Signal Processing EE2S31

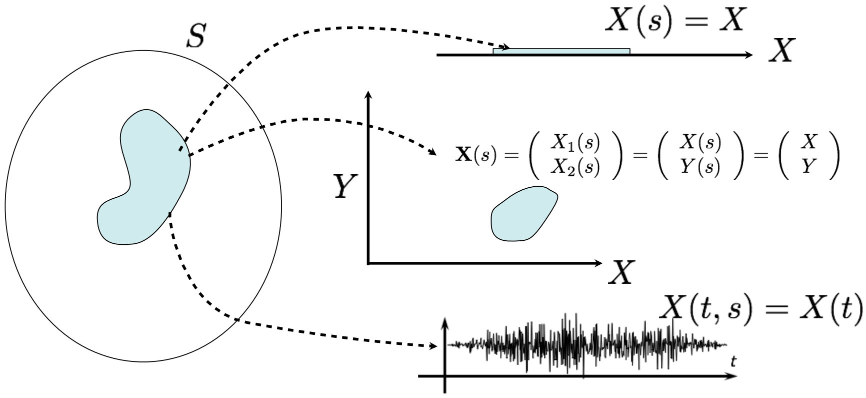
Stochastic Processes for EE

Lecture 3

1

TU Delft
Delft University of Technology

Summarizing



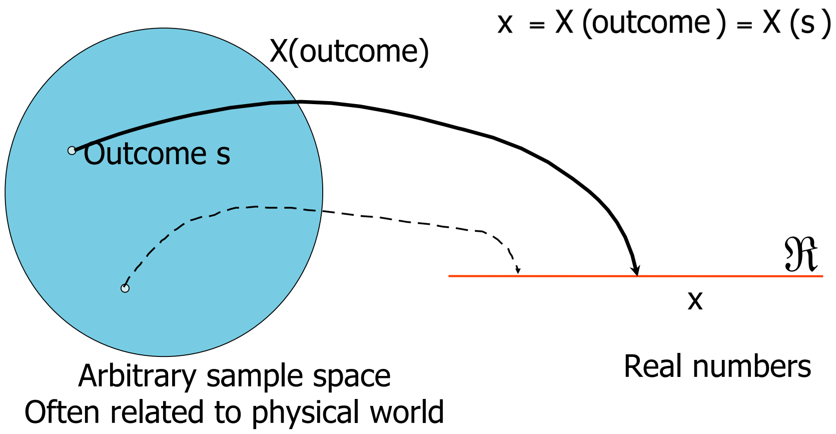
One experiment ⇨ one outcome or realization of the stochastic/random process

2

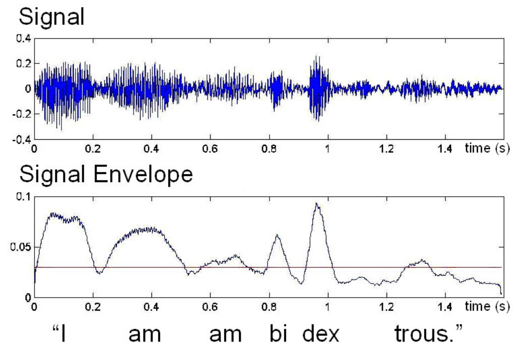
TU Delft
Delft University of Technology

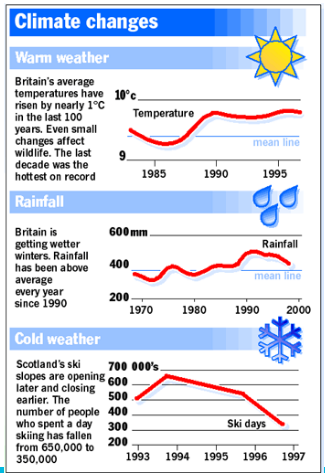
Definition of Random Variable

$$x = X(\text{outcome}) = X(s)$$



Voice signal, weather, stock market





What is a process?

- Outcome of an experiment is a
  - single value: random variables
  - multiple values: random vectors
  - series of (in one way or the other) ordered values (we will mainly consider ordering in time)

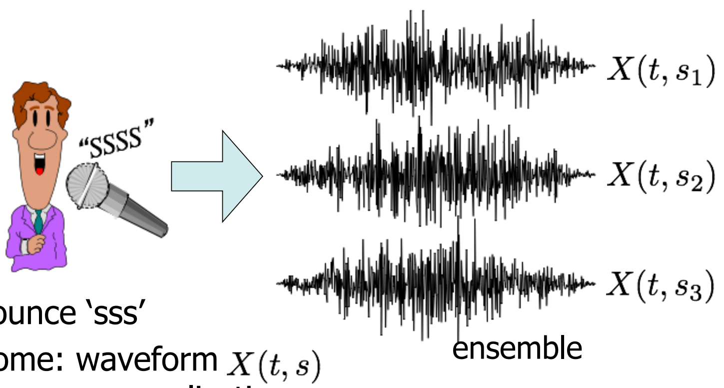
- Example: binary expansion of a random number of the interval [0,1]

random process

0.7187....	=	(0.)10111...
0.3130....	=	(0.)01001...

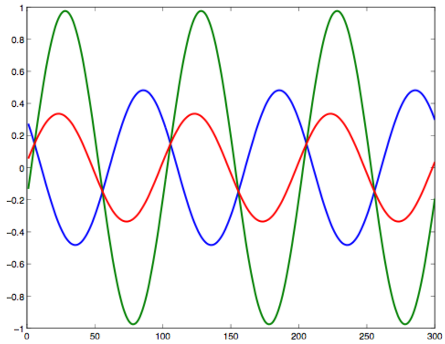
Example speech

- Experiment 's':



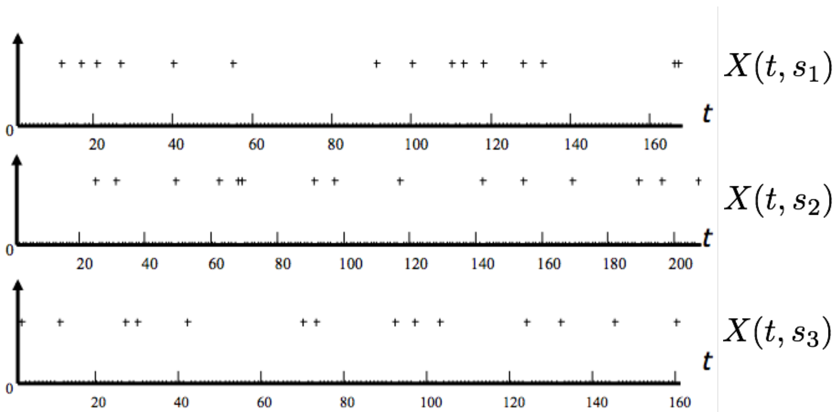
- Pronounce 'sss'
- Outcome: waveform $X(t, s)$
  one realization

Random sinusoidal process

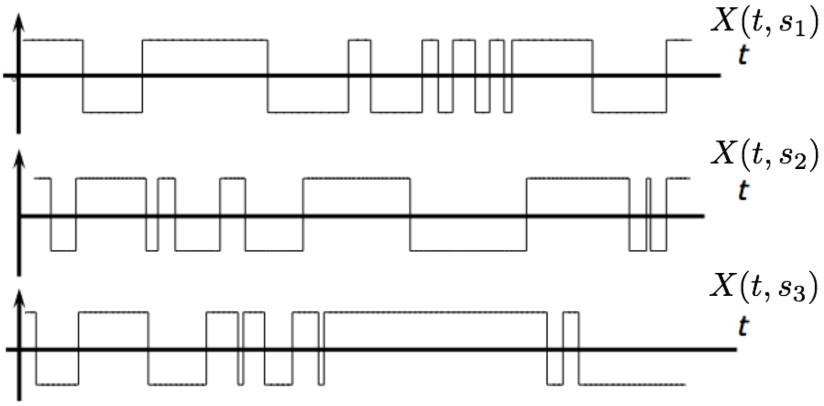


- Amplitude A and phase ϕ are random variables
- Random process: $X_n = A \sin(2\pi f n + \phi)$

Arrival times of packets in a network



Random telegraph signal



Bit patterns in binary images

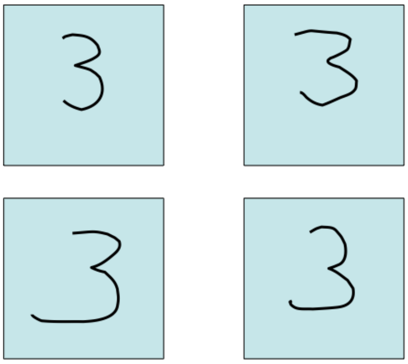
$$X_3 = 0000111111000001111010100011100000000000111010$$

$$X_{17} = 011110011100110101111010100011100010010100001$$

$$X_{55} = 000000001111111111111010011111111111100111110$$

Handwritten digits

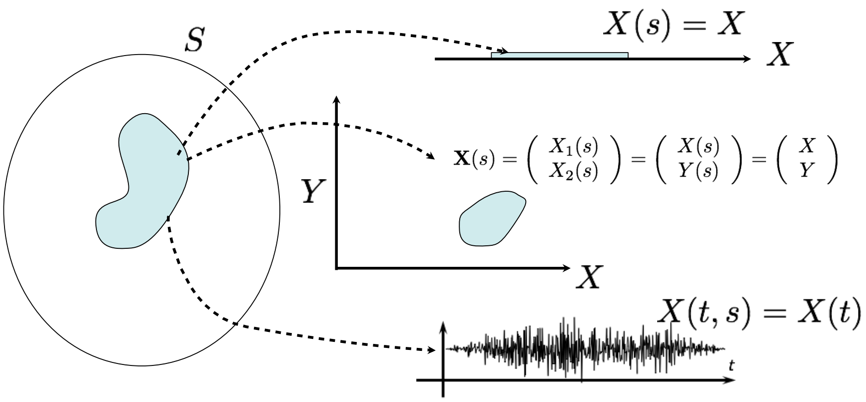
- Model trajectories, gestures



11

TU Delft
Delft University of Technology

Summarizing



One experiment → one outcome or realization of the stochastic/random process

12

TU Delft
Delft University of Technology

Notation and types

- Notation $X(t,s) = X(t)$

		amplitude of X(t)	
		continuous	discrete
time axis	continuous	continuous time continuous value	continuous time discrete value
	discrete	discrete time continuous value	discrete time discrete amplitude

Notation and types

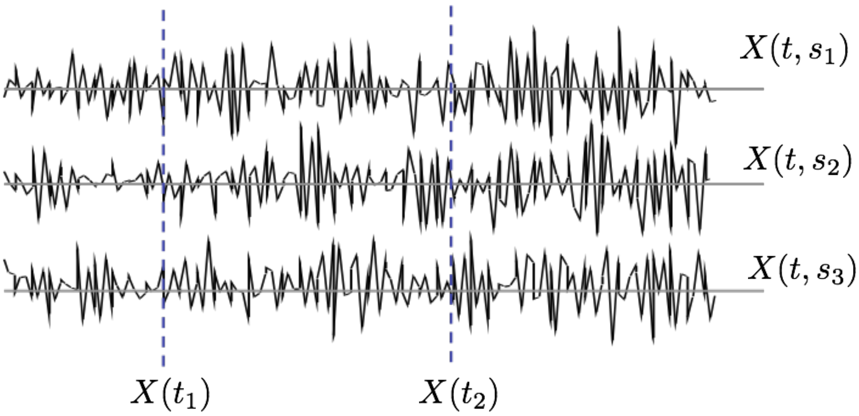
- Notation $X(t,s) = X(t)$

		amplitude of X(t)	
		continuous	discrete
time axis	continuous	continuous time continuous value	continuous time discrete value
	discrete	discrete time continuous value	discrete time discrete amplitude

$$X(t) = X(t_k) = X_k$$

$$X_n = X(n), \quad n = 1, 2, 3, \ldots$$

Description random process



Description of a random process

- At any (fixed) time t_k the stochastic process $X(t_k)$ can be regarded as a random variable

$$X(t_k) \rightarrow f_{X_{t_k}}(x_{t_k}) = f_{X_k}(x_k)$$

- This pdf may be different for each t_k !!

- The joint behavior for all t is given by the joint-PDF:

$$X(t_1)...X(t_k)... \rightarrow f_{X_1,X_2,...,X_k,...}(x_1, x_2, ..., x_k, ...)$$

Notice that ...

$$X(t_1)...X(t_k)... \to f_{X_1,X_2,...,X_k,...}(x_1,x_2,...,x_k,...)$$

resembles a vector random variable

- ... but can be of infinite dimensionality
- ... and **ordering** (in time) of $X(t_k)$ is **essential**

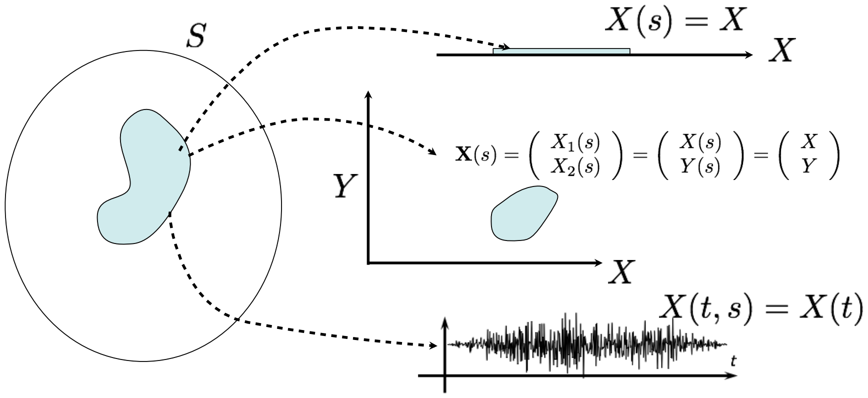
With the exception of a few "special cases"

- iid random sequence/process
- Gaussian stochastic process
- Poisson stochastic process

this joint PDF is **very** difficult to get in practice

TU Delft
Delft University of Technology

Summarizing



One experiment $\Rightarrow$ one outcome or realization of the stochastic/random process

TU Delft
Delft University of Technology

This week:

- We discuss now
  - IID process
  - Bernoulli process
  - Counting process
  - Poisson process
  - Interarrival times with exponential pdf
  - Gaussian process
- How to characterize processes?
- Autocovariance function and autocorrelation function
- Stationarity and wide-sense stationarity

IID random process (sequence)

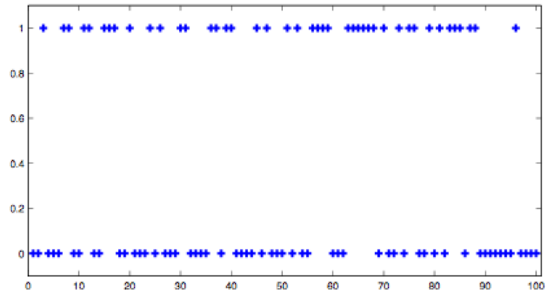
- IID means
  - Independent
  - Identically
  - Distributed random process
- This means that
  - all $X(t_k)$ are mutually independent random variables for all t_k
  - all $X(t_k)$ have the **same** pdf for all t_k

$$f_{X_1,X_2,\ldots}(x_1,x_2,\ldots) = f_{X_1}(x_1)f_{X_2}(x_2)\ldots$$
$$= f_X(x_1)f_X(x_2)\ldots$$

Bernoulli Process (time discrete)

- Bernoulli, for one time instance:

$$P_{X_k}(x_k) = \begin{cases} p & x_k = 1 \\ 1-p & x_k = 0 \\ 0 & \text{otherwise} \end{cases}$$



Bernoulli Process (time discrete)

- This

$$P_{X_k}(x_k) = \begin{cases} p & x_k = 1 \\ 1-p & x_k = 0 \\ 0 & \text{otherwise} \end{cases}$$

- can be rewritten as

$$P_{X_k}(x_k) = \begin{cases} p^{x_k}(1-p)^{1-x_k} & x = 0, 1 \\ 0 & \text{otherwise} \end{cases}$$

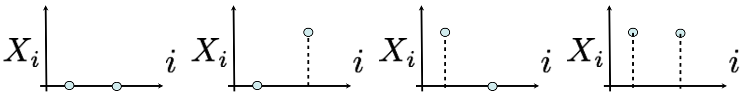
- So for two time instances:

$$\begin{aligned} P_{X_1,X_2}(x_1, x_2) &= P_{X_1}(x_1)P_{X_2}(x_2) \\ &= p^{x_1}(1-p)^{1-x_1}p^{x_2}(1-p)^{1-x_2} \\ &= p^{x_1+x_2}(1-p)^{2-x_1-x_2} \end{aligned}$$

Bernoulli Process

$$P_{X_1,X_2}(x_1,x_2) = p^{x_1+x_2}(1-p)^{2-x_1-x_2}$$

- Four possible situations:

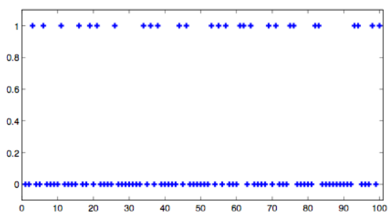


- Say, p=0.3

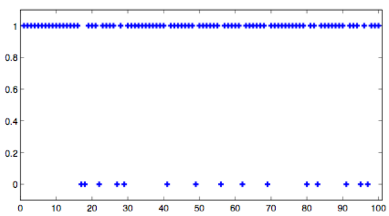
$$P_{X_1,X_2}(0,0) = p^{0+0}(1-p)^{2-0-0} = 0.7^2 = 0.49$$
$$P_{X_1,X_2}(0,1) = p^{0+1}(1-p)^{2-0-1} = 0.7 \cdot 0.3 = 0.21$$
$$P_{X_1,X_2}(1,0) = p^{1+0}(1-p)^{2-1-0} = 0.3 \cdot 0.7 = 0.21$$
$$P_{X_1,X_2}(1,1) = p^{1+1}(1-p)^{2-1-1} = 0.3^2 = 0.09$$

Bernoulli Process

$$P_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \begin{cases} p^{x_1+\ldots+x_n}(1-p)^{n-(x_1+\ldots+x_n)} & x_i = 0,1 \\ 0 & \text{otherwise} \end{cases}$$



$$P(x_1, \ldots, x_n) = 1.6 \cdot 10^{-26}$$

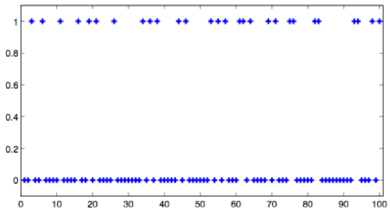


$$P(x_1, \ldots, x_n) = 1.7 \cdot 10^{-47}$$

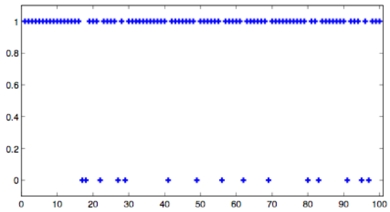
- Assume n=100, p=0.3

Bernoulli Process

$$P_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \begin{cases} p^{x_1+\ldots+x_n}(1-p)^{n-(x_1+\ldots+x_n)} & x_i = 0,1 \\ 0 & \text{otherwise} \end{cases}$$



$P(x_1, \ldots, x_n) = 1.6 \cdot 10^{-26}$

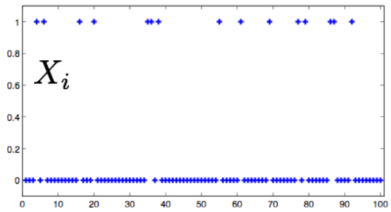


$P(x_1, \ldots, x_n) = 1.7 \cdot 10^{-47}$

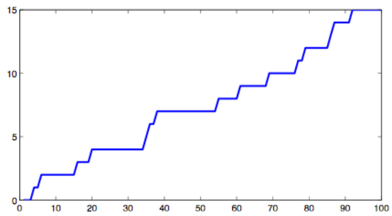
- Assume n=100, p=0.3
- Note: number of possible sequences is $2^{100} = 1.3 \cdot 10^{30}$

Counting process

- A process N(k) is a counting process if every sample function (realization) of N(k) is integer valued and non-decreasing

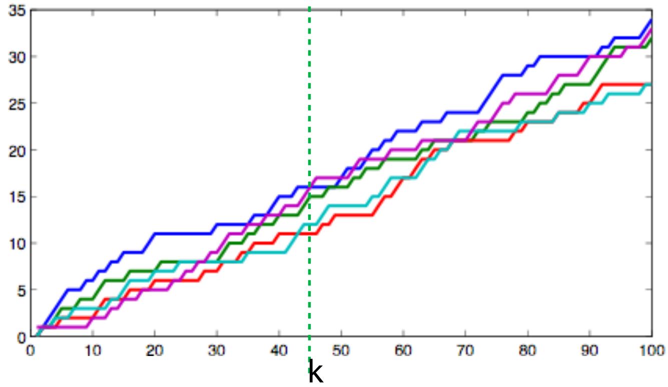


$N(k)$



- A simple counting process can be derived using a Bernoulli process

Counting process



- What is a proper description of the random process N(k)?

Counting

- Since we are counting the number of Bernoulli random variables with $X_i = 1$ we like to know

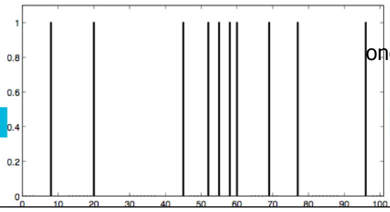
$$P_{N(k)}(j) = P(N(k) = j) \qquad j = 0, 1, 2, ..., k$$

- This is the probability of j successes in k tries (binomial random var.!):

$$P_{N(k)}(j) = \binom{k}{j} p^j (1-p)^{k-j}$$

Continuous time counting process

- Instead of a fixed number of 'tries', we take look at the number of arrivals in fixed amount of time.
  - N(t)=number of "arrivals" (successes, events, data packets, customers in queue, ...) in time interval [0,t].
- Instead of probability of success (p as in the binomial case), we take the average number of arrivals per time unit.
  - Rate λ (arrivals) per second



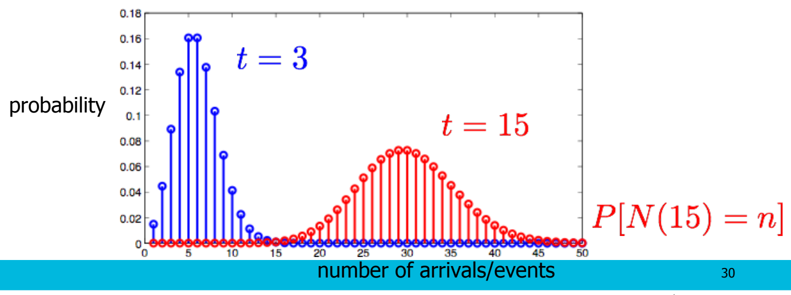
TU Delft
Delft University of Technology

Poisson process

- Number of events in [0,t] has a Poisson PMF with rate λ per time unit

$$P[N(t)=n]=P_{N(t)}(n)=\frac{(\lambda t)^n}{n!}\exp(-\lambda t)$$

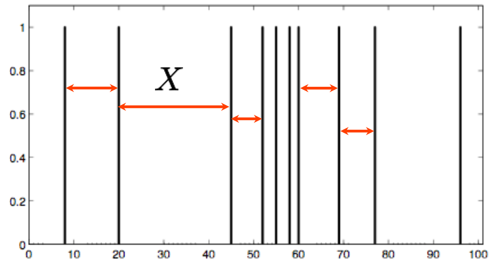
- Assume $\lambda = 2$



TU Delft
Delft University of Technology

Interarrival times

- In a Poisson process, the interarrival times X_i form an IID process



- Marginal pdf of the arrival times:

$$f_X(x) = \lambda \exp(-\lambda x)$$

Gaussian random variable vector

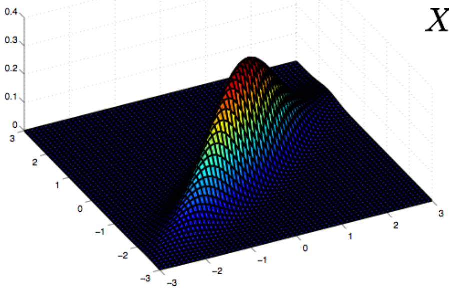
- When we have a **pair** of Gaussian random variables, and we write the pair $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

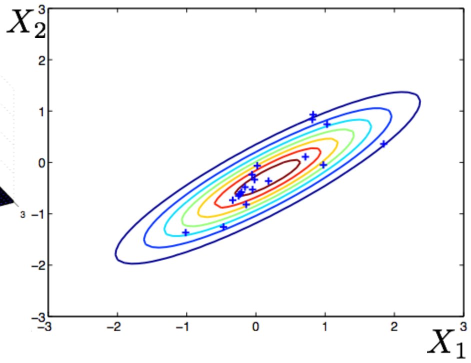
$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi \det(C_{\mathbf{X}})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu_{\mathbf{X}})' C_{\mathbf{X}}^{-1}(\mathbf{x}-\mu_{\mathbf{X}})\right)$$

- Here, the mean vector $\mu_{\mathbf{X}} = \begin{pmatrix} E[X_1] \\ E[X_2] \end{pmatrix}$
- and the **covariance matrix**:

$$C_{\mathbf{X}} = \begin{pmatrix} Var[X_1] & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var[X_2] \end{pmatrix}$$

Samples from a (2D) Gaussian pdf



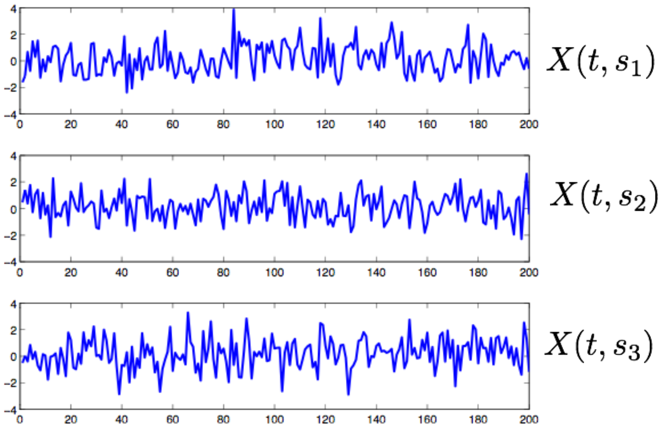


- 20 samples drawn from a Gaussian pdf with

$$\mu_{\mathbf{x}} = \begin{pmatrix} 0.2 \\ -0.3 \end{pmatrix} \qquad C_{\mathbf{x}} = \begin{pmatrix} 1.2 & 0.8 \\ 0.8 & 0.7 \end{pmatrix}$$

Gaussian process

- Each X_i is drawn from a (very high dim.) Gaussian pdf



Cov. matrix of a Gaussian process?

- In principle the size of the matrix is $\infty \times \infty$ (pretty large): mathematicians don't like that.
- Mathematically speaking, you define a Gaussian process by saying:

- Given **any** set of k time points $t_1, t_2, ..., t_k$

  the vector $\mathbf{X} = [X(t_1), X(t_2), ..., X(t_k)]$

  has a Gaussian distribution

Characterization of a Stochastic Process

Characterization of Random Process

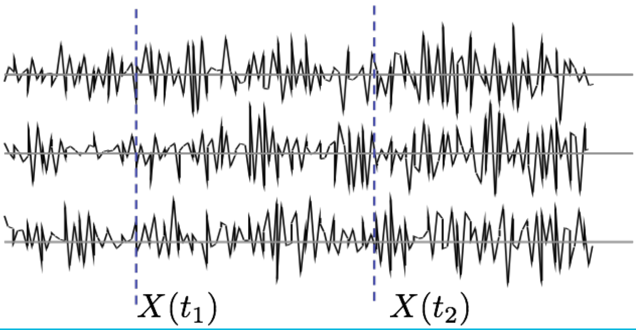
- Often we rely on properties, such as
  - Expected value
  - Variance
  - Covariance and correlation
    to summarize the behavior of a stochastic process, because the joint-pdf can not realistically be obtained

- Apply these (known) concepts to stochastic processes
- In practical cases of interest, we will use estimates of expected value, variance, and correlation

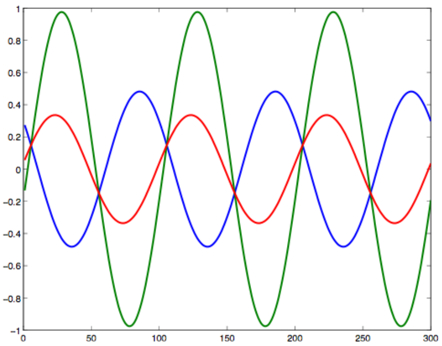
Expected value of X(t)

- Expected value of $X(t_k)$ at time t_k

$$\mu(t_k) = E[X(t_k)] = \int_{-\infty}^{\infty} x f_{X(t_k)}(x)dx$$



Example sinusoidal process



$$X(t) = A\sin(\omega t + \phi)$$

- Amplitude and phase are independent random variables:
- A is uniformly distributed on [-1,+1]
- ϕ is uniformly distributed on $[0, 2\pi]$

TU Delft
Delft University of Technology

Example sinusoidal process

- Expected value of this process:

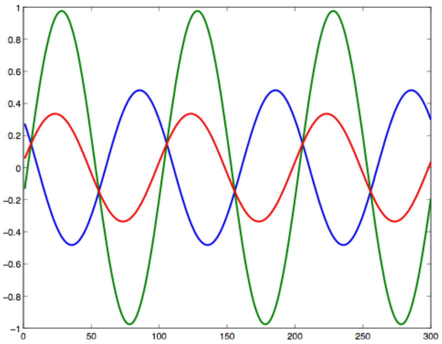
$$\begin{aligned}\mu_X(t) &= E[X(t)] = E[A\sin(\omega t + \Phi)]\\ &= E[A]\cdot E[\sin(\omega t + \Phi)]\end{aligned}$$

- We have:

$$\begin{aligned}E[A] &= 0\\ E[\sin(\omega t + \Phi)] &= \int_0^{2\pi} \sin(\omega t + u) f_\Phi(u) du\\ &= \frac{1}{2\pi}\int_0^{2\pi} \sin(\omega t + u) du = 0\end{aligned}$$

TU Delft
Delft University of Technology

Variant of sinusoidal process



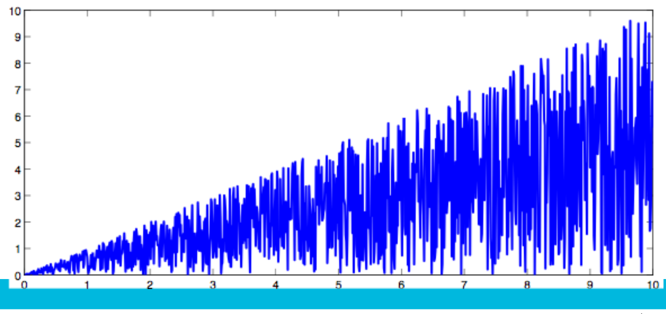
$$X(t) = A\sin(\omega t + \phi)$$

- Only the amplitude is a random variable:
- A is now uniformly distributed on [0,+1] (instead of [-1,+1])
- What is E[X] ? (do it yourself!)

Another example

$$X(t) = \begin{cases} 0 & t < 0 \\ A_t \cdot t & t \geq 0 \end{cases} \qquad f_{A_t}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Random variables A_t are stochastically independent for all t:



Another example

$$X(t) = \begin{cases} 0 & t < 0 \\ A_t \cdot t & t \geq 0 \end{cases} \qquad f_{A_t}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Expected value

$$E[X(t)] = E[A_t t] = tE[A_t] = 0.5t$$

(uhm, for t>0)

- Variance

$$Var[X(t)] = Var[A_t t] = t^2 Var[A_t] = \frac{1}{12}t^2$$

43

TU Delft
Delft University of Technology

Another variant

- Assume that the amplitude is a random variable, but it does NOT depend on t:

$$X(t) = \begin{cases} 0 & t < 0 \\ A \cdot t & t \geq 0 \end{cases} \qquad f_A(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

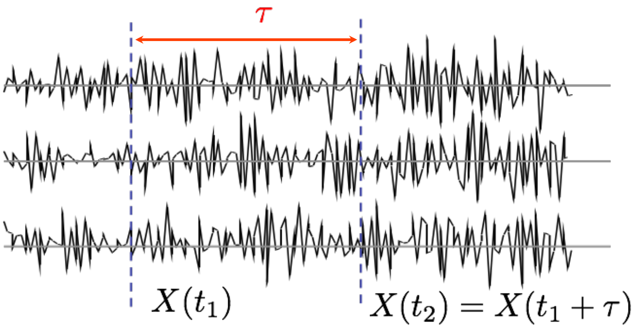
- How do the realizations of this process look like?
- What is the expected value and variance of this process?

44

TU Delft
Delft University of Technology

Autocovariance

- Joint behavior of X(t) at time t_1 and t_2 can be described by the auto-covariance function, the covariance of $X(t_1)$ and $X(t_2)$ for all t_1 and t_2



Auto-covariance/-correlation function

- Write down the covariance for two time instances

$$\begin{aligned} C_X(t,\tau) &= Cov[X(t), X(t+\tau)] \\ &= E\left[(X(t) - \mu_X(t))(X(t+\tau) - \mu_X(t+\tau))\right] \\ &= E\left[X(t)X(t+\tau)\right] - E[X(t)]E[X(t+\tau)] \end{aligned}$$

similar to $E[XY]$ similar to $E[X]E[Y]$

correlation $R_X(t,\tau)$

the autocorrelation function

Autocovariance function

- Write down the covariance for two time instances

$$\begin{aligned} C_X(t,\tau) &= Cov[X(t), X(t+\tau)] \\ &= E\left[(X(t)-\mu_X(t))(X(t+\tau)-\mu_X(t+\tau))\right] \\ &= E\left[X(t)X(t+\tau)\right] - E[X(t)]E[X(t+\tau)] \end{aligned}$$

- What is $C_X(t, \tau = 0)$?

$$\begin{aligned} C_X(t,\tau=0) &= E[X(t)X(t)] - E[X(t)]E[X(t)] \\ &= Var[X(t)] \end{aligned}$$

47

TUDelft
Delft University of Technology

Autocorrelation function

- Is a function of two variables:

$$R_X(t,\tau) = E[X(t)X(t+\tau)] = \iint xy f_{X(t)X(t+\tau)}(x,y)dxdy$$

amplitude and time continuous

$$\begin{aligned} R_X(n,k) &= E[X(n)X(n+k)] = E[X_n X_{n+k}] \\ &= \iint xy f_{X_n X_{n+k}}(x,y)dxdy \end{aligned}$$

amplitude continuous and time discrete

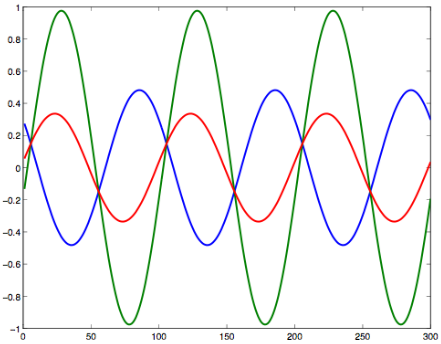
$$\begin{aligned} R_X(n,k) &= E[X_n X_{n+k}] \\ &= \sum_x \sum_y xy P[X_n = x, X_{n+k} = y] \end{aligned}$$

amplitude and time discrete

48

TUDelft
Delft University of Technology

Example sinusoidal process



$$X(t) = A\sin(\omega t + \phi)$$

- Amplitude and phase are independent random variables:
- A is uniformly distributed on [-1,+1]
- ϕ is uniformly distributed on $[0, 2\pi]$

Example sinusoidal process

$$\begin{aligned} R_X(t,\tau) &= E[X(t)X(t+\tau)] \\ &= E[A^2 \sin(\omega t + \phi)\sin(\omega(t+\tau)+\phi)] \\ &= E[A^2]E[\sin(\omega t + \phi)\sin(\omega(t+\tau)+\phi)] \end{aligned}$$

$$E[\sin(\omega t + \phi)\sin(\omega(t+\tau)+\phi)] =$$ (use math fact B.2)

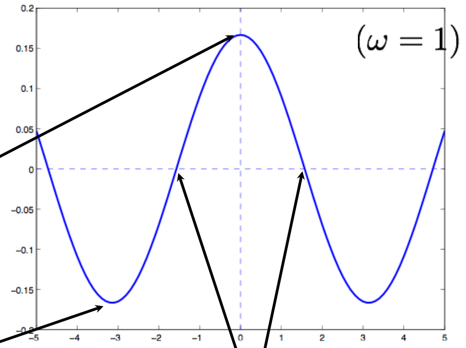
$$\begin{aligned} &= \frac{1}{2}E[\cos(-\omega\tau) - \cos(2\omega t + \omega\tau + 2\phi)] \\ &= \frac{1}{2}E[\cos(-\omega\tau)] - \frac{1}{2}E[\cos(2\omega t + \omega\tau + 2\phi)] \\ &= \frac{1}{2}\cos(\omega\tau) \end{aligned}$$

$$R_X(t,\tau) = \frac{1}{6}\cos(\omega\tau)$$

The autocorrelation function

- So we found:

$$R_X(t,\tau) = \frac{1}{6}\cos(\omega\tau)$$



high correlation for $\tau = 0$

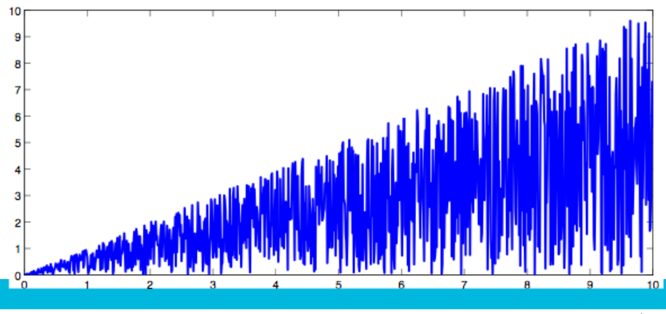
very negative correlation for $\tau = \pm\pi$

zero correlation for $\tau = \pm\pi/2$

Another example

$$X(t) = \begin{cases} 0 & t < 0 \\ A_t \cdot t & t \geq 0 \end{cases} \qquad f_{A_t}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Random variables A_t are stochastically independent for all t



Another example

$$X(t) = \begin{cases} 0 & t < 0 \\ A_t \cdot t & t \geq 0 \end{cases} \qquad f_{A_t}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} R_X(t,\tau) &= E[X(t)X(t+\tau)] \qquad\qquad t \geq 0 \\ &= E[(A_t t)(A_{t+\tau}(t+\tau))] \\ &= t(t+\tau)E[A_t A_{t+\tau}] \\ &= t(t+\tau)E[A_t]E[A_{t+\tau}] = \frac{1}{4}t(t+\tau) \end{aligned}$$

for $\tau \neq 0$

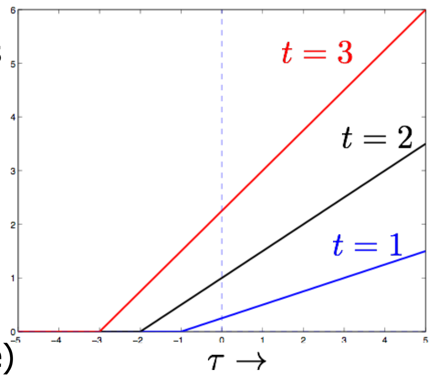
53

TU Delft
Delft University of Technology

Pictures of autocorrelation f.

$$R_X(t,\tau) = \frac{1}{4}t(t+\tau)$$

- Function of **two** variables
- Try a few values for variable t
- This $R_X(t,\tau)$ looks very strange indeed... (the signal is also very strange)



54

TU Delft
Delft University of Technology

Uncorrelated processes
Stationary processes
Wide-sense stationary processes

Uncorrelated process

- If all pairs $X(t)$, $X(t+\tau)$ are uncorrelated, i.e.

$$C_X(t,\tau) = \begin{cases} var(t) & \text{for all } t \text{ and } \ \tau = 0 \\ 0 & \text{for all } t \text{ and } \ \tau \neq 0 \end{cases}$$

  then $X(t)$ is called an uncorrelated process

- If all pairs $X(t)$, $X(t+\tau)$ are orthogonal, i.e.

$$R_X(t,\tau) = \begin{cases} E[X^2(t)] & \text{for all } t \text{ and } \ \tau = 0 \\ 0 & \text{for all } t \text{ and } \ \tau \neq 0 \end{cases}$$

  then $X(t)$ is called an orthogonal process.

Stationary process

- A stochastic process is stationary if and only if every joint-pdf is shift invariant:

$$f_{X(t_1),X(t_2),..,X(t_k)}(x_1,x_2,..,x_k) = f_{X(t_1+\Delta t),X(t_2+\Delta t),..,X(t_k+\Delta t)}(x_1,x_2,..,x_k)$$

- **Consequence I**
  - The marginal pdf's are independent of t:

$$f_{X(t)}(x) = f_{X(t+\Delta t)}(x) = f_X(x)$$

  - The marginal pdf's are identical for all t_k !!!

Stationary process

- Therefore:
  - Expected value is independent of time:

$$\mu_X(t) = E[X(t)] = \mu_X$$

  - Variance is independent of time:

$$Var_X(t) = Var[X(t)] = Var[X] = \sigma_X^2$$

TUDelft
Delft University of Technology

Stationary process

- Consequence II
  - The 2D joint-pdf is shift invariant

$$f_{X(t_1),X(t_2)}(x_1,x_2) = f_{X(t_1+\Delta t),X(t_2+\Delta t)}(x_1,x_2)$$
$$= f_{X(0),X(t_2-t_1)}(x_1,x_2)$$

  - ... only the 'distance' between t_2 and t_1 matters

- Therefore

$$R_X(t,\tau) = R_X(\tau)$$
$$C_X(t,\tau) = C_X(\tau) = R_X(\tau) - \mu_X^2$$

Stationary processes

- Example of stationary processes
  - iid process
  - Bernoulli process
  - Poisson process

- Non-stationary processes are difficult to model and to handle in practice

Wide-Sense stationary processes

- To show that a process is stationary, we need the overall joint-pdf
  - Pretty impossible to get, except for special cases
- We can often estimate the process'
  - expected value
  - correlation function
- If (only) these functions satisfy the property of stationarity, we call this process wide sense stationary (WSS)
  - Don't know anything about other properties of the process!

'zwak stationair'

61

TU Delft
Delft University of Technology

WSS Process

- A process is wide-sense stationary, if and only if

$$\mu_X(t) = \mu \qquad \text{for all } t$$
$$R_X(t,\tau) = R_X(\tau) \quad \text{for all } t$$

$$\mu_X(n) = \mu \qquad \text{for all } n$$
$$R_X(n,k) = R_X(k) \quad \text{for all } n$$

- Example: sinusoidal random process:

$$C_X(\tau) = R_X(\tau) = \frac{1}{6}\cos(\omega\tau)$$

62

TU Delft
Delft University of Technology

Autocorrelation function

- We will work a lot with the assumption of WSS
- With random time signals, we often (sometimes even implicitly) assume $E[X(k)] = 0$

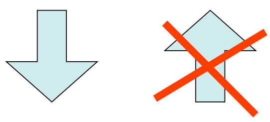
- The autocorrelation function is the most important property used in random signal processing. When WSS:

$$R_X(0) \geq 0$$
$$R_X(k) = R_X(-k)$$
$$|R_X(k)| \leq R_X(0)$$
$$\lim_{k \to \infty} R_X(k) = \mu_X^2$$

63

NOTE

Stationary Process
reflects properties of the joint-pdf



Wide-Sense Stationary process
reflects properties (only) of
•expected value
•autocorrelation function

64

Covered Today

- Chapter 10
- Key terms
  - IID process
  - Bernoulli process, counting process, Poisson process, Gaussian process
  - Auto-covariance function
  - Auto-correlation function
  - Uncorrelated processes
  - Stationarity and Wide-sense stationarity