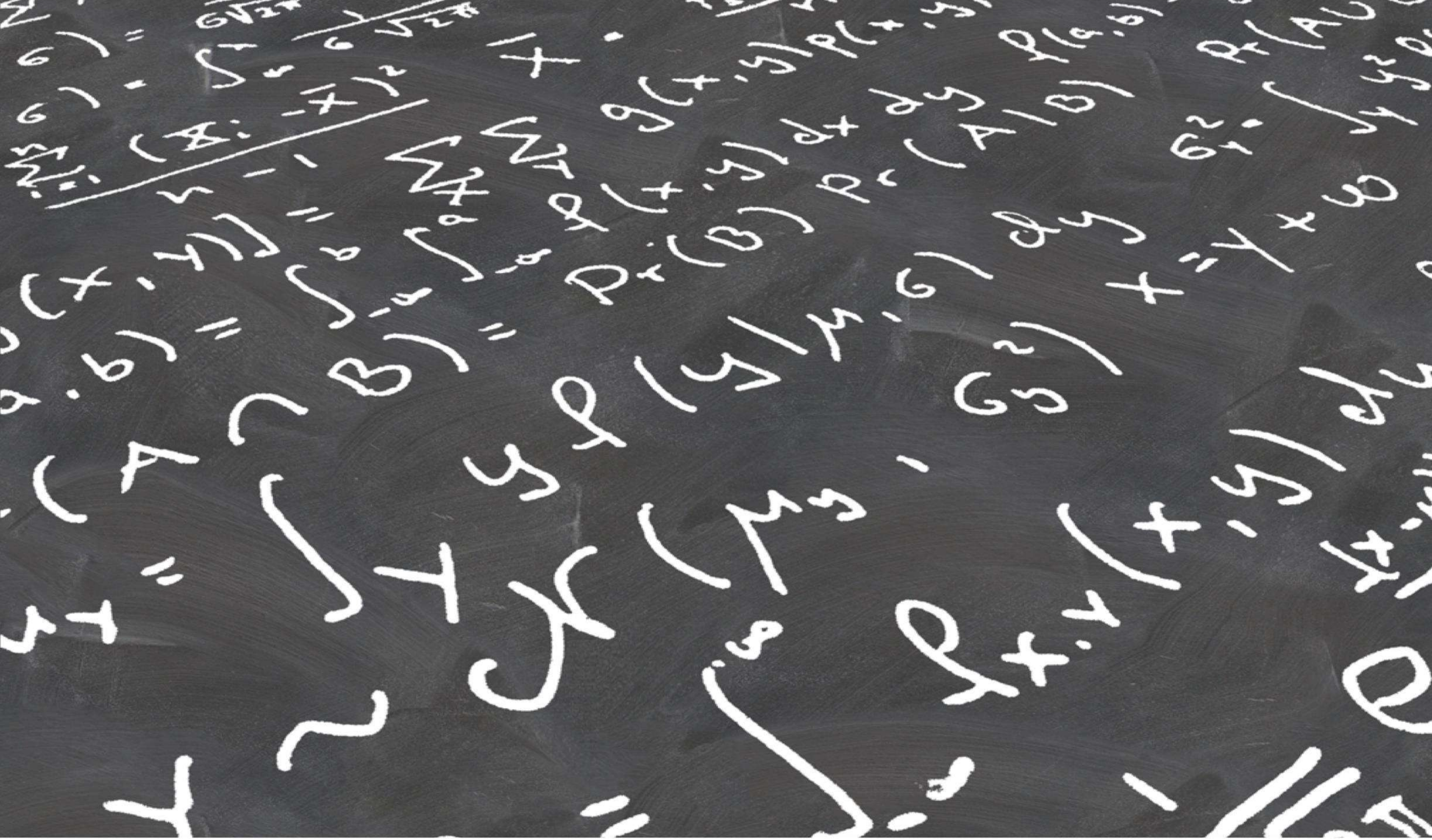


6. Introduction to Stochastic Processes

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

Two Random Variables

Joint (Simultaneous) pdf: $f_{X,Y}(x,y) \geq 0$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

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

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

The Conditional PDF and Independence

Conditional pdf:

- For a two dimensional pdf $f_{X,Y}(x,y)$, we can find the conditional pdf with Bayes rule:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Independence:

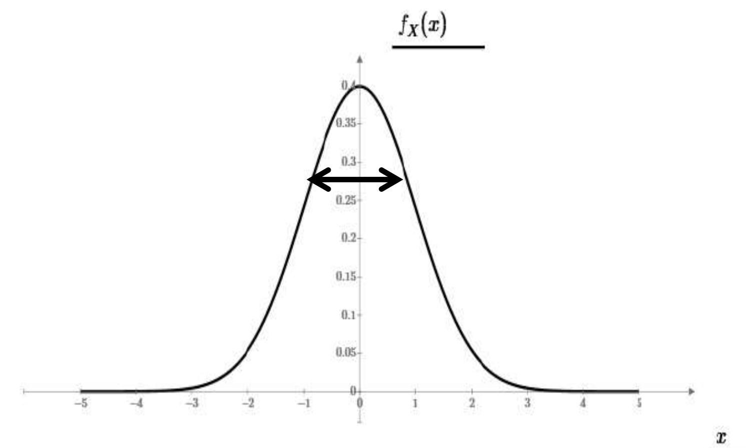
- X and Y are independent if and only if:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \quad \text{or} \quad f_{X|Y}(x|y) = f_X(x)$$

for all x and y

Expectations

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



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

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

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

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

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: $\rho = \frac{E[XY]-E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$ $-1 \leq \rho \leq 1$

- Sum: $\mathrm{Z} = X + Y$
- Expectation: $\mathrm{E}[\mathrm{Z}] = E[X] + E[Y]$
- Variance: $\mathrm{Var}[\mathrm{Z}] = Var[X] + Var[Y]$ if <u>independent</u>

  $\mathrm{Var}[\mathrm{Z}] = Var[X] + Var[Y] + 2cov(X,Y)$ if <u>dependent</u>

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

Example:

- Quantisation noise.

Central Limit Theorem

- Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean μ and variance σ^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 = μ and variance = $\frac{\sigma^2}{n}$.

The variance is reduced with a factor 1/n

Central Limit Theorem

- Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean μ and variance σ^2
- Let X be the random variable:

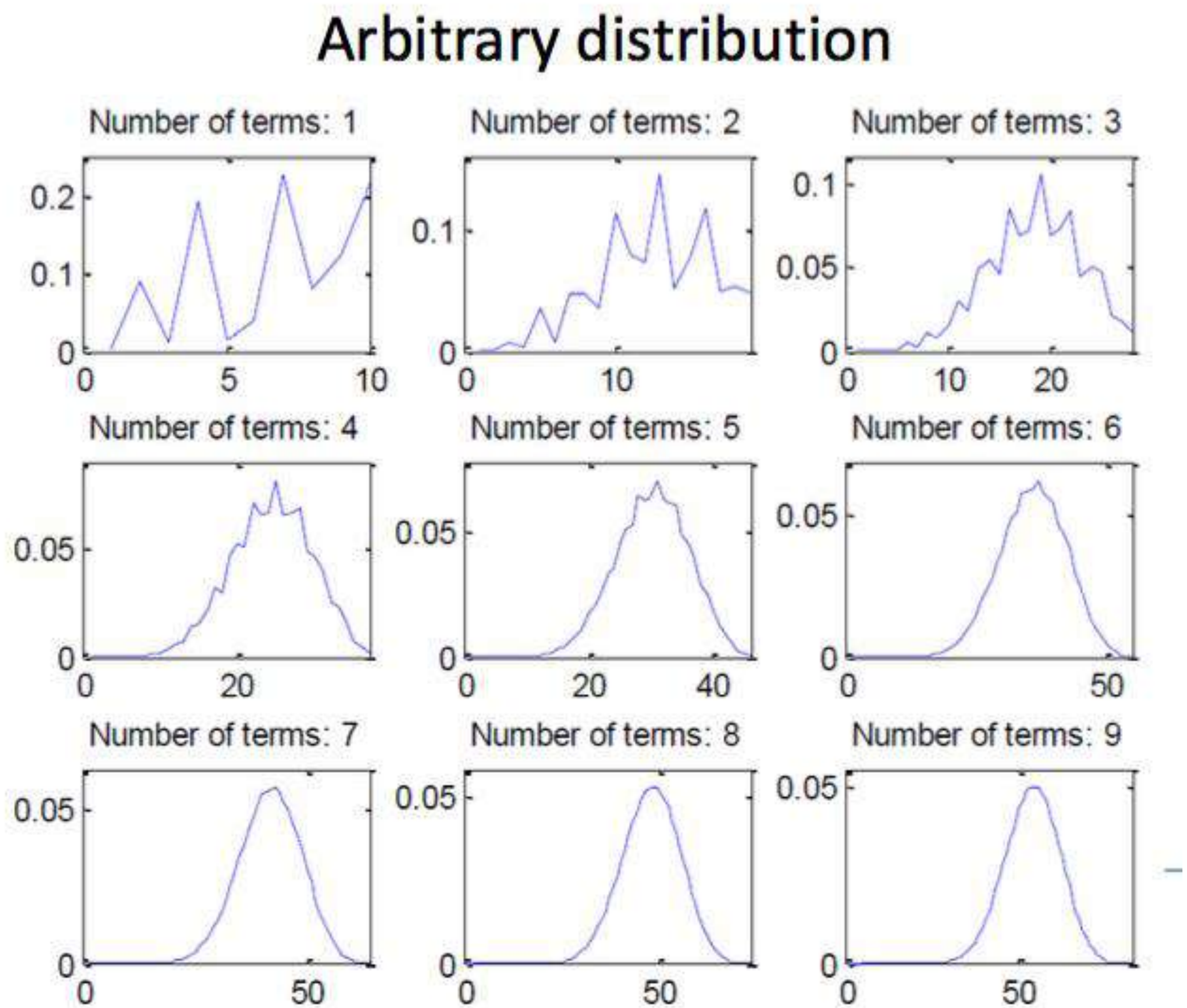
$$X = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} = \frac{\sum_{i=1}^{n} \frac{1}{n} X_i - \mu}{\sqrt{\sigma^2/n}} = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$$

- Then in the limit: $n \to \infty$ we have that: $X \sim \mathcal{N}(0,1)$

  i.e. in the limit X will be normally distributed with mean = 0 and variance = 1 (standard normal distributed).

Sum of Random Variables

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

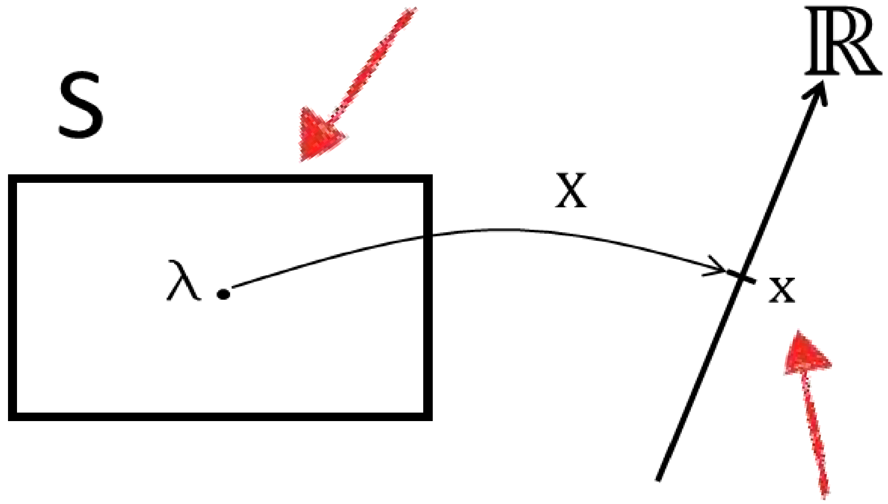


Stochastic Processes

Stochastic Variables

- Sample space for stochastic experiment

Sample space for
stochastic experiment



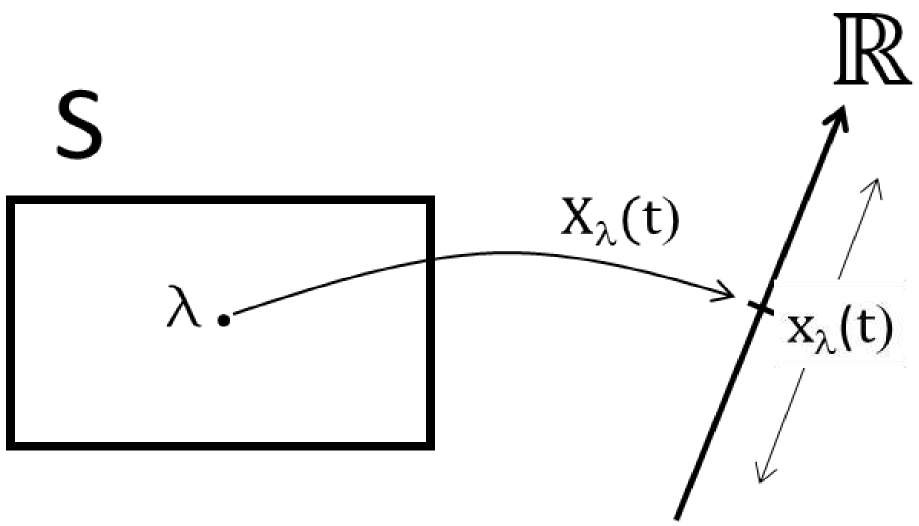
Sample space for
stochastic variable

Time dependent

Stochastic Processes (signals)

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

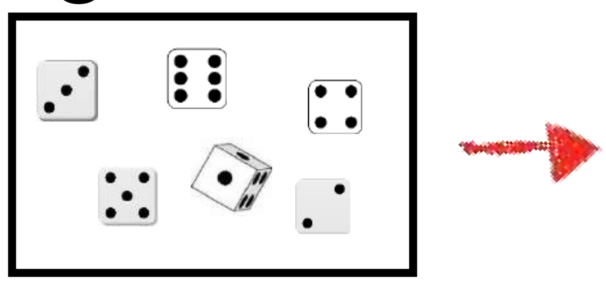
Sample space for
stochastic experiment



Sample space for
stochastic proces

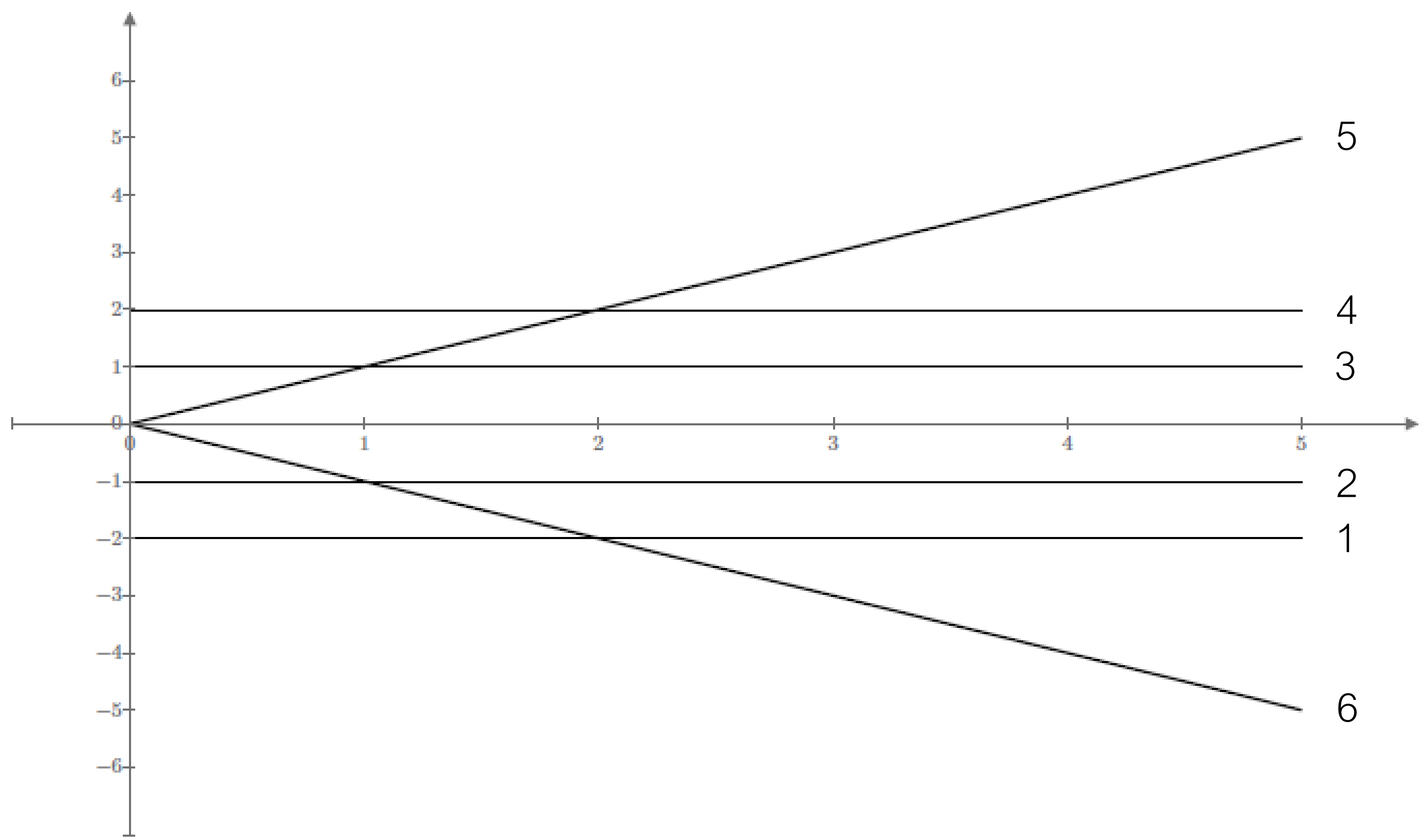
Stochastic Processes – Example

S



$$X_n(t): \quad x_1(t) = -2 \quad x_2(t) = -1 \quad x_3(t) = 1 \quad x_4(t) = 2 \quad x_5(t) = t \quad x_6(t) = -t$$

$x_1(t)$ ______ $x_2(t)$ ______ $x_3(t)$ ______ $x_4(t)$ ______ $x_5(t)$ ______ $x_6(t)$ ______

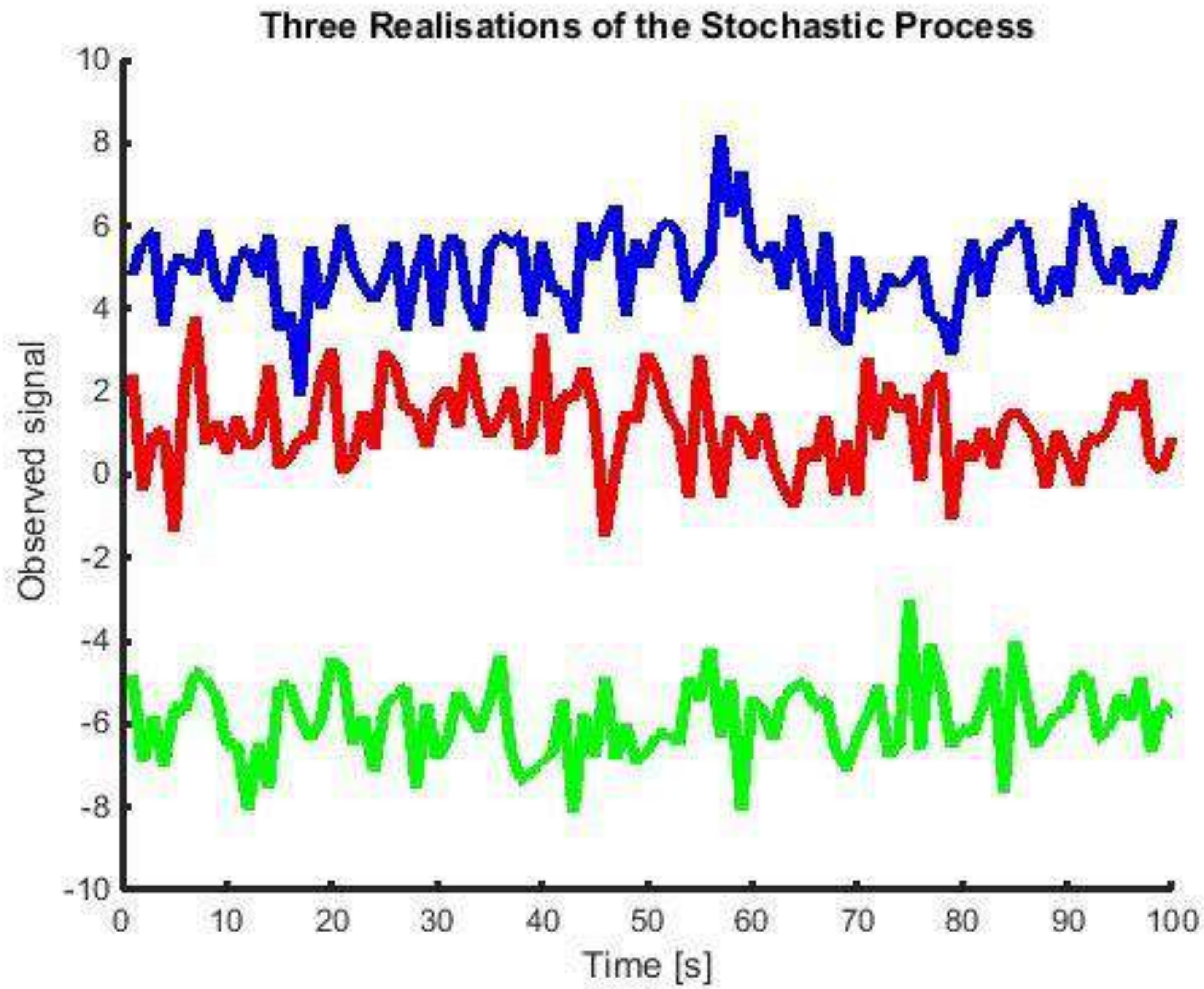


t

Stochastic Processes – Signals

Additive Noisemodel

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



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

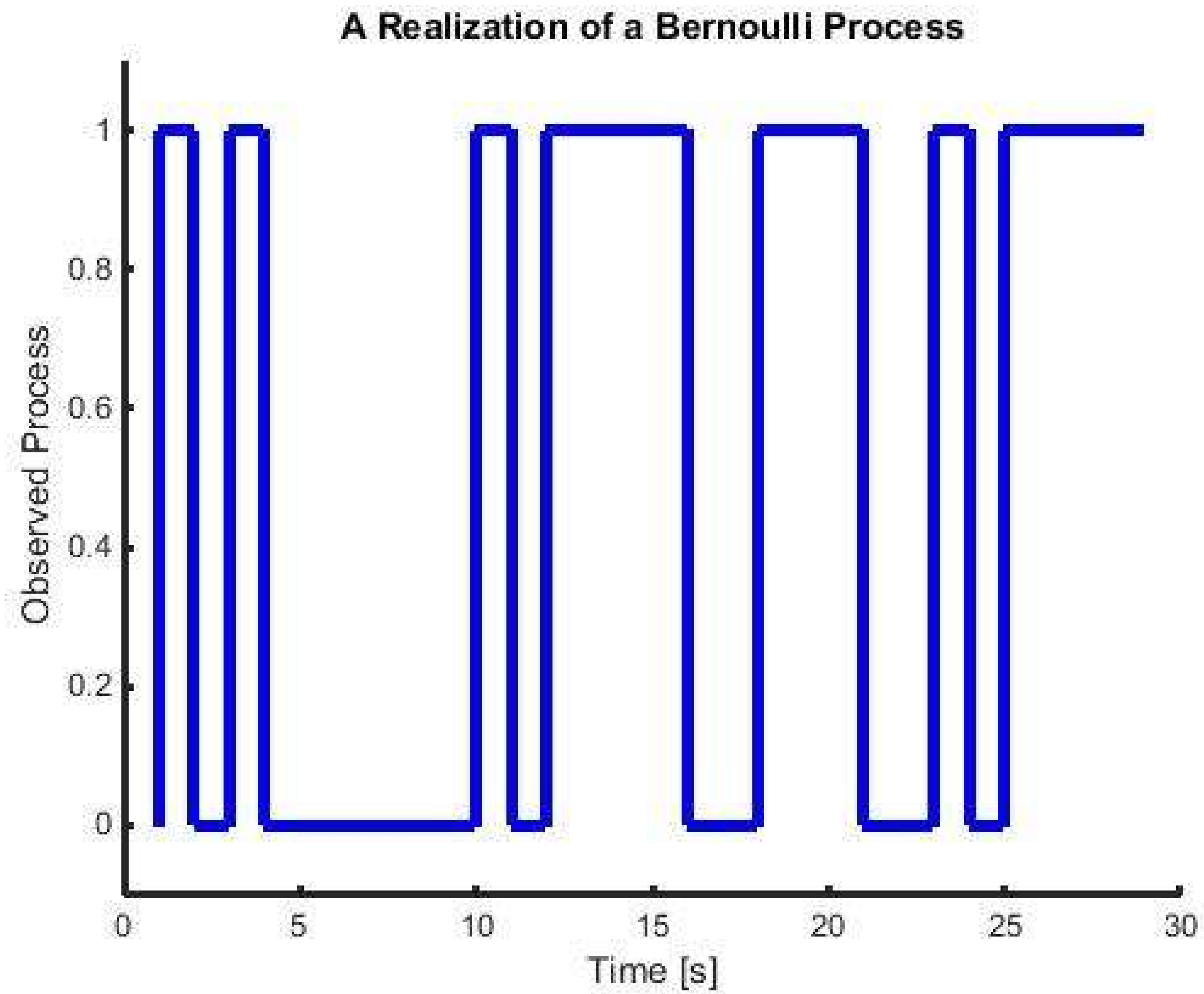
Example:

- 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 0s.
- Is a sequence of i.i.d of Bernoulli trials.



Time Dependent Probability Functions

- Probability density function (pdf):

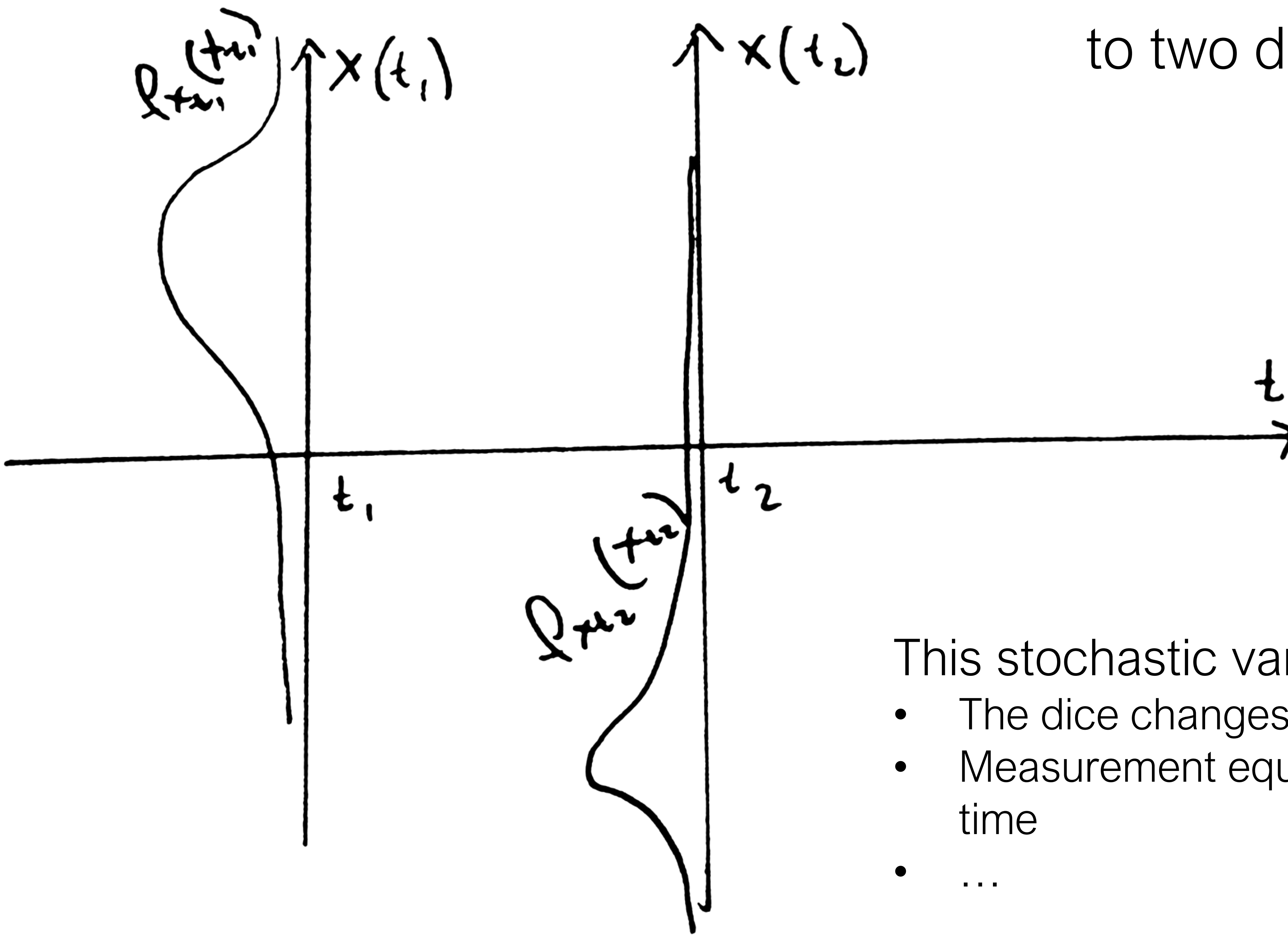
$$f_{X(t)}(x(t))$$

- 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

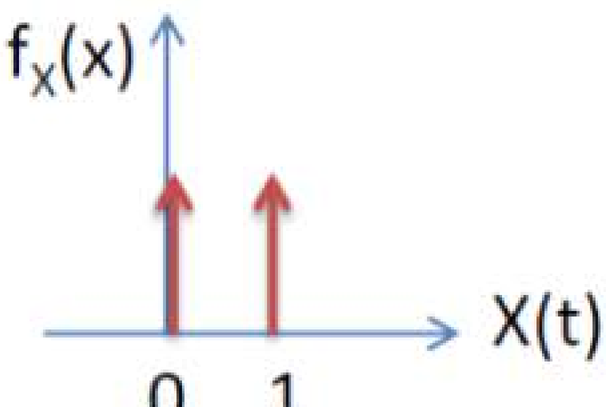


This stochastic variable is not i.i.d.:

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

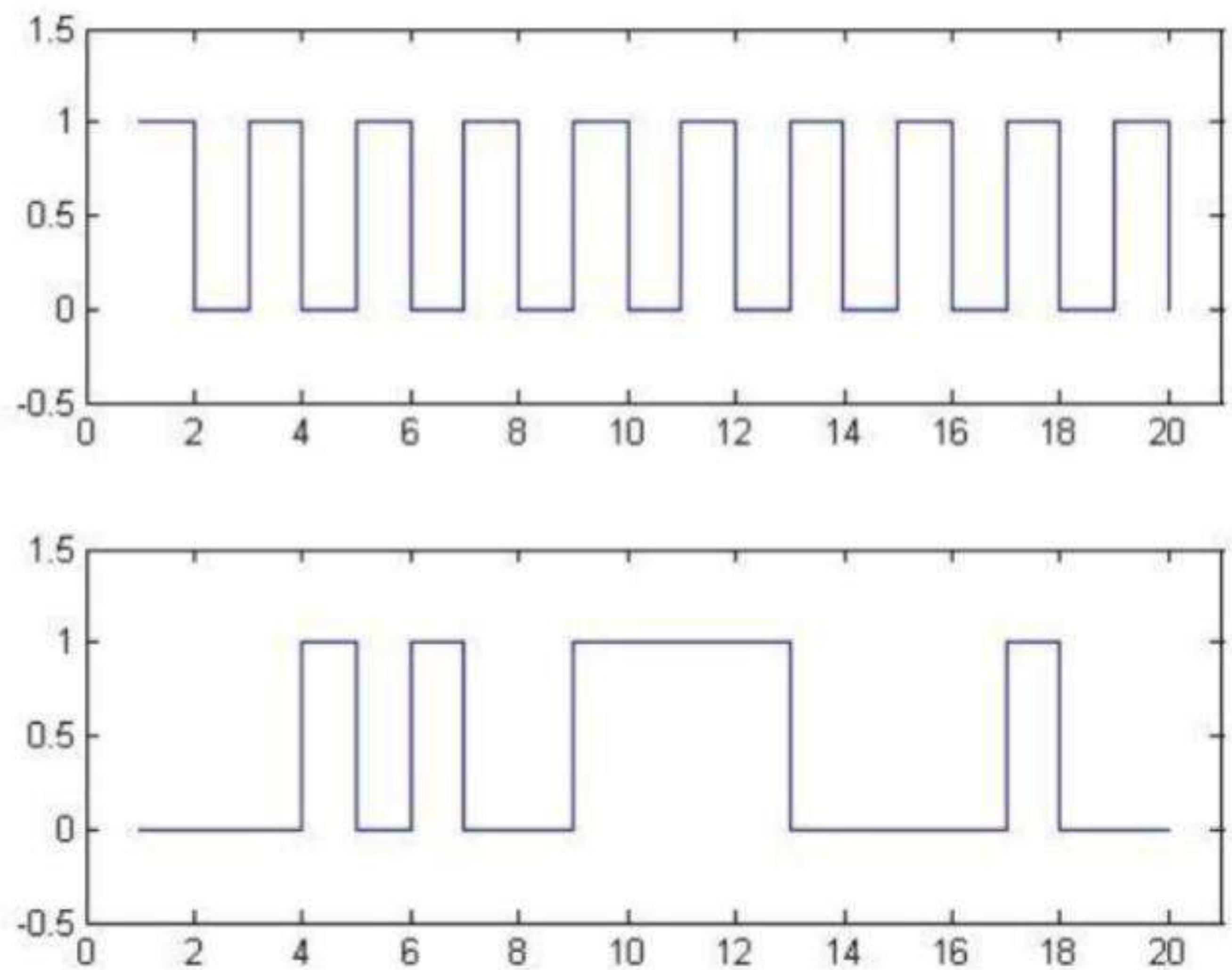
Deterministic Functions

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



pmf

Deterministic



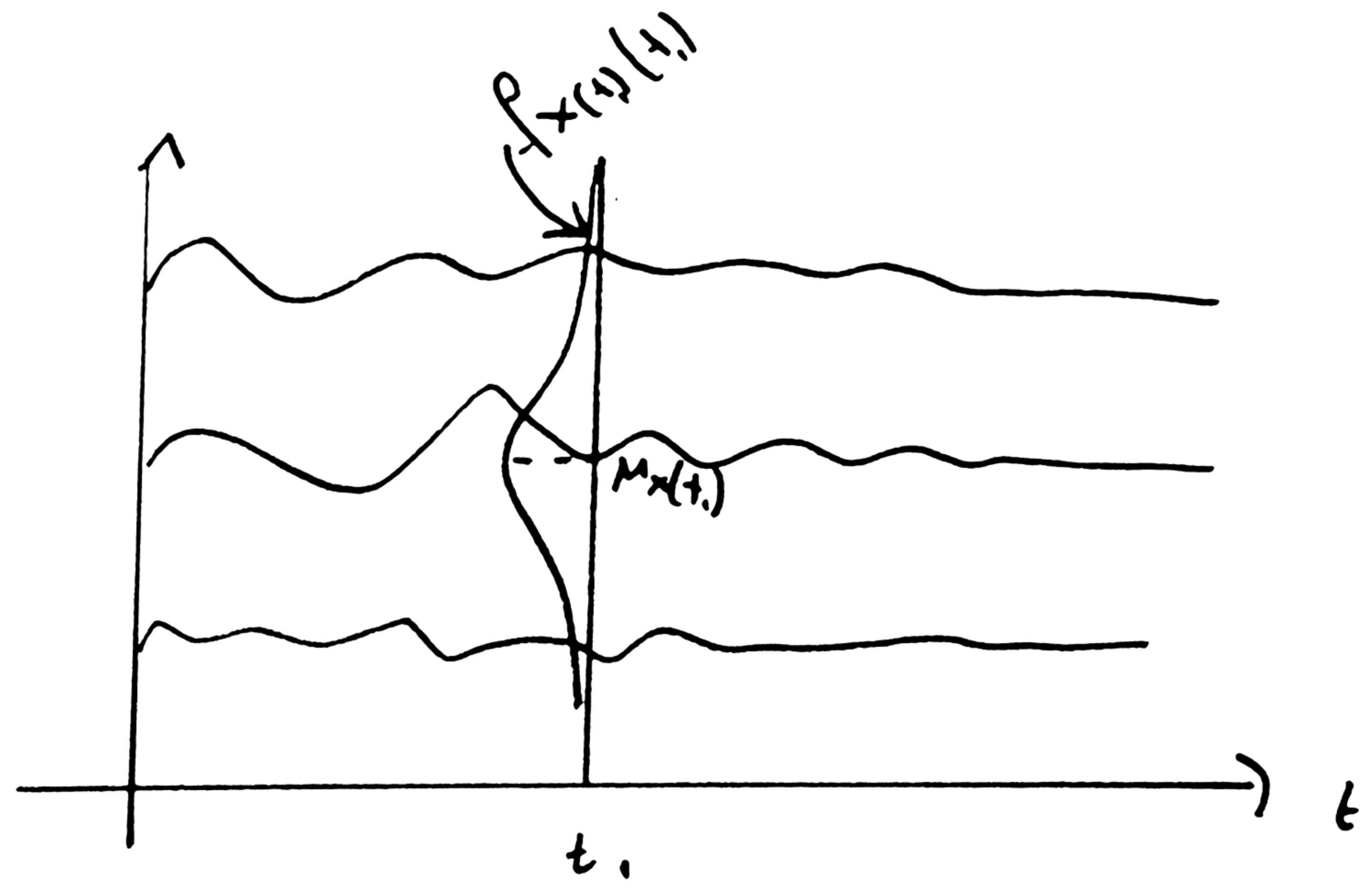
Random (non-deterministic)

Ensemble mean

- The mean value function:

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

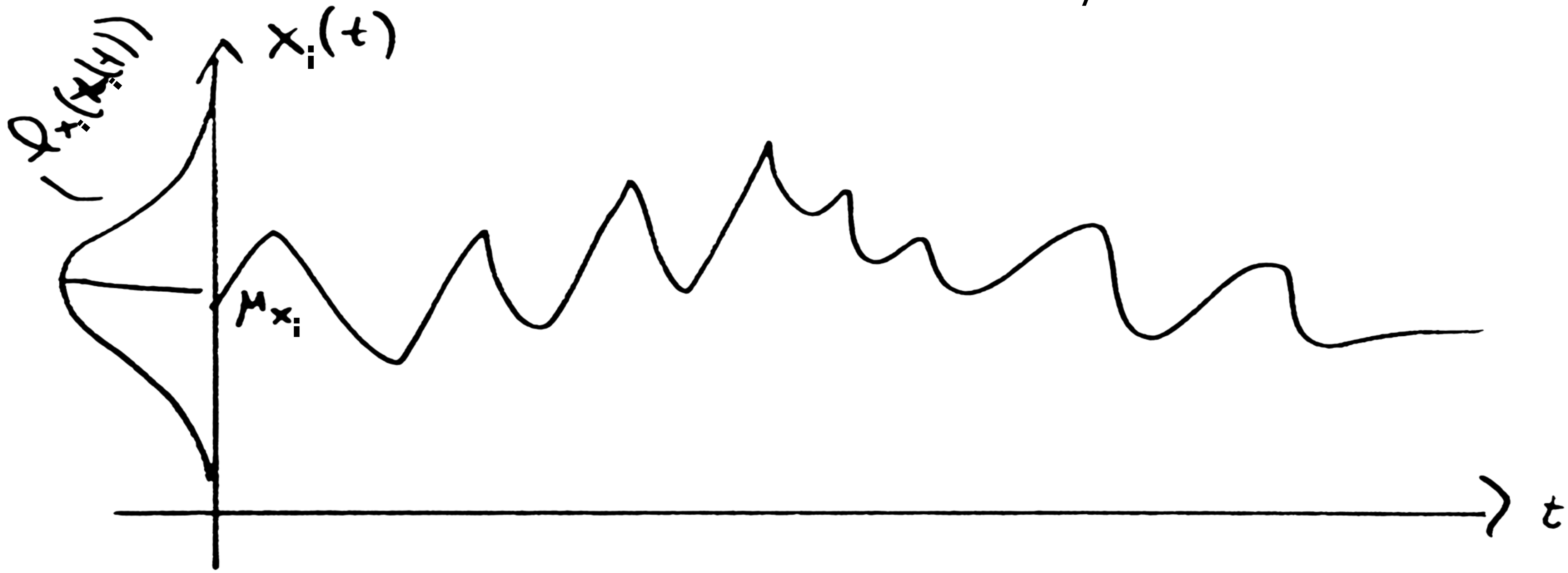
- The mean of all possible realizations to time t



Temporal Mean

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

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

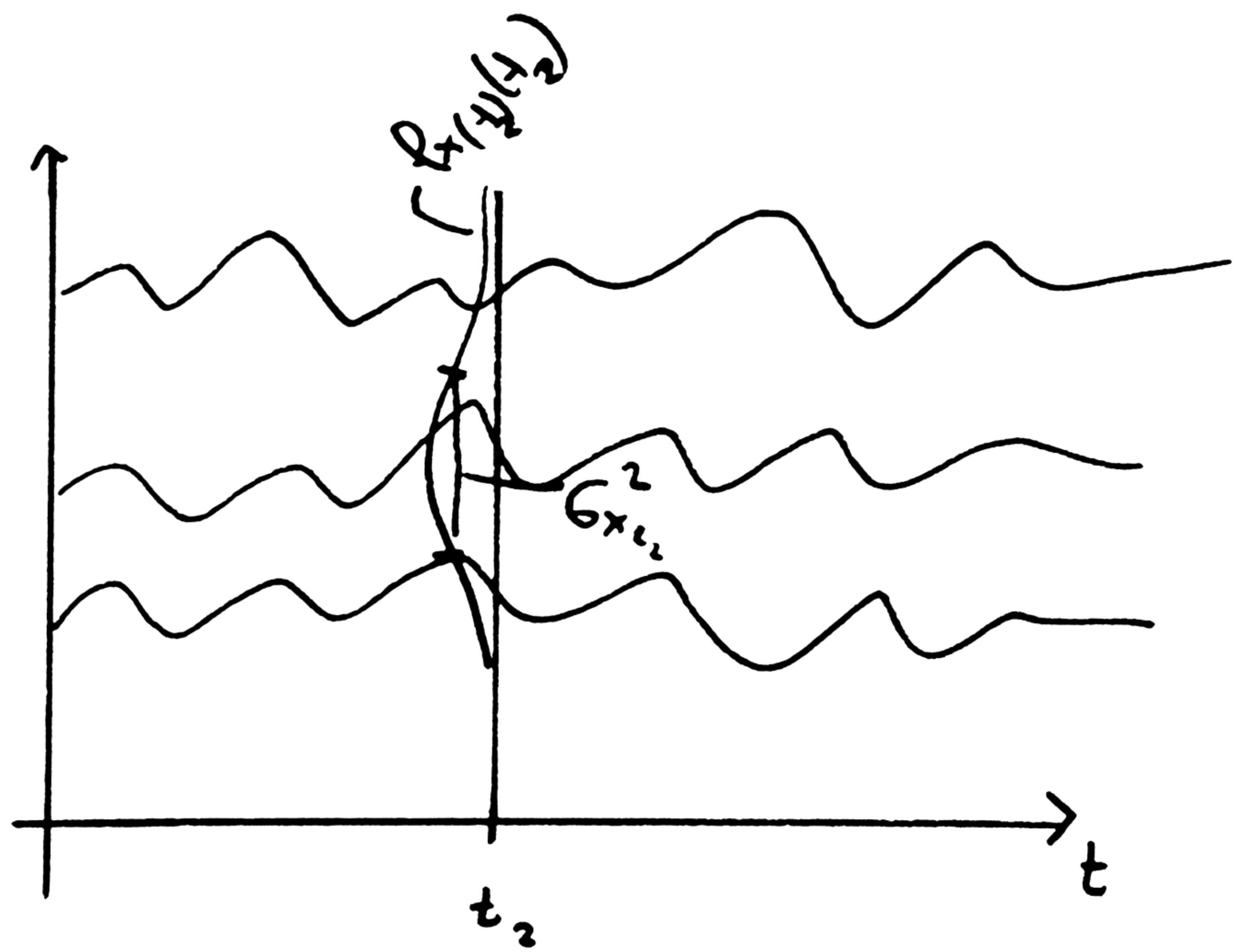


Ensemble Variance

- The variance function:

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

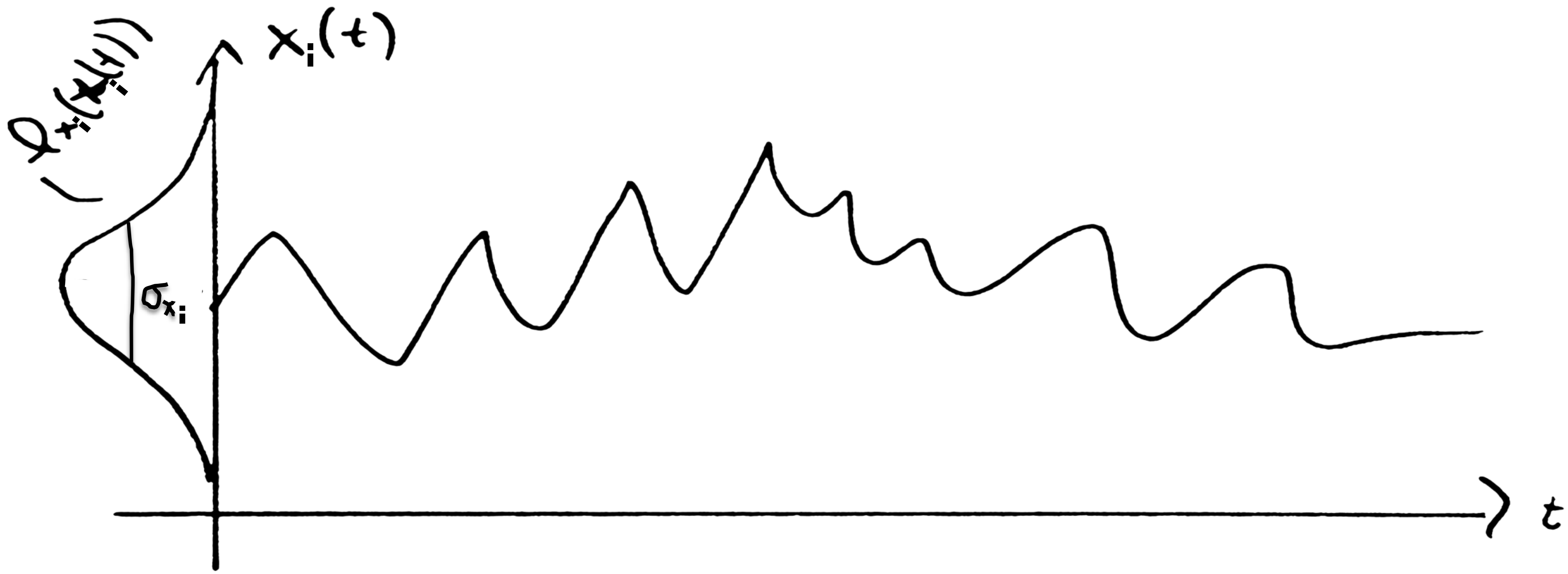
- The variance of all possible realizations to time t



Temporal Variance

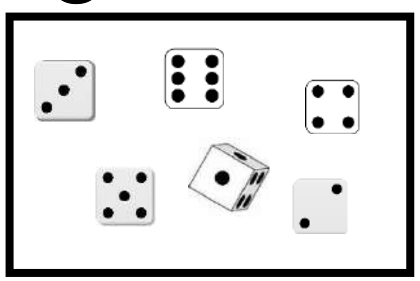
- The variance over time for one realization of the stochastic process
- The temporal variance can differ from the ensemble variance

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

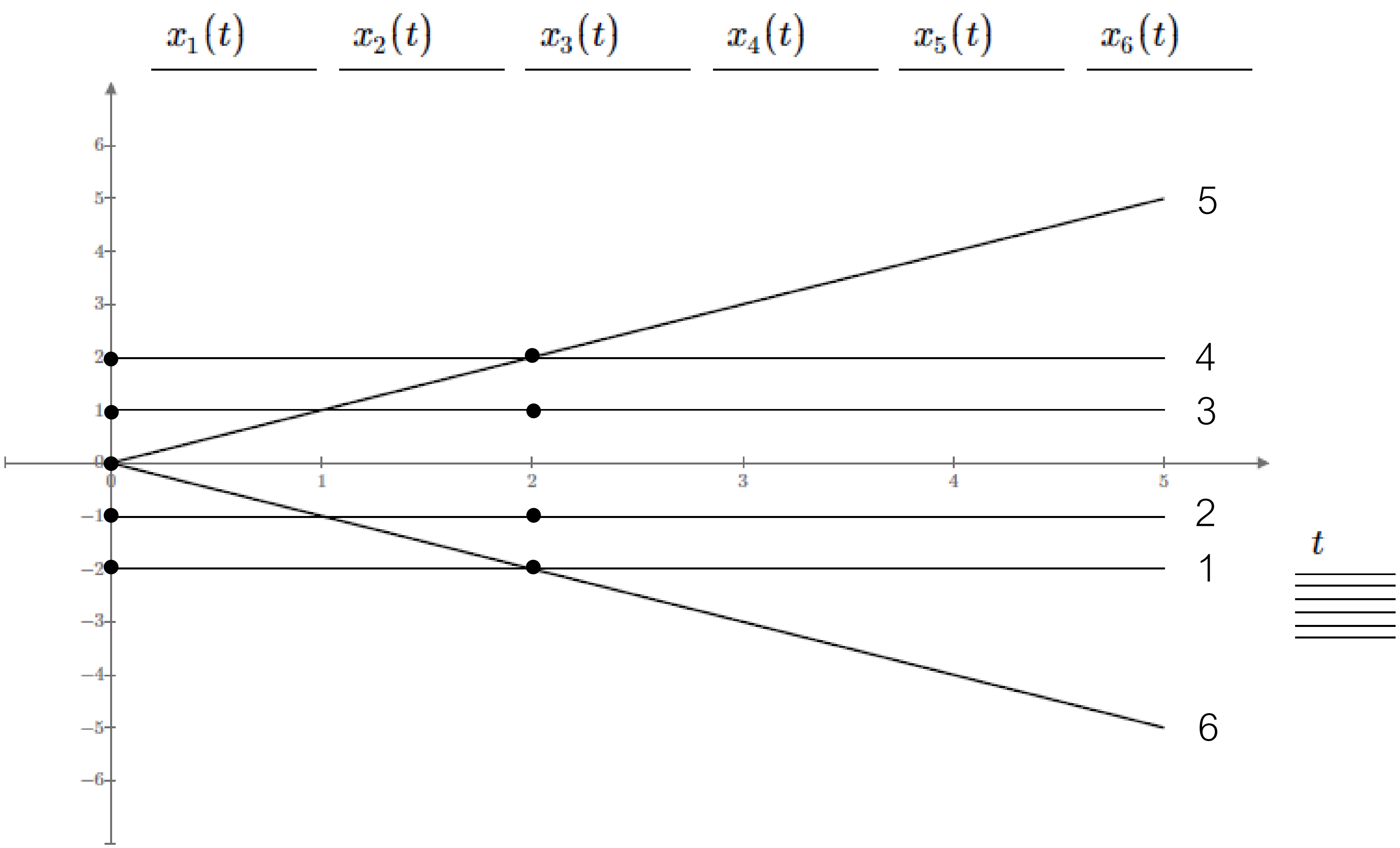


Stochastic Process - Example

S



$$X_n(t): \quad x_1(t) = -2 \quad x_2(t) = -1 \quad x_3(t) = 1 \quad x_4(t) = 2 \quad x_5(t) = t \quad x_6(t) = -t$$



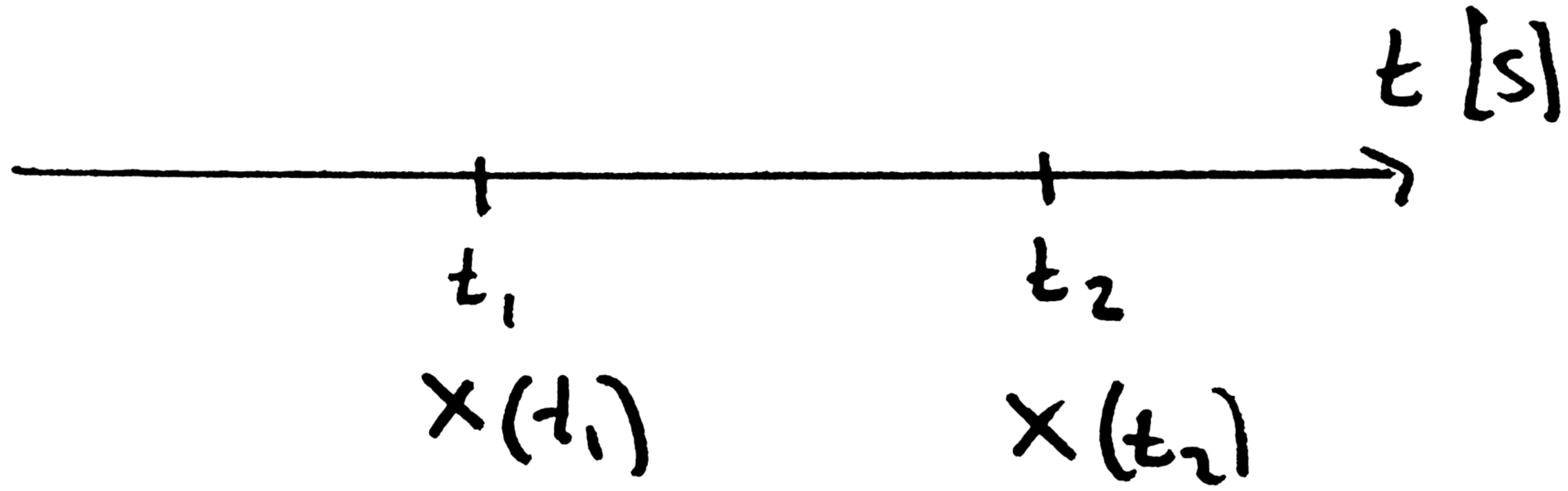
$$X(0) = \{-2, -1, 0, 1, 2\}$$

$$X(2) = \{-2, -1, 1, 2\}$$

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

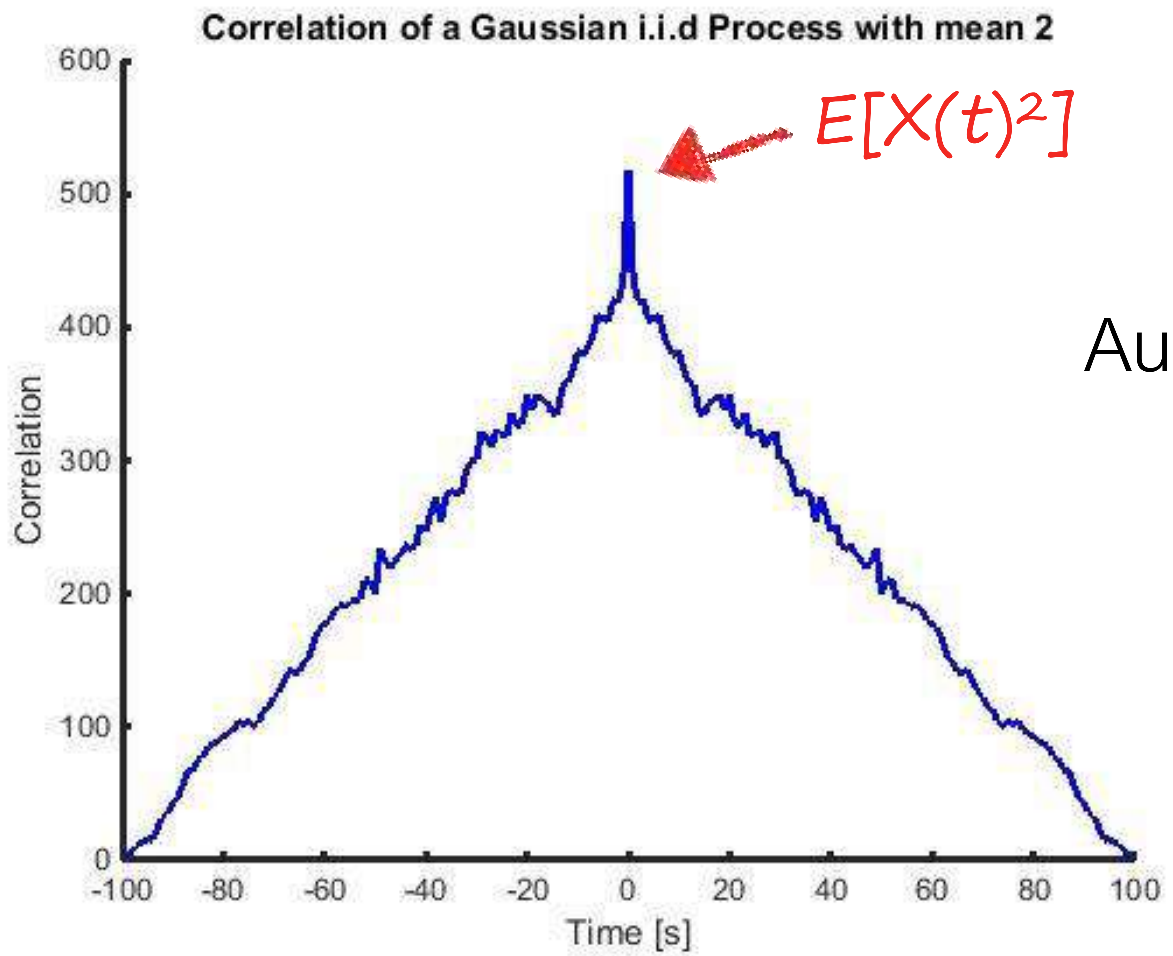
Tells about the connection at two different times

- 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:

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

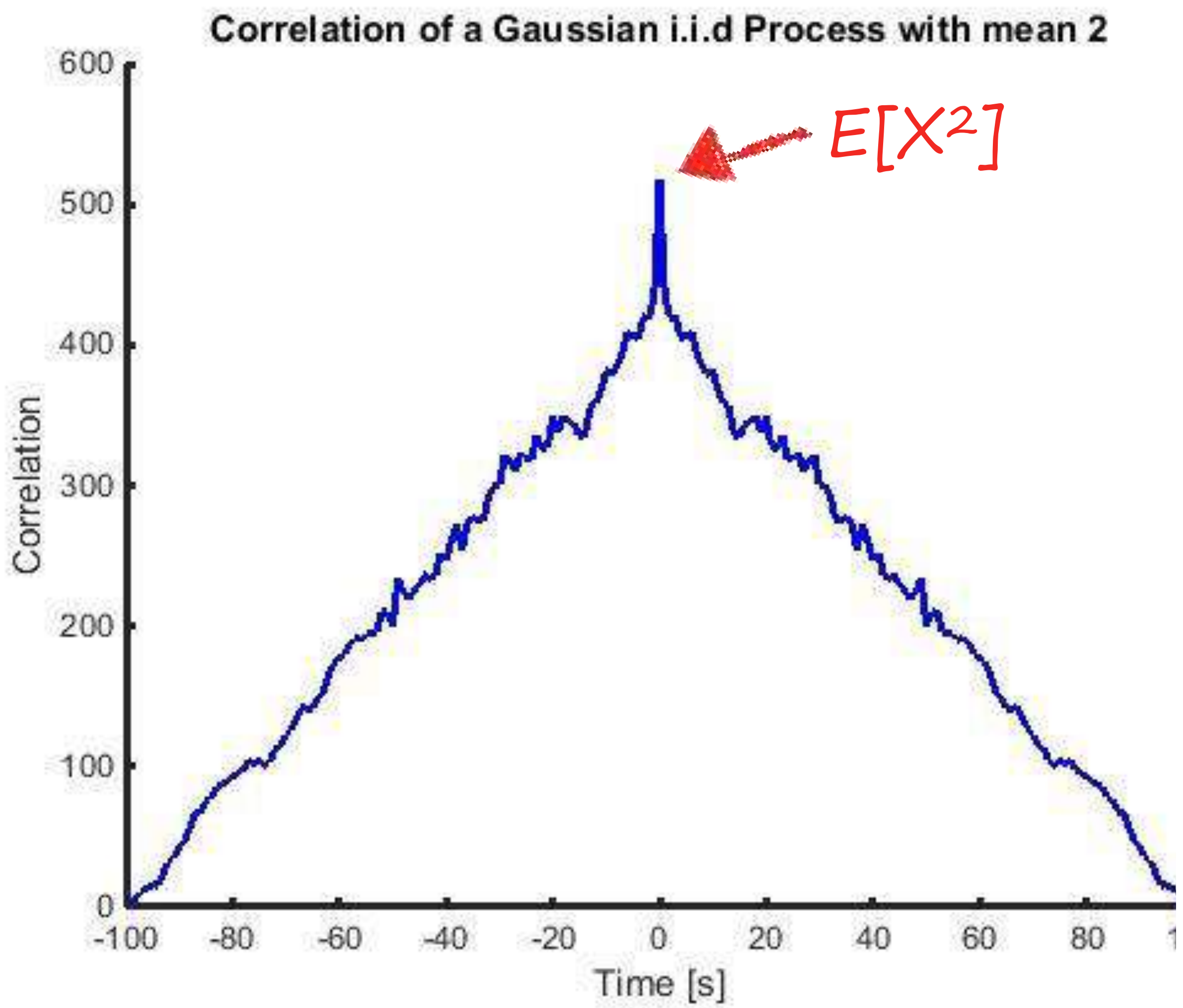
- Autocorrelation coefficient:

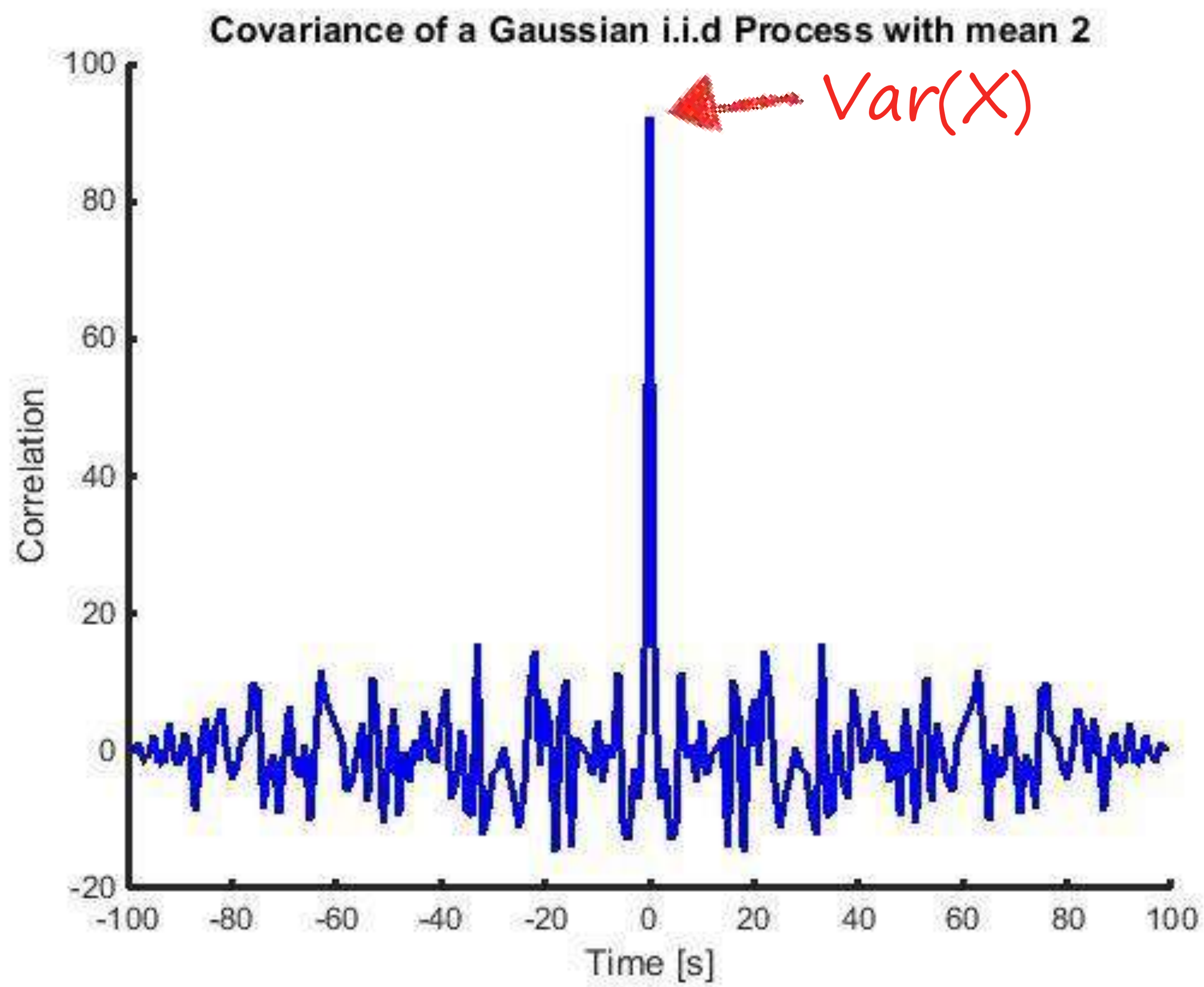
$$r_{XX}(t_1,t_2) = \frac{C_{XX}(t_1,t_2)}{\sqrt{C_{XX}(t_1,t_1)C_{XX}(t_2,t_2)}}; \qquad 0 \le r_{XX}(t_1,t_2) \le 1$$

Autocovariances

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

- Autocorrelation and autocovariance

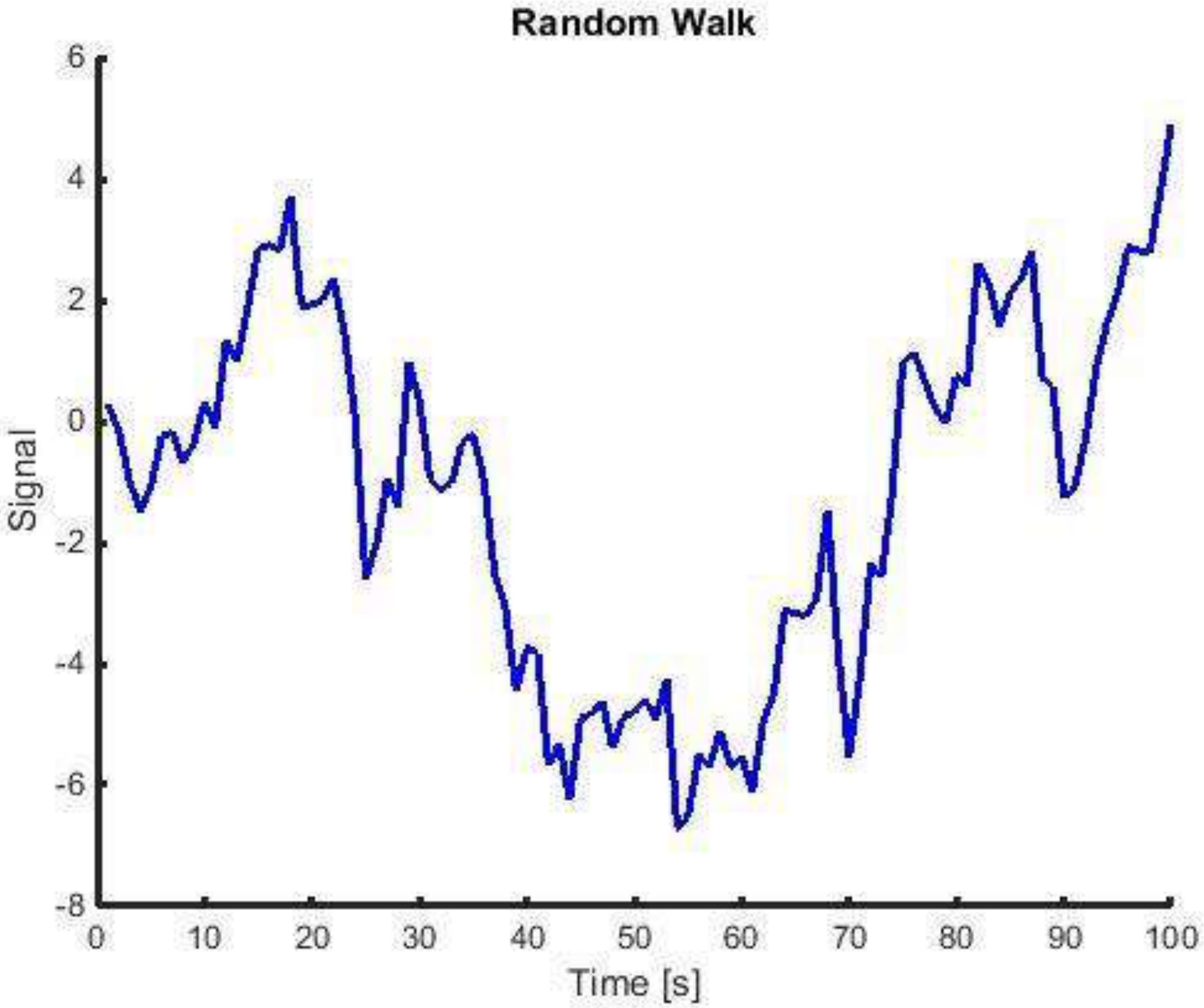




Random Walk – Example

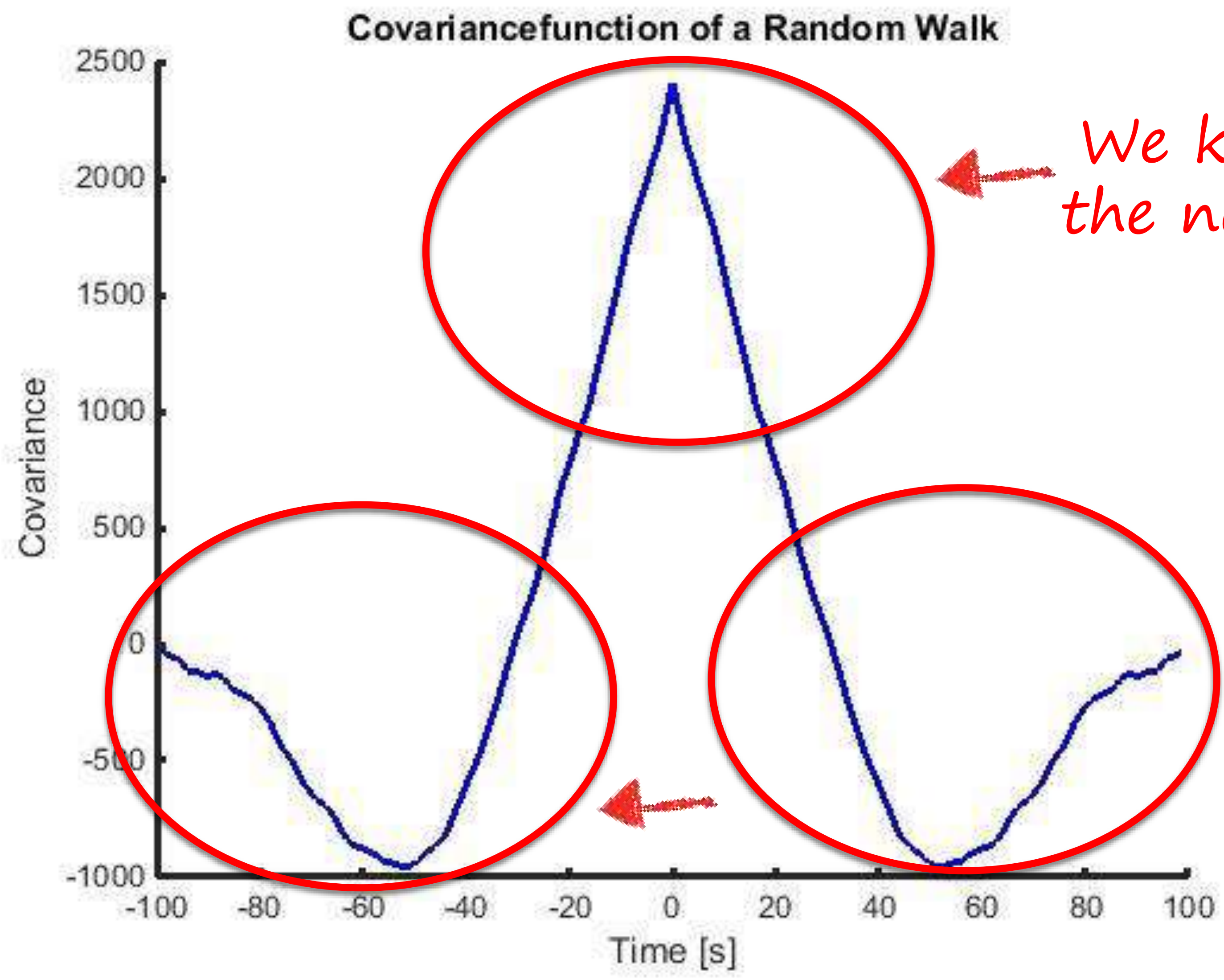
Brownish motions

- We consider a random walk.



Random Walk – Example

- Sample of the autocovariance function:



We know something about the nearest past and future

We know nothing about the far past and future

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 Δt_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 $\Delta \mathrm{t}$, 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

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

- $R_{XX}(\tau)$ decreases fast from 0, if $x(t)$ changes fast
- $R_{XX}(\tau)$ decreases slowly, if $x(t)$ changes slowly
- If $R_{XX}(\tau)$ contains periodic functions, $x(t)$ contains periodic functions

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)

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

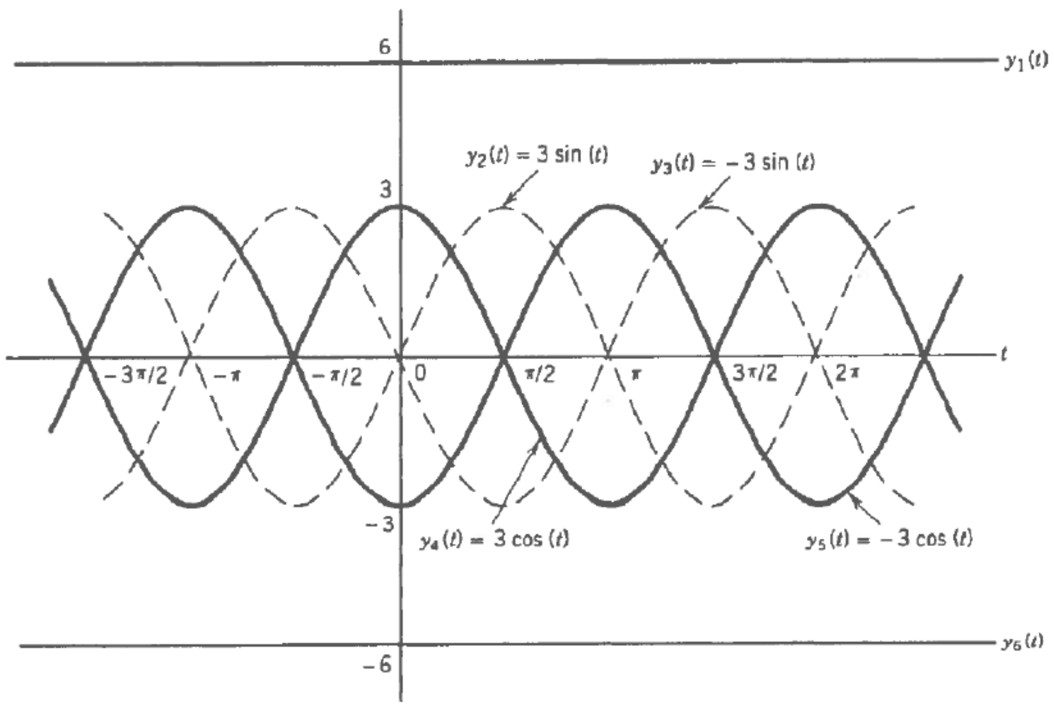
One realization

Ensemple (WSS)

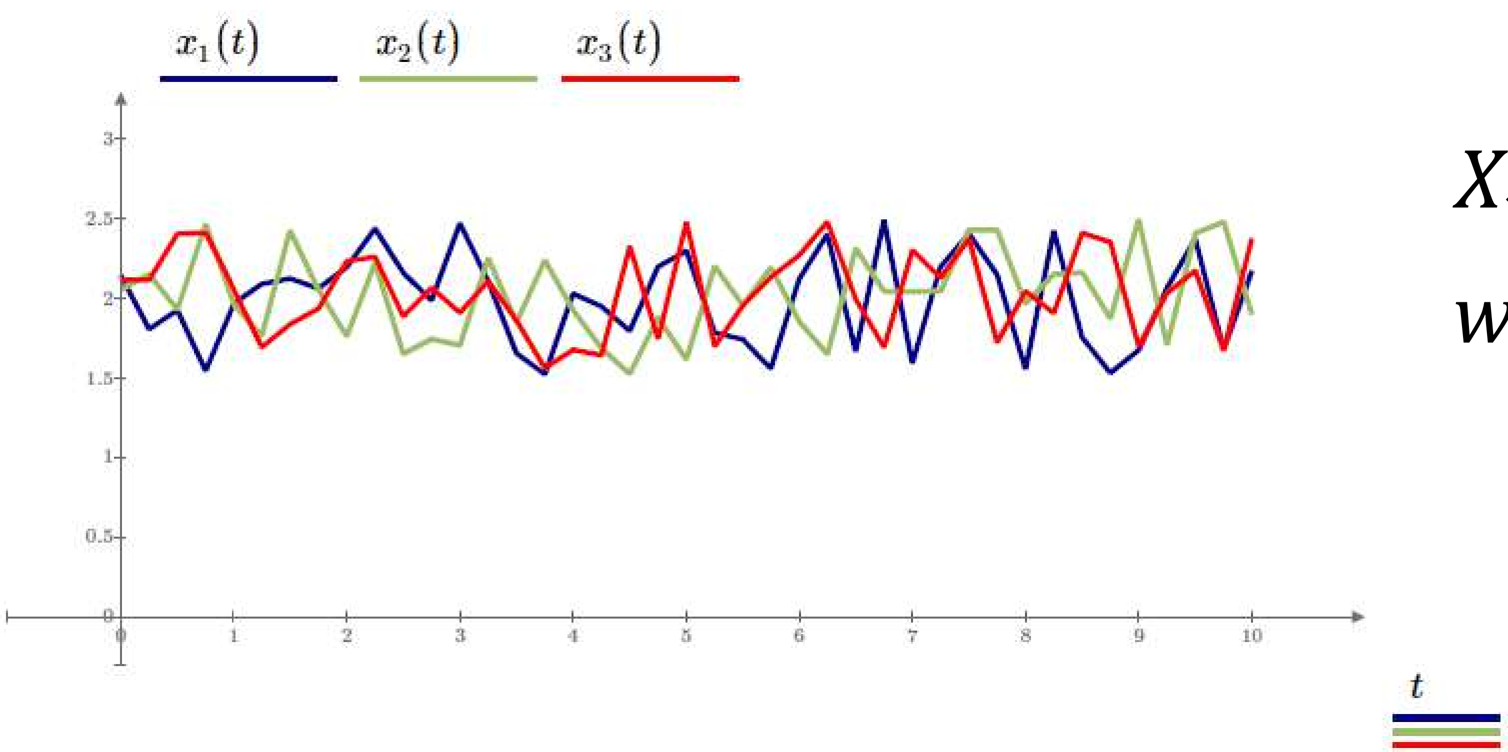
$$\left.\begin{aligned} \langle X_i \rangle_T &= \mu_X \\ \hat{\sigma}_{X_i}^2 &= \sigma_X^2 \end{aligned}\right\} \rightarrow Ergodic$$

All information is achieved with one measurement (realization)

WSS and Ergodicity – Examples



- Not SSS
- WSS
- Not ergodic



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

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

- WSS
- Ergodic

Words and Concepts to Know

Non-deterministic

Stochastic Processes

Ensemple variance

SSS

Stationarity

Temporal variance

Deterministic

Ergodicity

Autocovariance

WSS

Ensemple mean

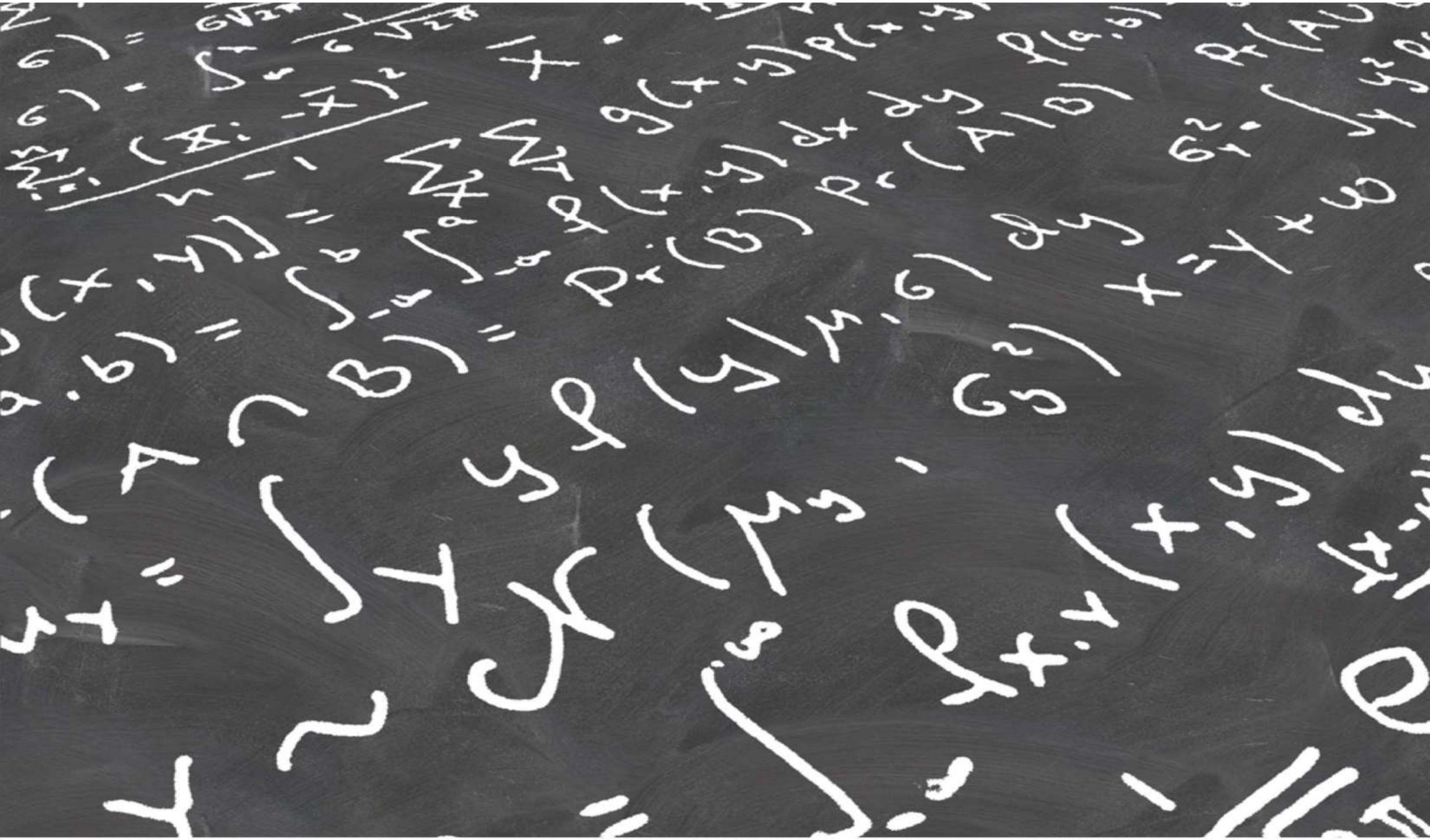
Autocorrelation

Strict Sense Stationary

Realization

Temporal mean

Wide Sense Stationary



7. Stochastic Processes and Correlation Functions

Gunvor Elisabeth Kirkelund
Lars Mandrup

Agenda for Today

- Stochastic Processes (repetition)
  - Mean and variance
  - Stationarity (WSS, SSS)
  - Ergodic Processes
- Correlation functions
  - Autocorrelation functions
  - Cross-correlation functions
- Power spectrum density

Stochastic Processes

Definitions:

- A stochastic process is a time dependent stochastic variable:

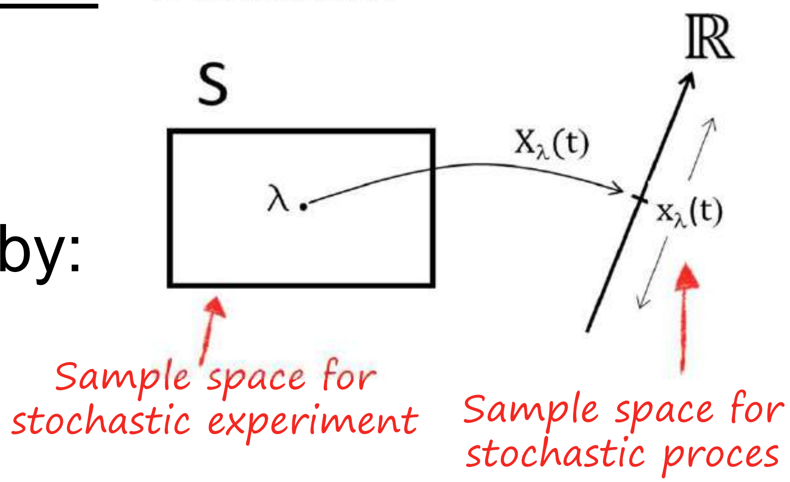
$$X(t)$$

- A discrete stochastic process is given by:

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

where n is an integer.

- Random events that develops in time



Notice:

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

Sample Functions

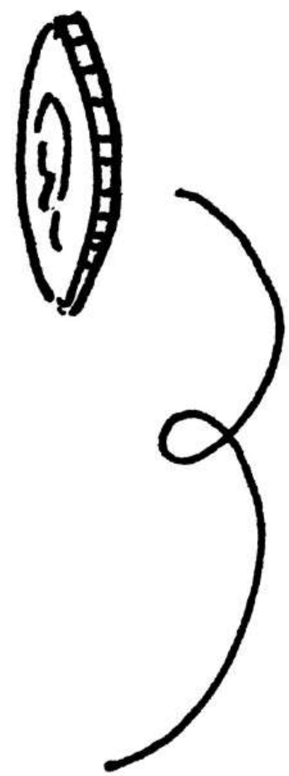
Definition:

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

Example:

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

$$HTHT$$



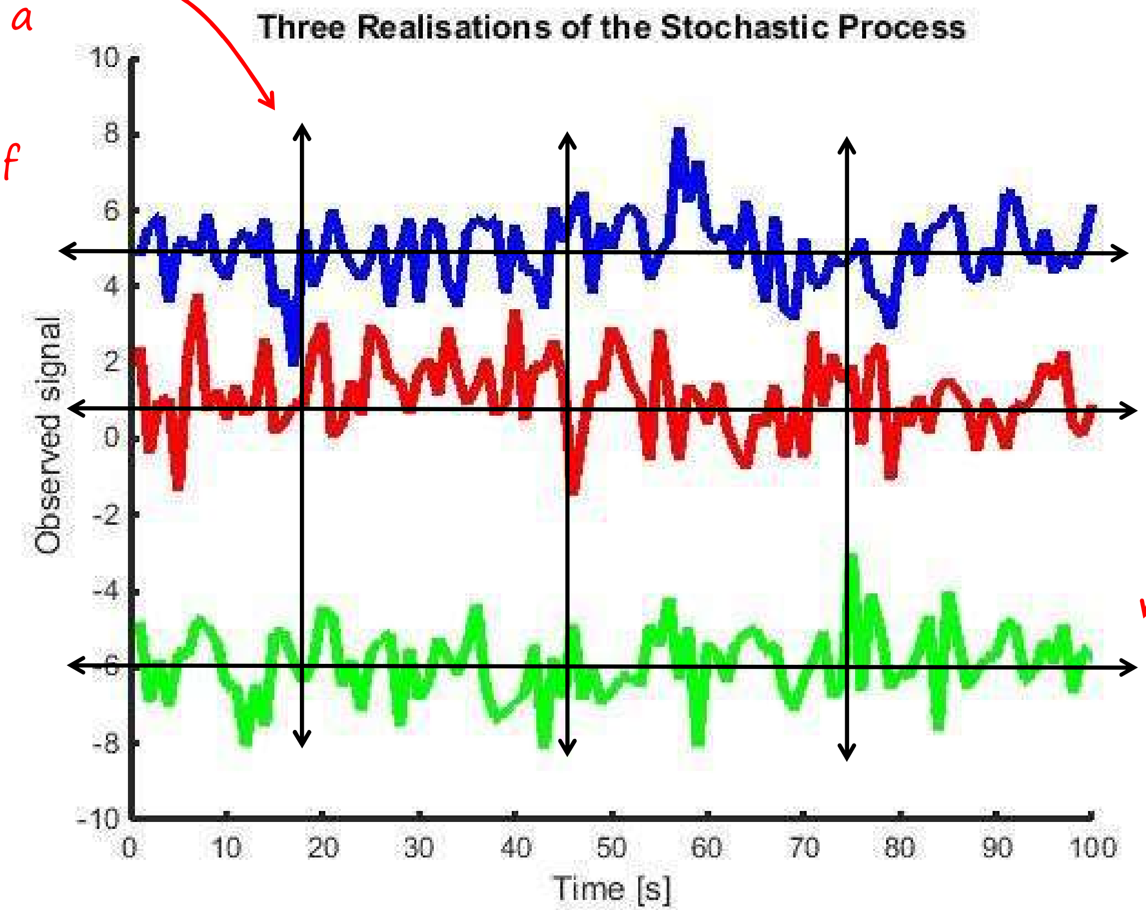
Stochastic Processes (signals)

Additive Noisemodel

$$observed\ signal = signal + noise$$

Ensemble mean and variance (to a specific time).

If indepentent of time: WSS



Time average and variance of each realization.

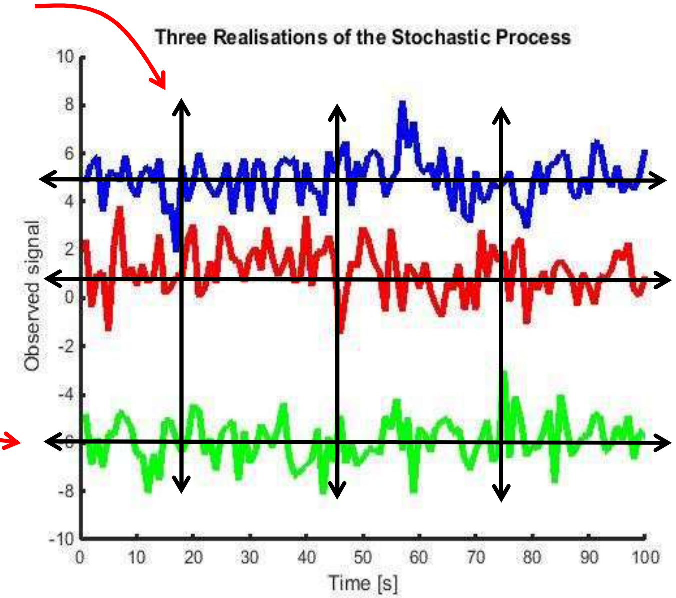
If equal (for all realizations): Ergodic

The Mean Functions

- Ensemple mean:

The mean of all possible realizations to time t

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



The time average for one realization of the stochastic process

- Temporal mean:

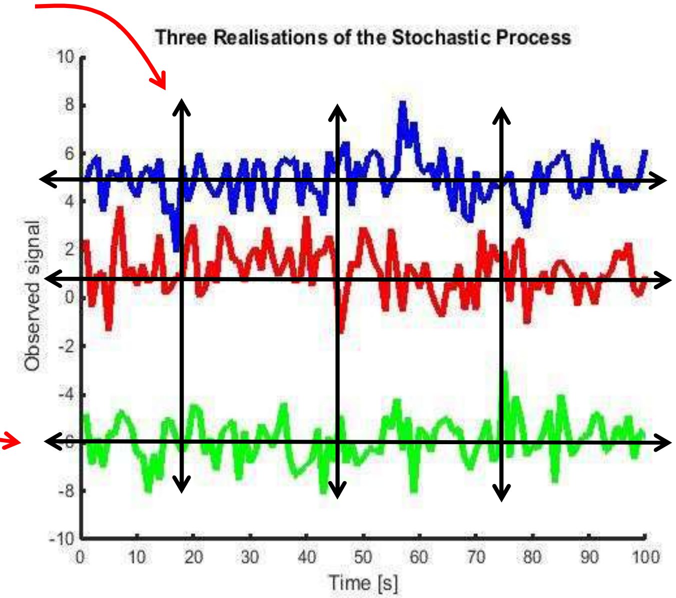
$$\hat{\mu}_{X_i} = \langle X_i \rangle_T = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)\, dt$$

The Variance Functions

- Ensemple variance:

The variance of all possible realizations to time t

$$Var\big(X(t)\big) = \sigma^2_{X(t)}(t) = E\left[\left(X(t) - \mu_{X(t)}(t)\right)^2\right]$$



The variance over time for one realization of the stochastic process

- Temporal variance:

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

Stationarity in the Wide Sense (WSS)

- Ensemble mean is a constant

Can be tested.

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

- Ensemble variance is a constant

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

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

Ergodicity

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

- 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 realization Ensemple (WSS)

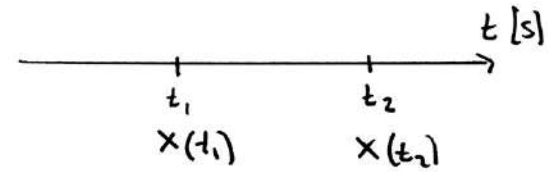
$$\left.\begin{matrix} \langle X_i \rangle_T = \mu_X \\ \hat{\sigma}_{X_i}^2 = \sigma_X^2 \end{matrix}\right\} \rightarrow Ergodic$$

All information is achieved with one measurement (realization)

Comparing realizations

Correlations

- We compare the process at two different times.



Correlation of a realization with itself

- Autocorrelation: $R_{XX}(t_1, t_2) = E[X(t_1)X(t_2)^*]$
  - ➢ Says something about how much the signal $X(t_1)$ resembles itself at time t_2
  - ➢ Must depent on how rapidly the signal changes over time
  - ➢ Larger if $|t_1 - t_2|$ is small

Correlation of two realizations

- Crosscorrelation: $R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*]$
  - ➢ Can be used to look for places where the signal $X(t)$ is similar to the signal $Y(t)$

Autocorrelation

Tells about the connection at two different times

- In general:

Complex conjugated

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

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

- For a stationary process (WSS):

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

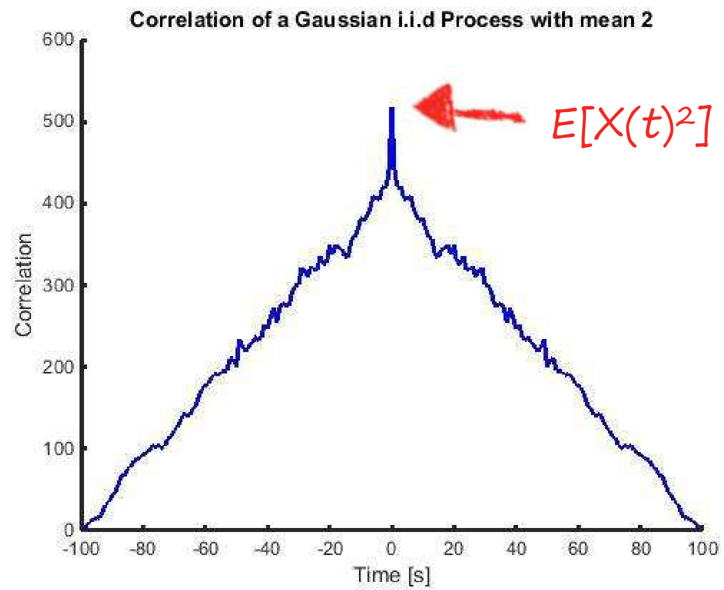
- We rewrite to: $R_{XX}(\tau) = E[X(t)X(t + \tau)^*]$

tau is the lag!

Autocorrelation

- For Real WSS: $R_{XX}(\tau) = E[X(t)X(t+\tau)]$

- Properties of the autocorrelation function $R_{XX}(\tau)$:
  - An even function of τ $(R_{XX}(\tau) = R_{XX}(-\tau))$
  - Bounded by: $|R_{XX}(\tau)| \leq R_{XX}(0) = E[X^2]$ (max. in $\tau = 0$)
  - If X(t) changes fast, then $R_{XX}(\tau)$ decreases fast from $\tau = 0$
  - If X(t) changes slowly, then $R_{XX}(\tau)$ decreases slowly from $\tau = 0$
  - if X(t) is periodic, then $R_{XX}(\tau)$ is also periodic



Temporal Autocorrelation

Temporal only looks at one realization of the stochastic process.

Convolution

- Temporal autocorrelation:

$$\mathcal{R}_{XX}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t) \cdot x\,(t+\tau)dt$$

- If the process is ergodic the temporal autocorrelation is equal to the ensemble autocorrelation:

$$R_{XX}(\tau) = \mathcal{R}_{XX}(\tau)$$

Ensemble

Temporal

Estimate Autocorrelation

We only have measurements of one realization of X(t)

Autocorrelation function:

- In practise, with respect to the lag:

temporal

$$\mathcal{R}_{XX}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t) \cdot x\,(t+\tau) dt$$

N+1 measurements

$$x(0), x(\Delta t), x(2\Delta t), \dots, x(N\Delta t)$$

- The estimated autocorrelation function:

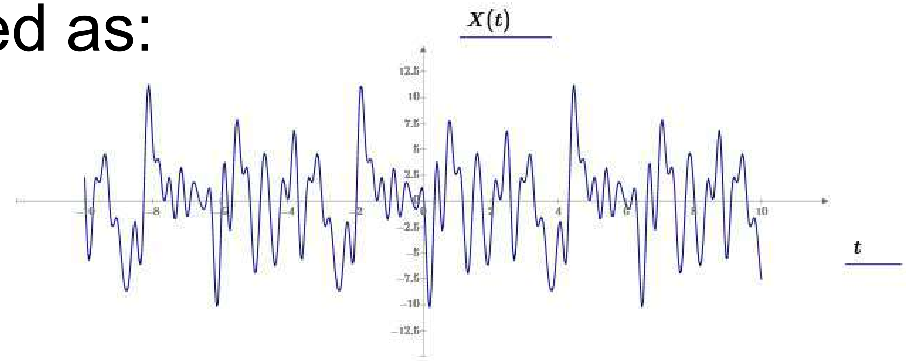
hat = estimation

$$\hat{R}_{XX}(n\Delta t) = \frac{1}{\underbrace{N-n+1}_{\text{Number of terms } (T/\Delta t)}} \sum_{k=0}^{N-n} \underbrace{x(k\Delta t)}_{t} \cdot \underbrace{x((k+n)\Delta t)}_{t+\tau}$$

Autocorrelation Functions – Example

- Let a stochastic process be defined as:

$$X(t) = \sum_{i=1}^{n} (A_i cos\omega_i t + B_i sin\omega_i t)$$



- where A_i , $B_i \sim \mathcal{N}(0, \sigma^2)$ and i.i.d., and $\omega_i = i \cdot \omega_0$

- Find the autocorrelation:

$$E[X(t)X(t+\tau)] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} (A_i cos\omega_i t + B_i sin\omega_i t) \ldots \right.$$
$$\left. \cdot (A_j cos\omega_j (t+\tau) + B_j sin\omega_j (t+\tau) \right]$$

Autocorrelation Functions – Example (cont'd)

$$E[X(t)X(t+\tau)] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(A_i cos\omega_i t + B_i sin\omega_i t) \ldots \right.$$

$$\left. \cdot (A_j cos\omega_j (t+\tau) + B_j sin\omega_j (t+\tau)\right]$$

- Since A and B are i.i.d. (and $E[A_i] = E[B_i] = 0$):

$$i \neq j : E[A_i A_j] = 0, E[A_i B_j] = 0, E[B_i A_j] = 0, E[B_i B_j] = 0$$

- We get: $E[X(t)X(t+\tau)] = \sum_{i=1}^{n}(E\left[A_i^2\right] \cdot cos\,\omega_i t \cdot cos\,\omega_i (t+\tau) \ldots$

$$+E\left[B_i^2\right] \cdot sin\,\omega_i t \cdot sin\,\omega_i (t+\tau))$$

Autocorrelation Functions – Example (cont'd)

- We can rewrite to:

$$R_{XX}(\tau) = E[X(t)X(t+\tau)]$$

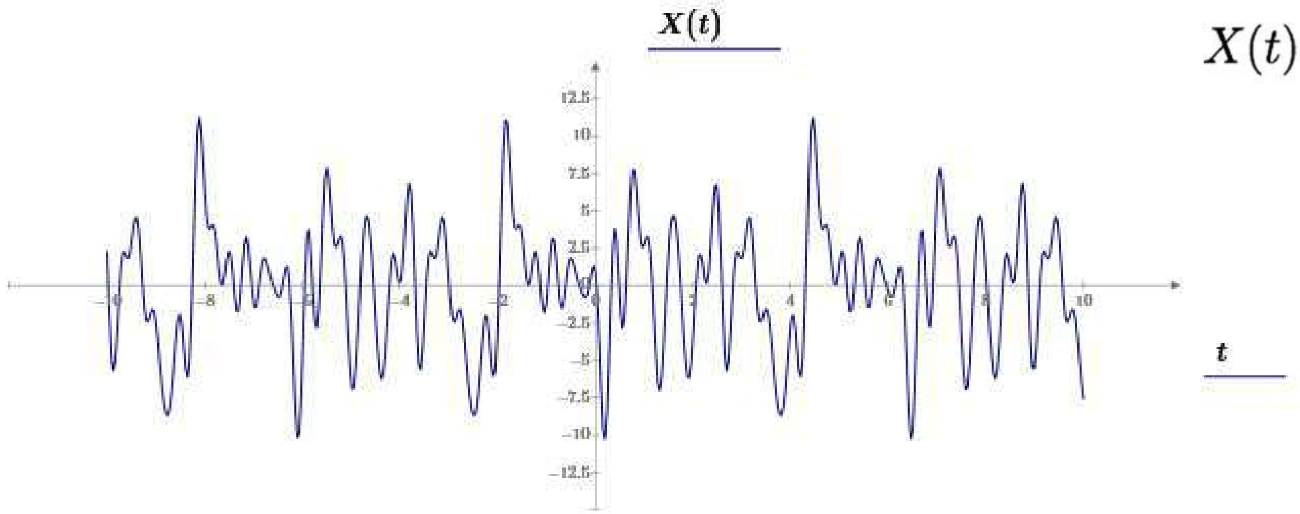
$$= \sum_{i=1}^{n} \left(E\left[A_i^2\right] \cdot cos\,\omega_i t \cdot cos\,\omega_i (t+\tau) + E\left[B_i^2\right] \cdot sin\,\omega_i t \cdot sin\,\omega_i (t+\tau)\right)$$

$$= \sigma^2 \sum_{i=1}^{n} cos\,\omega_i \tau$$

(since $E\left[A_i^2\right] = E\left[B_i^2\right] = \sigma^2$ and $\cos(\theta_1 - \theta_2) = cos\theta_1 \cdot cos\theta_2 + sin\theta_1 \cdot sin\theta_2$)

- We have: $R_{XX}(0) = n\sigma^2$

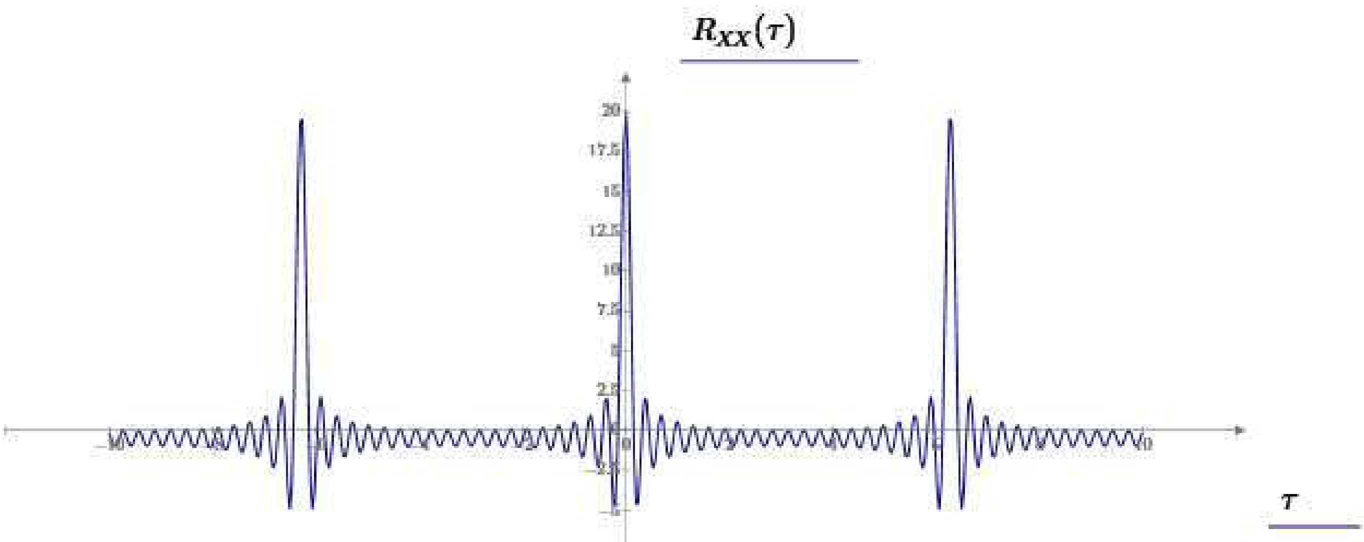
Autocorrelation Functions – Example (cont'd)



$$X(t) = \sum_{i=1}^{n} A_i \cos \omega_i t + B_i \sin \omega_i t$$

$$A_i, B_i \sim \mathcal{N}(0, \sigma^2)$$

$$\omega_i = i \cdot \omega_0$$
$$\omega_0 = 1$$



$$\sigma = 1, n = 20$$

$$R_{XX}(0) = n\sigma^2 = 20$$

Important Rules

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

- $E[aX + bY] = a \cdot E[X] + b \cdot E[Y]$ → Linearity of the mean
- $Var[aX + bY] = a^2 \cdot Var[X] + b^2 \cdot Var[Y] + 2ab \cdot Cov(X, Y)$

Correlation

- $Corr(X, Y) = E[XY]$ $(= E[X] \cdot \mathrm{E[Y]}$ if X and Y are independent)
- $Cov(X, Y) = E[(X - \bar{X})(Y - \bar{Y})] = E[XY] - E[X] \cdot E[Y]$
- $\rho = E\left[\frac{X-\bar{X}}{\sigma_X} \cdot \frac{Y-\bar{Y}}{\sigma_Y}\right] = \frac{E[XY]-E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$

Correlation coefficient

Notice that correlation and correlation coefficient are different, but can have same name and same notation!!

Tells about how much we can predict the future

Autocovariances

- Autocovariance function:

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

Especially: $C_{XX}(t, t\ ) = E[(X(t) - \mu_X(t))^2] = E[X(t)^2] - E[X(t)]^2 = \sigma_X^2(t)$

- Autocorrelation coefficient:

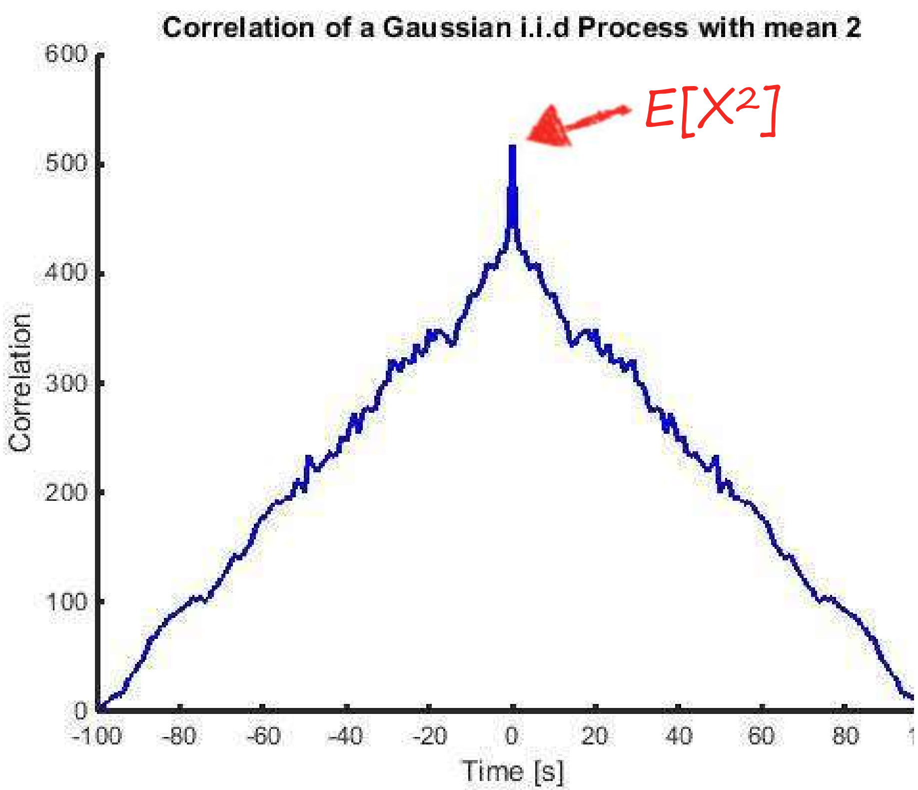
$$r_{XX}(t_1, t_2) = \frac{C_{XX}(t_1,t_2)}{\sqrt{C_{XX}(t_1,t_1)C_{XX}(t_2,t_2)}}; \qquad 0 \le r_{XX}(t_1, t_2) \le 1$$

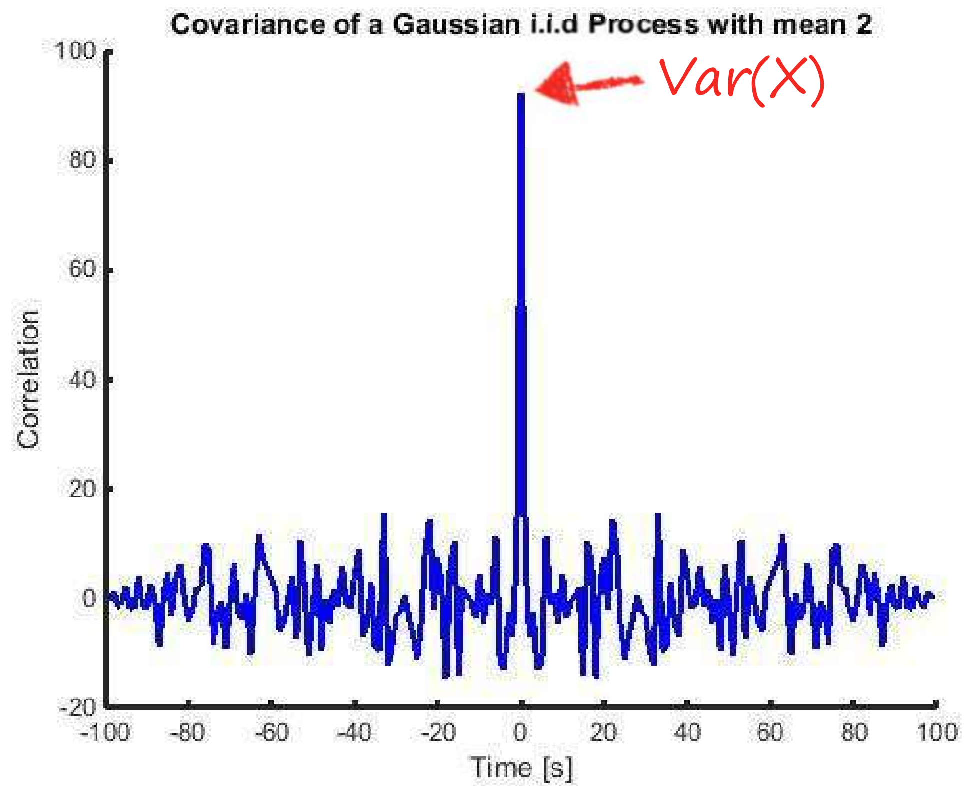
Especially: $r_{XX}(t, t) = 1$ (X(t) is totally correlated to itself!)

Autocovariances

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

- Autocorrelation and autocovariance





Two Stochastic Processes

- If we have two stochastic processes $X(t)$ and $Y(t)$
  - We can compare them by looking at the cross-correlation and cross-covariance:

Cross-correlation $R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*]$

Cross-covariance $C_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*] - E[X(t_1)]E[Y(t_2)]$

Ensemble Cross-correlation

Ensemble means that it applied for the ensemble of the two processes

- In general:

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*]$$
$$= \iint_{-\infty}^{\infty} x(t_1)\, y(t_2)^* f_{X(t_1),Y(t_2)}\big(x(t_1), y(t_2)\big) dx(t_1) dy(t_2)$$

- For two WSS stationary processes:

$$R_{XY}(t_1, t_2) = R_{XY}(t_1 + T, t_2 + T) = E[X(t_1 + T)Y(t_2 + T)^*]$$

- We write: $R_{XY}(\tau) = E[X(t) \cdot Y(t + \tau)^*]$

Cross-Correlation Functions

- For Real WSS processes $X(t)$ and $Y(t)$:

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

- Properties of the cross-correlation function $R_{XY}(\tau)$:
  - $R_{XY}(\tau) = R_{YX}(-\tau))$
  - $|R_{XY}(\tau)| \leq \sqrt{R_{XX}(0)R_{YY}(0)} = \sqrt{E[X^2]E[Y^2]}$
  - $|R_{XY}(\tau)| \leq \frac{1}{2}(R_{XX}(0) + R_{YY}\ (0))$
  - If $X(t)$ and $\mathrm{Y}(t)$ are orthogonal, then $R_{XY}(\tau) = 0$
  - If $X(t)$ and $\mathrm{Y}(t)$ are independant, then $R_{XY}(\tau) = \mu_X \cdot \mu_Y$

Temporal Cross-correlation

Temporal only looks at one realization
of the two stochastic processes.

- The temporal cross-correlation between X and Y:

Convolution

$$\mathcal{R}_{XY}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x\,(t) \cdot y\,(t+\tau) dt$$

- If the two processes are ergodic the temporal cross-correlation is equal to the ensemble cross-correlation:

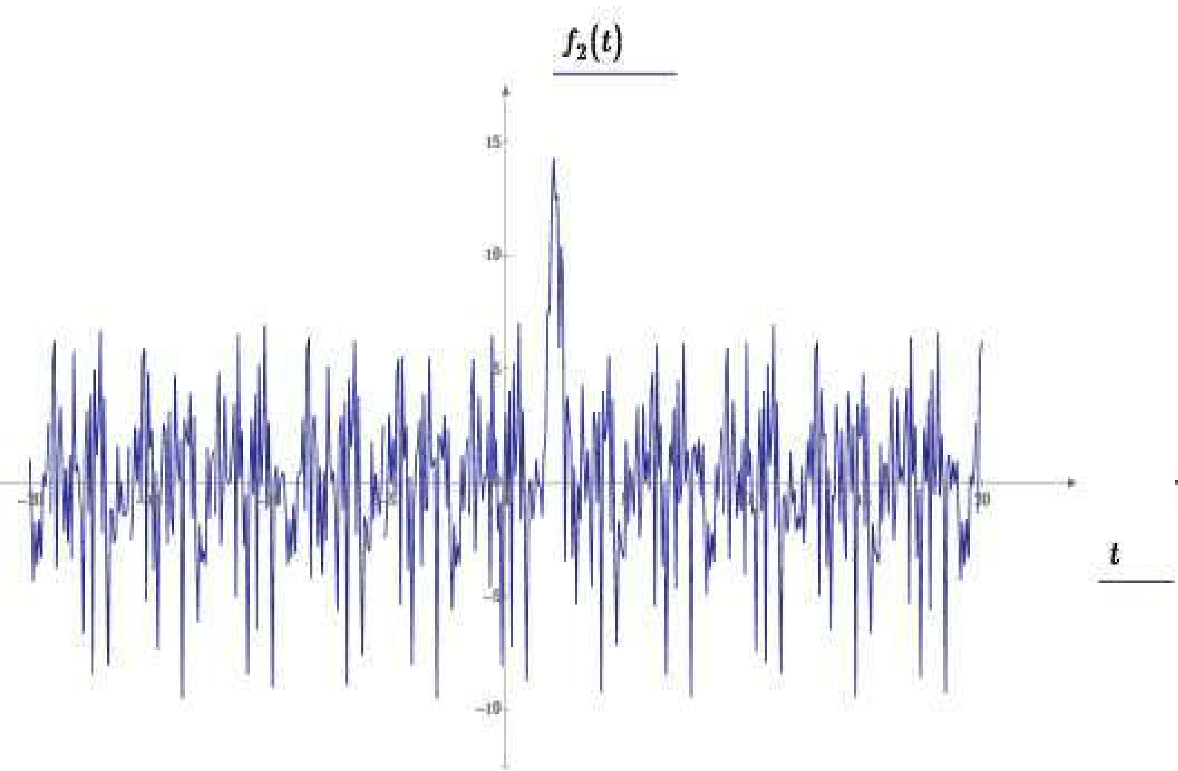
$$R_{XY}(\tau) = \mathcal{R}_{XY}(\tau)$$
$$R_{YX}(\tau) = \mathcal{R}_{YX}(\tau)$$

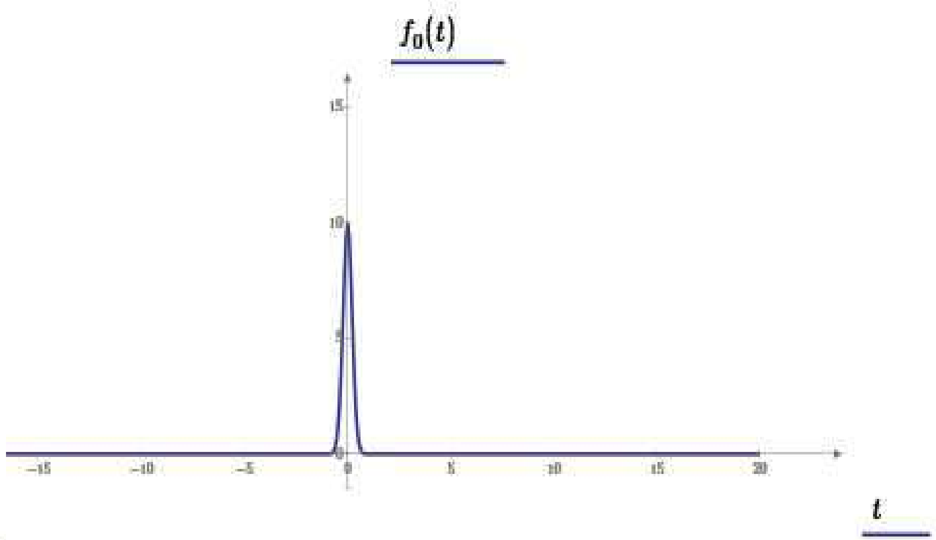
Ensemble

Temporal

Cross-correlation – Uncalibrated noisy signal

- Comparing two signals:
  - An uncalibrated and noisy signal: $f_2(t)$
  - Reference signal: $f_0(t) = 10 \cdot e^{-10t^2}$

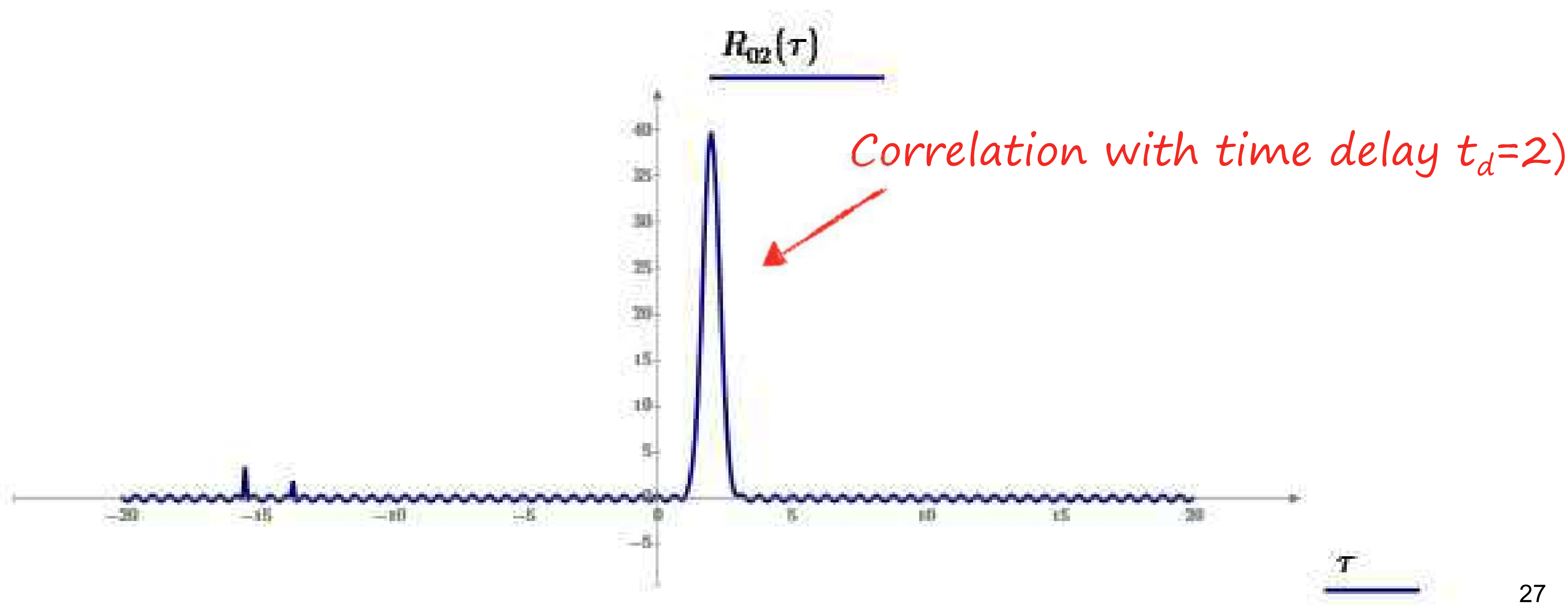




Cross-correlation – Uncalibrated noisy signal

- Comparing two signals:
  - An uncalibrated and noisy signal $f_2(t)$
  - Reference signal $f_0(t) = 10 \cdot e^{-10t^2}$

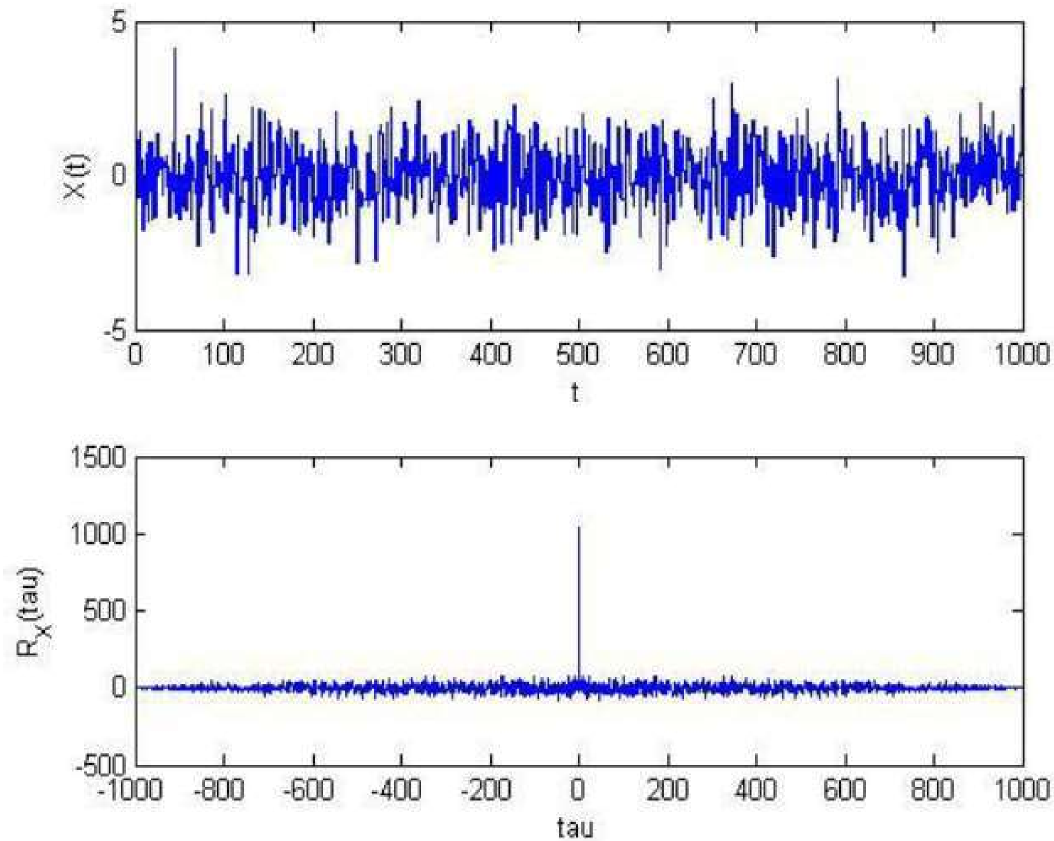
- Cross-correlation: $R_{02}(\tau) = \int_{-\infty}^{\infty} f_0(t) \cdot f_2(t+\tau)dt$



Autocorrelation of White Noise

Correlation is equal to covariance (mean=0)

```
% Autokorrelation af hvid
t = 0:999;
tau = -999:999;
x = randn(1,1000);
Rx = conv(x,fliplr(x));
figure
subplot(2,1,1)
stairs(t,x)
xlabel('t')
ylabel('X(t)')
subplot(2,1,2)
plot(tau,Rx)
xlabel('tau')
ylabel('R_X(tau)')
```



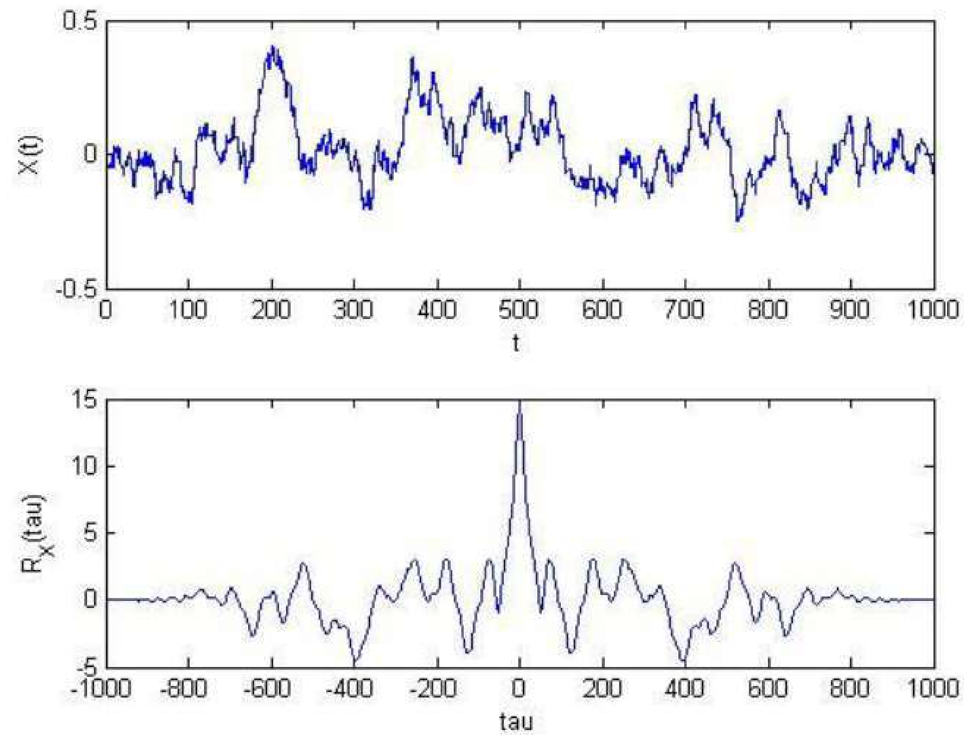
Uncorrelated for lag different from 0

Indicates independence – but not with 100% centainty

Autocorrelation of LP Filtered White Noise

Correlation is equal to covariance (mean=0)

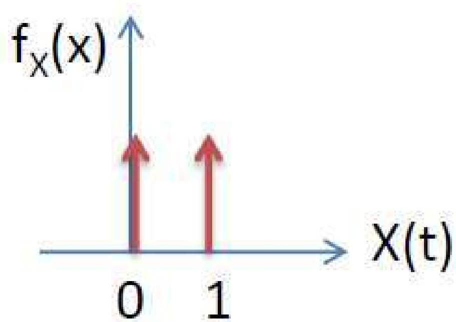
```
% Autokorrelation af
filtreret hvid støj
t = 0:999;
tau = -999:999;
x = randn(1,1000);
h = ones(1,51)/51;
x = conv(x,h,'same');
Rx = conv(x,fliplr(x));
figure
subplot(2,1,1)
stairs(t,x)
xlabel('t')
ylabel('X(t)')
subplot(2,1,2)
plot(tau,Rx)
xlabel('tau')
ylabel('R_X(tau)')
```

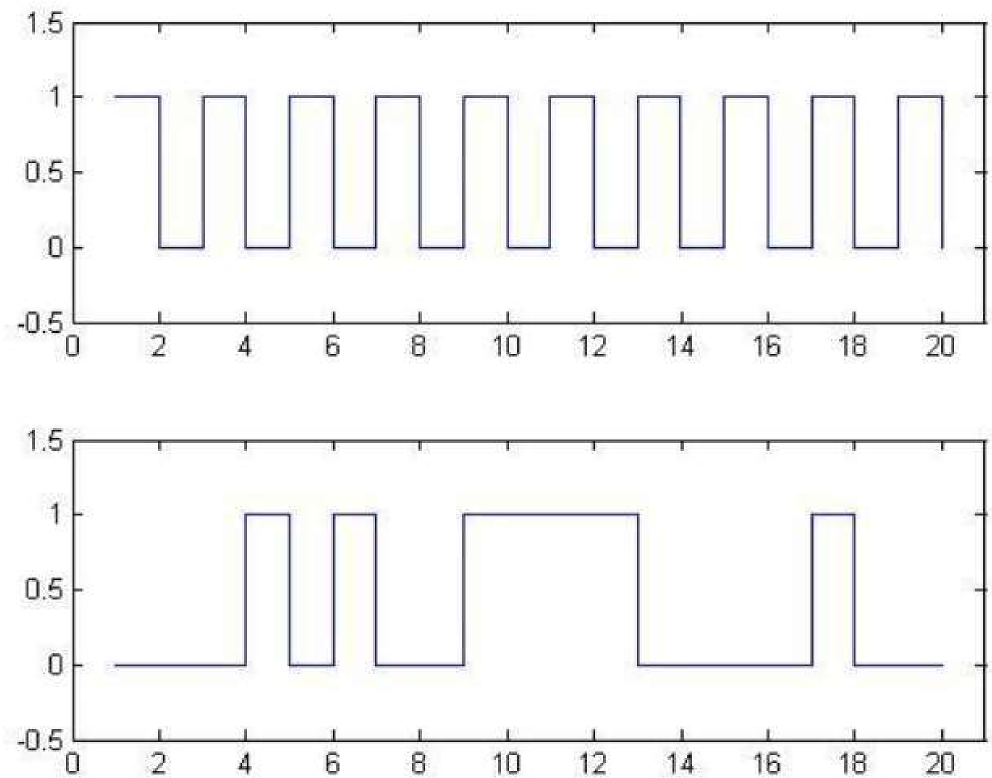


Correlated for lag different from 0

Deterministic vs. Stochastic

The probability mass function:

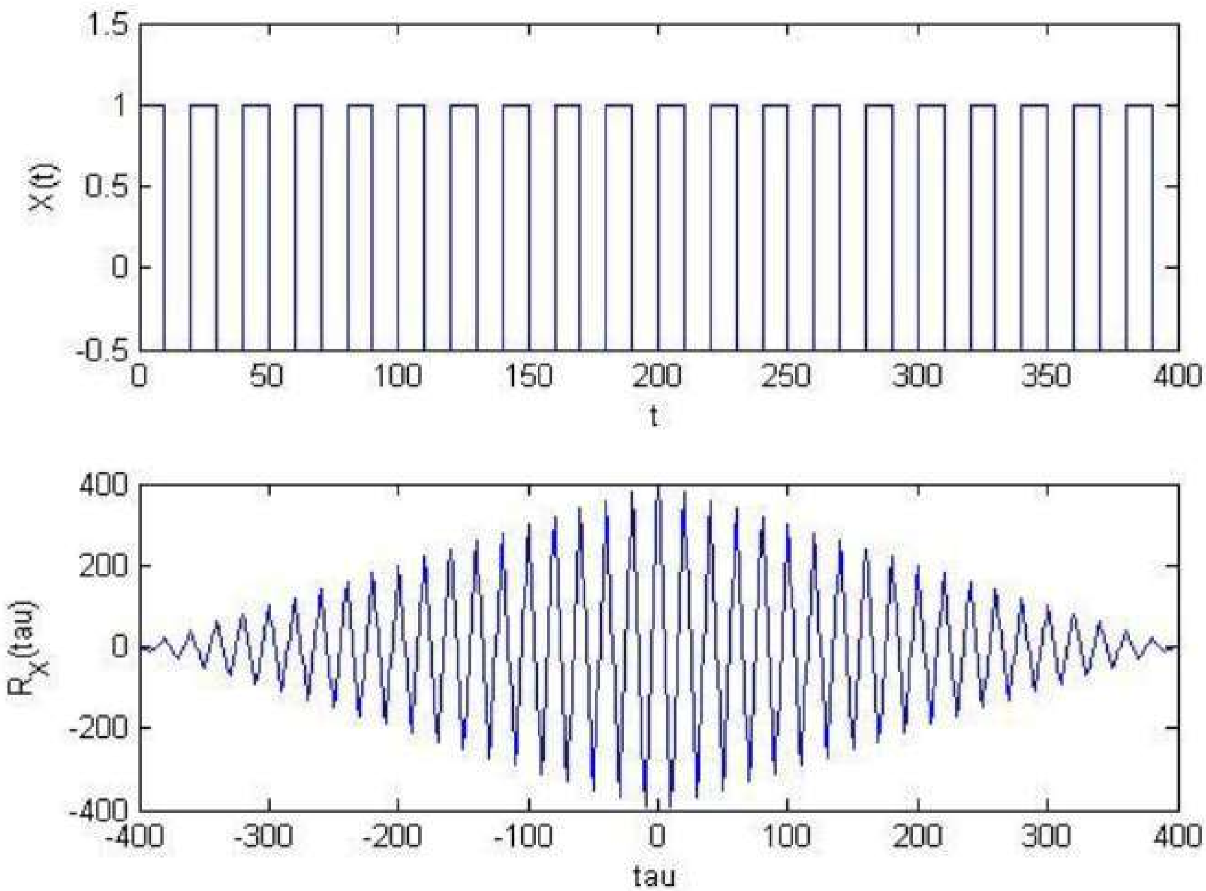




The two random processes have the same pmf.

Deterministic

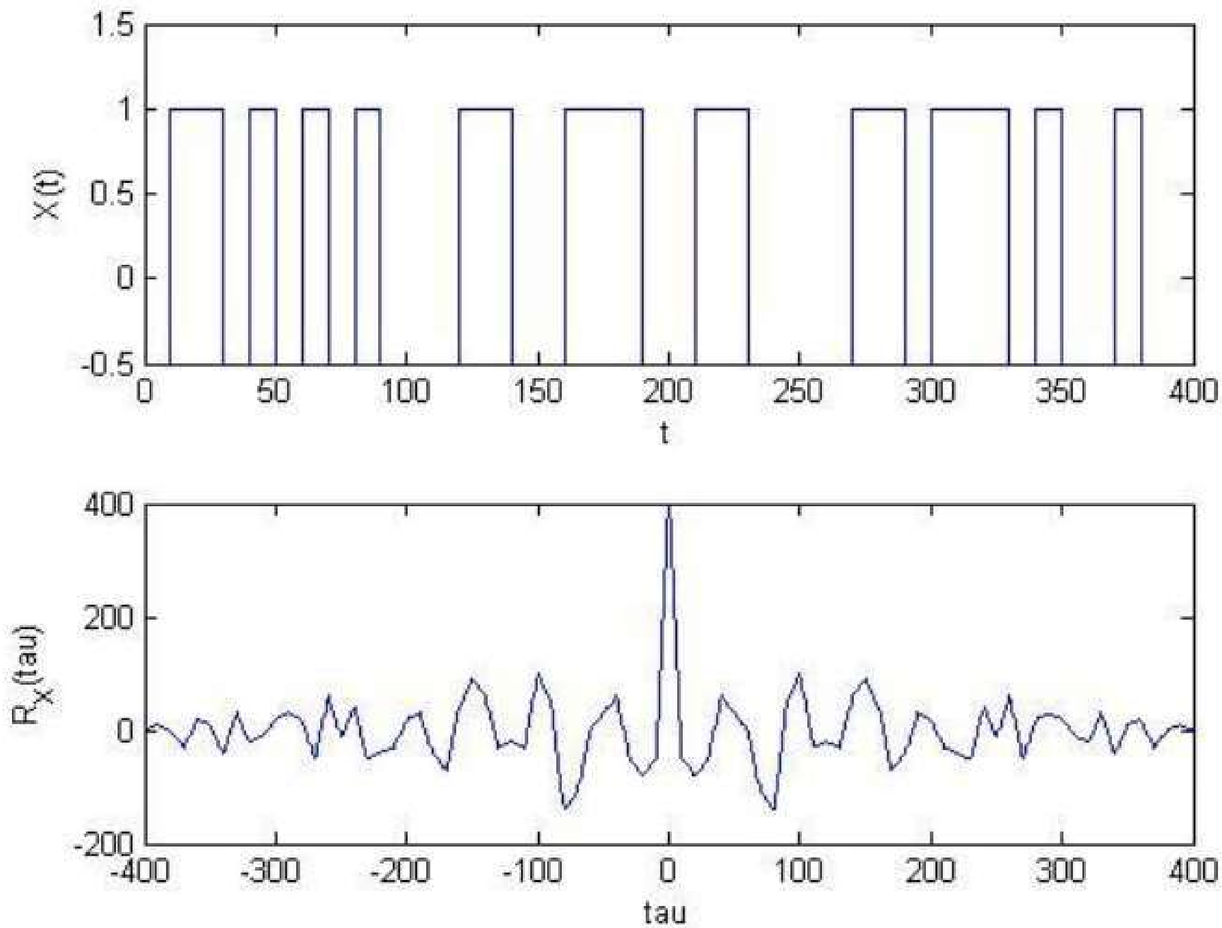
Periodic signal
↓
R_{XX} periodic




```
Rx = conv(x,fliplr(x));
```

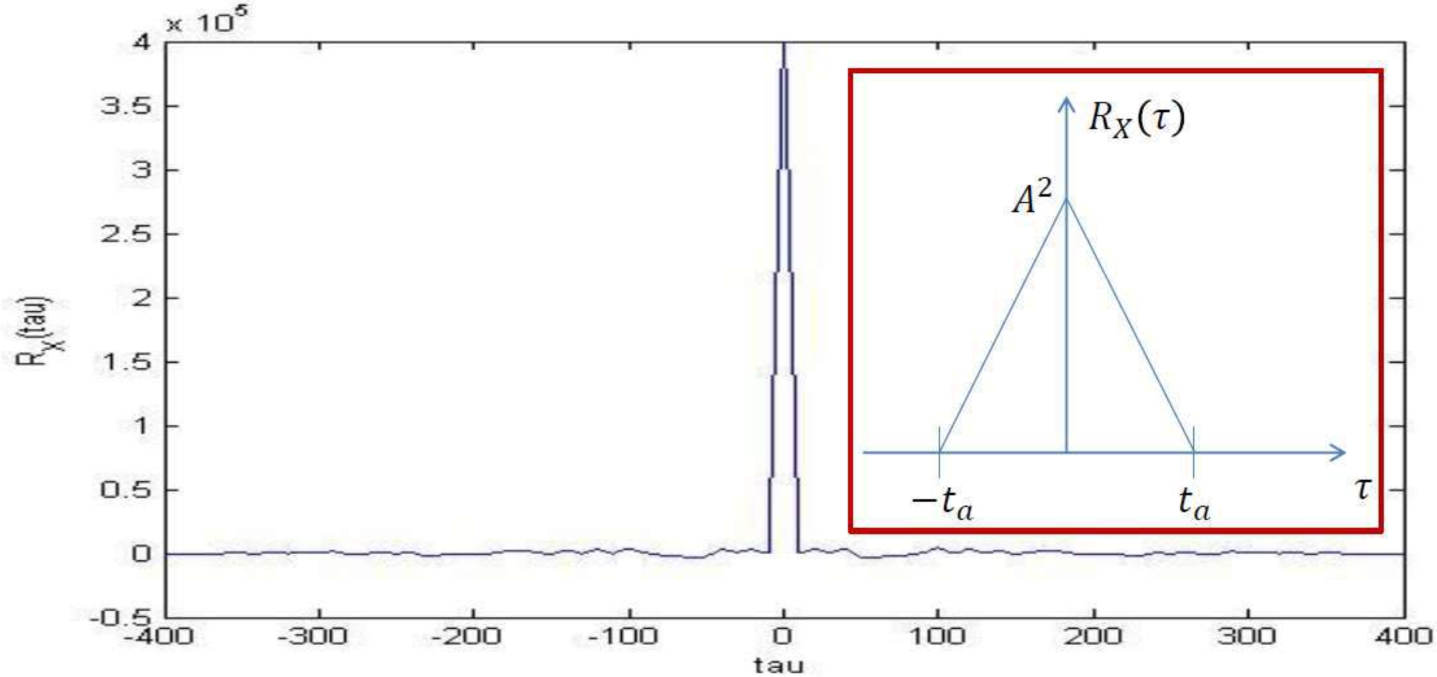
Stochastic

Also called Non-deterministic

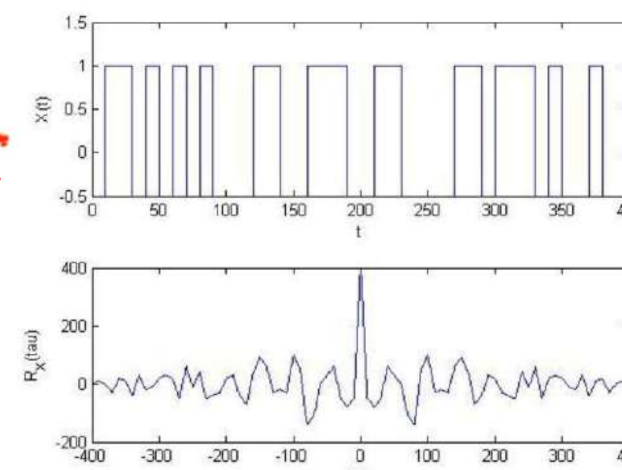



```
Rx = conv(x,fliplr(x));
```

Autocorrelation for Stochastic



Autocorrelation function averaged over
1000 simulations.



Power Spectral Density (psd)

- Frequency domain:
  - ➢ Deterministic signals f(t) → Fourier-transformation $\mathcal{F}(\mathrm{f}(\mathrm{t}))$
  - ➢ Random signals X(t) → ÷Fourier-transformation

- For Real WSS:
- Properties of the autocorrelation function $R_{XX}(\tau)$:
  - ➢ If X(t) changes fast, then $R_{XX}(\tau)$ decreases fast from $\tau = 0$
  - ➢ If X(t) changes slowly, then $R_{XX}(\tau)$ decreases slowly from $\tau = 0$
  - ➢ If X(t) is periodic, then $R_{XX}(\tau)$ is also periodic

→ $R_{XX}(\tau)$ contain information about the frequency content in X(t)

Power Spectral Density (psd)

- Deterministic signals $x(t)$:

Time-average

- Average power: $P_X = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)^2 dt$
- $x(t)$ periodic T_0: $\langle R_{XX}(\tau)\rangle_{T_0} = \frac{1}{T_0}\int_0^{T_0} x(t)x(t+\tau)dt$
- Power Spectral Density Function (psd):

$$S_{XX}(f) = \mathcal{F}(\langle R_{XX}(\tau)\rangle_{T_0}) \Rightarrow P_X = \int_{-\infty}^{\infty} S_{XX}(f)df$$

Fourier-transform

Average power in $x(t)$

Power Spectral Density (psd)

- WSS random signals X(t):

- Power Spectral Density Function (psd):

Fourier-transform

$$S_{XX}(f) = \mathcal{F}(\langle R_{XX}(\tau)\rangle_{T_0}) = \int_{-\infty}^{\infty} R_{XX}(\tau)e^{-j\cdot 2\pi f\cdot \tau}d\tau$$

Invers Fourier-transform

$$\Rightarrow R_{XX}(\tau) = \mathcal{F}^{-1}(\langle R_{XX}(\tau)\rangle = \int_{-\infty}^{\infty} S_{XX}(f)e^{j\cdot 2\pi f\cdot \tau}df$$

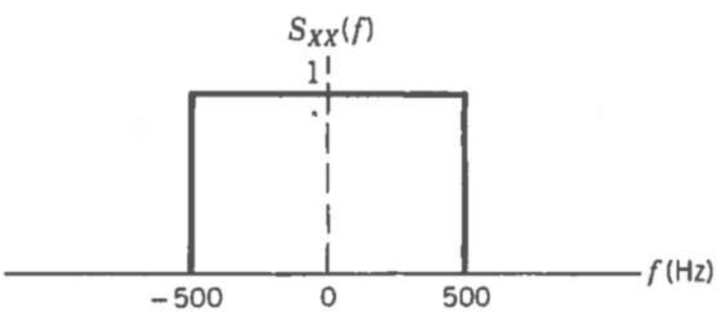


Figure 3.19a Psd of a lowpass random process $X(t)$.

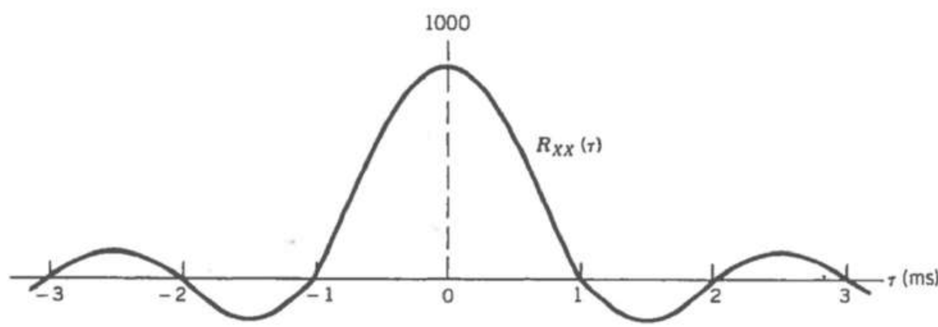


Figure 3.19b Autocorrelation function of $X(t)$.

Figures from "Random Signals"

Power Spectral Density (psd)

- Properties of psd $S_{XX}(f)$ (spectrum of $X(t)$):
  - $S_{XX}(f) \in \mathbb{R}$
  - $S_{XX}(f) \geq 0$
  - If $X(t) \in \mathbb{R}$: $R_{XX}(-\tau) = R_{XX}(f)$ and $S_{XX}(-f) = S_{XX}(f) \rightarrow$ even functions
  - If $X(t)$ periodic components: $S_{XX}(f)$ will have impulses (δ−functions)
  - $[S_{XX}(f)] = \frac{W}{Hz} \rightarrow$ Distribution of power with frequency (power spectral density of the stationary random process X(t)
  - $P_X = E[X(t)^2] = R_{XX}(0) = \int_{-\infty}^{\infty} S_{XX}(f)df$

    i.e. if $X(t) = V(t)$ (voltage signal)

    $\rightarrow P_X =$ power in 1Ω-resistor

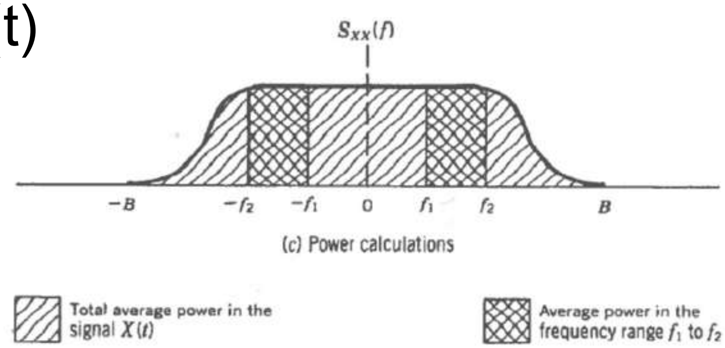


Figure from ”Random Signals”

  - $P_X[f_1, f_2] = 2\int_{f_1}^{f_2} S_{XX}(f)df \rightarrow$ Power in the frequency-interval $[f_1, f_2]$

Power Spectral Density – Random Binary Signal

Figures from ”Random Signals”

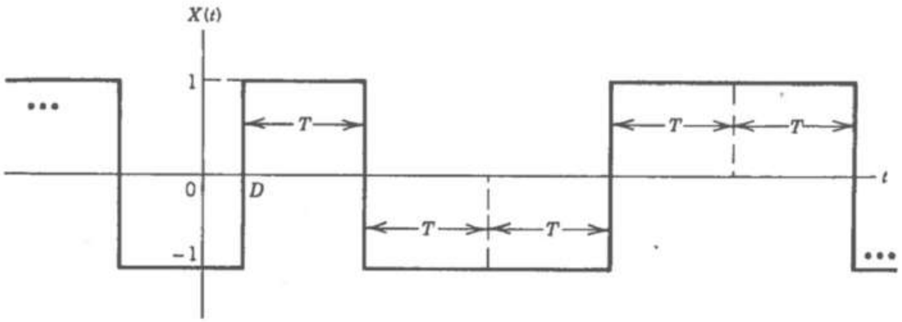


Figure 3.7 Random binary waveform.

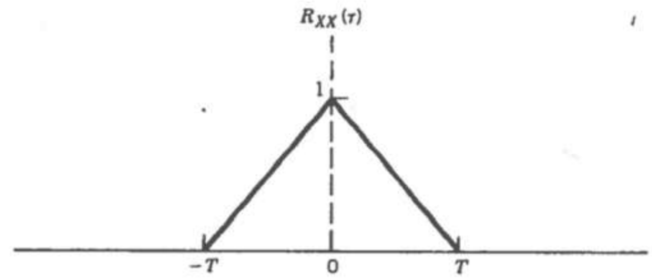


Figure 3.18a Autocorrelation function of the random binary waveform.

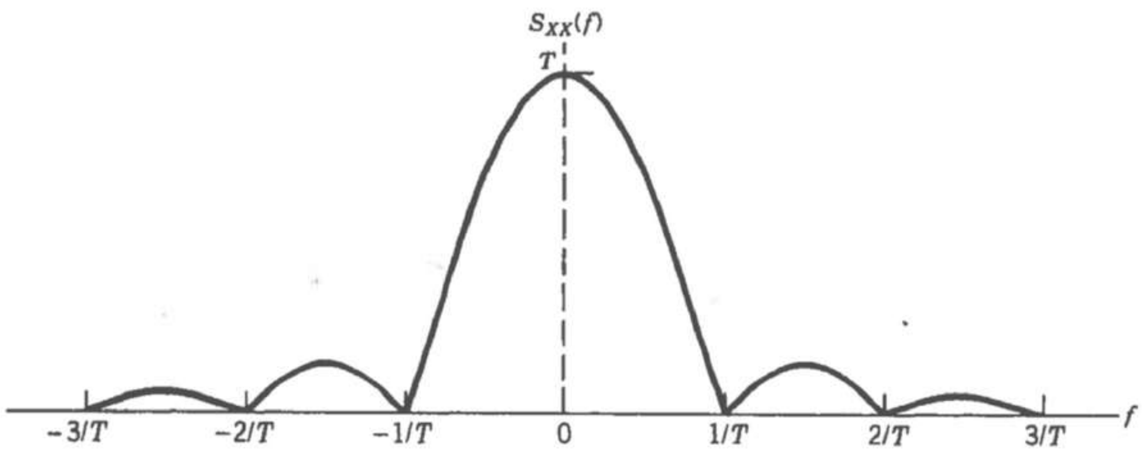


Figure 3.18b Power spectral density function of the random binary waveform.

Words and Concepts to Know

Cross-correlation

Power Spectral Density

Deterministic

Cross-covariance

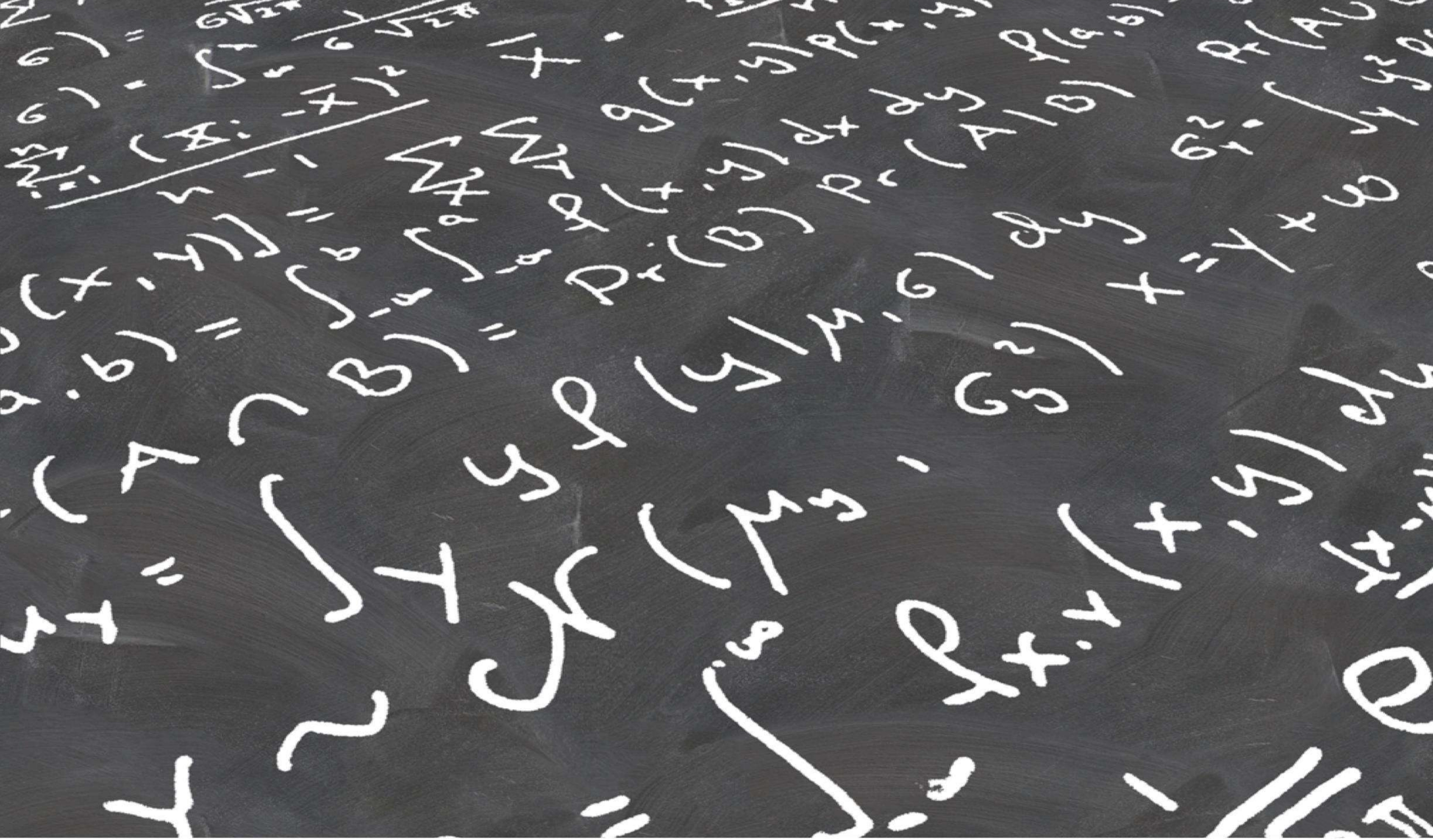
psd

Temporal Autocovariance

Autocorrelation Coefficient

Temporal cross-correlation

Non-deterministic



8. Resume of Probability and Stochastic Processes

Gunvor Elisabeth Kirkelund
Lars Mandrup

Agenda for Today

Resume of stochastic processes:

- Probability
  - Bayes rule
  - Conditional
  - Total
- Stochastic variables
  - pmf/pdf/cdf
  - Joint/marginal/conditional
  - Mean/Variance/Correlation
- Stochastic Processes
  - Ensemble/Sample functions
  - Stationarity and Ergodic Processes
  - Auto- and Cross-correlation functions
  - Power Spectrum Density

Basic Probability

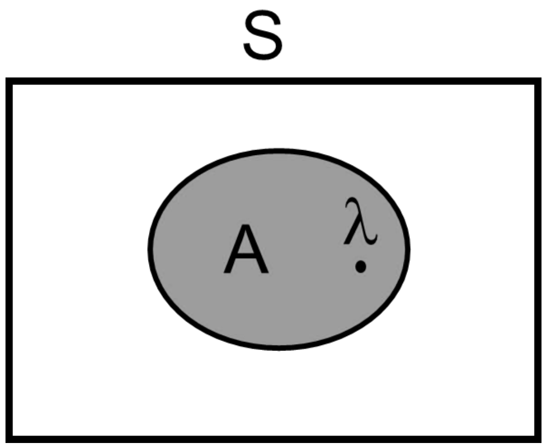
- Probability theory tells us what is in the sample given nature.

- Basic Axions:

$$\text{Axion 1: } \boldsymbol{0 \le Pr(A) \le 1}$$

$$\text{Axion 2: } \boldsymbol{Pr(S) = 1}$$

S: Sample space
A: Event
λ: Sample point

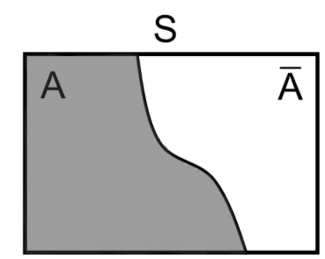


- Often (but not always) we use the relative frequency:

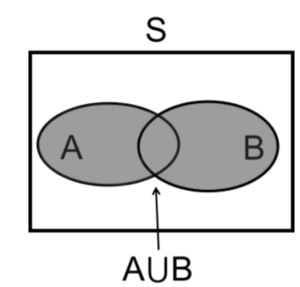
$$\Pr(A) = \frac{N_A}{N}$$

Basic Probability

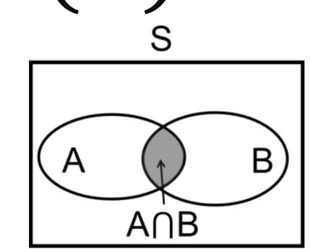
- Complement: $Pr(A) = 1 - Pr(\bar{A})$



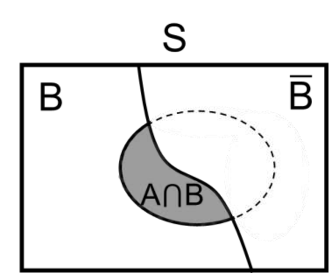
- Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$



- Joint: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$



- Conditional: $Pr(A|B)$



Bayes Rule and Independence

- Bayes Rule:

$$\boxed{Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B)}}$$

- A and B independent:

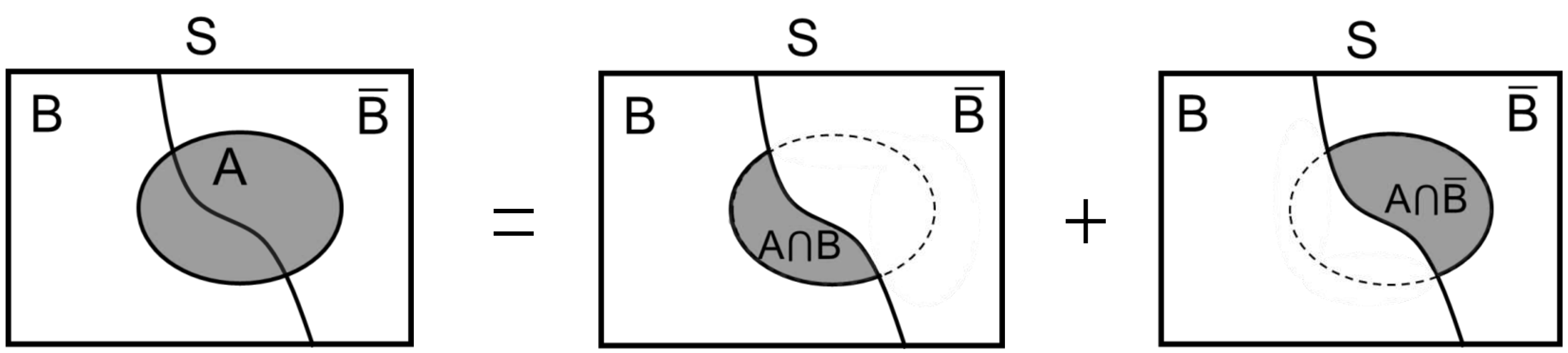
$$\boxed{Pr(A \cap B) = Pr(A) \cdot Pr(B)}$$

$$\boxed{Pr(B|A) = Pr(B) \quad \text{and} \quad Pr(A|B) = Pr(A)}$$

Total Probability

We sometime call it the marginal

- Pr(A) of an event is the total probability of that event.

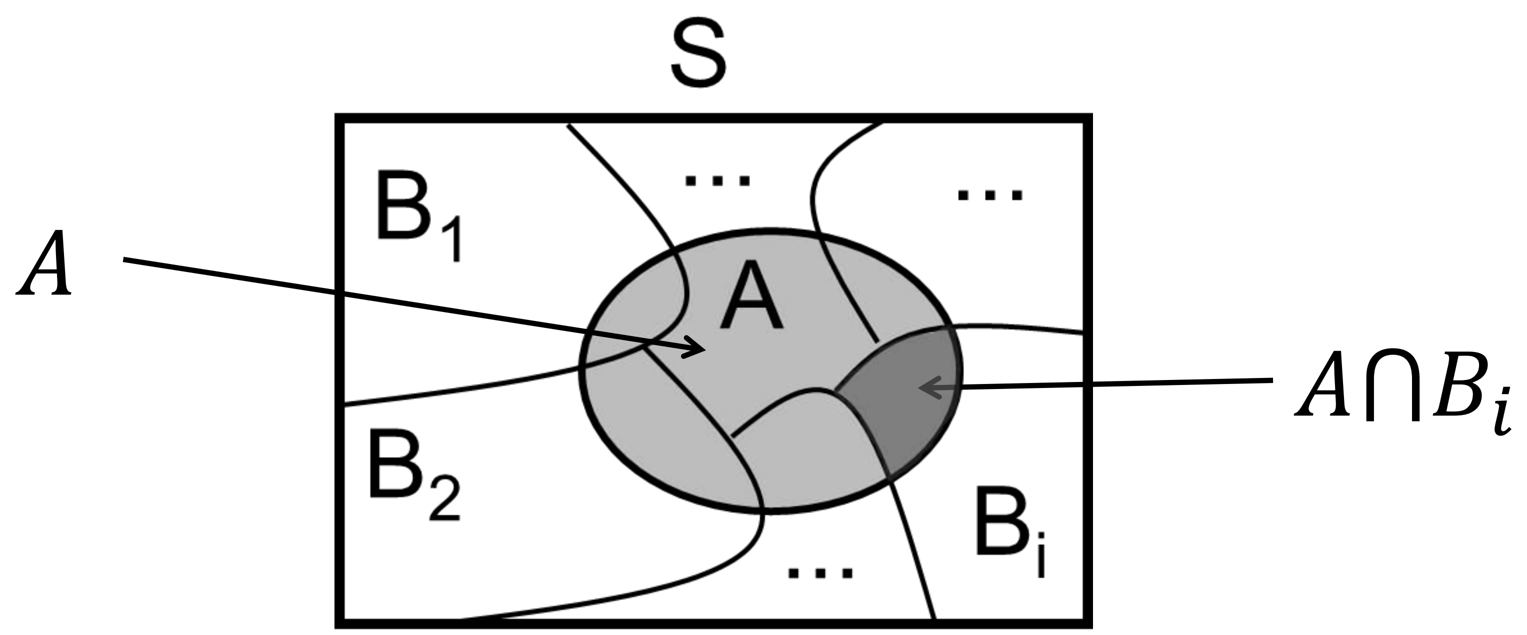


$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap \bar{B})$$
$$= \Pr(A|B) \cdot \Pr(B) + \Pr(A|\bar{B}) \cdot \Pr(\bar{B})$$

Total Probability

We sometime call it the marginal

- Pr(A) of an event is the total probability of that event.

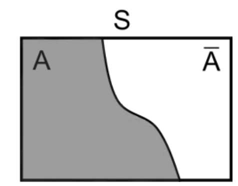


$$Pr(A) = Pr(A \cap B_1) + Pr(A \cap B_2) + \cdots + Pr(A \cap B_i) + \cdots$$
$$= Pr(A|B_1) \cdot Pr(B_1) + Pr(A|B_2) \cdot Pr(B_2) + \cdots$$

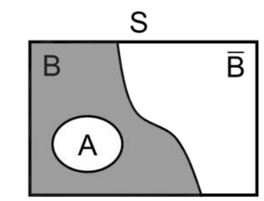
where the B_i's are mutually exclusive ($B_i \cap B_j = \varnothing$ for $i \neq j$)
and $S = B_1 \cup B_2 \cup \ldots \cup B_i \cup \ldots$

Summary of Probability

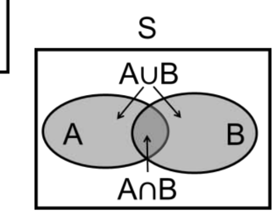
Relative frequency: $Pr(A) = \frac{N_A}{N_S}$



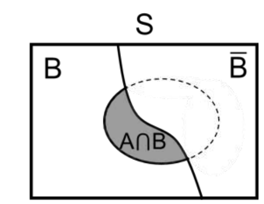
Complement: $Pr(\bar{A}) = 1 - Pr(A)$



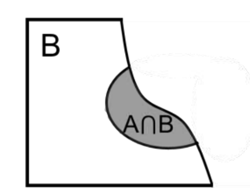
Exclusive: $Pr(\bar{A} \cap B) = Pr(B) - Pr(A)$ if $A \subset B$



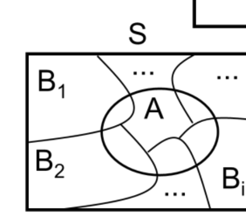
Union: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$



Joint: $Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$



Conditional: $Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}$ if $Pr(B) \neq 0$



Total probability: $Pr(A) = \sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)$

Bayes rule: $Pr(B|A) = \frac{Pr(A|B) \cdot Pr(B)}{Pr(A)}$

Bayes formula: $Pr(B_i|A) = \frac{Pr(A|B_i) \cdot Pr(B_i)}{\sum_{i=1}^{n} Pr(A|B_i) \cdot Pr(B_i)}$

Independence: $Pr(A \cap B) = Pr(A) \cdot Pr(B)$

Combinatorics

- The number of possible outcomes of k trials, sampled from a set of n objects.

Types of Experiments:

- With or without replacement
- Ordered or unordered

		Replacement	
		With	Without
Sampling	Ordered	n^k	$P_k^n = \dfrac{n!}{(n-k)!}$
	Unordered	$\binom{n+k-1}{k} = \dfrac{(n+k-1)!}{k!\,(n-1)!}$	$\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$

The Binomial Distribution

- We have n repeated trials.
- Each trial has two possible outcomes
  - **Success** — probability p
  - **Failure** — probability q=1-p

Bernoulli trial

- What is the probability of having k successes out of n trials?
- We write this question as:

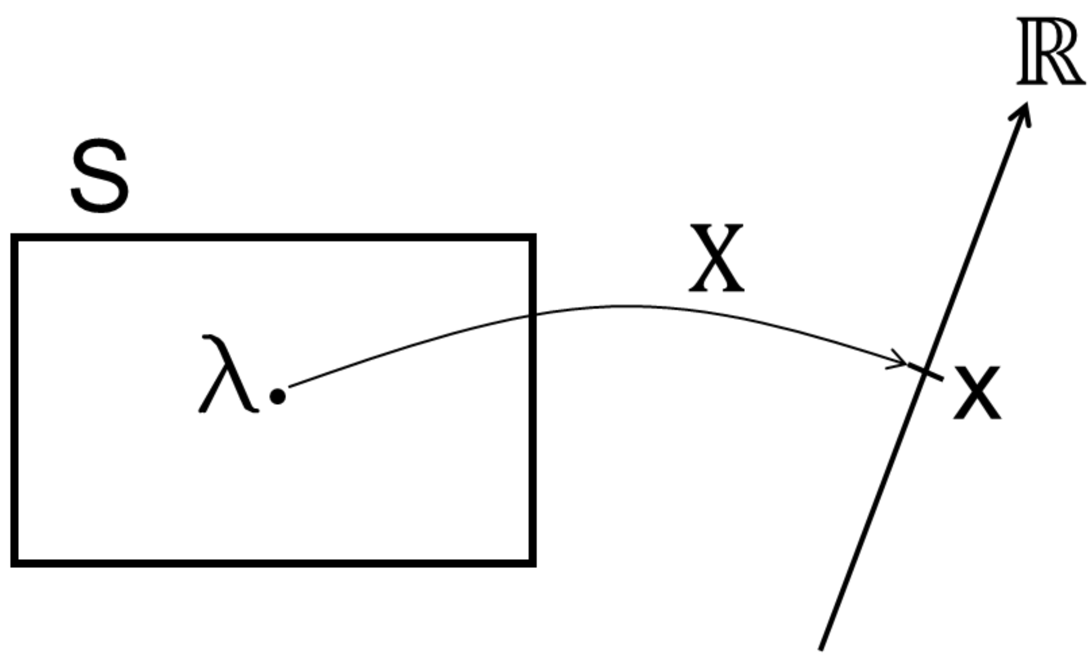
$$\boxed{Pr_n(k) = \frac{n!}{k!\,(n-k)!} p^k q^{n-k} = \binom{n}{k} p^k q^{n-k}}$$

- Faculty: $n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$
  $0! = 1$

Also just called a random variables

Stochastic Random Variables

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



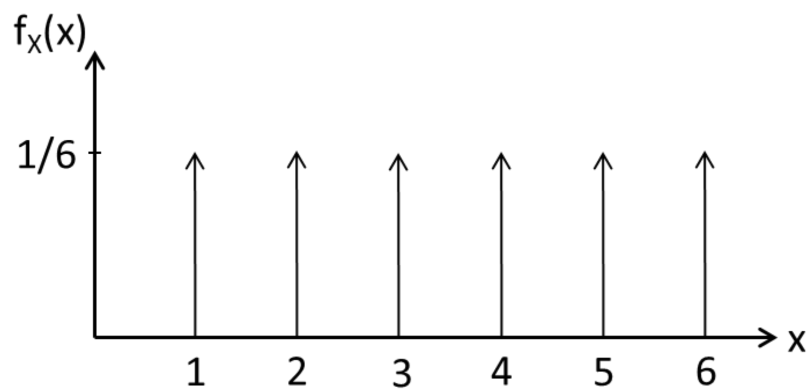
Examples:

- The numbers on a dice (discrete):
  - Sample space for variable X is : $\{1, 2, 3, 4, 5, 6\}$
  - Sample space for variable Y "Even (1)/Uneven (-1)": $\{1, -1\}$
- The hight of students at IHA (continous):
  - Sample space for variable H is all real numbers: [100;250] cm.

One Stochastic Variable – Discrete

- Probability mass function (pmf):

$$f_X(x) = \begin{cases} Pr(X = x_i) & for\ X = x_i \\ 0 & otherwise \end{cases}$$

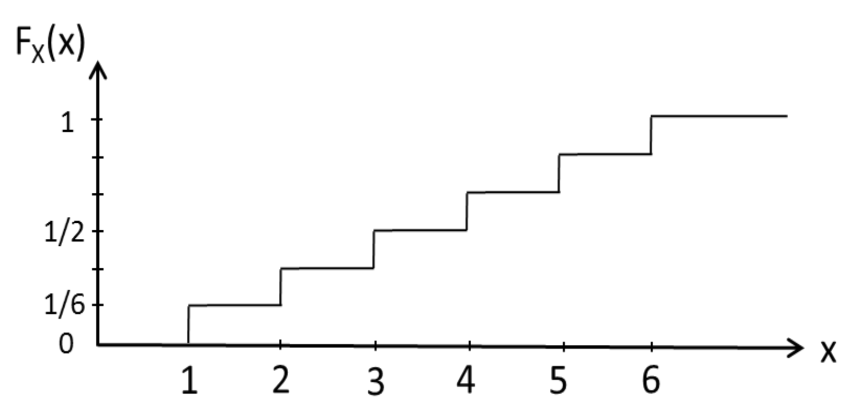


$$0 \le f_X(x) \le 1$$

$$\sum_{i=1}^{n} f_X(x_i) = \sum_{i=1}^{n} Pr(X = x_i) = 1$$

- Cumulative distribution function (cdf):

$$F_X(x) = P\,r(X \le x) = \sum_{i=1}^{n_x} f_X(x_i)$$



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

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

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

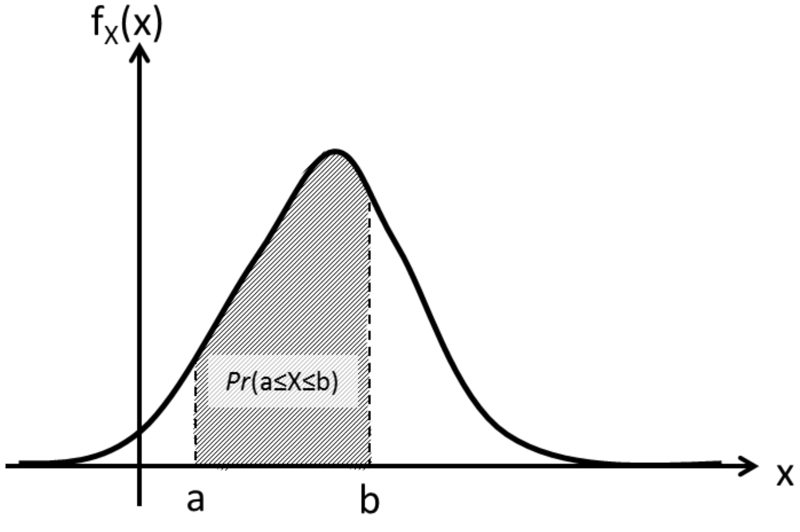
One Stochastic Variable – Continuous

- Probability density function (pdf):

$$Pr(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

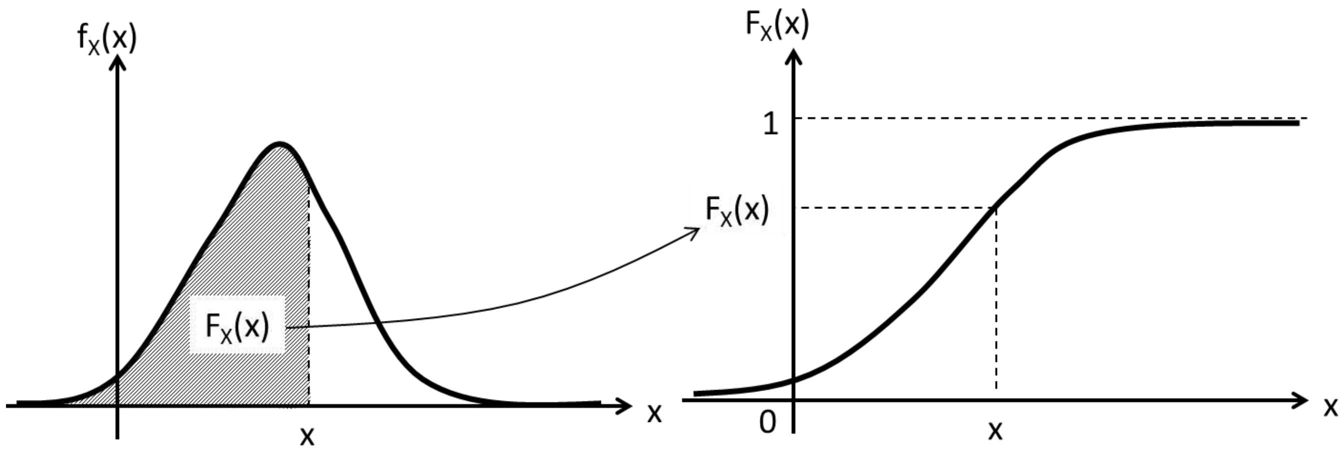
$$f_X(x) \geq 0$$

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$



- Cumulative distribution function (cdf):

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



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

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

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

Transformation of Variable X to Y

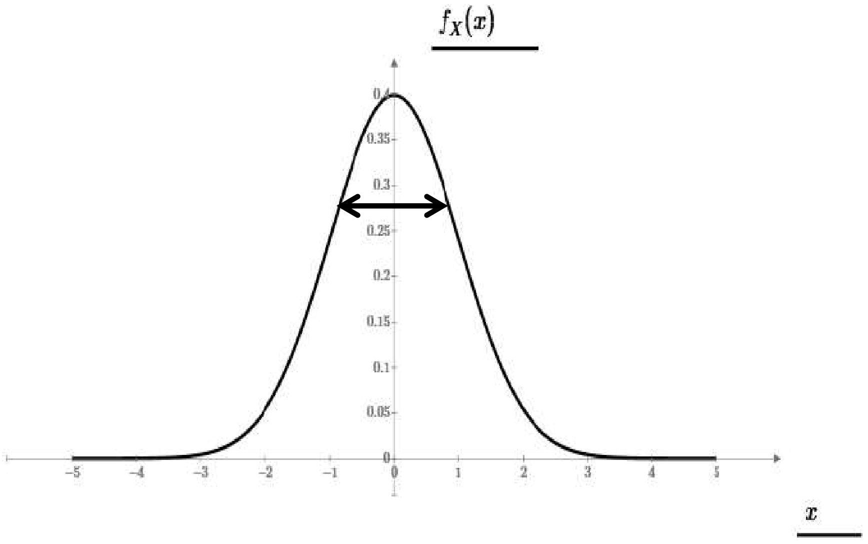
- Given:
  - Pdf: $f_X(x)$
  - Function/Transformation: $\mathrm{Y} = g(X)$
  - Limits: a $\leq X \leq b$

- 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 a $\leq X \leq b$
  4. New pdf: $f_Y(y) = \sum \left|\frac{dx(y)}{dy}\right| f_X\left(g^{-1}(y)\ \right) = \sum \frac{f_X(x)}{\left|\frac{dy}{dx}\right|}$

Expectations

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

- Variance: $Var(\mathrm{X}) = \sigma_X{}^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 \cdot f_X(x)dx = E[X^2] - E[X]^2$



- Standard deviation: $\sigma_X = \sqrt{Var(\mathrm{X})}$

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

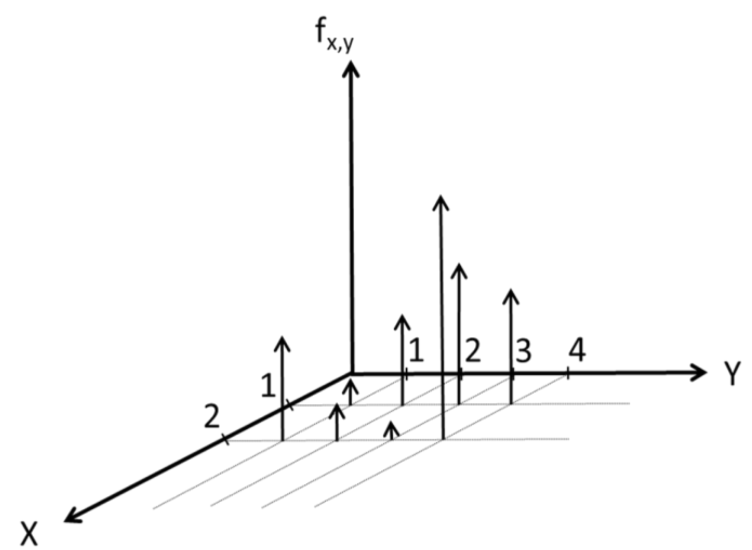
$$Var[aX + b\,] = a^2(E[X^2] - E[X]^2) = a^2 \cdot Var(X)$$

Two Stochastic Variables X, Y – Discrete

Joint (Simultaneous) pmf:

$$f_{X,Y}(x,y) = \begin{cases} Pr\Big((X = x_i) \cap (Y = y_j)\Big) & for\ X = x_i \wedge Y = y_j \\ 0 & otherwise \end{cases}$$

$$0 \leq f_{X,Y}(x,y) \leq 1 \qquad \sum_{i=1}^{m}\sum_{j=1}^{n} f_{X,Y}(x_i, x_j) = 1$$



Marginal pmfs:

$$f_X(x) = \sum_{y} f_{X,Y}(x,y) \qquad f_Y(y) = \sum_{x} f_{X,Y}(x,y)$$

Cumulative Distribution Function cdf:

$$F_X(x_j) = Pr(X \leq x_j) = \sum_{i=1}^{j} f_X(x_j)$$

Two Stochastic Variables X, Y – Continuous

Joint (Simultaneous) pdf:

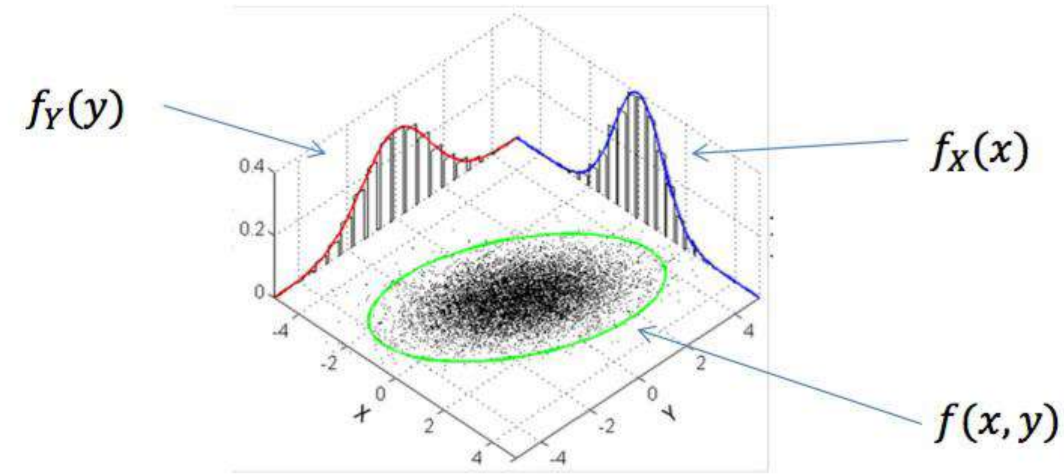
$$f_{X,Y}(x,y) \geq 0$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

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

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

Bayes Rule, Conditional PDF and Independence

Bayes rule:

- The joint/simultaneous pmf/pdf for two stochastic variables:

$$f_{X,Y}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)$$

Conditional pdf:

- For a two dimensional pmf/pdf $f_{X,Y}(x,y)$, we can find the conditional pdf with Bayes rule:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Independence:

- X and Y are independent if and only if:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \quad \text{or} \quad f_{X|Y}(x|y) = f_X(x) \quad \text{for all x and y}$$

Correlation and Covariance

Correlation tells of the (biased) coupling between variables

- Correlation: $$corr(X,Y) = E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot y \cdot f_{X,Y}(x,y)dxdy$$

Covariance is without bias from the mean

- Covariance: $$cov(X,Y) = E[(X-\bar{X})(Y-\bar{Y})] = E[XY] - E[X] \cdot E[Y]$$

Correlation Coefficient is the normalized Covariance

- Correlation coefficient: $$\rho = E\left[\frac{X-\bar{X}}{\sigma_X} \cdot \frac{Y-\bar{Y}}{\sigma_Y}\right] = \frac{E[XY] - E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$$
$$-1 \leq \rho \leq 1$$

- If X and Y are independent:
$$E[XY] = E[X] \cdot E[Y] \quad \text{and} \quad cov(X,Y) = \rho = 0$$

Important Rules

- $E[aX + b\,] = a \cdot E[X] + b$
- $Var[aX + b\,] = a^2 \cdot Var(X)$
- $E[aX + bY] = a \cdot E[X] + b \cdot E[Y]$ → Linearity of the mean
- $Var[aX + bY] = a^2 \cdot Var[X] + b^2 \cdot Var[Y] + 2ab \cdot Cov(X, Y)$

Correlation

- $Corr(X, Y) = E[XY]$ $\quad (= E[X] \cdot \mathrm{E}[\mathrm{Y}]$ if X and Y are independent)
- $Cov(X, Y) = E[(X - \bar{X})(Y - \bar{Y})] = E[XY] - E[X] \cdot E[Y]$
- $\rho = E\left[\frac{X - \bar{X}}{\sigma_X} \cdot \frac{Y - \bar{Y}}{\sigma_Y}\right] = \frac{E[XY] - E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$

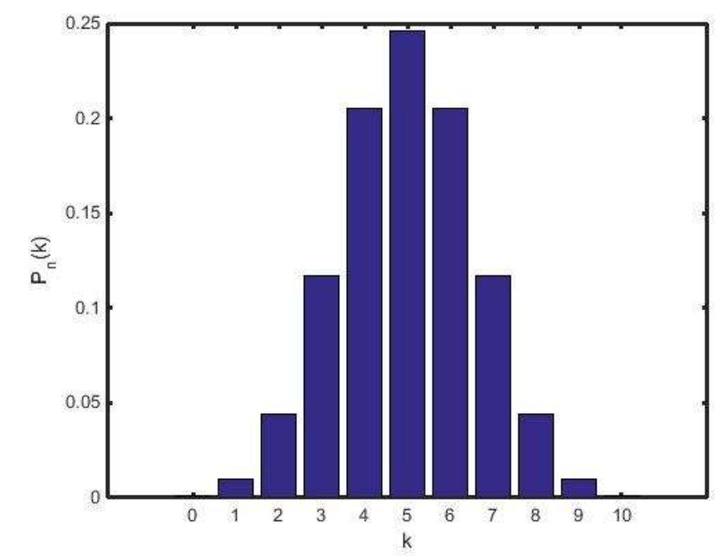
Correlation coefficient

Notice that correlation and correlation coefficient are different, but can have same name and same notation!!

The Binomial Distribution

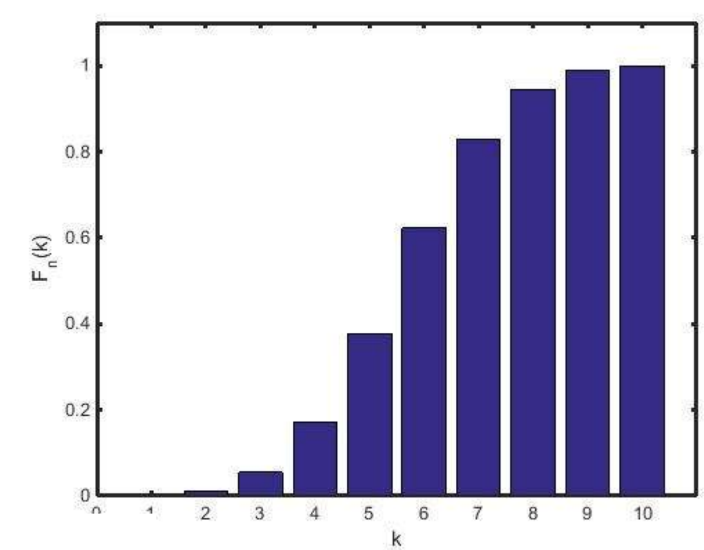
Also called a Bernoulli trial

- n repeated trials – each with two possible outcomes
  - **Success** — probability p
  - **Failure** — probability 1-p



- Probability mass function (pmf):

$$f(k|n,p) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$



- Cumulative distribution function (cdf):

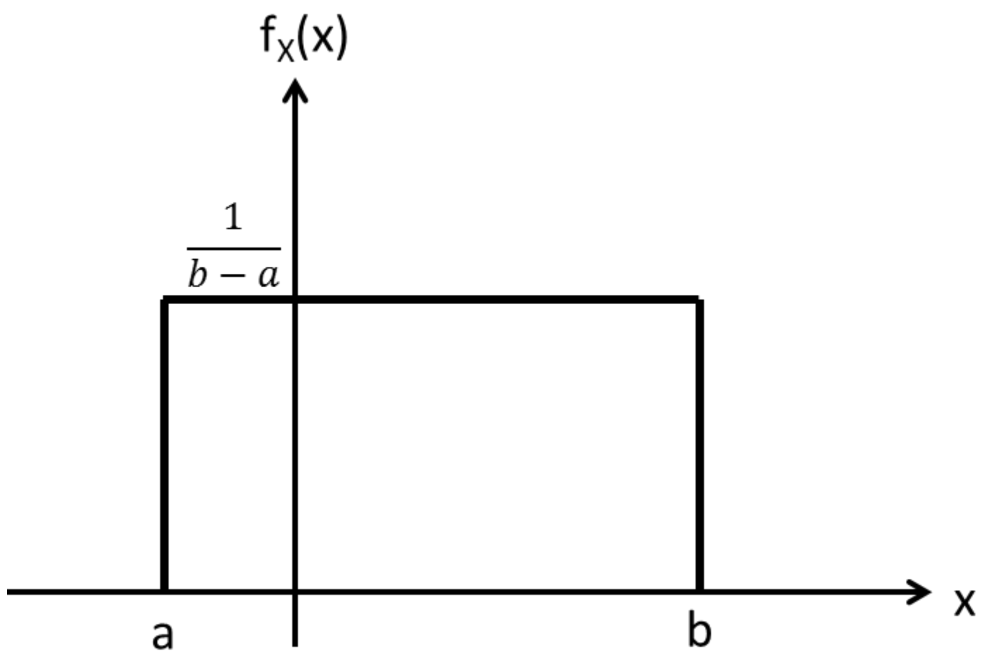
$$F(k|n,p) = \sum_{i=0}^{k} f(i|n,p)$$

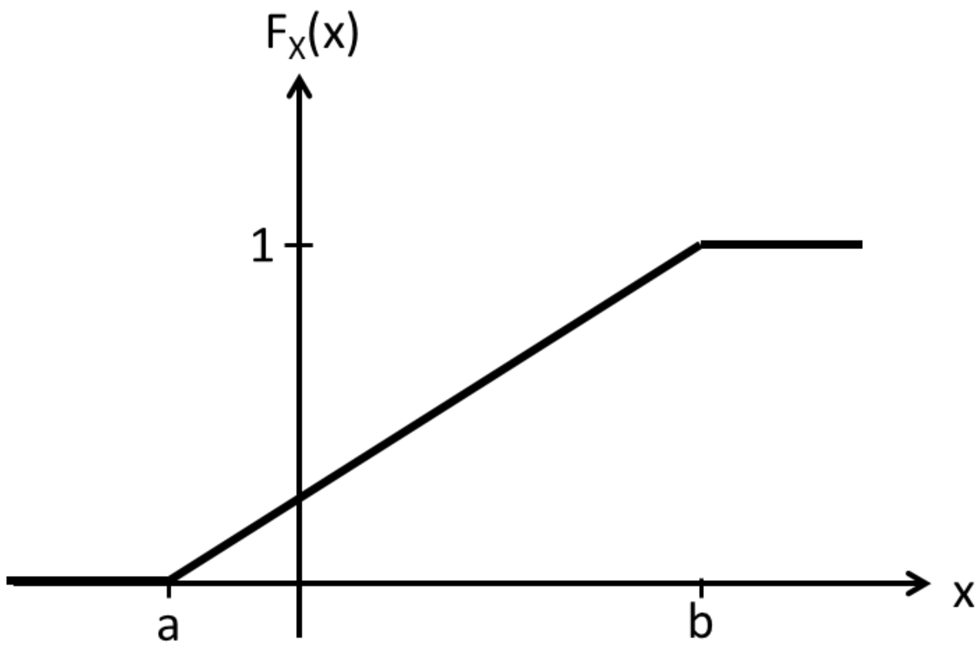
- Mean and variance:

$$E[k] = n \cdot p$$
$$Var(X) = n \cdot p \cdot (1-p)$$

Uniform Distribution

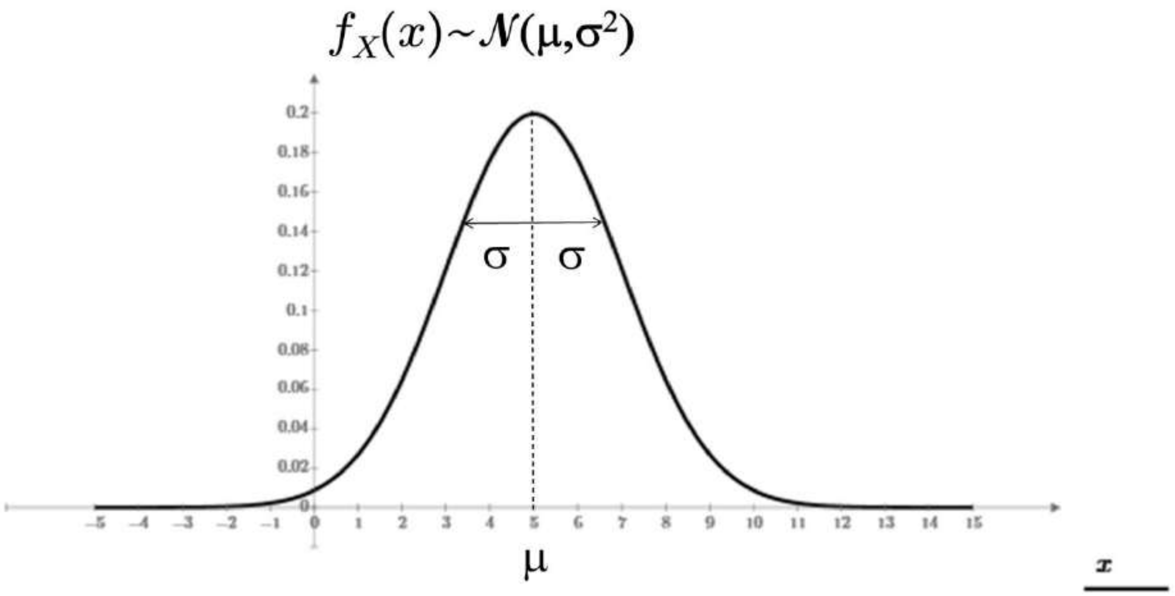
- $\mathcal{U}$(a,b)
- Mean value: $\mu = \frac{a+b}{2}$
- Variance: $\sigma^2 = \frac{1}{12}(b-a)^2$
- pdf: $f_X(x) = \begin{cases} \frac{1}{b-a} & for\ a \leq x \leq b \\ 0 & otherwise \end{cases}$
- cdf: $F_X(x) = \begin{cases} 0 & for\ x \leq a \\ \frac{x-a}{b-a} & for\ a \leq x \leq b \\ 1 & for\ x \geq b \end{cases}$



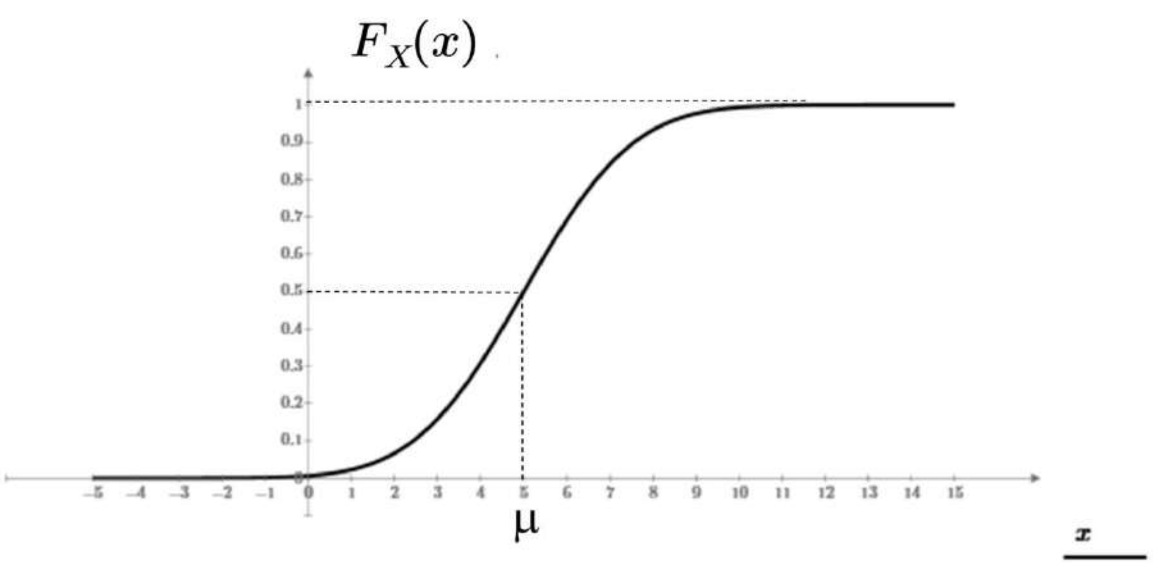


Gaussian Distribution = Normal Distribution

- $\mathcal{N}(\mu,\sigma^2)$
- Mean value: μ
- Variance: σ^2



- pdf: $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- cdf: $F_X(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$



No closed expression for the cdf

erf= error-function: $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$

Gaussian Distribution = Normal Distribution

- Beregninger med normalfordelinger: Tabelopslag og Matlab:
- $X \sim \mathcal{N}(\mu, \sigma^2) \;\rightarrow\; Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$ (Standard Normal Distribution)
- $F_X(x) = Pr(X \leq x) = Pr\left(Z \leq \frac{x-\mu}{\sigma}\right) = F_Z(z)$ hvor $\mathrm{z} = \frac{x-\mu}{\sigma}$

$$= \begin{cases} \Phi(z) & Tabel\ 1\ (\text{"}Statistik\ og\ Sandsynlighedsregning\text{"}) \\ 1 - \mathrm{Q}(z) & App.D\ (\text{"}Random\ Signals\text{"}) \end{cases}$$

- $\Phi(z) = Pr(Z \leq z)$
- $\Phi(-z) = 1 - \Phi(z)$
- $\mathrm{Q}(z) = Pr(Z \geq z) = 1 - Pr(Z \leq z) = 1 - \Phi(z)$
- $\mathrm{Q}(-z) = 1 - \mathrm{Q}(z)$

- Matlab:
  - $\Pr(X \leq x) = F_X(x) = normcdf(x, \mu, \sigma)$
  - $\Pr(Z \leq z) = F_Z(z) = normcdf(z, 0{,}1) = normcdf(z)$

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.

> i.i.d. = Independent and Identically distributed

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

Example:

- Quantisation noise.

Very important!

Central Limit Theorem

- Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables with mean μ and variance σ^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 = μ and variance = $\frac{\sigma^2}{n}$.

The variance is reduced with a factor 1/n

Very important!

Central Limit Theorem

- Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean μ and variance σ^2
- Let X be the random variable:

$$X = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} = \frac{\sum_{i=1}^{n} \frac{1}{n} X_i - \mu}{\sqrt{\sigma^2/n}} = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$$

- Then in the limit: $n \to \infty$ we have that: $X \sim \mathcal{N}(0,1)$

  i.e. in the limit X will be normally distributed with mean = 0 and variance = 1 (standard normal distributed).

Sampling From Any Distribution

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

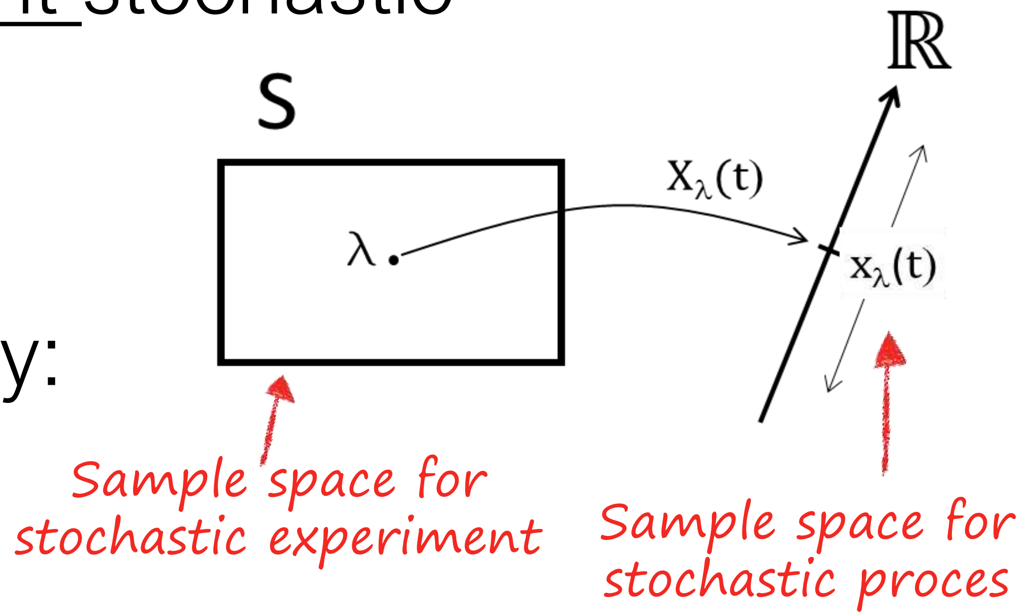
- Find the cdf of the distribution: $F_X(x)$
- Find the inverse of the cdf: $y = F_X(x) \Rightarrow x = F_X^{-1}(y)$
- Draw a ramdom sample: $y \sim \mathcal{U}[0;1]$
- Insert into the inverse cdf: $x = F_X^{-1}(y)$
- The samples $X = x$ is distributed according to: $F_X(x)$

Stochastic Processes

Definitions:

- A stochastic process is a <u>time dependent</u> stochastic variable:

$$X(t)$$



- A discrete stochastic process is given by:

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

  where n is an integer.

- Random events that develops in time

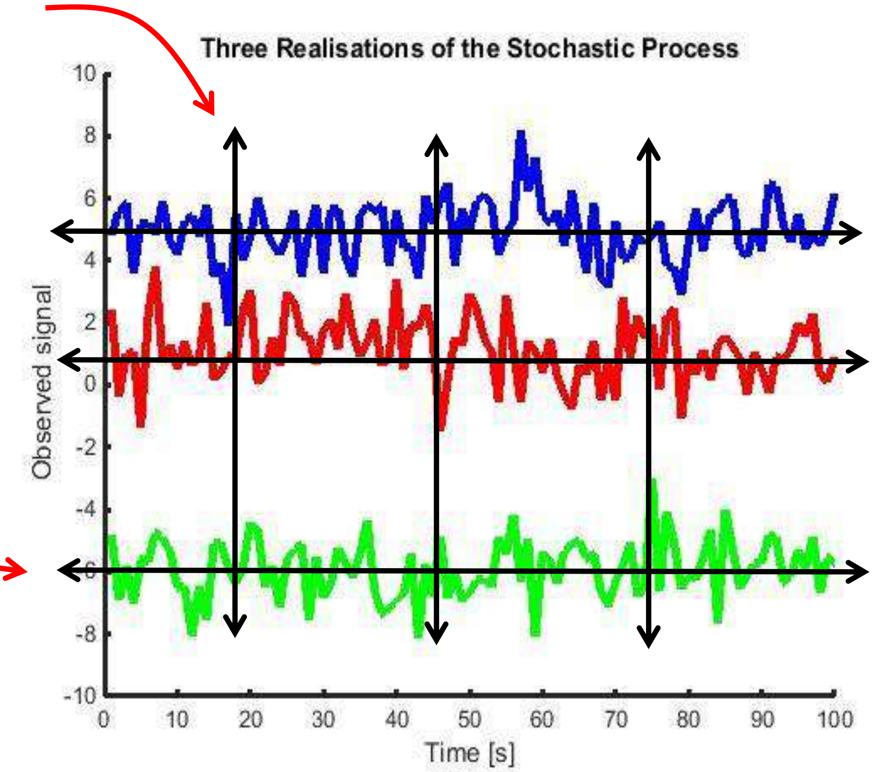
- A sample function (observed signal) is a realization of a stochastic process $x(t)$

The Mean Functions

- Ensemple mean:

The mean of all possible realizations to time t

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



The time average for one realization of the stochastic process

- Temporal mean:

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

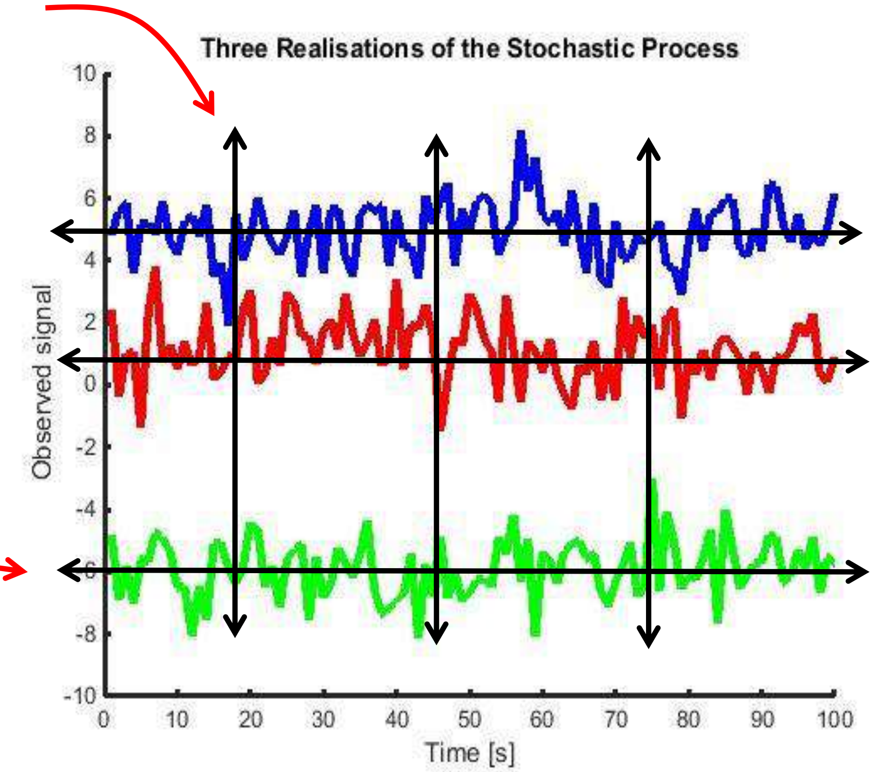
$$\left(\lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} x_i(t)\, dt \right)$$

The Variance Functions

- Ensemple variance:

The variance of all possible realizations to time t

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



The variance over time for one realization of the stochastic process

- Temporal variance:

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

$$\left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \big(x_i(t)^2 - \hat{\mu}_{X_i}{}^2\big)dt\right)$$

Stationarity in the Wide Sense (WSS)

- Ensemble mean is a constant

Can be tested.

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

- Ensemble variance is a constant

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

Stationarity in the Strict Sense (SSS):

- The density function $f_{X(t)}\big(x(t)\big)$ do not change with time

Difficult to test in reality.

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.

- 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 realization   Ensemple (WSS)

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

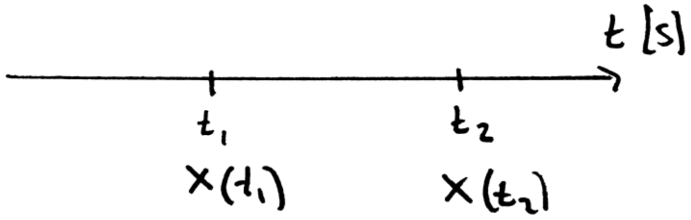
All information is achieved with one measurement (realization)

Correlations

Comparing realizations

- We compare the process at two different times

Correlation of a realization with itself



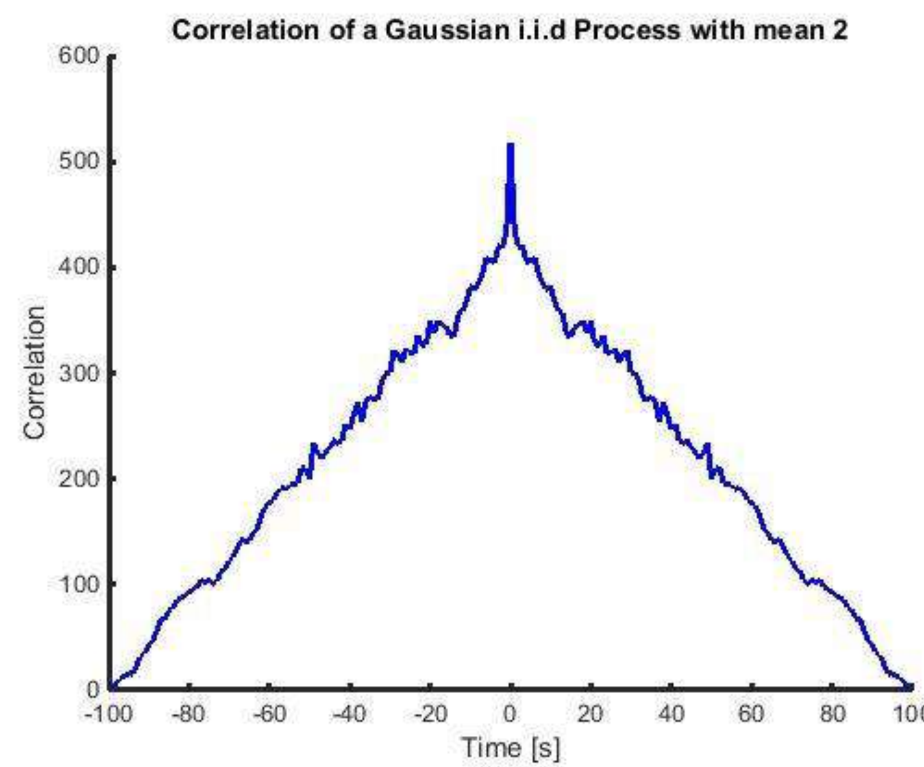
- <u>Autocorrelation</u>: $R_{XX}(t_1, t_2) = E[X(t_1)X(t_2)^*]$
  - ➢ Says something about how much the signal $X(t_1)$ resembles itself at time t_2

Correlation of two realizations

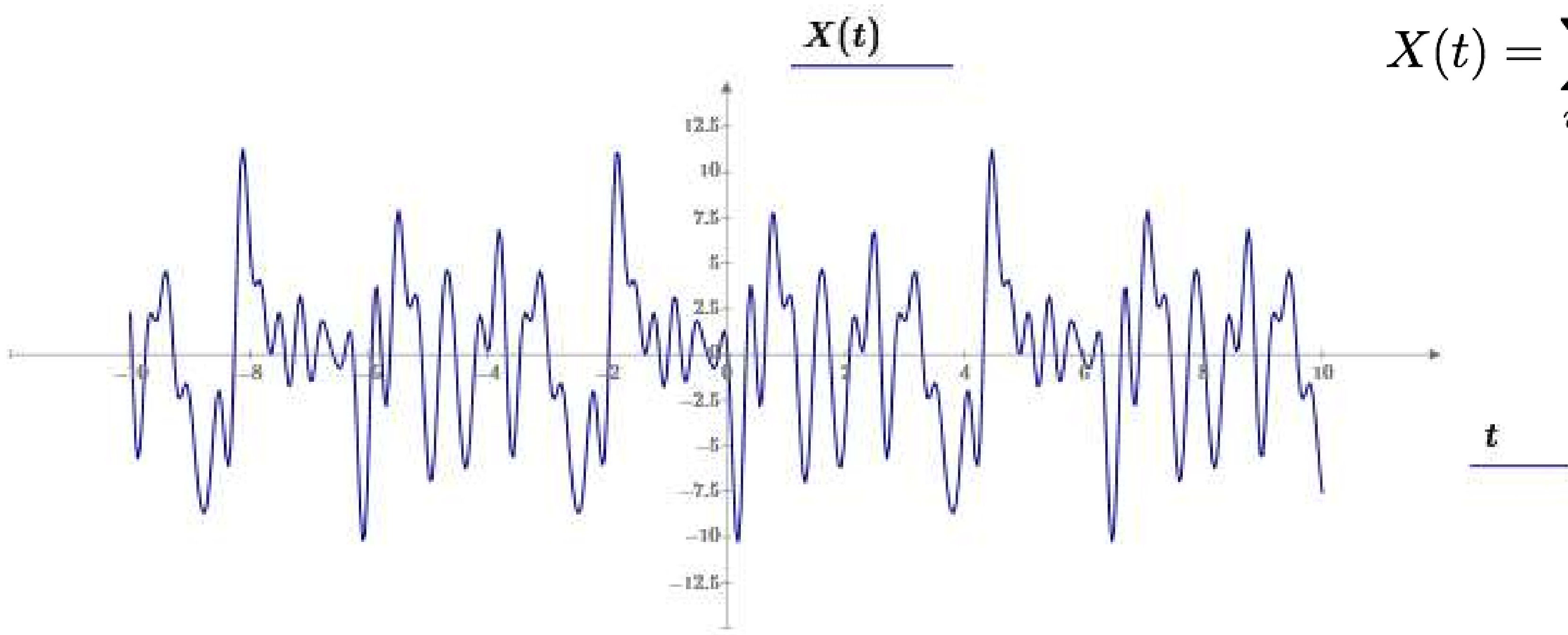
- <u>Crosscorrelation</u>: $R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*]$
  - ➢ Can be used to look for places where the signal $X(t)$ is similar to the signal $Y(t)$

Autocorrelation

- For Real WSS: $R_{XX}(\tau) = E[X(t)X(t+\tau)]$
- Properties of the autocorrelation function $R_{XX}(\tau)$:
  - An even function of τ $(R_{XX}(\tau) = R_{XX}(-\tau))$
  - Bounded by: $|R_{XX}(\tau)| \leq R_{XX}(0) = E[X^2]$ (max. in $\tau = 0$)
  - If X(t) changes fast, then $R_{XX}(\tau)$ decreases fast from $\tau = 0$
  - If X(t) changes slowly, then $R_{XX}(\tau)$ decreases slowly from $\tau = 0$
  - if X(t) is periodic, then $R_{XX}(\tau)$ is also periodic



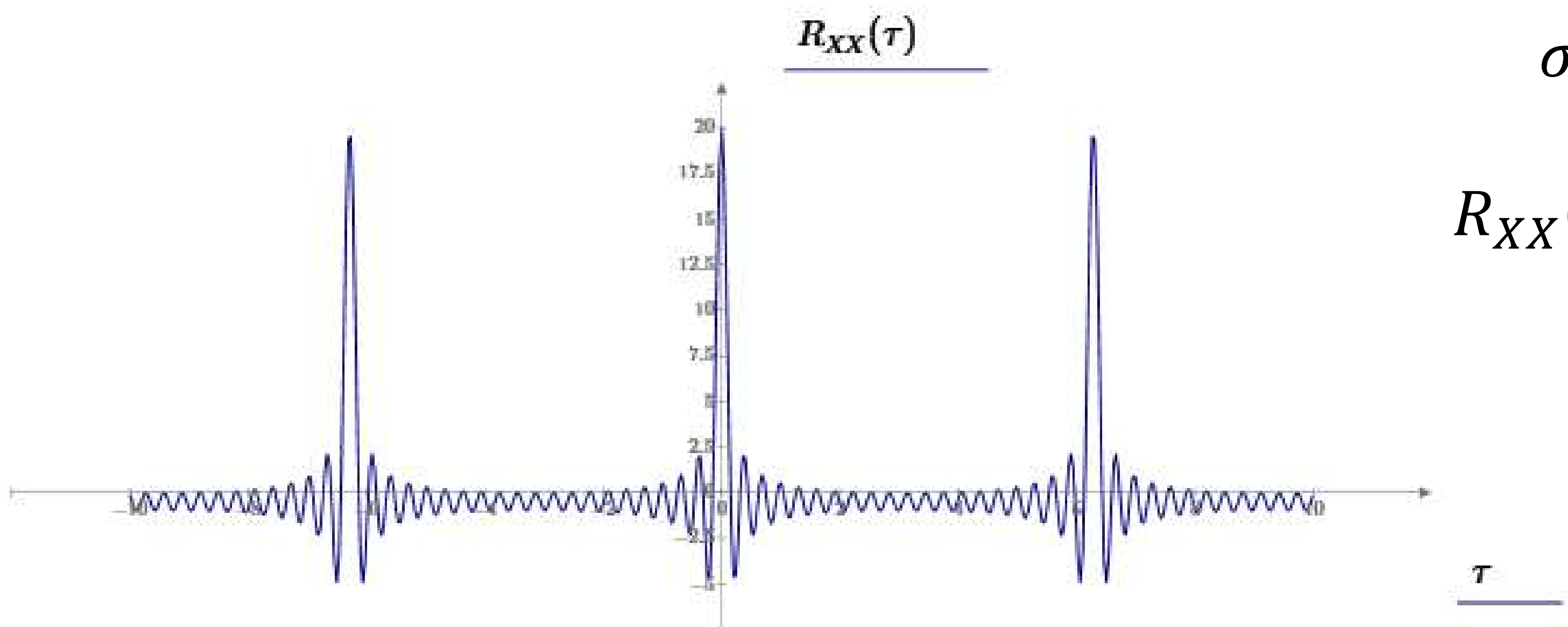
Uncalibrated Noisy Signal



$$X(t) = \sum_{i=1}^{n} A_i \cos \omega_i t + B_i \sin \omega_i t$$

$$A_i, B_i \sim \mathcal{N}(0, \sigma^2)$$

$$\omega_i = i \cdot \omega_0$$
$$\omega_0 = 1$$



$$\sigma = 1, n = 20$$

$$R_{XX}(0) = n\sigma^2 = 20$$

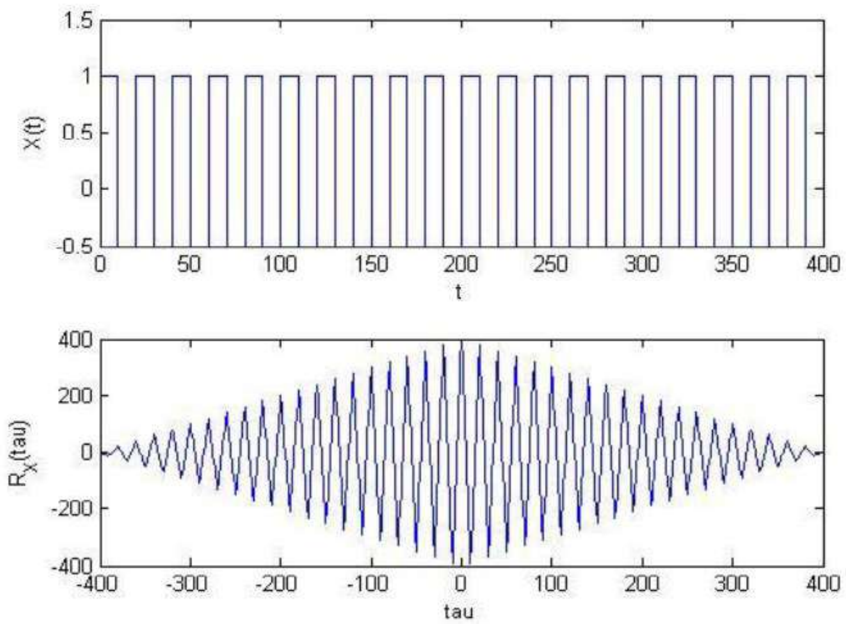
Random Binary (Digital) Signal

Deterministic:

Periodic signal

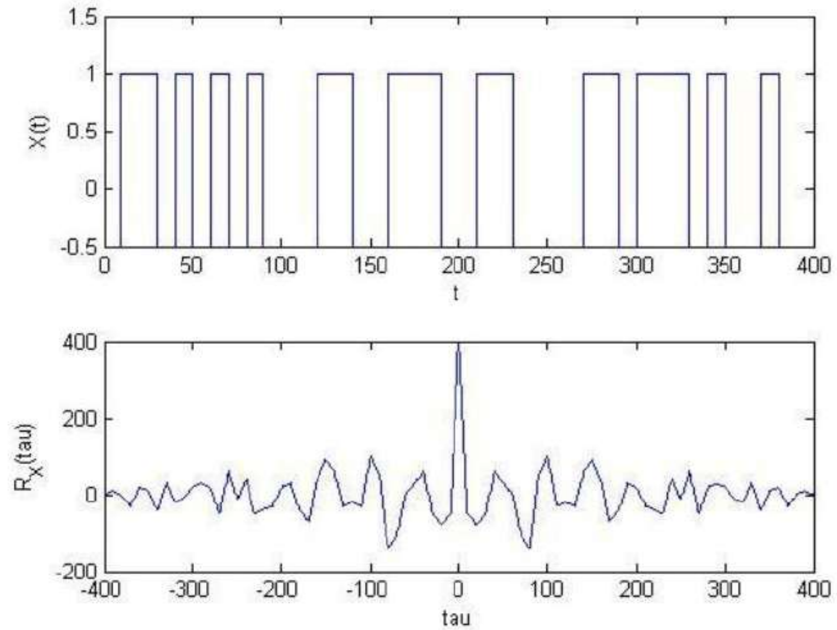
↓

R_{XX} periodic



`Rx = conv(x,fliplr(x));`

Non-deterministic
(Stochastic)



`Rx = conv(x,fliplr(x));`

Tells about how much we can predict the future

Autocovariances

- Autocovariance function:

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

Especially: $C_{XX}(t, t) = E[(X(t) - \mu_X(t))^2] = E[X(t)^2] - E[X(t)]^2 = \sigma_X^2(t)$

- Autocorrelation coefficient:

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

Especially: $r_{XX}(t, t) = 1$ ($X(\mathrm{t})$ is totally dependent of itself!)

Two Stochastic Processes

- If we have two stochastic processes $X(t)$ and $Y(t)$
  - We can compare them by looking at the cross-correlation and cross-covariance:

Cross-correlation $$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*]$$

Cross-covariance $$C_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)^*] - E[X(t_1)]E[Y(t_2)]$$

Cross-Correlation Functions

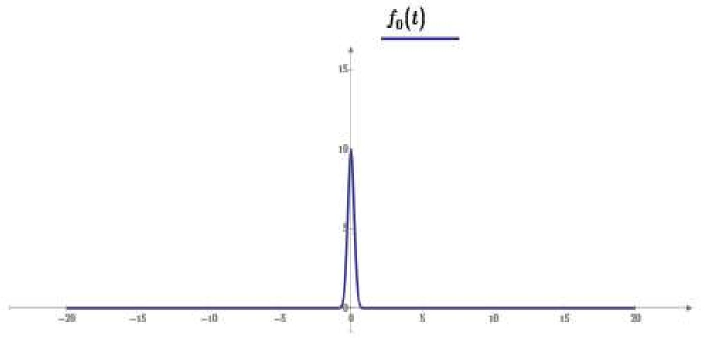
- For Real WSS processes $X(t)$ and $Y(t)$:

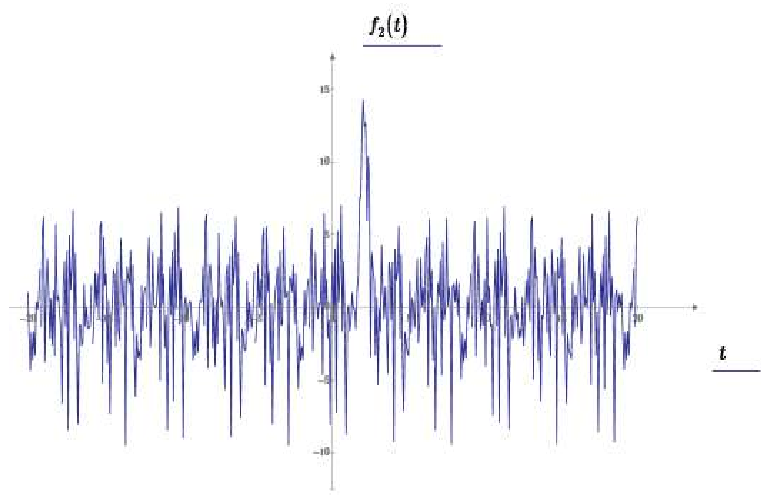
$$R_{XY}(\tau) = E[X(t)Y(t+\tau)]$$

- Properties of the cross-correlation function $R_{XY}(\tau)$:
  - $R_{XY}(\tau) = R_{YX}(-\tau))$
  - $|R_{XY}(\tau)| \leq \sqrt{R_{XX}(0)R_{YY}(0)} = \sqrt{E[X^2]E[Y^2]}$ (max. in $\tau = 0$)
  - $|R_{XY}(\tau)| \leq \frac{1}{2}(R_{XX}(0) + R_{YY}(0))$
  - If $X(t)$ and $\mathrm{Y}(t)$ are orthogonal, then $R_{XY}(\tau) = 0$
  - If $X(t)$ and $\mathrm{Y}(t)$ are independant, then $R_{XY}(\tau) = \mu_X \cdot \mu_Y$

Cross-correlation – Uncalibrated noisy signal

- Comparing two signals:
  - An uncalibrated and noisy signal $f_2(t)$
  - Reference signal $f_0(t) = 10 \cdot e^{-10t^2}$

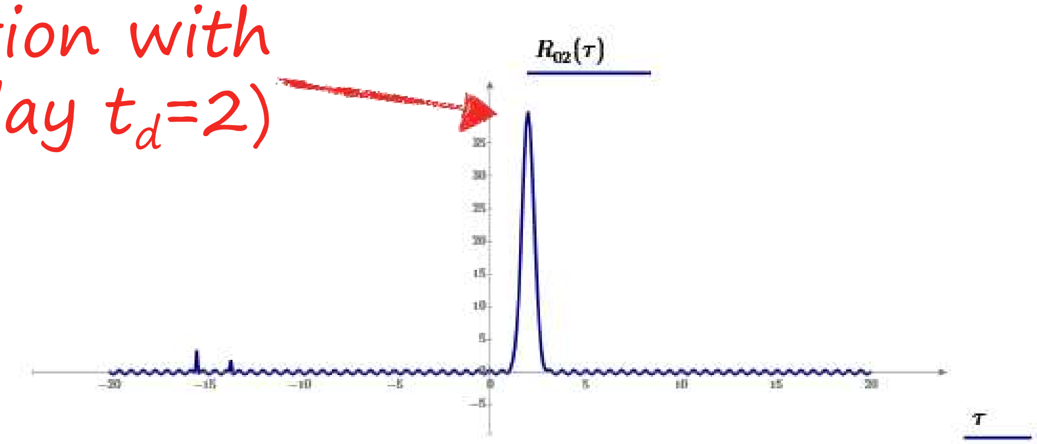




- Cross-correlation:

$$R_{02}(\tau) = \int_{-\infty}^{\infty} f_0(t) \cdot f_2(t+\tau)dt$$

Correlation with
time delay t_d=2)



Power Spectral Density (psd)

- WSS random signals X(t):
- Power Spectral Density Function (psd):

Fourier-transform

$$S_{XX}(f) = \mathcal{F}(\langle R_{XX}(\tau)\rangle_{T_0}) = \int_{-\infty}^{\infty} R_{XX}(\tau)e^{-j\cdot 2\pi f\cdot\tau}d\tau$$

Invers Fourier-transform

$$\Rightarrow R_{XX}(\tau) = \mathcal{F}^{-1}(\langle R_{XX}(\tau)\rangle = \int_{-\infty}^{\infty} S_{XX}(f)e^{j\cdot 2\pi f\cdot\tau}df$$

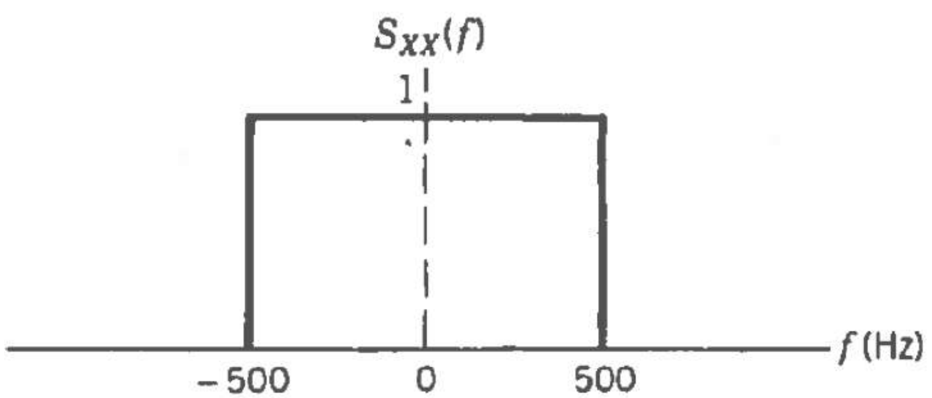


Figure 3.19a Psd of a lowpass random process $X(t)$.

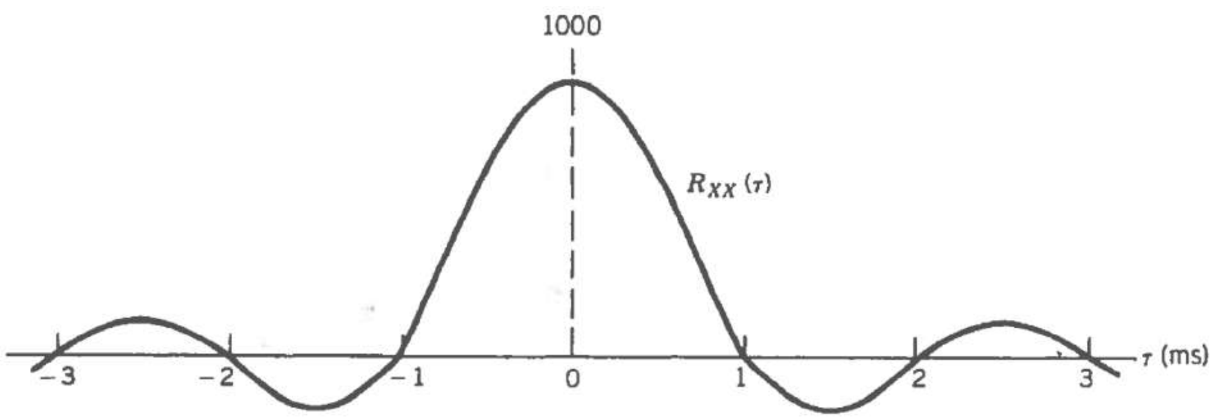


Figure 3.19b Autocorrelation function of $X(t)$.

Figures from ”Random Signals”

Power Spectral Density (psd)

- Properties of psd $S_{XX}(f)$ (spectrum of $X(t)$):
  - $S_{XX}(f) \in \mathbb{R}$
  - $S_{XX}(f) \geq 0$
  - If $X(t) \in \mathbb{R}$: $R_{XX}(-\tau) = R_{XX}(f)$ and $S_{XX}(-f) = S_{XX}(f)$ → even functions
  - If $X(t)$ periodic components: $S_{XX}(f)$ will have impulses (δ−functions)
  - $[S_{XX}(f)] = \frac{W}{Hz}$ → Distribution of power with frequency (power spectral density of the stationary random process X(t)
  - $P_X = E[X(t)^2] = R_{XX}(0) = \int_{-\infty}^{\infty} S_{XX}(f)df$

    i.e. if $X(t) = V(t)$ (voltage signal)

    → $P_X =$ power in 1Ω-resistor

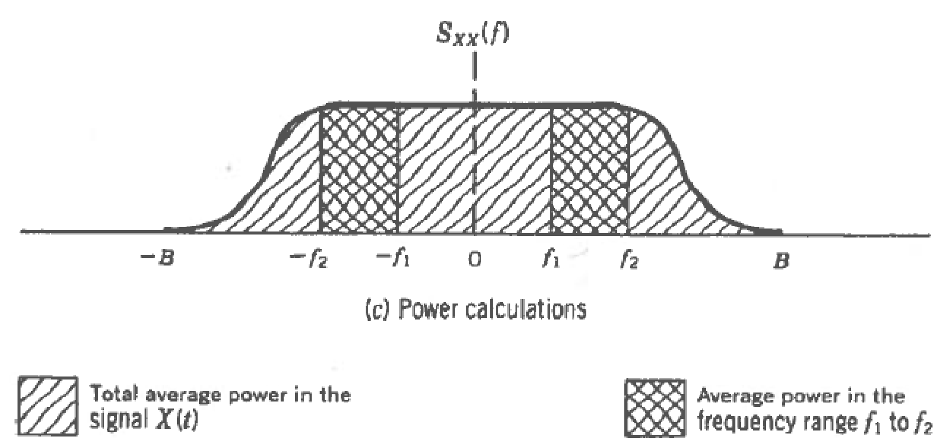


Figure from "Random Signals"

  - $P_X[f_1, f_2] = 2\int_{f_1}^{f_2} S_{XX}(f)df$ → Power in the frequency-interval $[f_1, f_2]$

Words and Concepts to Know

Probability density function Binomial coefficient Cross-covariance Convolution
Deterministic Rayleigh Distribution Deterministic Intersection Type I Error SSS
pdf Temporal cross-correlation Cross-correlation Correlation Markov chain
Probability Mass Function i.i.d. Temporal mean Continuous random variable
Randomly Sampled Data Temporal variance Marginal Correlation coefficient
Stochastic Processes Unordered Mutually Exclusive/Disjoint Ensemple variance
Uniform distribution Replacement Sampling Non-deterministic Ergodicity
Sample point Specificity Stationarity Gaussian distribution Sample space
Central Limit Theorem Experiment/Trial cdf Complement/not Joint pmf WSS
Covariance
Likelihood Simultanious pmf
Independent and Identically Distributed Event
Relative frequency Realization Independence Union Correlation coefficient
Normal distribution Sensitivity Combinatorics
Bivariate Normal Distribution
Transformation of stochastic variables Binomial distribution
Joint events
Empty set/Null set Binomial Mass Function Standard deviation
Total probability
Strict Sense Stationary Ordered Set Conditional probability
Mean Simultaneous density function Variance Bayes Rule pmf Ensemple mean
Autocovariance Type II Error Autocorrelation Coefficient Joint density function
Power Spectral Density Non-deterministic Stochastic Posterior Autocorrelation
Wide Sense Stationary Bernoulli Trial Prior Expectation Subset
Cumulative Distribution Function psd Marginal probability density function

Assignment 8

- Find a stochastic process in your area

  (discharge of a capacitor, bitrate, failure, hight, weight, …)
- Make a signal model: $X(t) = \cdots$
- Make three realizations
- Determine the ensemble mean and variance
- Determine the temporal mean and variance
- Determine stationarity and ergodicity