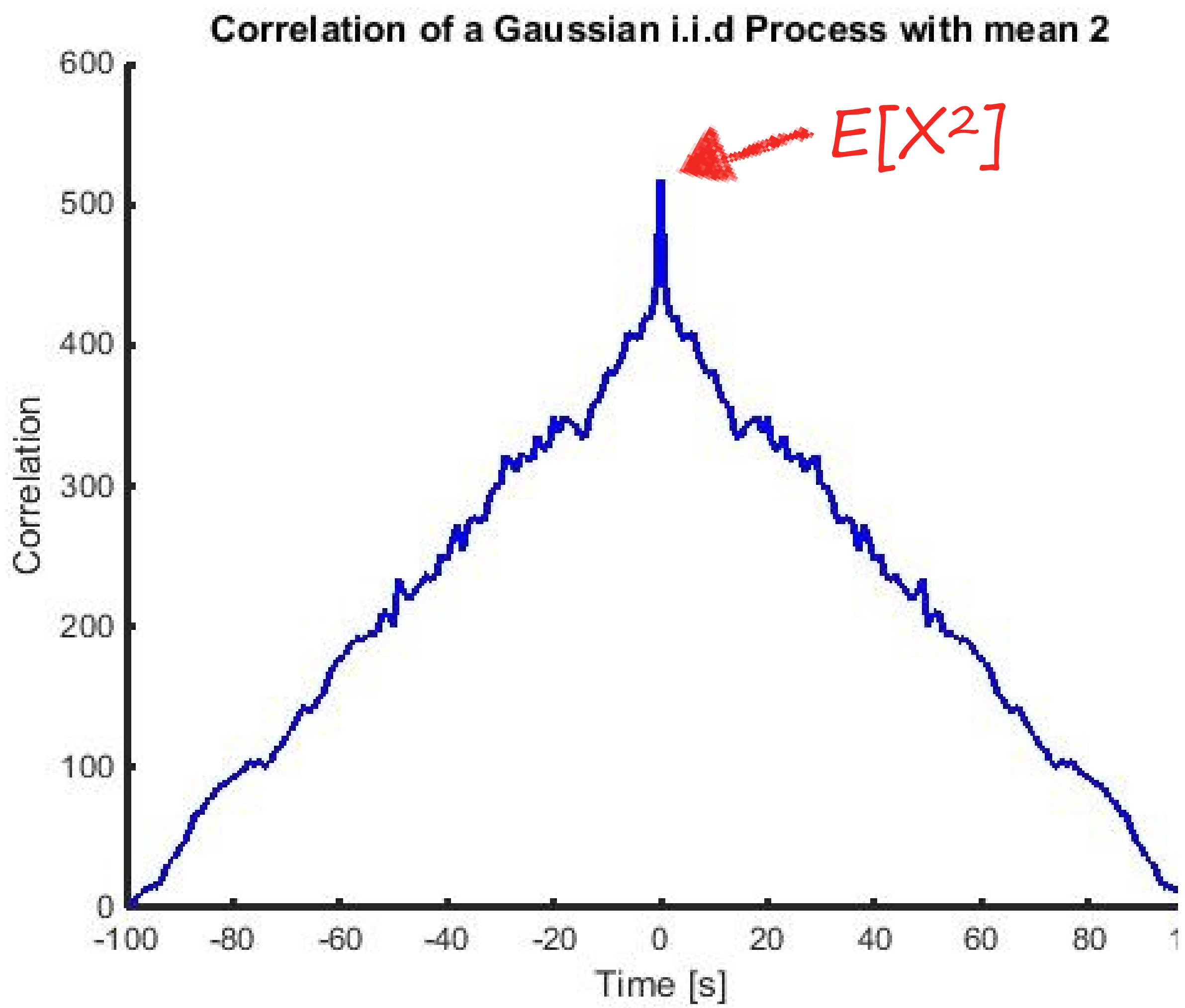
- Autocorrelation coefficient:

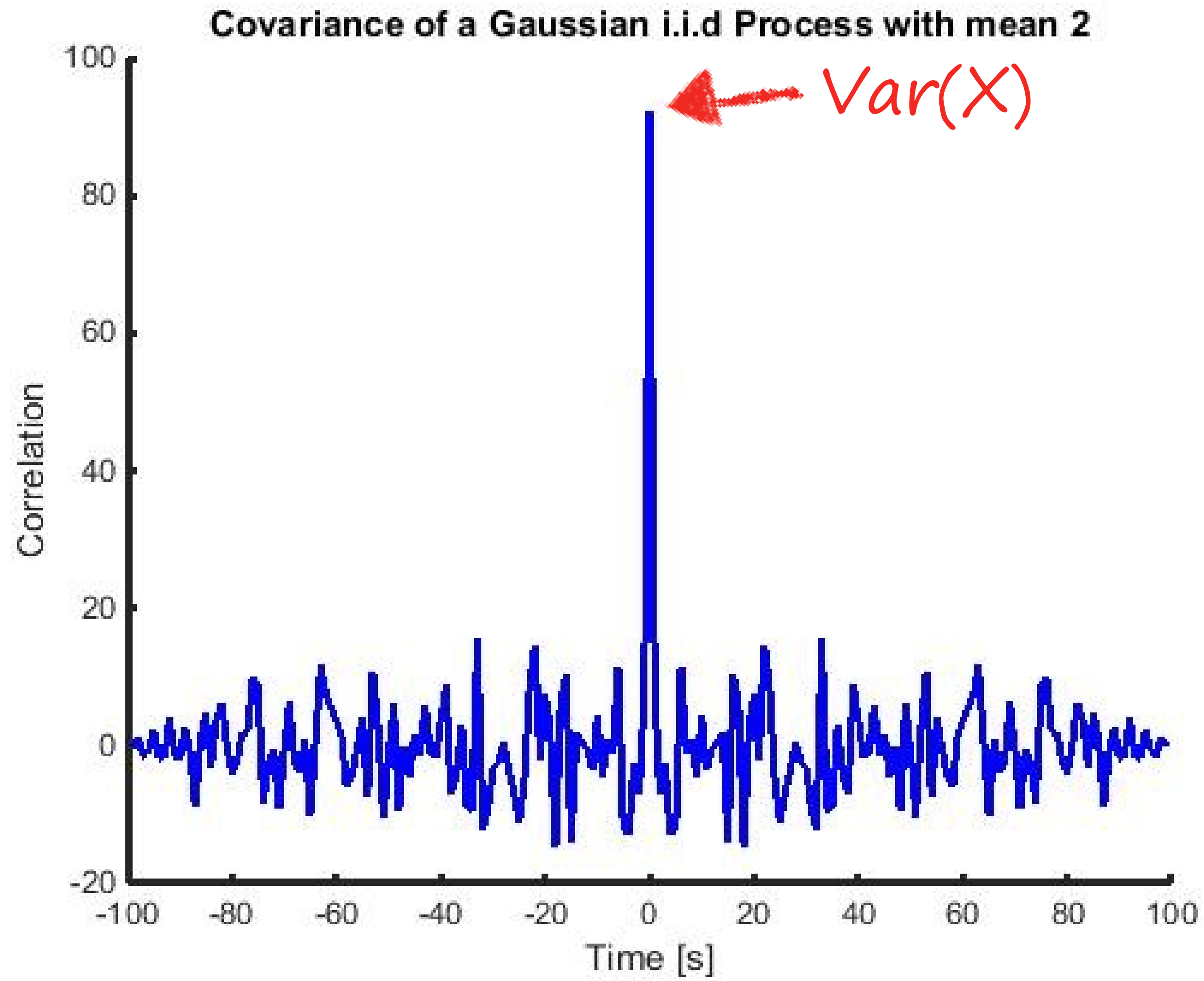
$$r_{XX}(t_1,t_2) = \frac{C_{XX}(t_1,t_2)}{\sqrt{C_{XX}(t_1,t_1)C_{XX}(t_2,t_2)}}; \qquad 0 \le r_{XX}(t_1,t_2) \le 1$$

# Autocovariances

For i.i.d. Gaussian (stationary) noise

- Autocorrelation and autocovariance

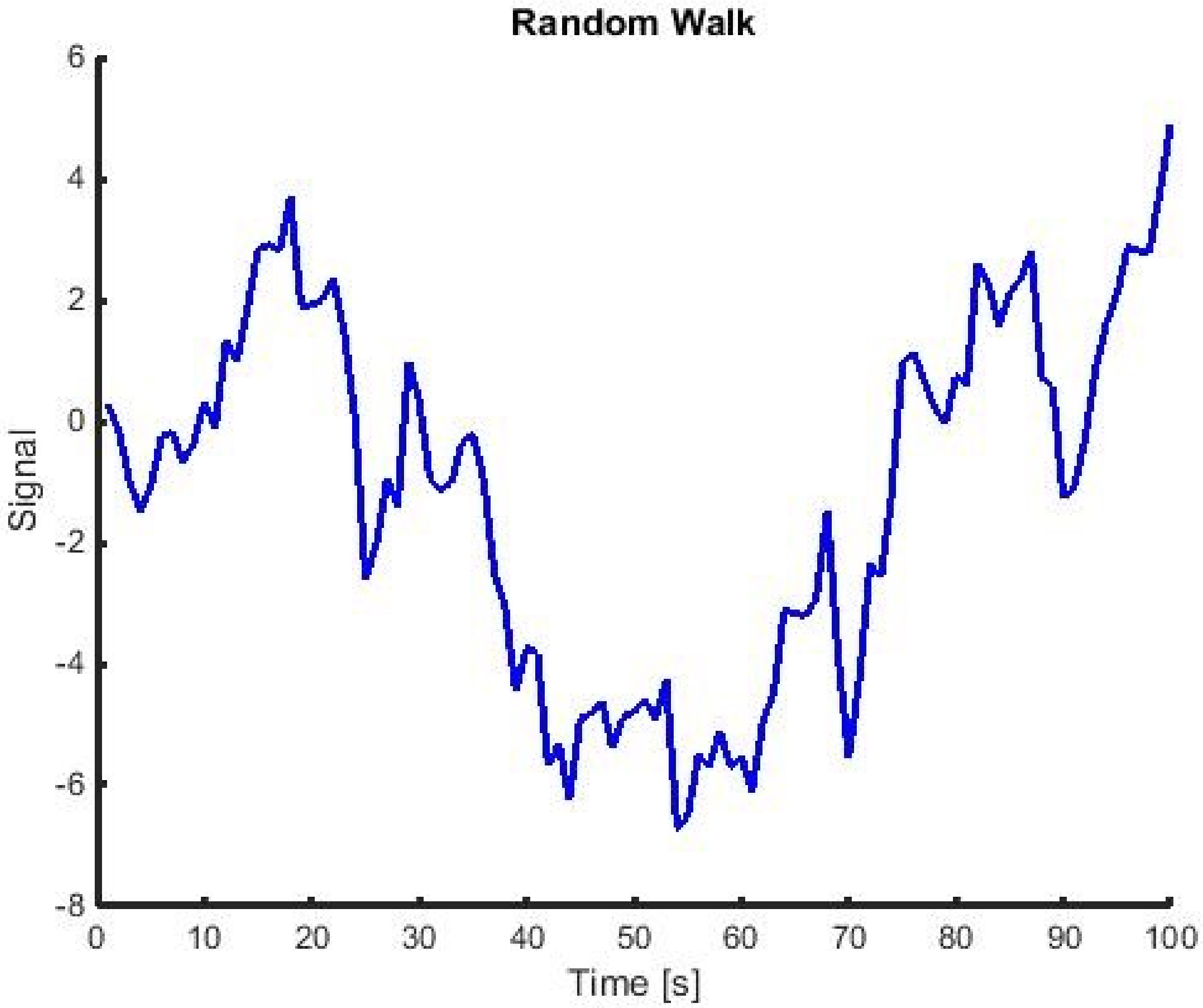


# Random Walk – Example

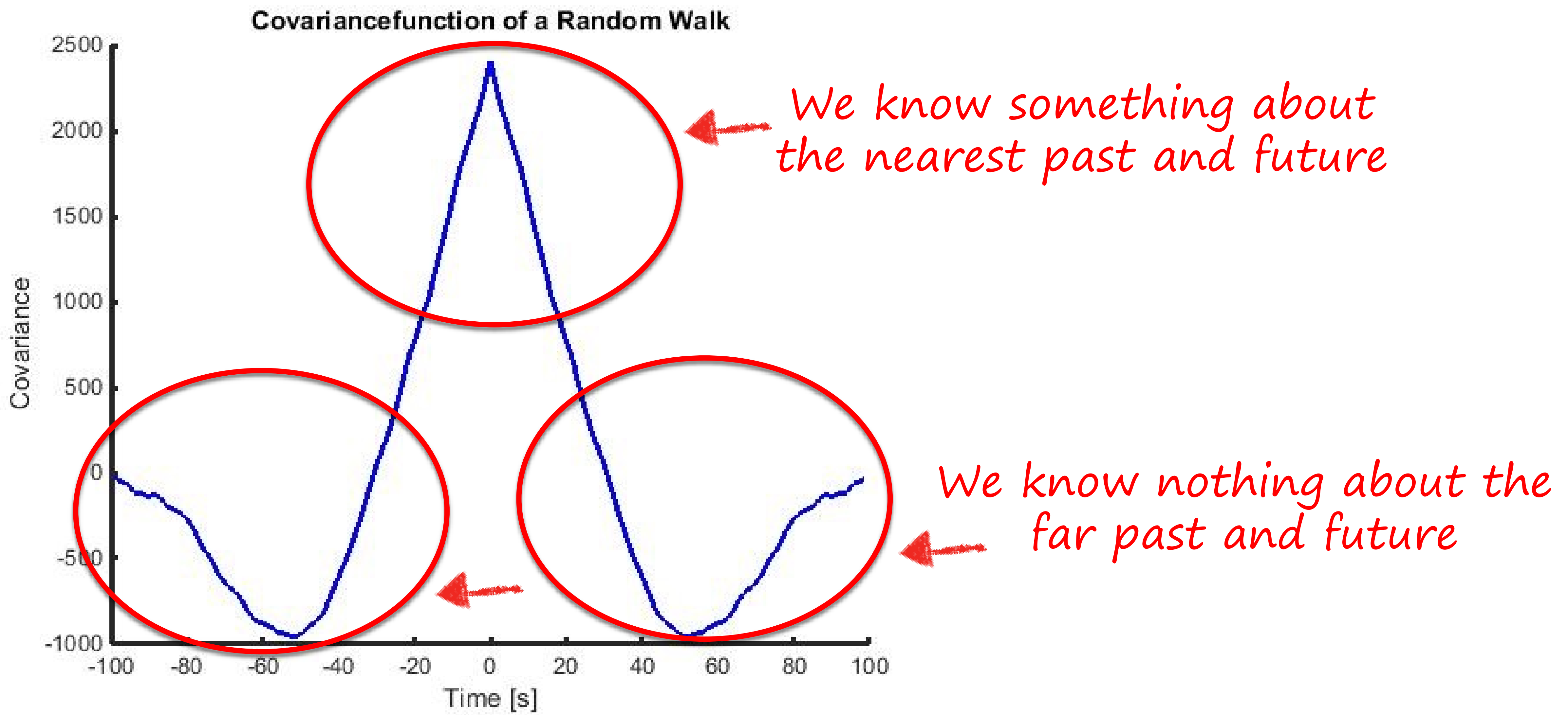
Brownish motions

- We consider a random walk.



# Random Walk – Example

- Sample of the autocovariance function:



# Stationarity in the Strict Sense (SSS)

Difficult to test in reality

- The density function $f_{X(t)}\big(x(t)\big)$ do not change with time
  - For all choices of $t_1$ and $\Delta t_1$, the marginal pdf:

$$f_{X(t_1)}\big(x(t_1)\big) = f_{X(t_1+\Delta t_1)}\big(x(t_1+\Delta t_1)\big)$$

  - For all choices of $t_1$, $t_2$ and $\Delta$t, the simultaneous pdf:

$$f_{X(t_1),X(t_2)}\big(x(t_1), x(t_2)\big) = f_{X(t_1+\Delta t),X(t_2+\Delta t)}\big(x(t_1+\Delta t), x(t_2+\Delta t)\big)$$

# Stationarity in the Wide Sense (WSS)

Can be tested

- Ensemble mean is a constant

$$\mu_X(t) = E[X(t)] = \mu_X$$ - independent of time

- Autocorrelation depends only on the time difference $\tau = t_2 - t_1$

$$R_{XX}(t, t+\tau) = E[X(t)X(t+\tau)] = R_{XX}(\tau)$$ - independent of time

→ Ensemble variance is a constant

$$\sigma_X^2(t) = E\left[X(t)^2\right] - E[X(t)]^2 = \sigma_X^2$$ - independent of time

# Ergodicity

- We can say something about the properties of the stochastic process in general based on one sample function, as long as we have observed it for long enough.

**Example:**

- An i.i.d Gaussian noise stream

# Ergodicity

- If ensemble averaging is equivalent to temporal averaging:

$$\mu_X(t) = \bar{X}(t) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)\, dt = \langle X_i \rangle_T = \hat{\mu}_{X_i}$$

- For any moment: In practice: n=2 (Variance)

$$\overline{X^n} = \int_{-\infty}^{\infty} x^n f_X(x)\, dx = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i^n\,(t)\, dt$$

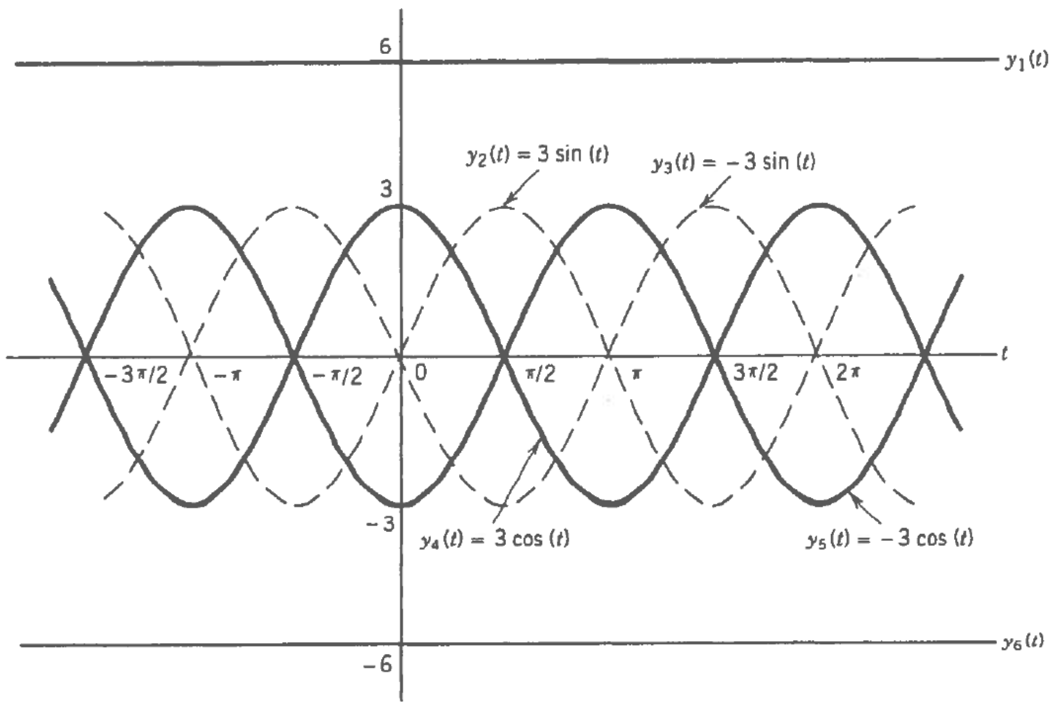
One (any) realization

Ensemple (WSS)

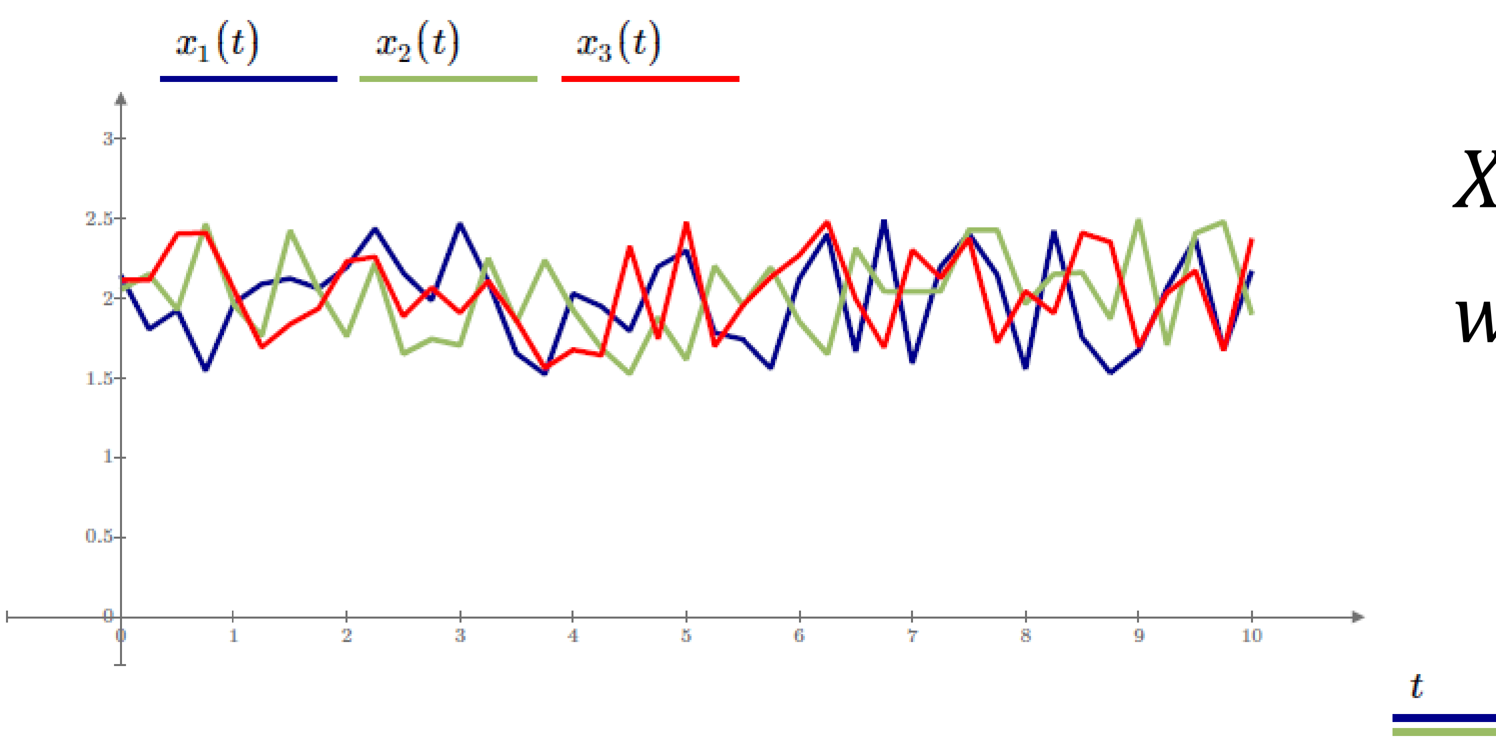
$$\left.\begin{aligned} \langle X_i \rangle_T &= \mu_X \\ \hat{\sigma}_{X_i}^2 &= \sigma_X^2 \end{aligned}\right\} \rightarrow Ergodic$$

All information is achieved with one measurement (realization)

# WSS and Ergodicity – Examples




% SSS
✓ WSS
% Ergodic

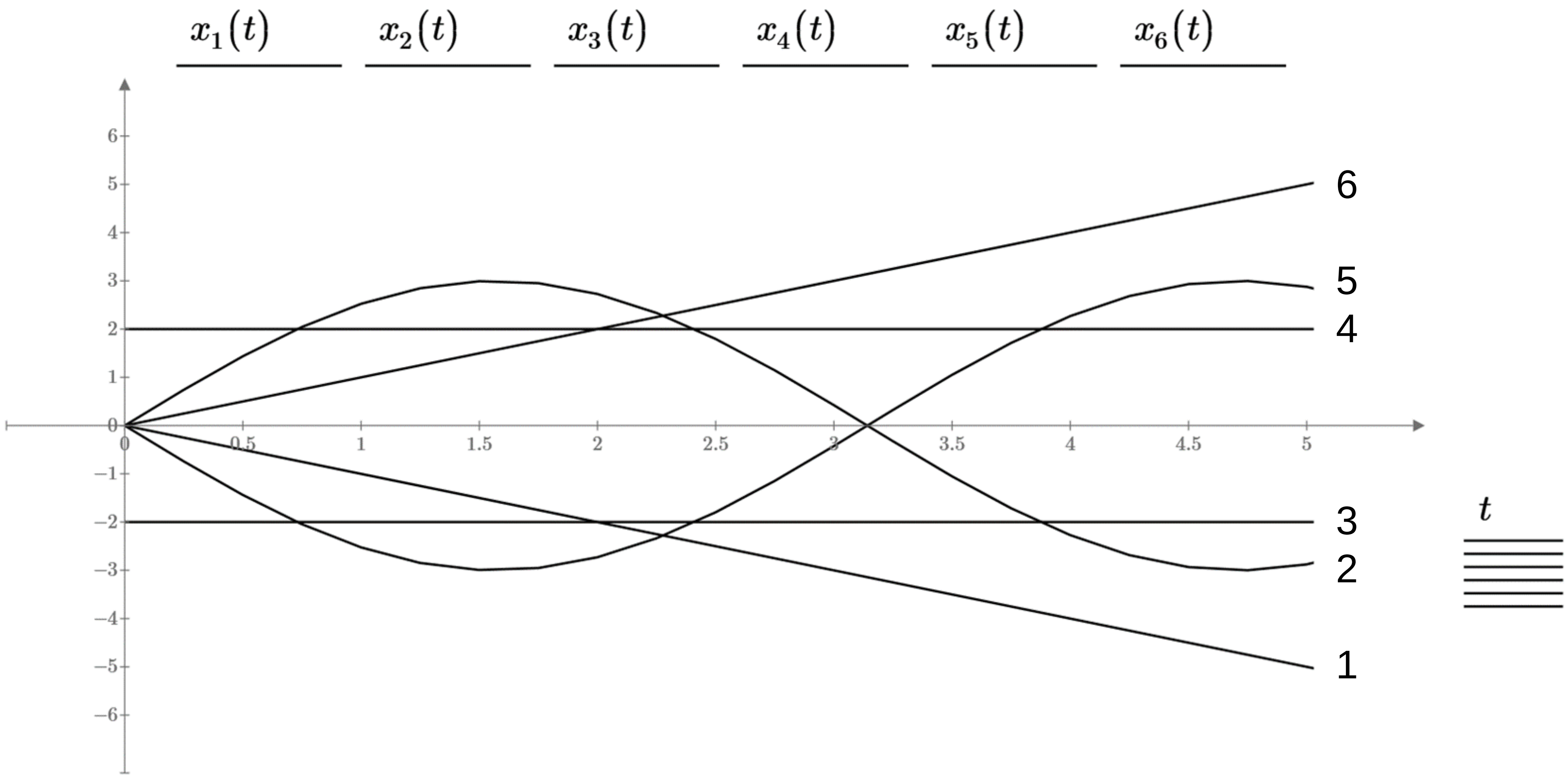



$$X_n(t) = 2 + w_n(t)$$

$$w_n(t) \sim \mathcal{U}[-0{,}5; 0{,}5]$$

✓ WSS
✓ Ergodic

# WSS and Ergodicity– Example




- SSS ???
- WSS ???
- Ergodic ???

# Words and Concepts to Know

Non-deterministic

Stochastic Processes

Ensemple variance

SSS

Stationarity

Temporal variance

Deterministic

Ergodicity

Autocovariance

WSS

Ensemple mean

Autocorrelation

Strict Sense Stationary

Realization

Temporal mean

Wide Sense Stationary