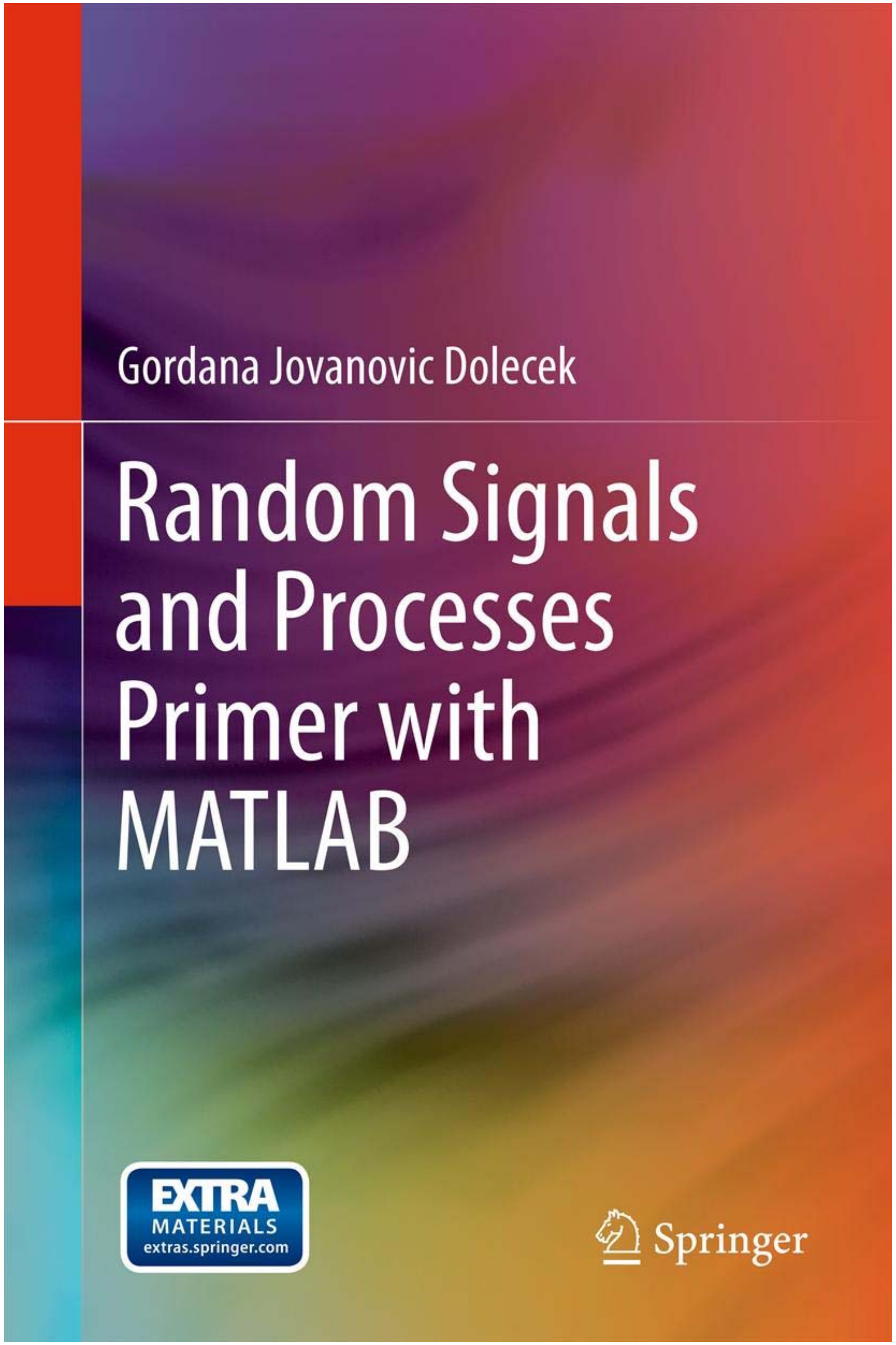
Everything you want to know about random variables and processes, but are afraid to ask

Gordana Jovanovic Dolecek
INAOE, Puebla, Mexico
gordana@ieee.org
http://www-elec.inaoep.mx/directorio/investigadores/gordana.php


North America
0 miles 250 500 750 1,000 miles
Puebla, Mexico
Text
ALASKA (U.S.A)
YUKON TERRITORY
NORTHWEST TERRITORIES
NUNAVUT TERRITORY
CANADA
BRITISH COLUMBIA
ALBERTA
SASKATCHEWAN
MANITOBA
ONTARIO
QUEBEC
NEWFOUNDLAND
NOVA SCOTIA
UNITED STATES OF AMERICA
MEXICO
CUBA
BAHAMAS
HAITI
DOMINICAN REP.
JAMAICA
BELIZE
Hudson Bay
Gulf of Alaska
Gulf of Mexico
NORTH PACIFIC OCEAN
LABRADOR SEA
Hawaiian Islands (U.S.A)
Rocky Mountains
Mackenzie Mountains
Sierra Madre Occidental
Sierra Madre du Sud
Baja California
Yucatán
TROPIC OF CANCER
ARCTIC CIRCLE


Gordana Jovanovic Dolecek
Random Signals and Processes Primer with MATLAB
EXTRA
MATERIALS
extras.springer.com
Springer

Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

Outline:

1.ONE RANDOM VARIABLE

1.1. Density and Distribution

1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE

2.1. Independence and Correlation

2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE

3.1. Properties and MATLAB generation

3.2. Central limit theorem

4. RANDOM PROCESS

4.1. What does AF tell us and why do we need it?

1.ONE RANDOM VARIABLE

1.1. Density and Distribution

- In order to completely describe a random variable, it is necessary to know not only all of the possible values it takes, but also how often it takes these values.
- In other words, it is necessary to know the corresponding probabilities.

1.ONE RANDOM VARIABLE
1.1. Density and Distribution

- The probability that random variable *X* has values less than or equal to *x* is called ***Distribution Function***, or ***Distribution***.
- Some authors also use the name ***Cumulative Distribution Function*** (***CDF***).

$$F_X(x) = P\{X \le x\}, \quad -\infty < x < \infty$$

1.ONE RANDOM VARIABLE
1.1. Density and Distribution

- ***A discrete random variable is only defined in discrete points. Does this mean that its distribution function only exists in discrete points?***

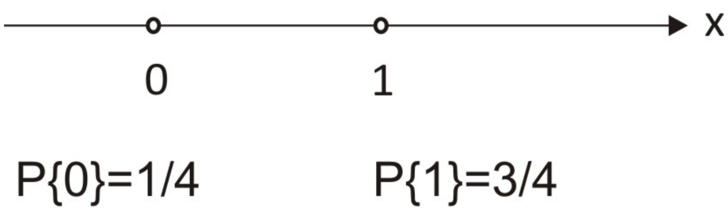
1.ONE RANDOM VARIABLE
1.1. Density and Distribution

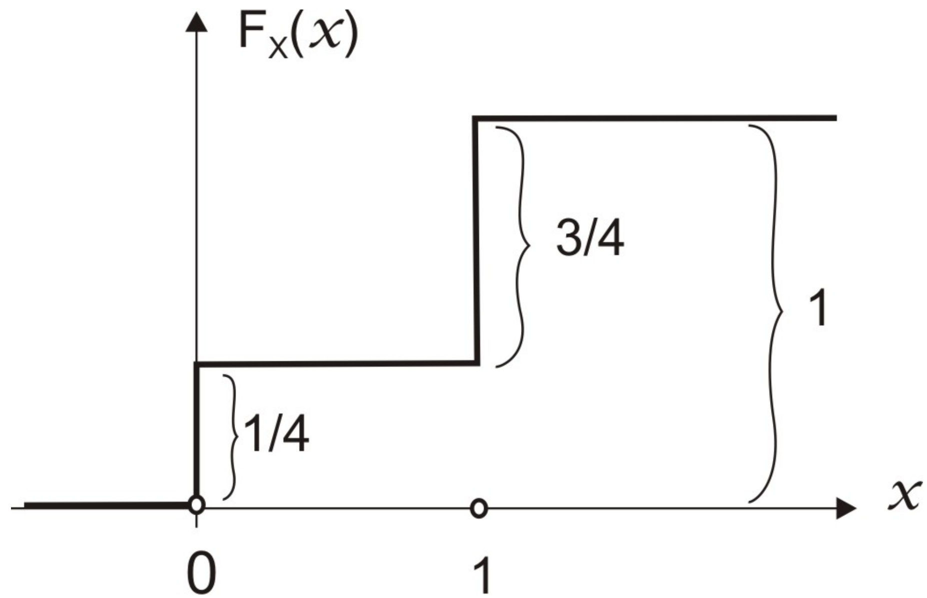
Example 1:

***P*{*X*=1}=*p*=3/4 ´,**
***P*{*X*=0}=*q*=1/4.**

$$F_X(x) = P\{X \le x\},$$

$$-\infty < x < \infty$$



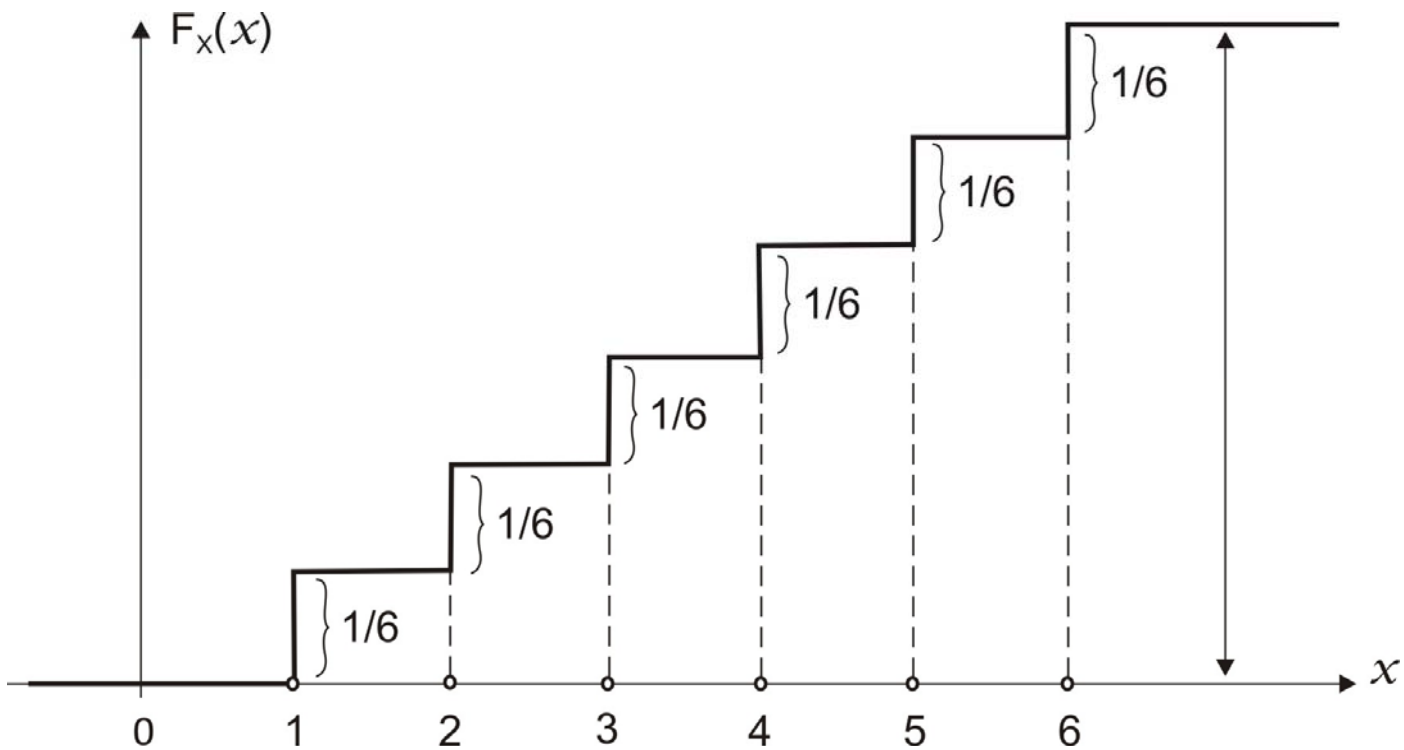


1.ONE RANDOM VARIABLE
1.1. Density and Distribution

$$F_X(x) = \sum_{k=-\infty}^{\infty} P(x_k)u(x - x_k)$$

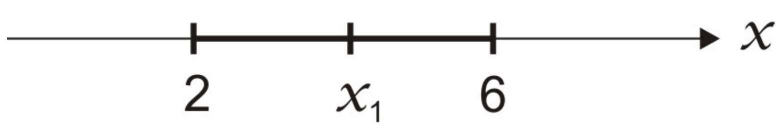
1.ONE RANDOM VARIABLE
1.1. Density and Distribution

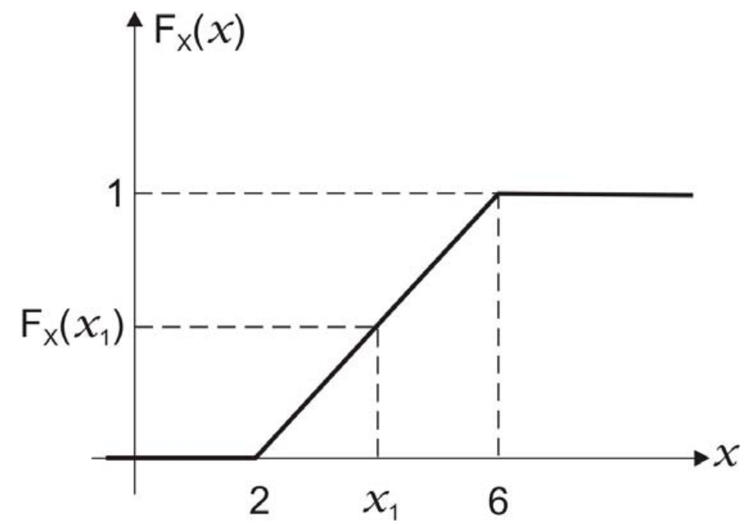
Example:



1.ONE RANDOM VARIABLE
1.1. Density and Distribution

- Continuous variable



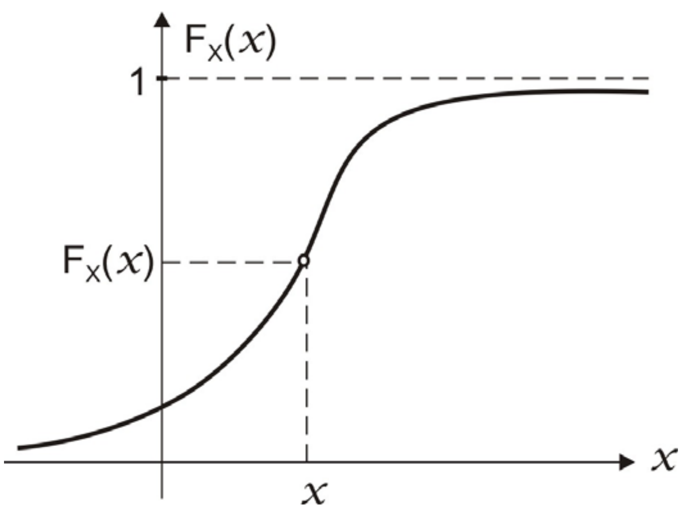


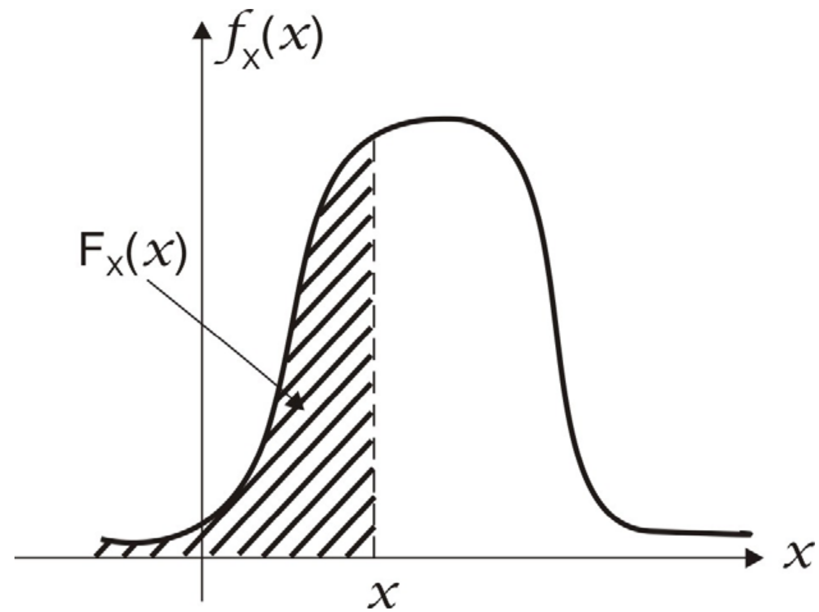
$$F_X(x) = \begin{cases} 0 & for \quad -\infty < x < 2 \\ \dfrac{x-2}{4} & for \quad 2 < x < 6 \\ 1 & for \quad 6 < x < \infty \end{cases}$$

1.ONE RANDOM VARIABLE
1.1. Density and Distribution

$$f_X(x) = \frac{dF_X(x)}{dx}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(v)dv$$





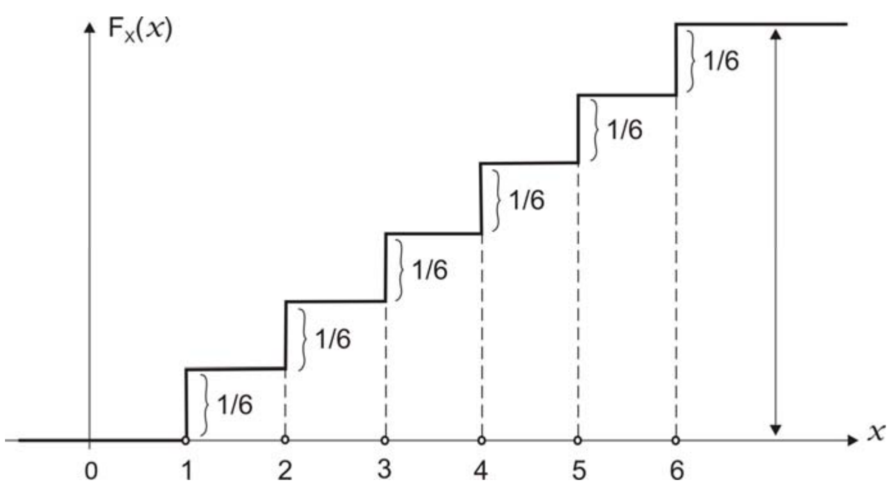
1.ONE RANDOM VARIABLE
1.1. Density and Distribution

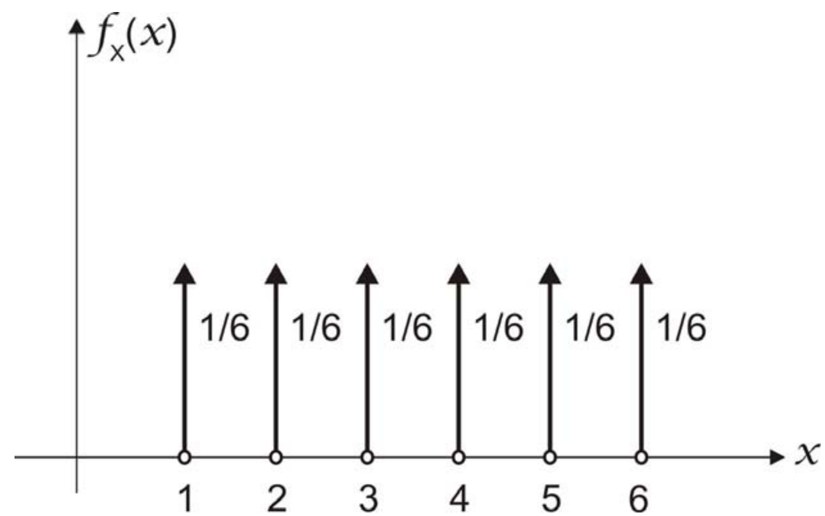
Example: Discret. Random variable

$$f_X(x) = \frac{d}{dx}\left(\sum_{i=-\infty}^{\infty} P\{x_i\}u(x - x_i)\right) = \sum_{i=-\infty}^{\infty} P\{x_i\}\frac{d}{dx}\left(u(x - x_i)\right)$$

$$f_X(x) = \sum_{i=-\infty}^{\infty} P\{x_i)\delta(x - x_i)$$

1.ONE RANDOM VARIABLE
1.1. Density and Distribution





1.ONE RANDOM VARIABLE
1.1. Density and Distribution

- ***Why must the PDF of a discrete random variable only have delta functions in the discrete values of a random variable?***

1.ONE RANDOM VARIABLE
1.1. Density and Distribution

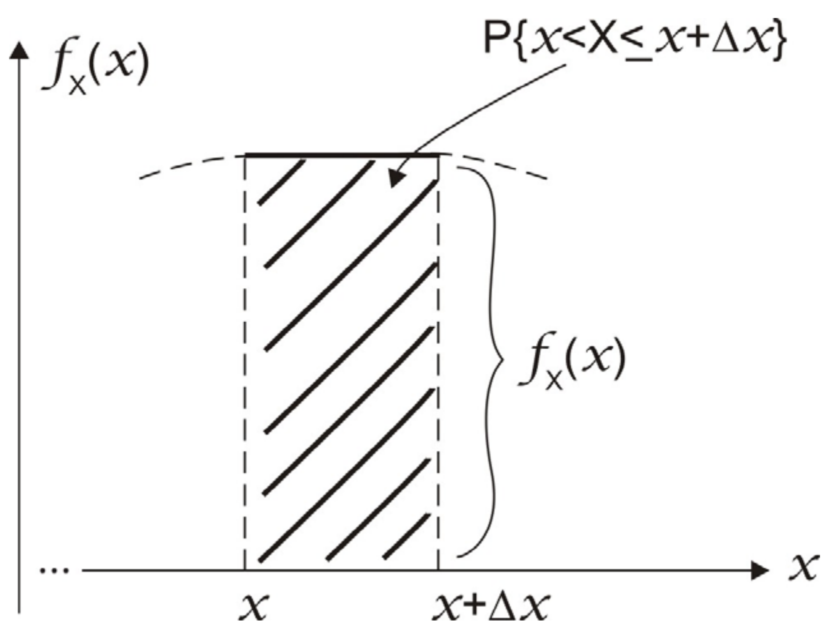
Density and probability

$$\frac{dF_X(x)}{dx} = \lim_{\Delta x \to 0} \frac{F_X(x + \Delta x) - F_X(x)}{\Delta x}$$

$$F_X(x + \Delta x) - F_X(x) = P\{x < X \le x + \Delta x\}$$

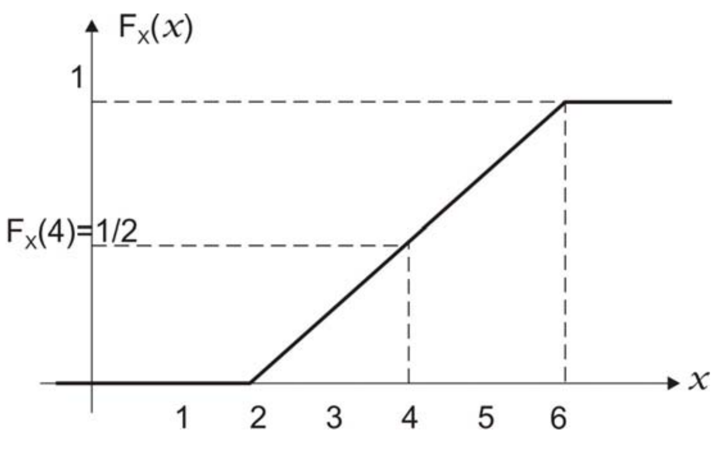
$$f_X(x) = \lim_{\Delta x \to 0} \frac{P\{x < X \le x + \Delta x\}}{\Delta x} = \frac{P\{x < X \le x + dx\}}{dx}$$

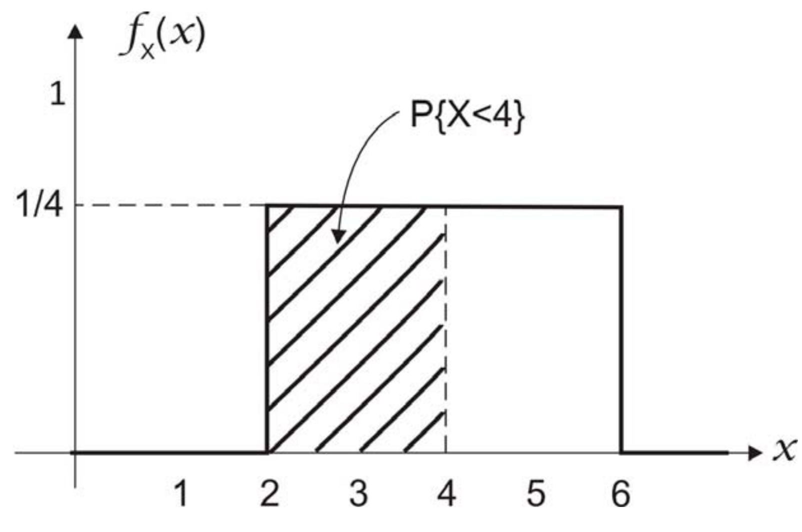
1.ONE RANDOM VARIABLE
1.1. Density and Distribution



1.ONE RANDOM VARIABLE
1.1. Density and Distribution

Example: Cont random variable





1.ONE RANDOM VARIABLE
1.1. Density and Distribution

- D_2_DEMO Probability, Density and Distribution

Outline:

1.ONE RANDOM VARIABLE

1.1. Density and Distribution

1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE

2.1. Independence and Correlation

2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE

3.1. Properties and MATLAB generation

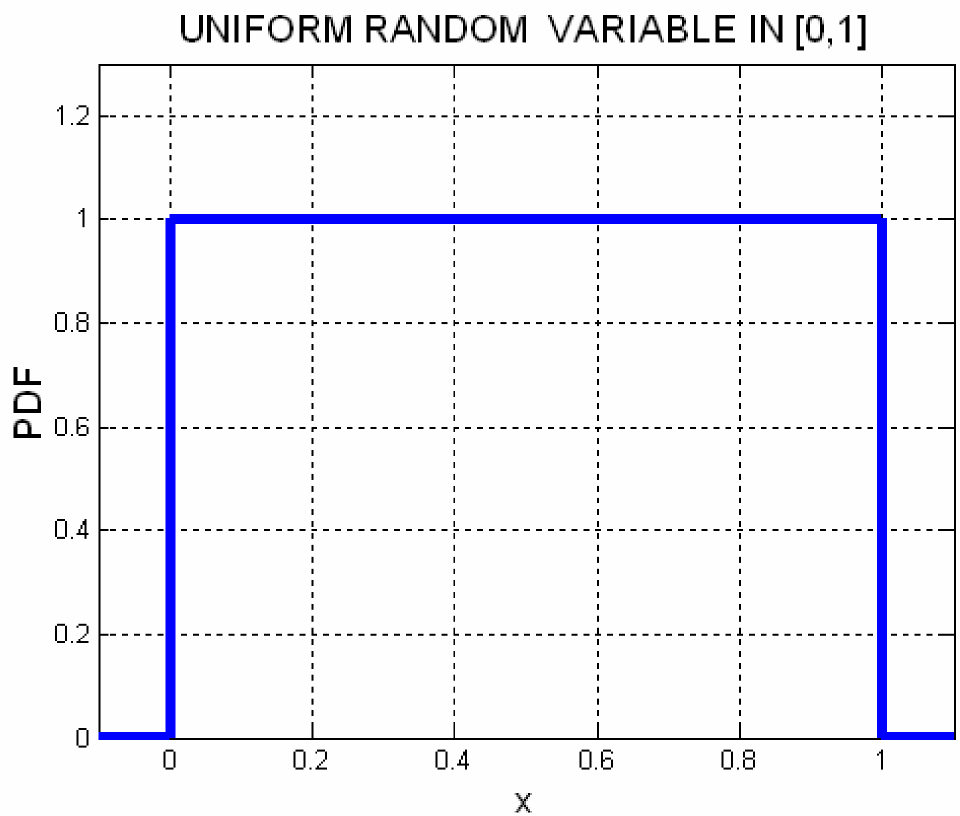
3.2. Central limit theorem

4. RANDOM PROCESS

4.1. What does AF tell us and why do we need it?

1.ONE RANDOM VARIABLE

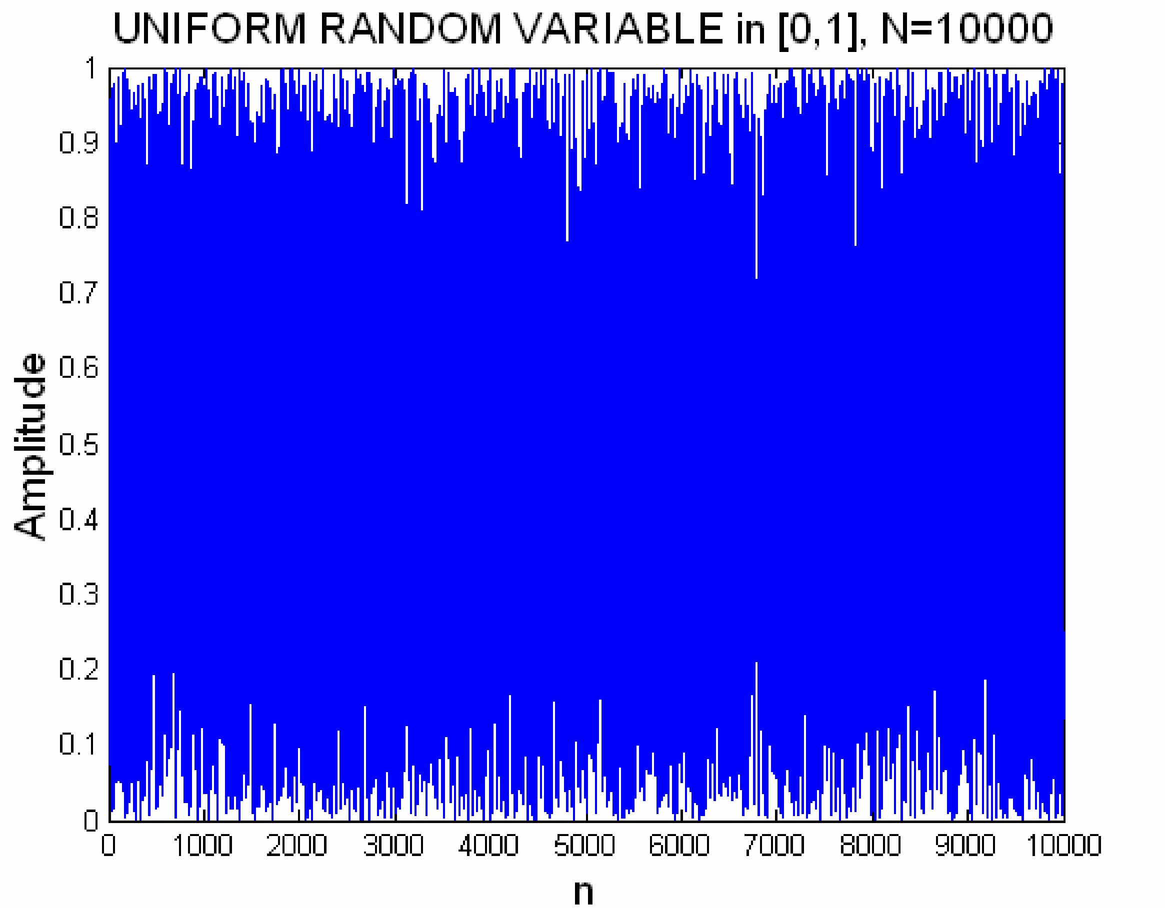
1.2. How to estimate Density and Distribution in MATLAB?



1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?

- X=rand(1,10000);



1.ONE RANDOM VARIABLE

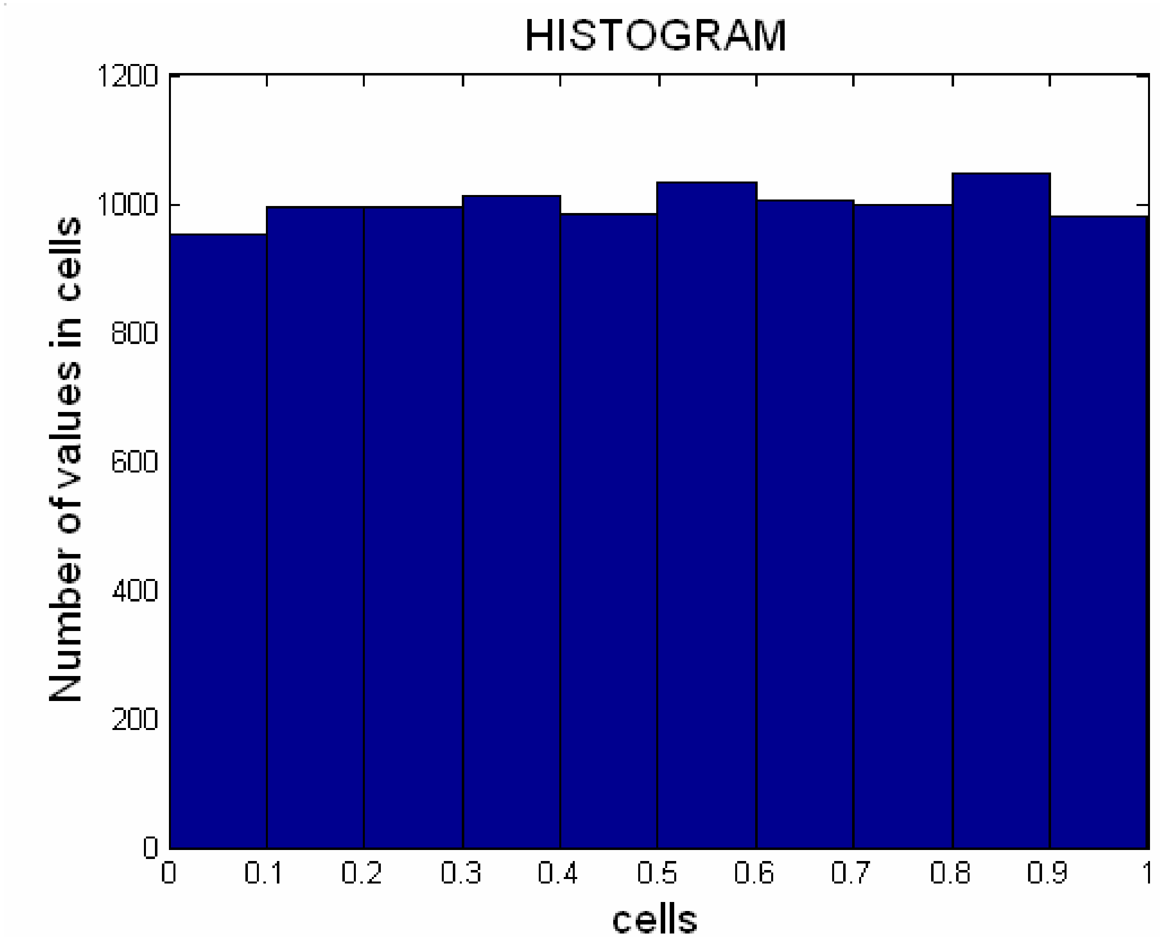
1.2. How to estimate Density and Distribution in MATLAB?

The ***histogram,*** *NN=hist(x,M)* shows the values N_i, $i=1,..,M$, (i.e., how many values of the random variable X are in each cell), where M is the number of cells. If we use $M=10$, we get the following result:

NN = [952, 994, 993, 1012, 985, 1034, 1006, 998, 1047, 979].

1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?



1.ONE RANDOM VARIABLE

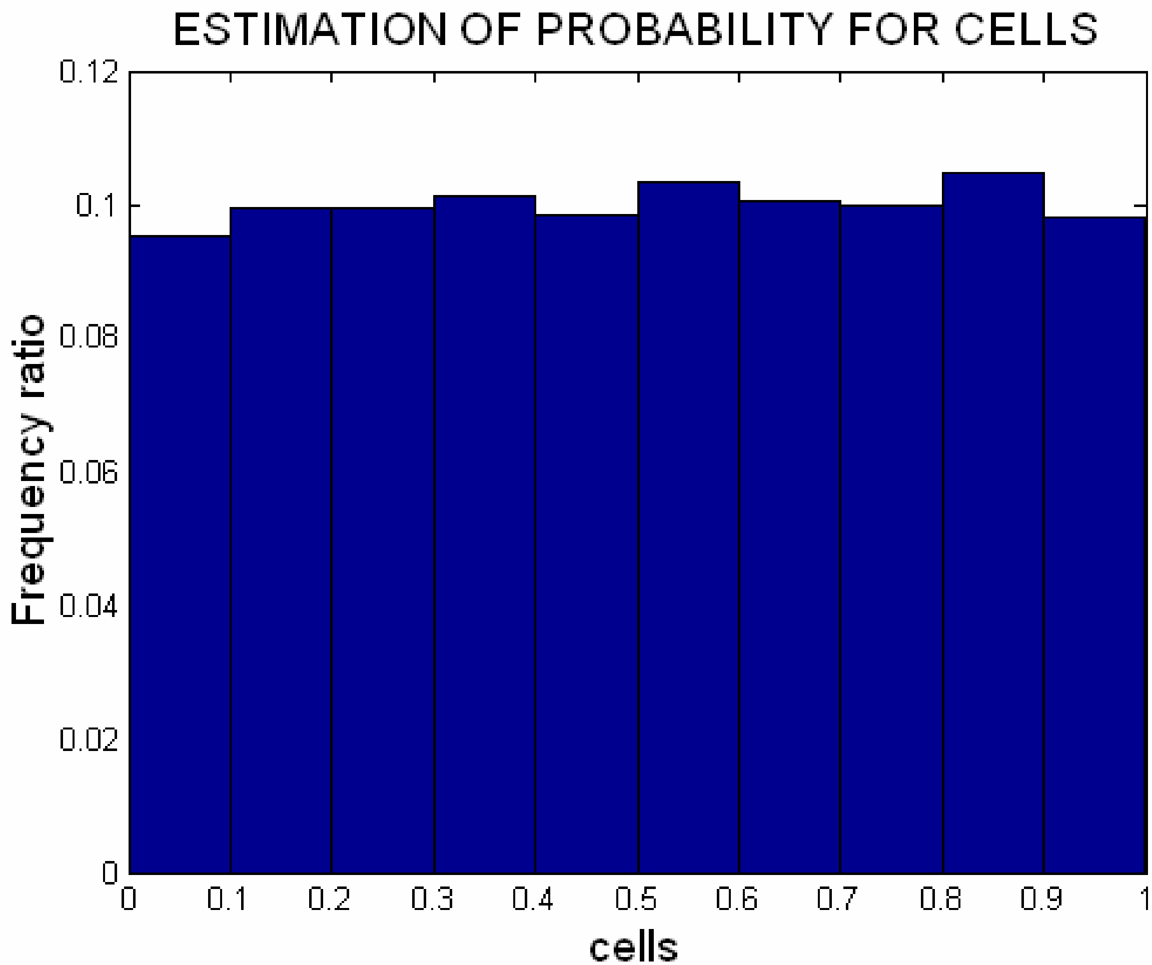
1.2. How to estimate Density and Distribution in MATLAB?

- Let *Ni* be the number of values of the random variable *X* in the *i*-th cell. Then, the probability that the random variable belongs to the *i*-th cell is approximated by a quantity *Ni/N*, called the ***frequency ratio***,

$$P\{(i-1)\Delta x < X \le i\Delta x\} \approx N_i / N = hist(X, M) / N$$

1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?



1.ONE RANDOM VARIABLE

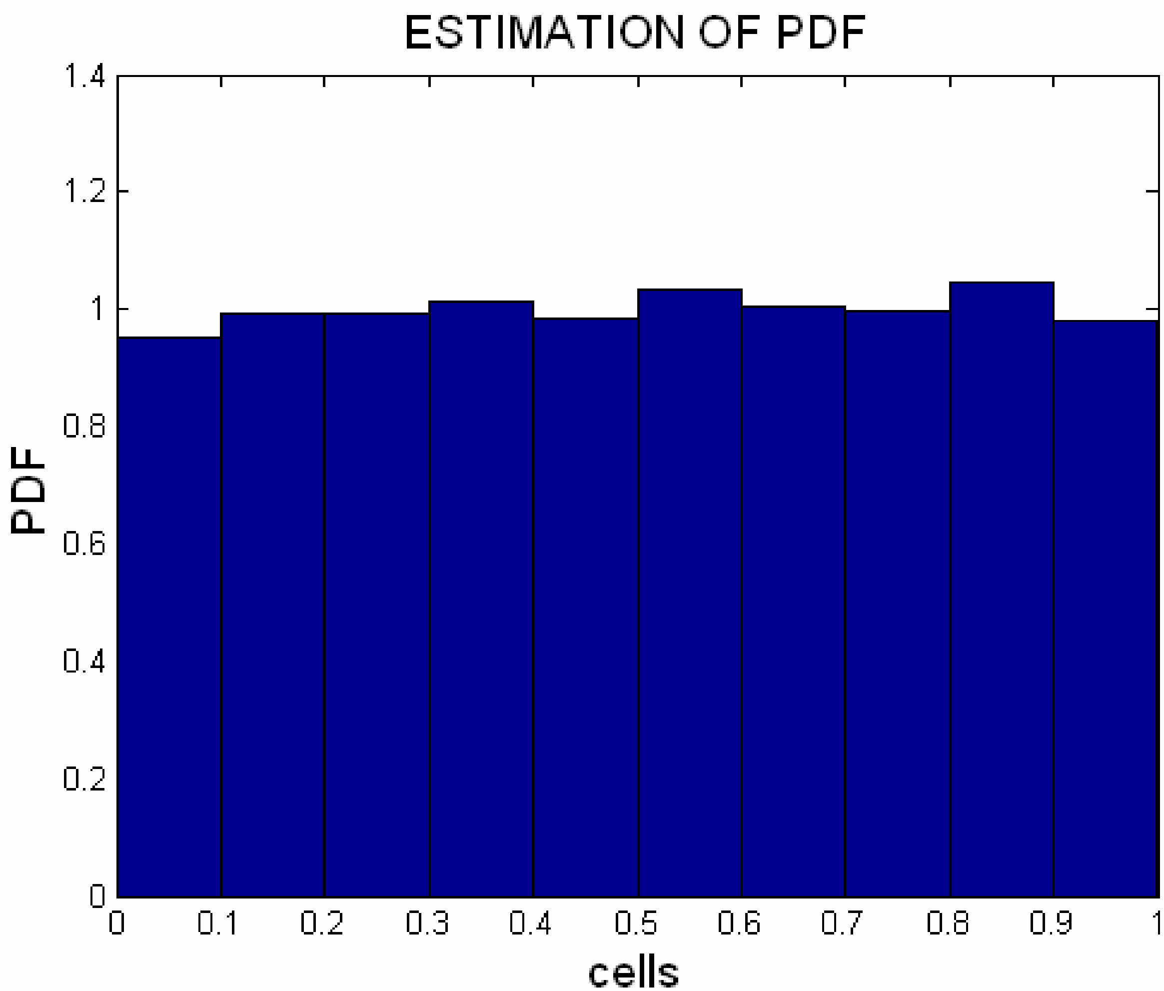
1.2. How to estimate Density and Distribution in MATLAB?

$$f_X(i\Delta x) \approx \frac{P\{(i-1)\Delta x < X \leq i\Delta x\}}{\Delta x} = \frac{N_i}{N}\frac{1}{\Delta x}$$

PDF$_i$=$P\{X$ belongs to the cell $i\}/\Delta x$=. $\dfrac{hist(x,M)}{N\Delta x}$

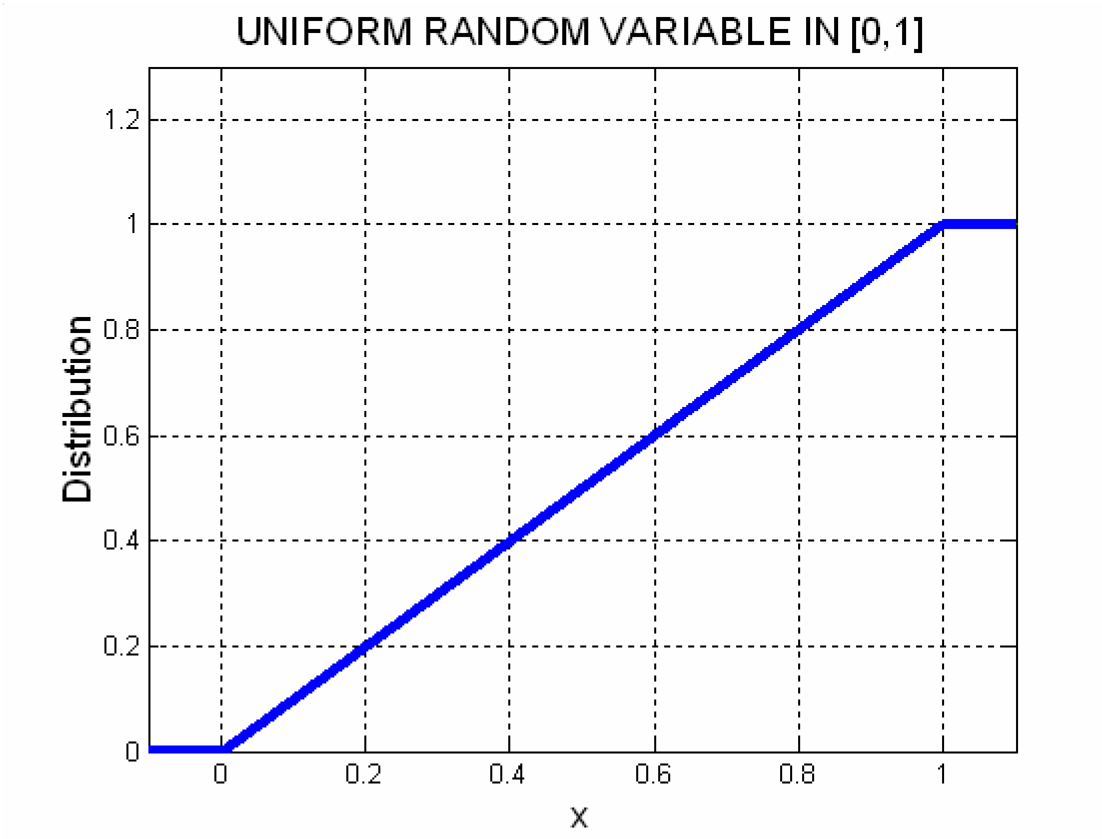
1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?



1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?



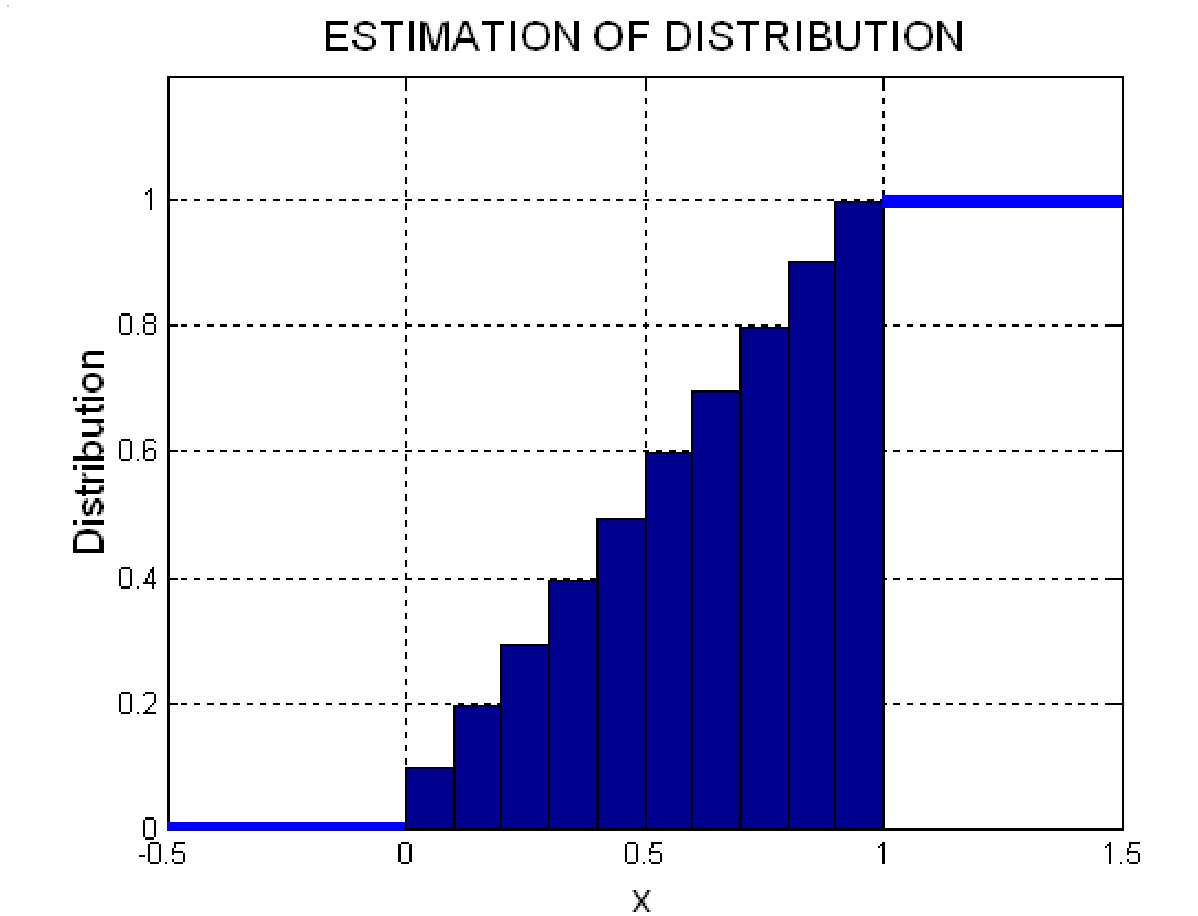
1.ONE RANDOM VARIABLE

1.2. How to estimate Density and Distribution in MATLAB?

- The MATLAB function *cumsum.m* performs this estimation .

1.ONE RANDOM VARIABLE

1.4. How to estimate Density and Distribution in MATLAB?



Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

- Random variables are independent if one variable does not have any influence over the values of another variable, and vice versa.
- For nonrandom variables dependency means that if we know one variable, we can find the exact values of another variable.
- However, dependence between random variables can have different degrees of dependency.

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

- To this end, the dependence between variables is expressed using some characteristics, like ***covariance***.
- However, the covariance contains the information, not only of dependency of random variables, but also the information about the dissipation of variables around their mean values.
- If, for example, the dissipations of random variables X_1 and X_2 around their mean values were very small, then the covariance $C_{X_1X_2}$ would be small for any degree of dependency in the variables.
- This problem is solved by introducing the ***correlation coefficient*** ($\rho_{X_1 X_2}$)

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

$$\rho_{X_1X_2} = \frac{C_{X_1X_2}}{\sigma_{X_1}\sigma_{X_2}} = \frac{\overline{(X_1 - \overline{X_1})(X_2 - \overline{X_2})}}{\sigma_{X_1}\sigma_{X_2}} = \frac{\overline{X_1X_2} - \overline{X_1}\,\overline{X_2}}{\sigma_{X_1}\sigma_{X_2}}$$

What are the values that the correlation coefficient can take?

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

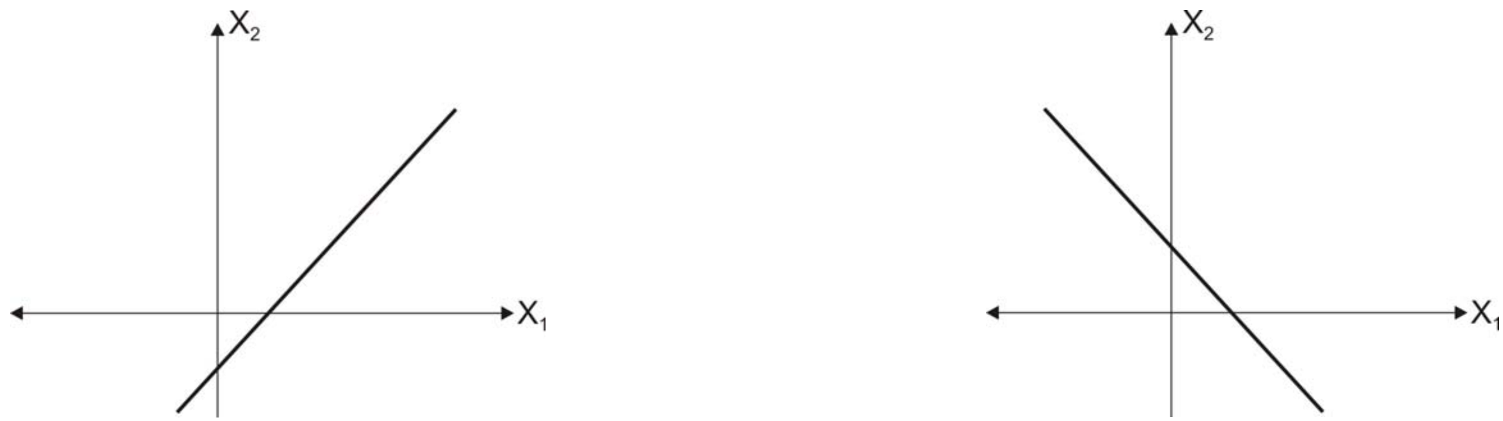
a. The variables are equal

$$\rho_{X_1X_2} = \frac{C_{X_1X_2}}{\sigma_{X_1}\sigma_{X_2}} = \frac{\overline{(X_1 - \overline{X_1})(X_1 - \overline{X_1})}}{\sigma_{X_1}\sigma_{X_1}} = \frac{\overline{(X_1 - \overline{X_1})^2}}{\sigma_{X_1}^2} = \frac{\sigma_{X_1}^2}{\sigma_{X_1}^2} = 1$$

b. Linear dependency of random variables

$X_2 = aX_1 + b$

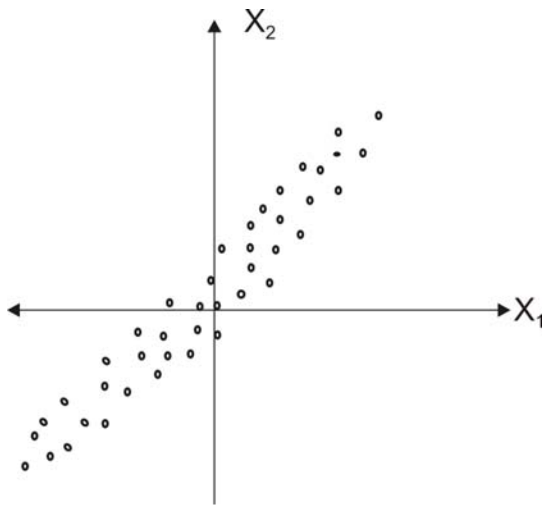
2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation



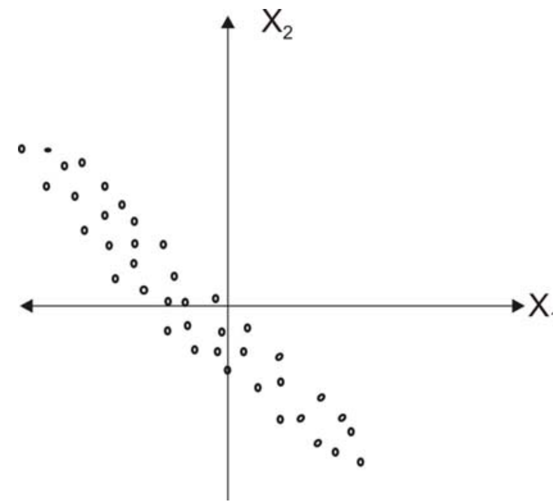
$$\rho_{X_1X_2} = \frac{C_{X_1X_2}}{\sigma_{X_1}\sigma_{X_2}} = \frac{a\sigma_{X_1}^2}{\sigma_{X_1}|a|\sigma_{X_1}} = \frac{a}{|a|} = \begin{cases} 1 & for \quad a > 0 \\ -1 & for \quad a < 0 \end{cases}$$

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

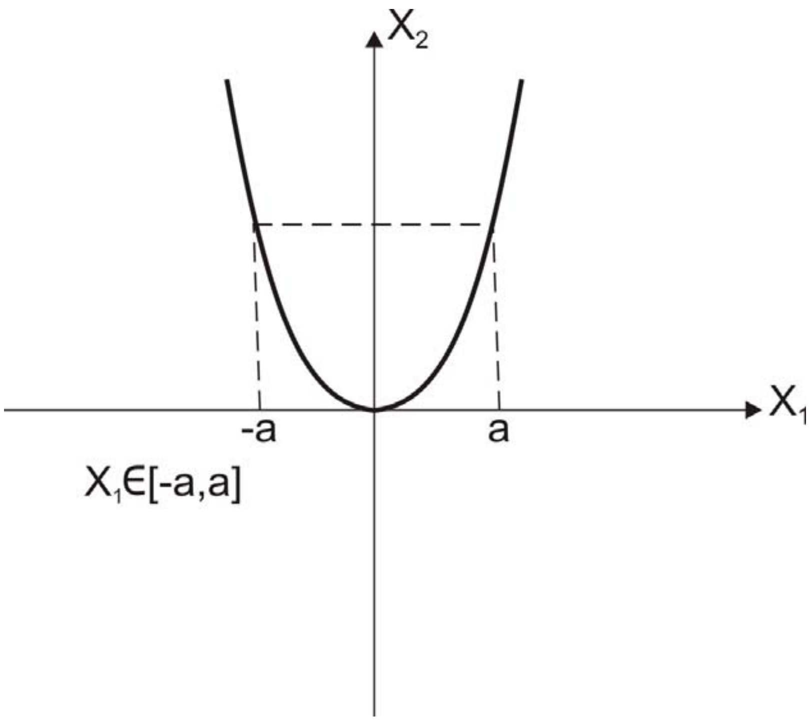
$\rho_{X_1 X_2} > 0$

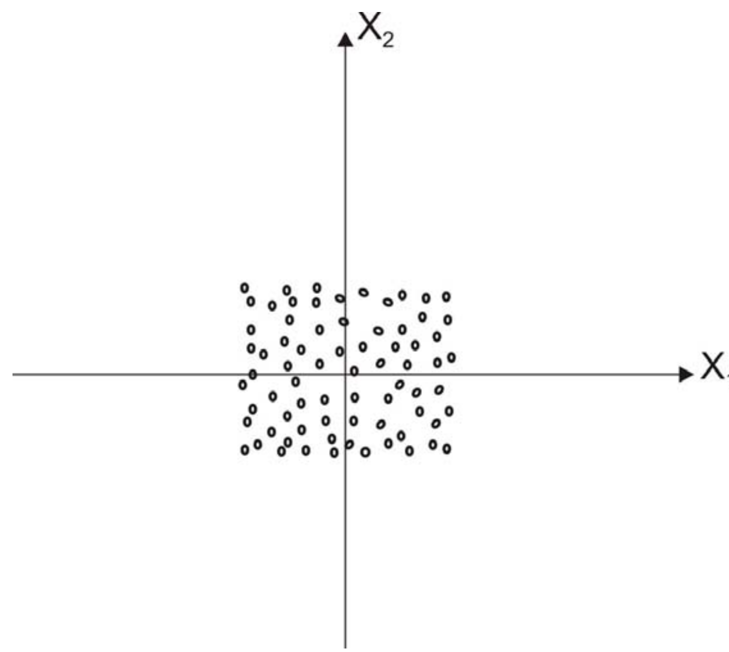


$\rho_{X_1 X_2} < 0$

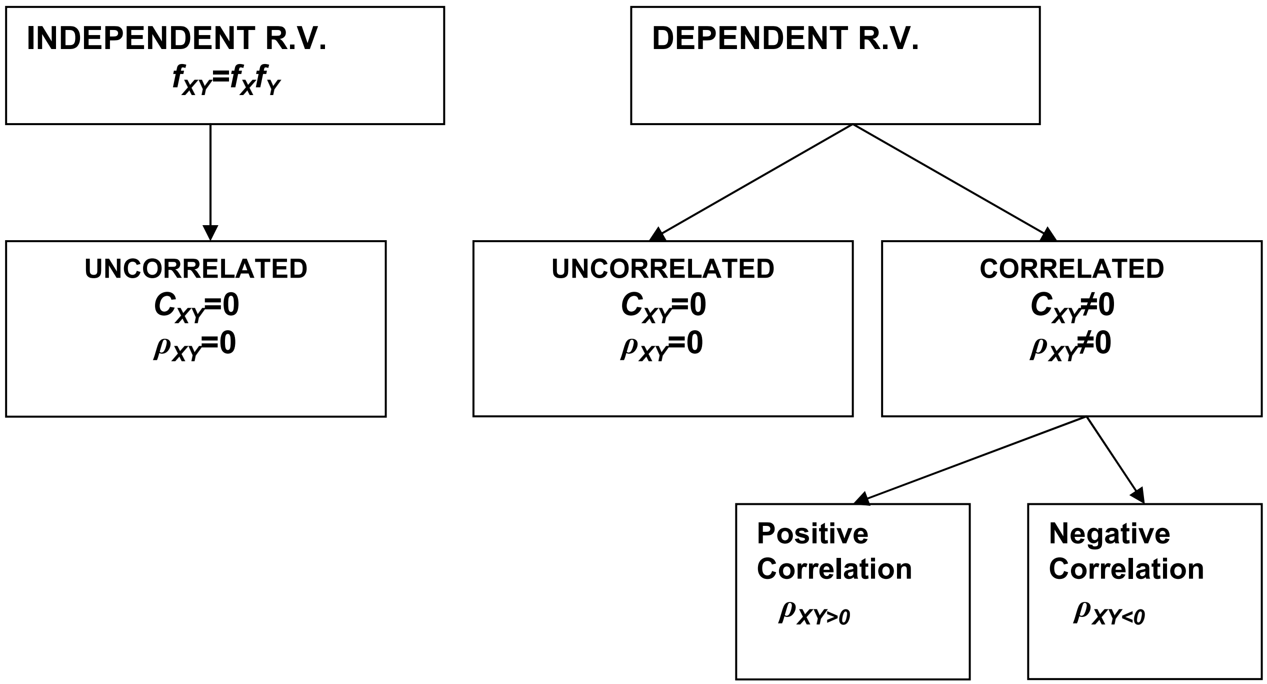


2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation





2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation



2.TWO-DIMENSIONAL RANDOM VARIABLE
2.1.Independence and Correlation

- ***Why is the covariance not zero if the variables are correlated?***

$$C_{X_1X_2} = E\{(X_1 - \overline{X_1})(X_2 - \overline{X_2})\}$$

Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

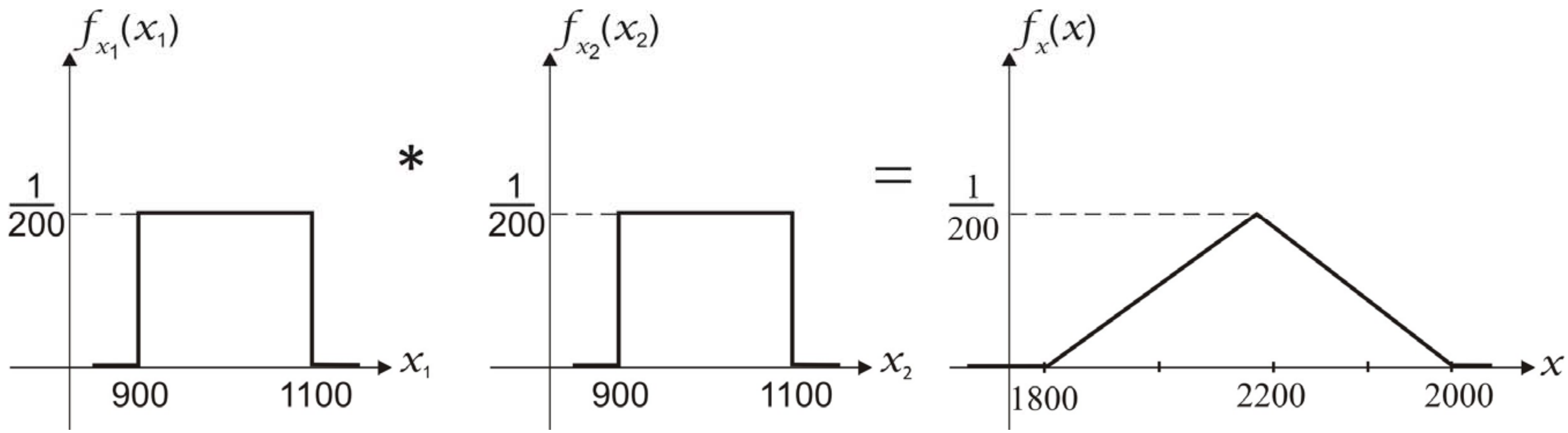
2.TWO-DIMENSIONAL RANDOM VARIABLE
2.2.PDF of sum of independent r.v.

- $Y = X_1 + X_2$.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2}(x_2)\, f_{X_1}(y - x_2)\, dx_2$$

$$f_Y(y) = f_{X_1}(x_1) * f_{X_2}(x_2) = f_{X_2}(x_2) * f_{X_1}(x_1)$$

2.TWO-DIMENSIONAL RANDOM VARIABLE
2.2.PDF of sum of independent r.v.



2.TWO-DIMENSIONAL RANDOM VARIABLE
2.2.PDF of sum of independent r.v.

D_3_1_DEMO.m: Sum of independent random variables

Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

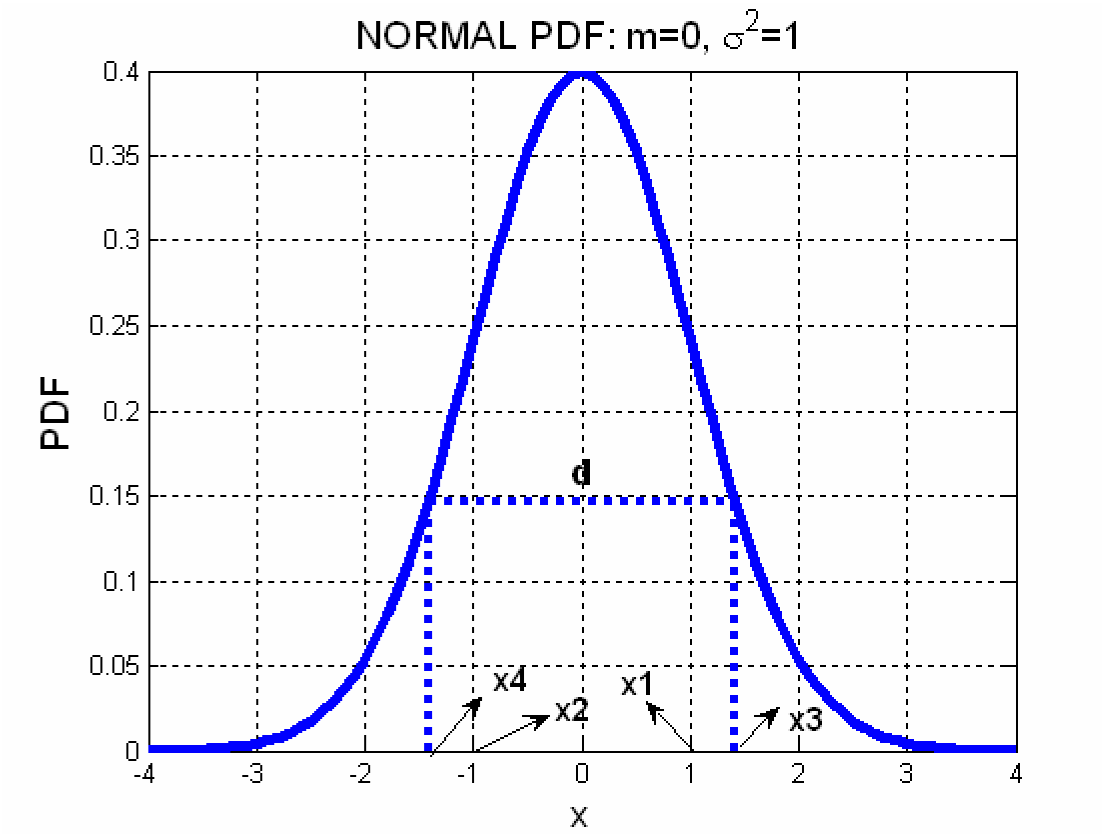
3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation

- PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

$$f_X(x) = N(m, \sigma^2)$$

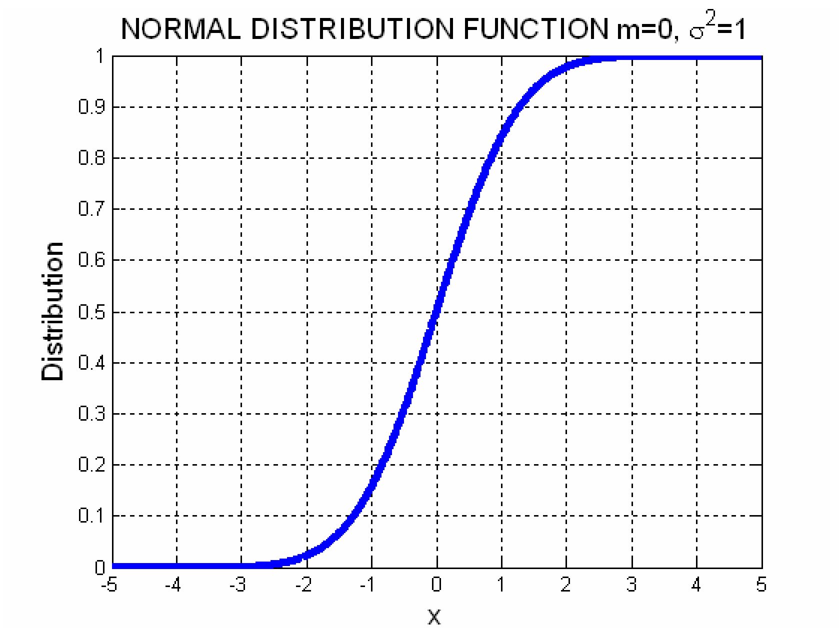
3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation



3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation

- Distribution

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(x)dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(x-m)^2}{2\sigma^2}}dx$$



3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du$$

$$F_X(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-m}{\sqrt{2}\sigma}\right)\right]$$

$$P\{m-k\sigma \le X \le m+k\sigma\} = \frac{1}{2}\left[erf\left(\frac{m+k\sigma-m}{\sqrt{2}\sigma}\right) - erf\left(\frac{m-k\sigma-m}{\sqrt{2}\sigma}\right)\right] = erf\left(\frac{k}{\sqrt{2}}\right).$$

3 3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation

- **P. 1:** ***The PDF is symmetrical around its mean value*.**
- **P. 2:** ***The PDF has a maximum value in x=m*** .
- ***P.3:The“3 σ rule”***

$$P\{m-3\sigma \le X \le m+3\sigma\} = 0.9973$$

3. NORMAL RANDOM VARIABLES
3.1. Properties and MATLAB generation

- ***P.4: Sum of normal variables are noncorrelated , they are independent.les is normal.***
- ***P.5. Linear transform of normal is normal.***
- ***P.6. If normal variables are noncorrelated, they are independent.***

3. NORMAL RANDOM VARIABLES
How to generate a normal r.v. in MATLAB?

We consider the generation in MATLAB of normal variable with an arbitrary mean value and variance.

However, the standard MATLAB file ***randn.m*** generates the normal random variable with the mean value 0 and variance 1.

3. NORMAL RANDOM VARIABLES
How to generate a normal r.v. in MATLAB?

$$Y = aX + b, X = randn(1, N)$$

$$m_Y = am_X + b$$

$$\sigma_Y^2 = a^2 \sigma_X^2$$

$m_Y = b; \quad \sigma_Y^2 = a^2$

$$Y = \sigma_Y X + m_Y.$$

3. NORMAL RANDOM VARIABLES

- D_4_1_DEMO.m: Normal density and distribution
- D_4_2_DEMO.m Linear transformation of normal random variable
- D_4_3_DEMO.m Sum of normal variables

Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

3. NORMAL RANDOM VARIABLES
3.2. Central Limit Theorem

- According to the CLT ***the sum of N independent random variables Xi, each of which contributes a small amount to the total, approaches the normal random variable.***

$$X = \sum_{i=1}^{N} X_i$$

$$f_X(x) = f_{X_1}(x) * f_{X_2}(x) * \ldots * f_{X_N}(x)$$

3. NORMAL RANDOM VARIABLES
3.2. Central Limit Theorem

$$m_X = \sum_{i=1}^{N} m_{X_i} \qquad \sigma_X^2 = \sum_{i=1}^{N} \sigma_{X_i}^2$$

$$\lim_{N \to \infty} f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

3. NORMAL RANDOM VARIABLES
3.2. Central Limit Theorem

- D_4_5_DEMO_1.m Central limit theorem

Outline:

1.ONE RANDOM VARIABLE
1.1. Density and Distribution
1.2. How to estimate Density and Distribution in MATLAB?

2. TWO-DIMENSIONAL RANDOM VARIABLE
2.1. Independence and Correlation
2.2. PDF of the sum of independent random variables

3. NORMAL RANDOM VARIABLE
3.1. Properties and MATLAB generation
3.2. Central limit theorem

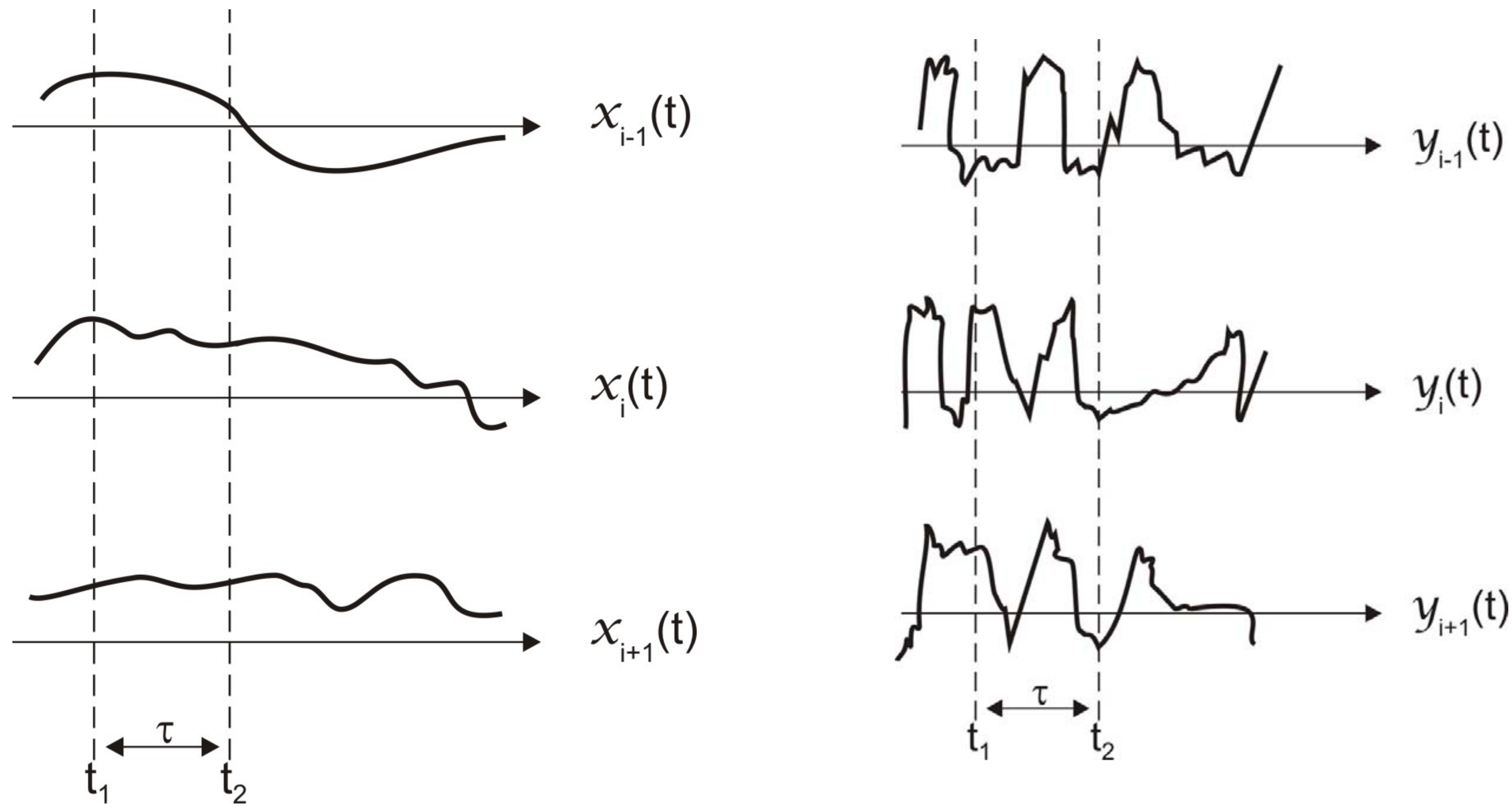
4. RANDOM PROCESS
4.1. What does AF tell us and why do we need it?

4. RANDOM PROCESS

4.1.What does AF tell us and why do we need it?

- From the definition of an autocorrelation function we know that an autocorrelation function is a statistical characteristic of a process that is obtained by observing the process in two points, the same as a joint density function of a second order.
- The question which arises is: “**why do we need an autocorrelation function if we already have the second order joint density function**?”

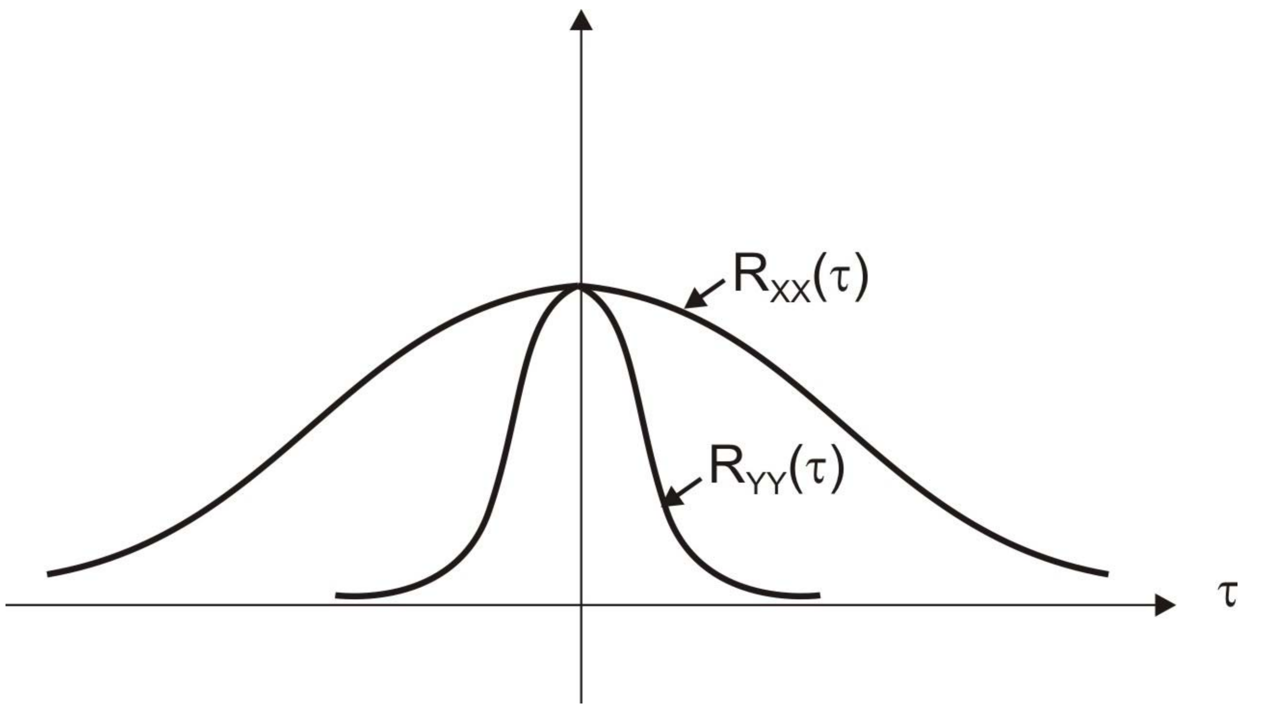
4.1.What does AF tell us and why do we need it?



4.1.What does AF tell us and why do we need it?

$$R_{XX}(\tau) = \overline{X(t)X(t+\tau)} = \int_{-\infty}^{\infty} x_1 x_2 f_{X_1X_2}(x_1, x_2; \tau) dx_1 dx_2$$

$$R_{YY}(\tau) = \overline{Y(t)Y(t+\tau)} = \int_{-\infty}^{\infty} y_1 y_2 f_{X_1X_2}(y_1, y_2; \tau) dy_1 dy_2$$



4.1.What does AF tell us and why do we need it?

- **P. 1:** ***The maximum value of the autocorrelation function is at τ=0***.
- **P. 2:** ***An autocorrelation function is an even function***.
- **P. 3:** ***The value of an autocorrelation function in τ=0 is a mean squared value of a process, that is a power of the process***,

$$R_{XX}(0) = \overline{X^2(t)}$$

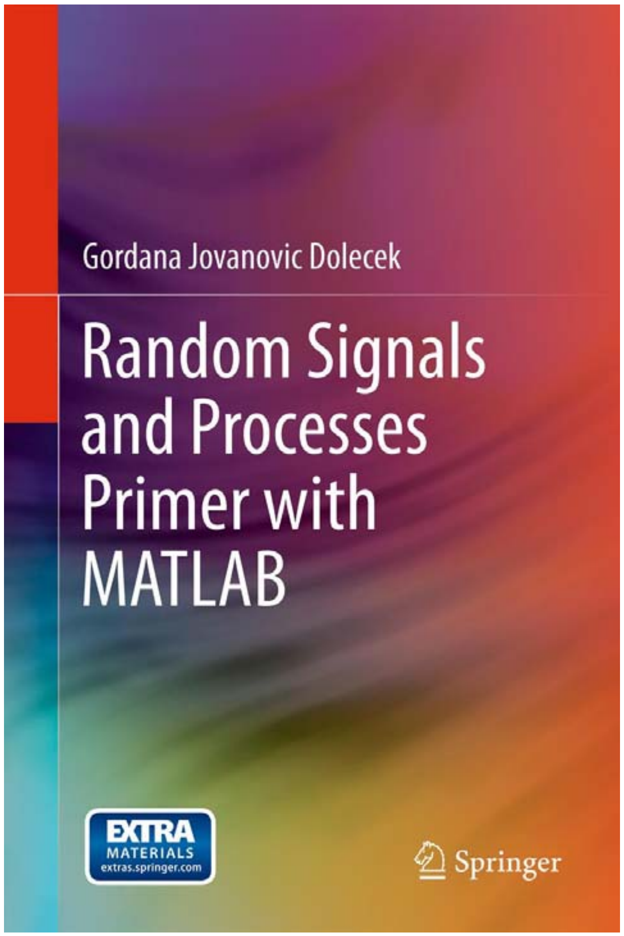
4.1.What does AF tell us and why do we need it?

- **P. 4:** ***If the mean value of the process is not zero, then the autocorrelation function has a constant term which is equal to the squared mean value.***
- ***P. 5: If a random process is a periodic process, then its autocorrelation function also has a periodic component of the same period as the process itself.***

4. RANDOM PROCESS
4.1.What does AF tell us and why do we need it?

- D_6_1_DEMO.m: Autocorrelation function for WSS processes

Everything you want to know about random variables and processes, but are afraid to ask



HVALA!!!!

gordana@ieee.org