Random signals

Jan Černocký

ÚPGM FIT VUT Brno, cernocky@fit.vutbr.cz

Random signals

- deterministic signals (can be represented by an equation) have one substantial drawback – they carry very little information (for example cosine: amplitude, frequency, phase shift).
- real-world signals are tough to be described as deterministic (for example physical model of speech production is very complex and anyway simplified).

$\Rightarrow$ in signal theory, we will consider these **useful** signals as **random signals (processes)** (for example speech, circulation of letters in agencies of Czech Posts, exchange rate CZK/EUR ...).

According to the character of the time axis, random signals are divided into **continuous time** random signals (the time is defined for all t) and **discrete time** random signals (only for discrete n).

These signals can not be represented in all time-points (in this case, they would be deterministic), we will rather look for characteristic properties of random signals such as mean value, probability density function, etc.

Definition of random process

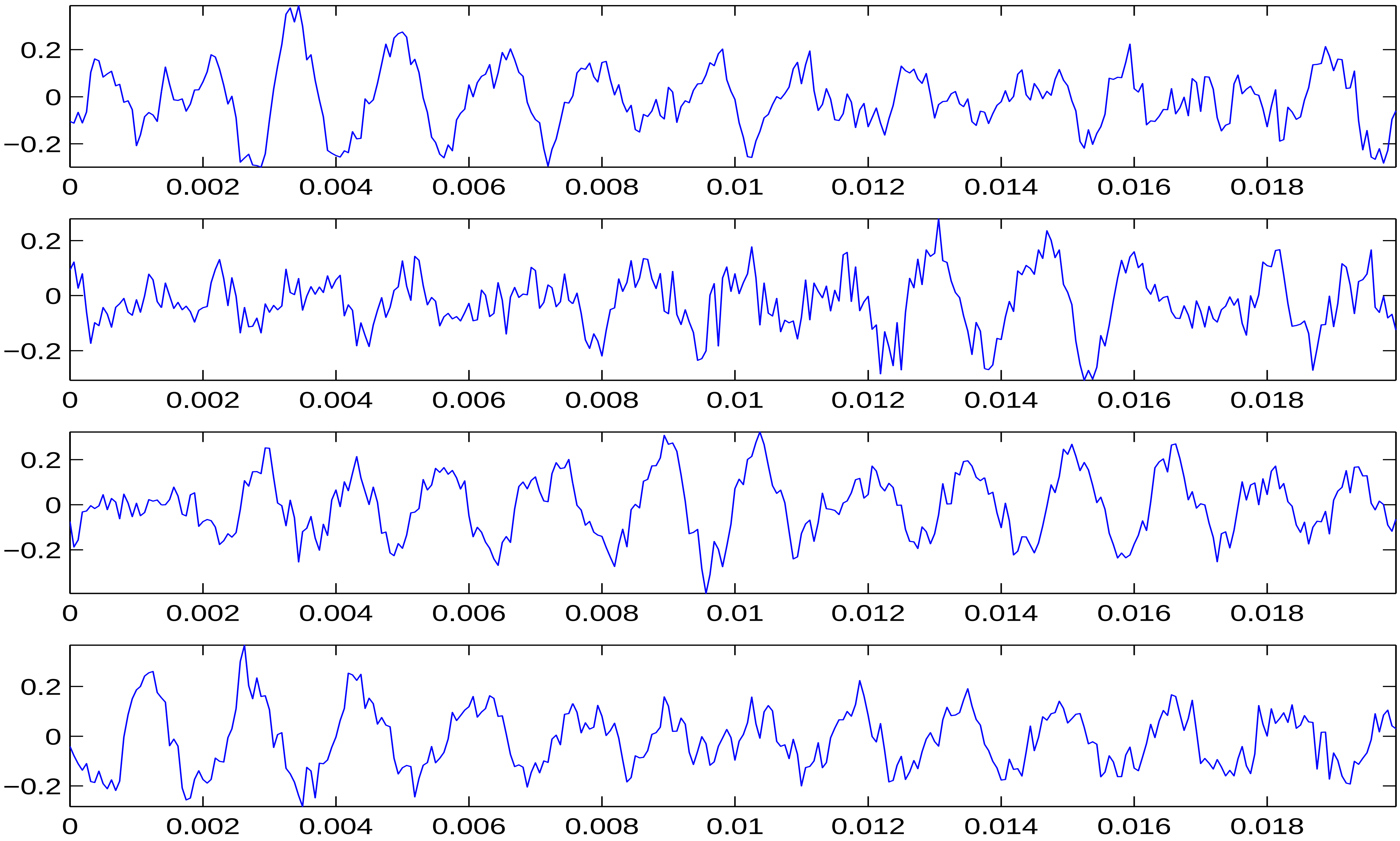
- continuous time: the system $\{\xi_t\}$ of random variables defined for all $t \in \Re$ is called random process, it is denoted $\xi(t)$.
- discrete time: the system $\{\xi_n\}$ of random variables defined for all $n \in N$ is called random process, it is denoted $\xi[n]$.

Set of realizations of random process

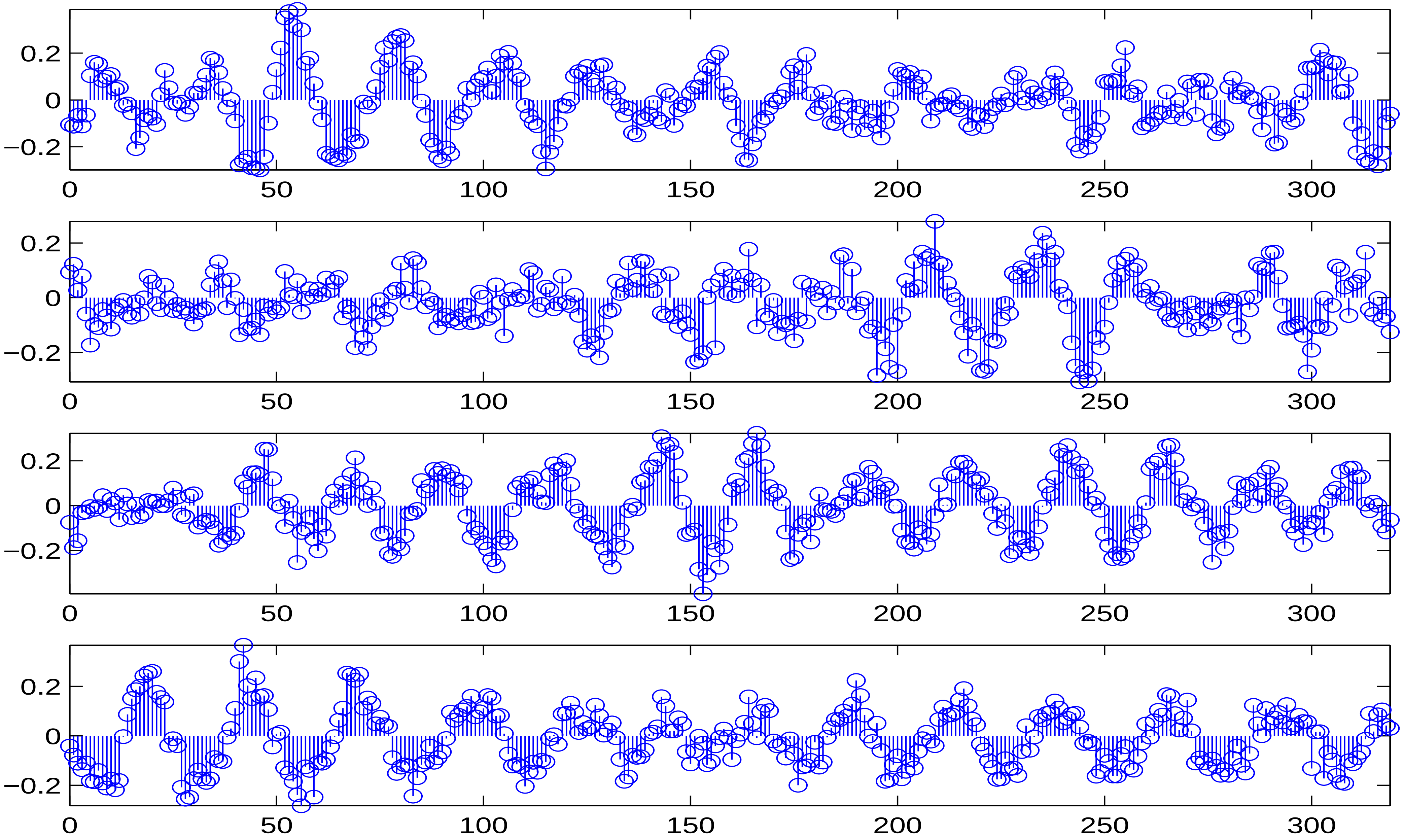
the set of realizations includes infinity of possible "runs" of random process - its **realizations**. We will limite ourselves to finite number Ω and denote each realization as $\xi_\omega(t)$, or $\xi_\omega[n]$. In case any parameters are estimated on this set, we will speak about **ensemble estimates**.

Example: the random signal is recording of water flowing through the water tube in my flat. 1068 realizations, each of 20 ms, were recorded. For the demonstration of continuous random signals, we will imagine this set as $\xi_\omega(t)$, for discrete random signals as $\xi_\omega[n]$.

$\xi_\omega(t)$ for $\omega = 1,\ 200,\ 500,\ 1000$



$\xi_\omega[n]$ for $\omega = 1,\ 200,\ 500,\ 1000$



Distribution function

is defined for one random variable: the random process for given time t or n **is** such a random variable. Definition:

$$F(x,t) = \mathcal{P}\{\xi(t) < x\},$$

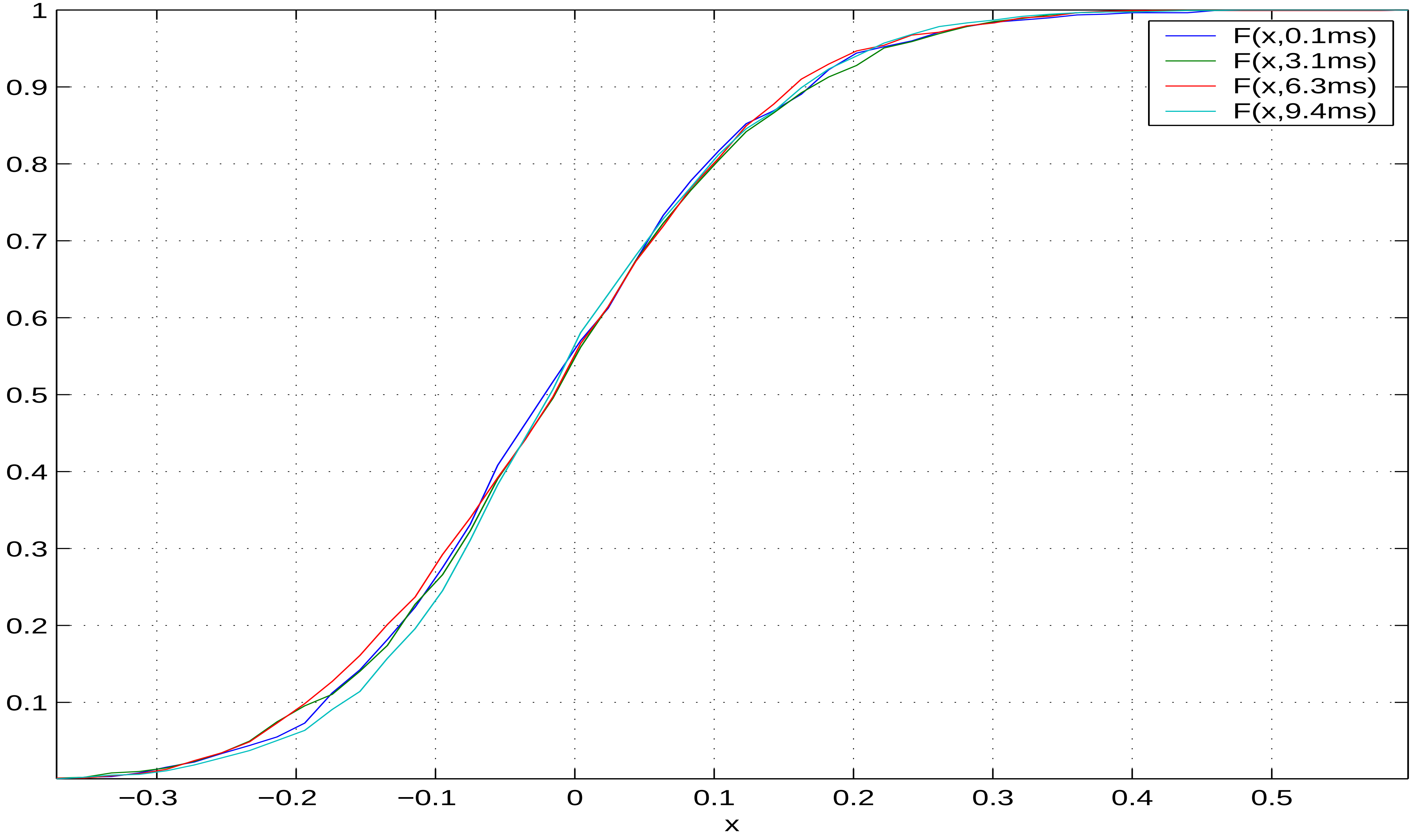
$$F(x,n) = \mathcal{P}\{\xi[n] < x\},$$

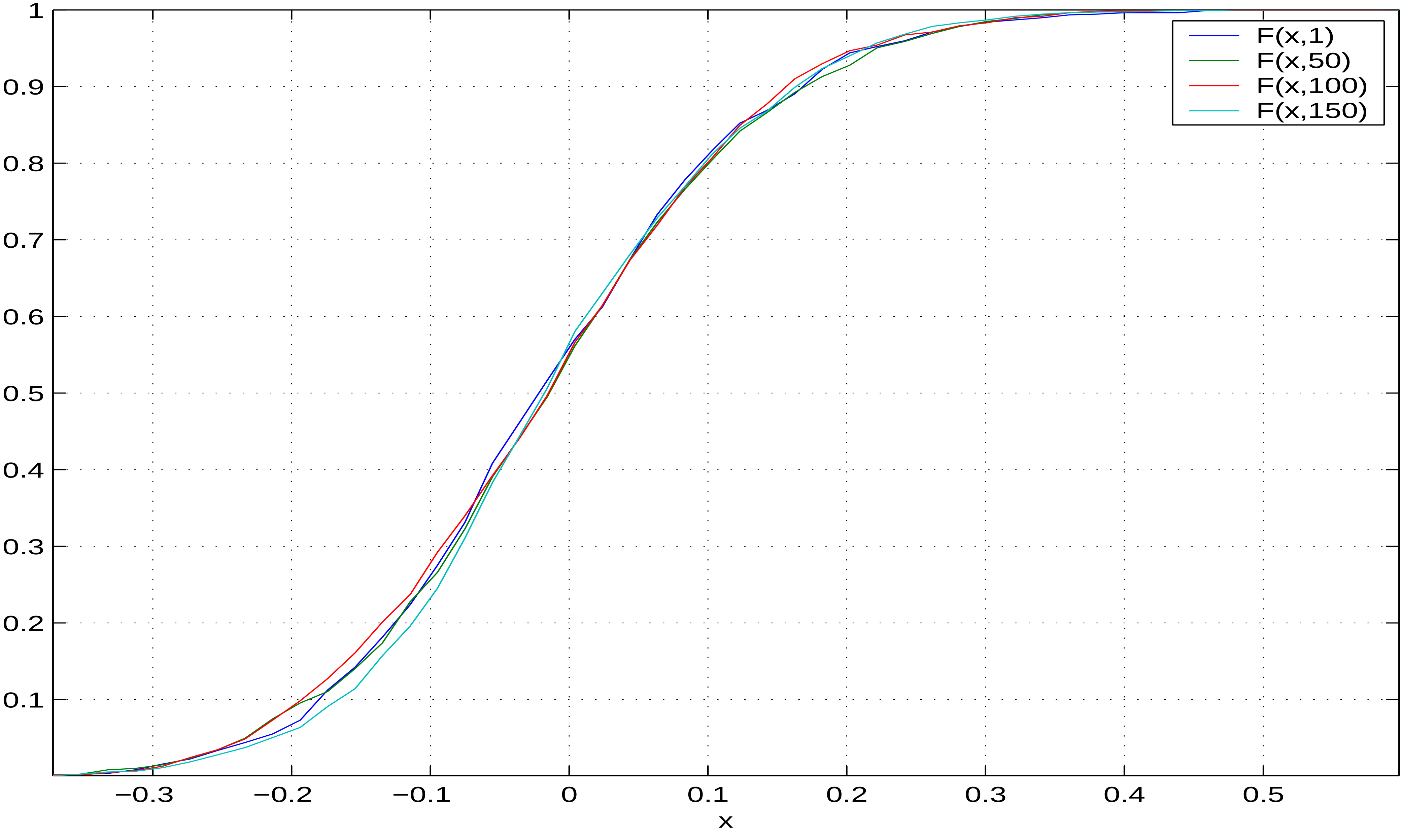
where $\mathcal{P}\{\xi(t) < x\}$ or $\mathcal{P}\{\xi[n] < x\}$ is the probability that random variable in given time will be smaller than x. Note, that x is nothing random, it is an auxiliary variable.

Ensemble estimation of distribution function: we will fix ourselves in given time t or n, and take Ω realizations. For given x we estimate:

$$\hat{F}(x,t) = \frac{\sum_{\omega=1}^{\Omega} 1 \;\; \text{if} \;\; \xi_\omega(t) < x, \;\;\; 0 \text{ else}}{\Omega}$$

$$\hat{F}(x,n) = \frac{\sum_{\omega=1}^{\Omega} 1 \;\; \text{if} \;\; \xi_\omega[n] < x, \;\;\; 0 \text{ else}}{\Omega}$$


F(x,0.1ms)
F(x,3.1ms)
F(x,6.3ms)
F(x,9.4ms)
x



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
−0.3
−0.2
−0.1
0
0.1
0.2
0.3
0.4
0.5
x
F(x,1)
F(x,50)
F(x,100)
F(x,150)

Probability density function

is again defined for one random variable (random process for a given time t or n is such a random variable). Definition:

$$p(x,t) = \frac{\delta F(x,t)}{\delta x}$$
$$p(x,n) = \frac{\delta F(x,n)}{\delta x}$$

Ensemble estimation of Probability density function: The function can be obtained by numeric derivation from estimated $\hat{F}(x,t)$ or $\hat{F}(x,n)$ or it can be estimated by a **histogram**:

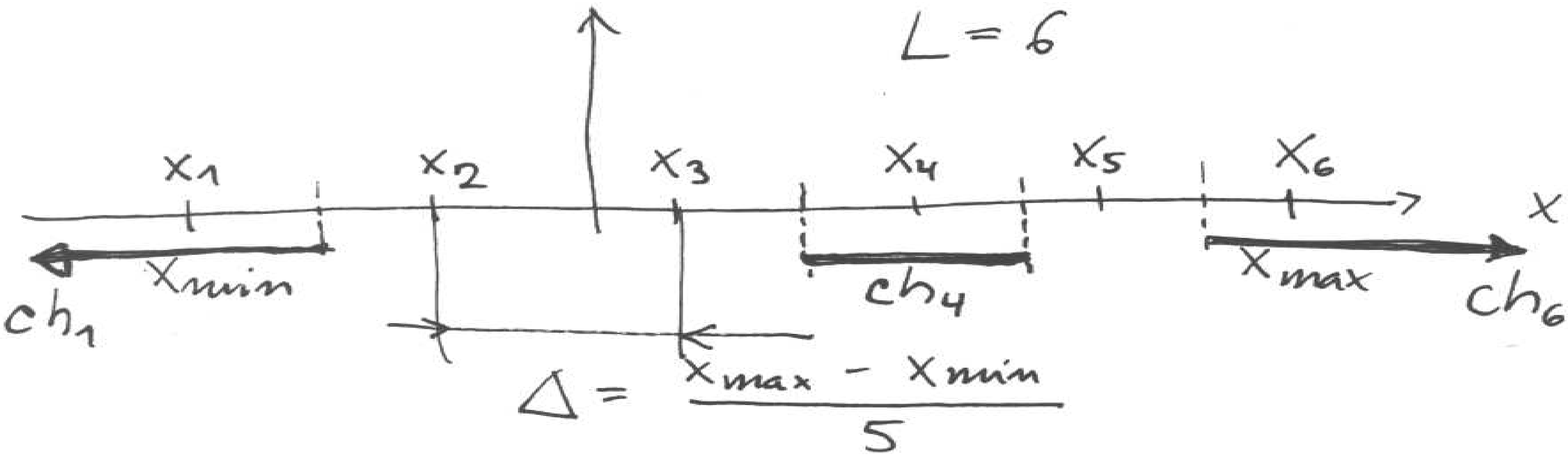
- choose given t or n
- CHoose L values x from x_{min} till x_{max}, with regular step $\Delta = \frac{x_{max}-x_{min}}{L-1}$:

$$x_1 = x_{min}, \quad x_2 = x_{min} + \Delta, \quad x_3 = x_{min} + 2\Delta \quad \ldots$$

$$\ldots \quad x_{L-1} = x_{min} + (L-2)\Delta, \quad x_L = x_{min} + (L-1)\Delta = x_{max}$$

In such a way, we'll obtain L "cages" with width Δ, for x_i, given cage ch_i is from

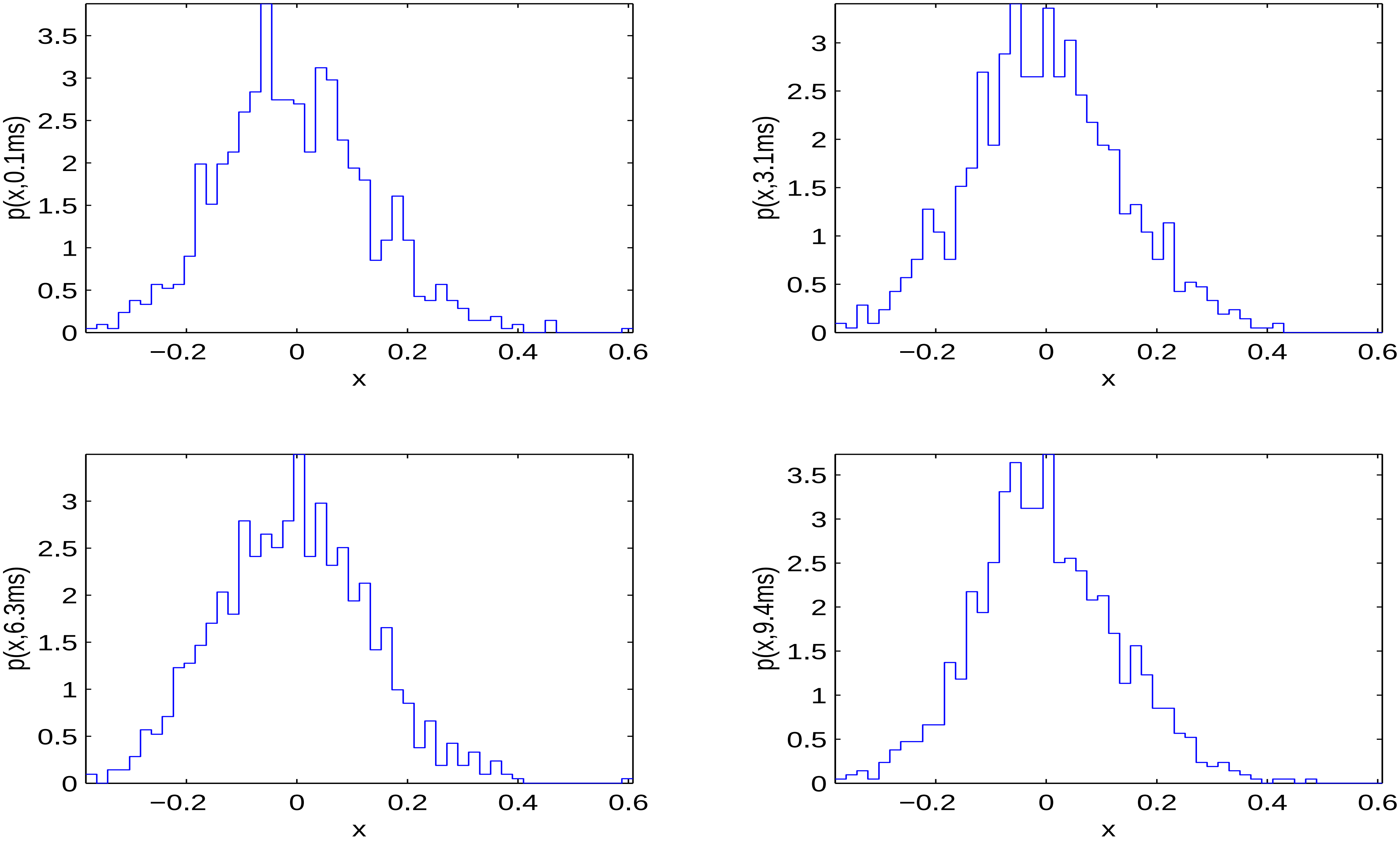
$x_i - \frac{\Delta}{2}$ till $x_i + \frac{\Delta}{2}$. The left edge of the left-most cage (1) will be stretched till $-\infty$, the right edge of the right-most (L) till $+\infty$.

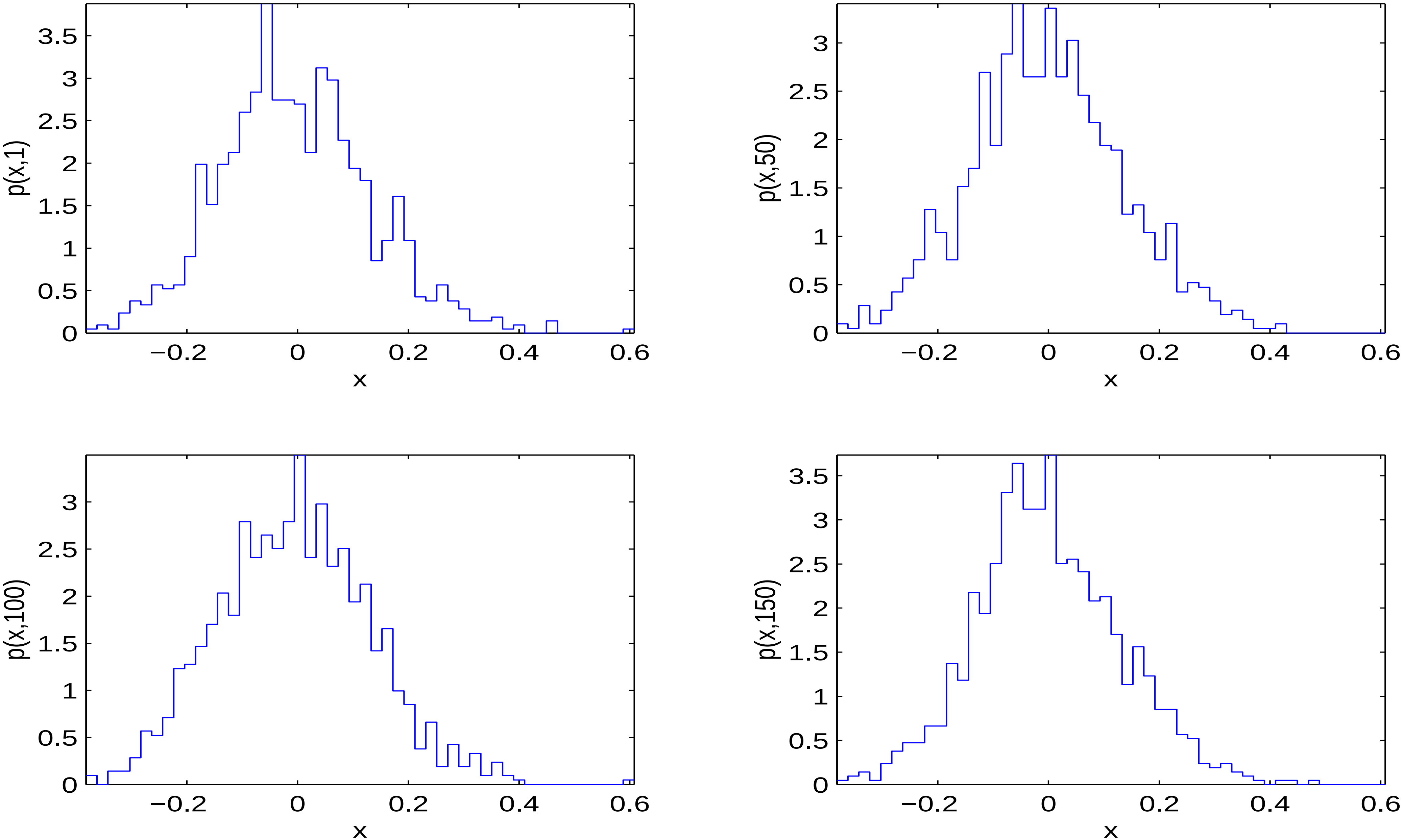


- the estimate for x_i is computed

$$\hat{p}(x_i, t) = \frac{\sum_{\omega=1}^{\Omega} 1 \text{ if } \xi_\omega(t) \in ch_i, \text{ 0 else}}{\Omega \Delta}$$

$$\hat{p}(x_i, n) = \frac{\sum_{\omega=1}^{\Omega} 1 \text{ if } \xi_\omega[n] \in ch_i, \text{ 0 else}}{\Omega \Delta}$$


p(x,0.1ms)
p(x,3.1ms)
p(x,6.3ms)
p(x,9.4ms)
x
x
x
x


p(x,1)
x
p(x,50)
x
p(x,100)
x
p(x,150)
x

$F(x,t)$, $p(x,t)$ **and probabilities**

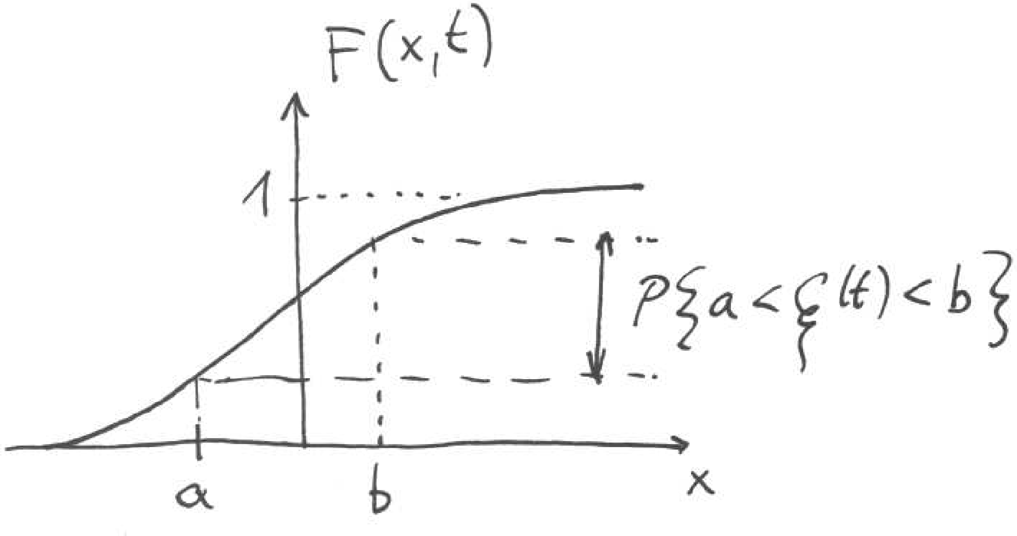
our task is to determine the probability, that the value of random process in time t or n is in the interval $[a,b]$. We can estimate this in two ways (shown only for continuous time):

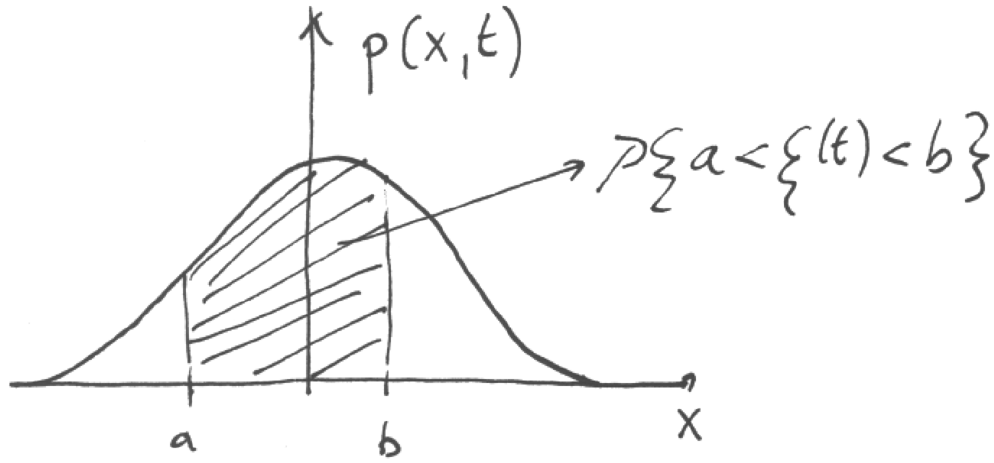
- from distribution function, having in mid its definition: $F(x,t) = \mathcal{P}\{\xi(t) < x\}$, then

$$\mathcal{P}\{a < \xi(t) < b\} = F(b,t) - F(a,t)$$

- from probability density function. Densities can be integrated:

$$\mathcal{P}\{a < \xi(t) < b\} = \int_a^b p(x,t)dx$$





Properties of distribution function and probability density function:

- the values of random process will hardly be smaller than $-\infty$, therefore $F(-\infty, t) = \mathcal{P}\{\xi(t) < -\infty\} = 0$.
- the values of random process will all be smaller than ∞, therefore $F(+\infty, t) = \mathcal{P}\{\xi(t) < +\infty\} = 1$.
- probability density function is derivation of distribution function, the inverse sense is given by an integral:

$$F(x,t) = \int_{-\infty}^{x} p(g,t)dg$$

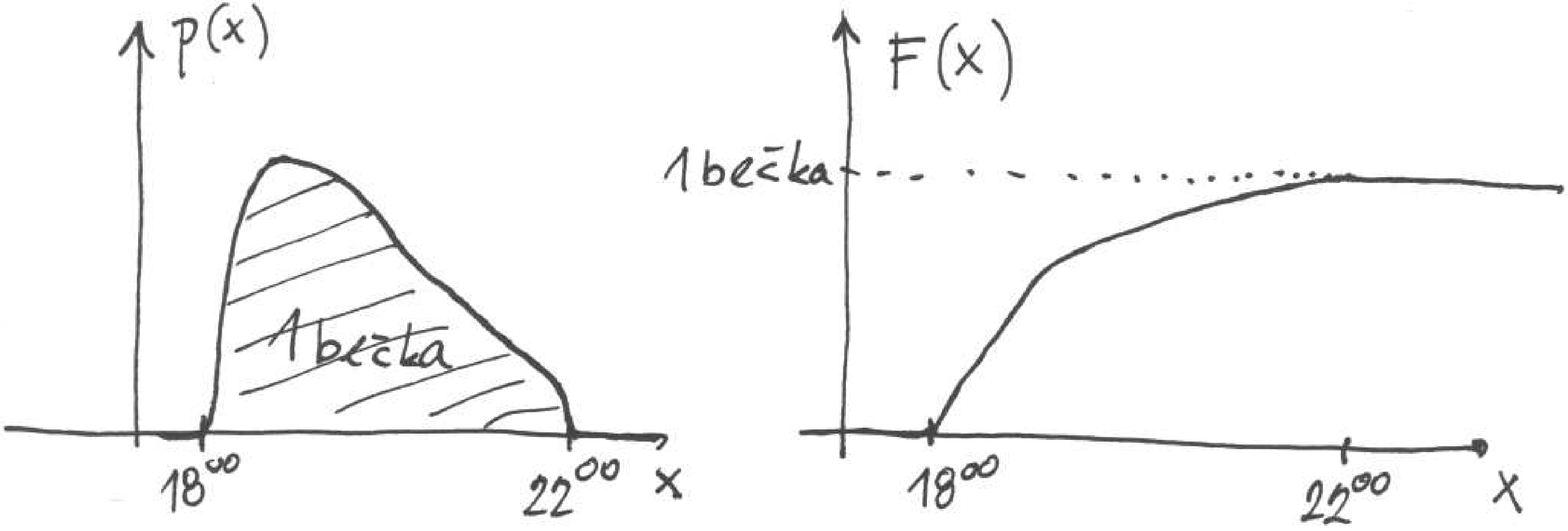
as $F(+\infty, t) = 1$, $p(x, t)$ must obey:

$$\int_{-\infty}^{+\infty} p(x,t)dx = 1$$

- the value of probability densifty function for given x **is not a probability !!!**

Illustration of $F(x,t)$, $p(x,t)$ – a beer keg ...

beer keg is being drunk from $x = 18.00$ till 22.00. We'll define function $p(x)$ as immediate beer consumption (drinking function) and $F(x)$ as function of drunk beer (attention, x is time in this example):



The behavior is simnilar to PDF and distribution functions, just replace 1 with "1 keg":

- $F(x)$ is zero in time $-\infty$ (it is zero till 18.00), as there was no beer.
- $F(x)$ is 1 keg in time $+\infty$ (actually already at 22.00), as all beer was drunk and

there'll be no more.

- amount of drunk beer at time x: $F(x) = \int_{-\infty}^{x} p(g)dg$.
- total amount of drunk beer: $F(+\infty) = \int_{-\infty}^{+\infty} p(x)dx = 1$ keg.
- amount of beer drunk from time x_1 till x_2 can be computed in as difference of 2 points on $F(x)$ or by integration of $p(x)$.
- one value of $p(x)$ (for example. $p(19.00)$) can NOT be called "amount of drunk beer" – noone can drink anything in infinitely short time.

Moments

on contrary to functions, moments are **values**, characterizing random process in time t or n:

mean value or "Expectation", is the 1st moment:

$$a(t) = E\{\xi(t)\} = \int_{-\infty}^{+\infty} xp(x,t)dx$$

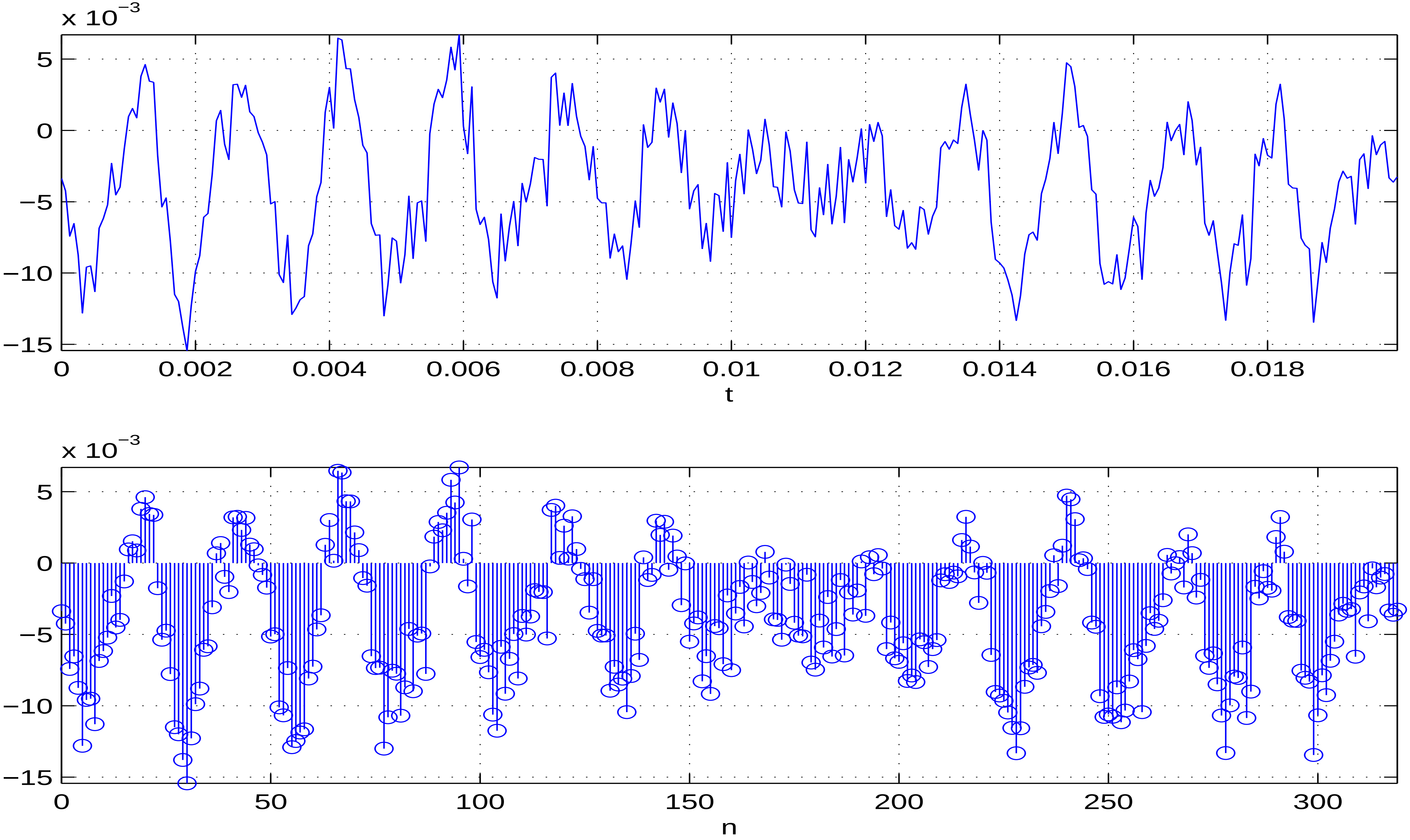
$$a[n] = E\{\xi[n]\} = \int_{-\infty}^{+\infty} xp(x,n)dx$$

ensemble estimate of mean value for each time t or n the estimate is simply the mean value of samples over all realizations:

$$\hat{a}(t) = \frac{1}{\Omega}\sum_{\omega=1}^{\Omega} \xi_\omega(t)$$

$$\hat{a}[n] = \frac{1}{\Omega}\sum_{\omega=1}^{\Omega} \xi_\omega[n]$$

For our signals:



variance (dispersion), standard deviation

$$D(t) = E\{[\xi(t) - a(t)]^2\} = \int_{-\infty}^{+\infty} [x - a(t)]^2 p(x,t) dx$$

$$D[n] = E\{[\xi[n] - a[n]]^2\} = \int_{-\infty}^{+\infty} [x - a[n]]^2 p(x,n) dx$$

std - standard deviation is the square root of variance:

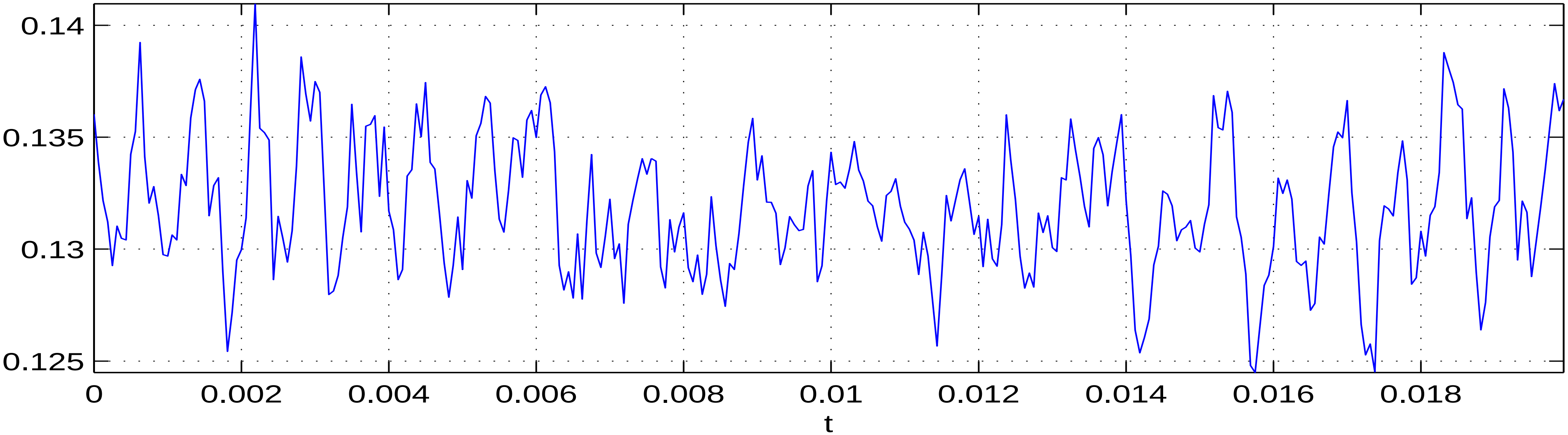
$$\sigma(t) = \sqrt{D(t)} \qquad \sigma[n] = \sqrt{D[n]}$$

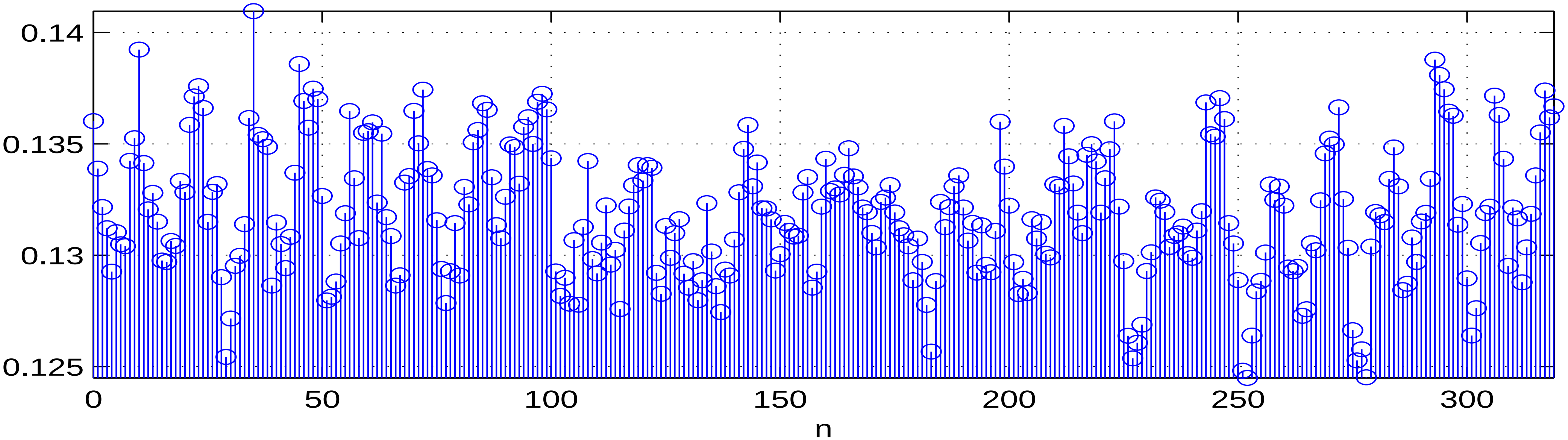
ensemble estimate of variance and std: for each t or n:

$$\hat{D}(t) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} [\xi_\omega(t) - \hat{a}(t)]^2, \quad \hat{\sigma}(t) = \sqrt{\hat{D}(t)}$$

$$\hat{D}[n] = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} [\xi_\omega[n] - \hat{a}[n]]^2, \quad \hat{\sigma}[n] = \sqrt{\hat{D}[n]}$$

For our signals:





Correlation function

is quantifying the similarity between values of random process in times t_1 (or n_1) and t_2 (or n_2):

$$R(t_1, t_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 p(x_1, x_2, t_1, t_2) dx_1 dx_2,$$

$$R(n_1, n_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 p(x_1, x_2, n_1, n_2) dx_1 dx_2,$$

where $p(x_1, x_2, t_1, t_2)$, resp. $p(x_1, x_2, n_1, n_2)$ is two-dimensional probability density function between times t_1 and t_2, resp. n_1 and n_2. Theoretically, it can be computed from two-dimensional distribution function:

$$F(x_1, x_2, t_1, t_2) = \mathcal{P}\{\xi(t_1) < x_1 \ \text{ a } \ \xi(t_2) < x_2\},$$

$$F(x_1, x_2, n_1, n_2) = \mathcal{P}\{\xi[n_1] < x_1 \ \text{ a } \ \xi[n_2] < x_2\}$$

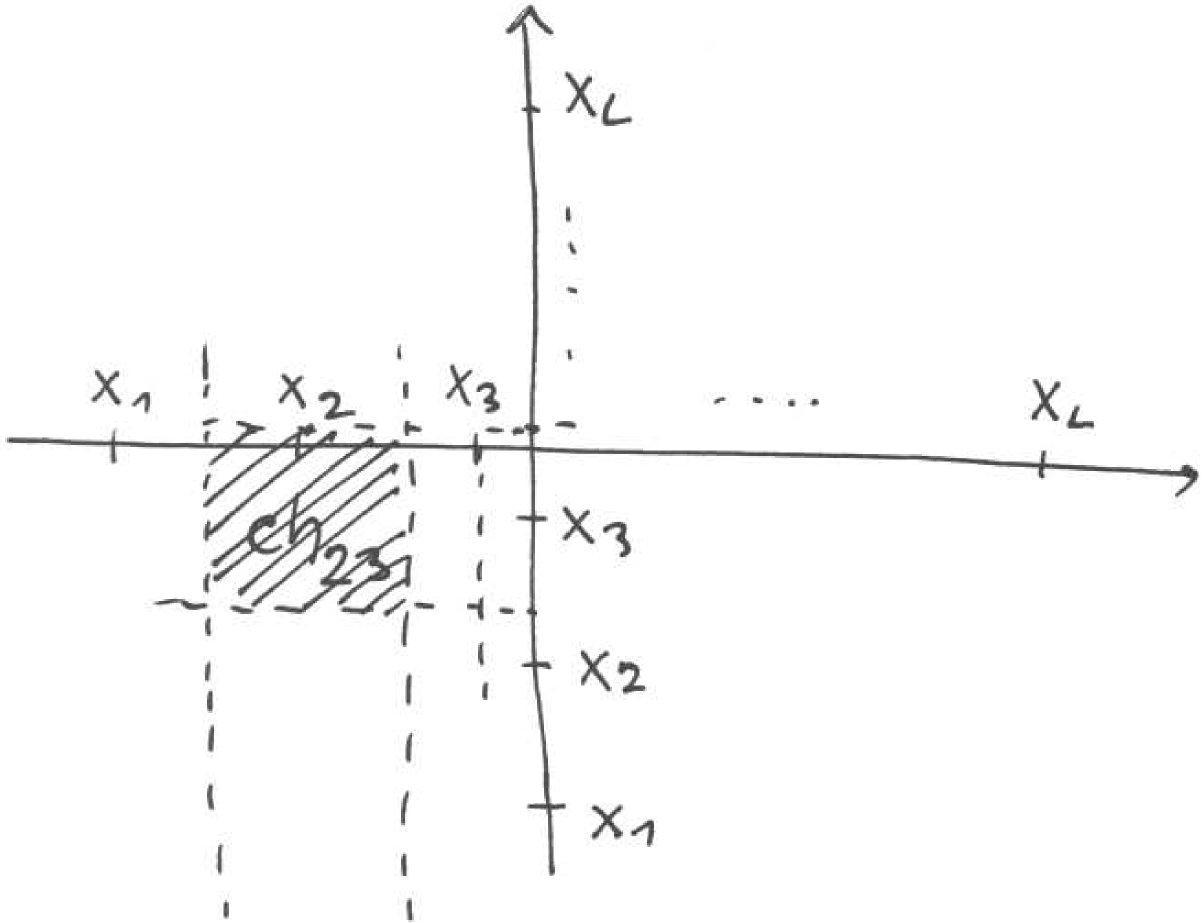
by deriving along x_1 and x_2:

$$p(x_1, x_2, t_1, t_2) = \frac{\delta^2 F(x_1, x_2, t_1, t_2)}{\delta x_1 \delta x_2}$$

$$p(x_1, x_2, n_1, n_2) = \frac{\delta^2 F(x_1, x_2, n_1, n_2)}{\delta x_1 \delta x_2}$$

We'll rather be interested in ensemble estimation using 2D-histogram:

- similarly as for 1D histogram, define "cages", in forms of squares: ch_{ij} is for the first dimension from $x_i - \frac{\Delta}{2}$ till $x_i + \frac{\Delta}{2}$ and for the second dimension from $x_j - \frac{\Delta}{2}$ till $x_j + \frac{\Delta}{2}$.



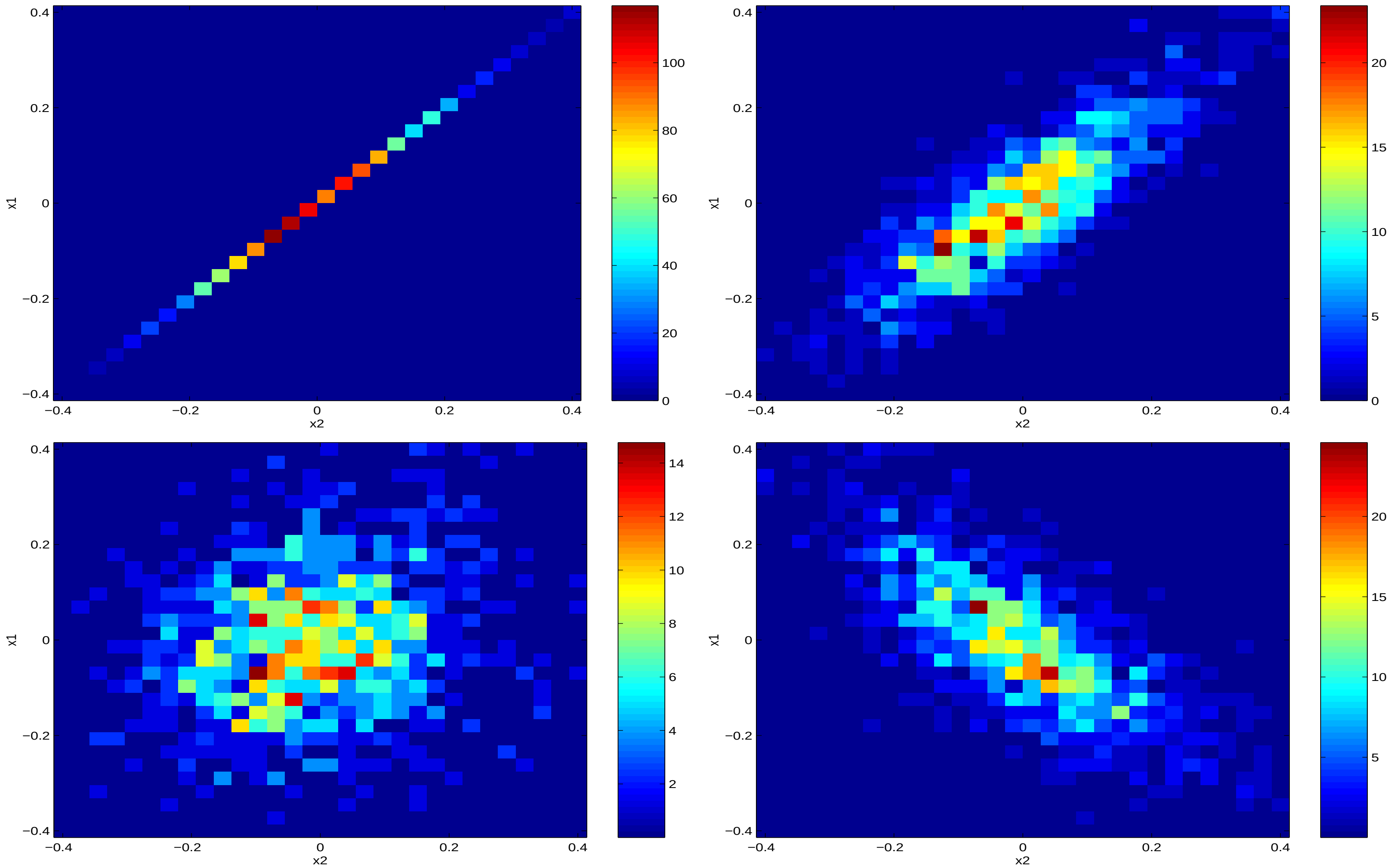
x_L
x_1
x_2
x_3
. . . .
x_L
ch_{23}
x_3
x_2
x_1

- the value of 2D-histogramu for cage ch_{ij} with values x_i and x_j will be:

$$\hat{p}(x_i, x_j, t_1, t_2) = \frac{\sum_{\omega=1}^{\Omega} 1 \ \ \text{if} \ \ \xi_\omega(t_1), \xi_\omega(t_2) \in ch_{ij}, \ \ 0 \text{ else}}{\Omega \Delta^2}$$

$$\hat{p}(x_i, x_j, n_1, n_2) = \frac{\sum_{\omega=1}^{\Omega} 1 \ \ \text{if} \ \ \xi_\omega[n_1], \xi_\omega[n_2] \in ch_{ij}, \ \ 0 \text{ else}}{\Omega \Delta^2}$$

For our signals (only discrete-time): $n_1 = 0$, $n_2 = 0, 1, 5, 11$.

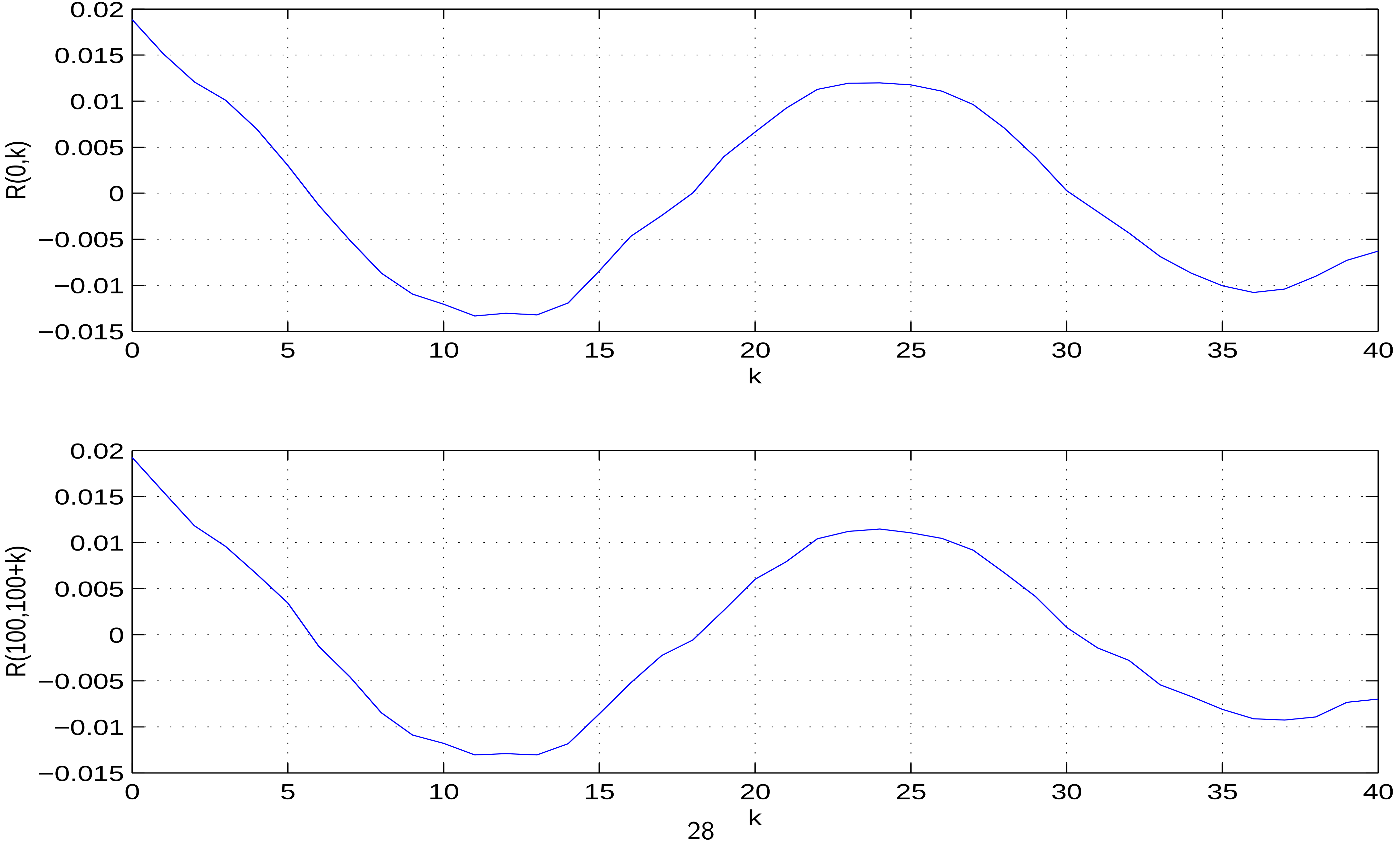


For $n_1 = 0$, $n_2 = 0, 1, 5, 11$ the following coefficients were obtain after numeric integration:

- $R(0, 0) = 0.0188$: the process is very similar to itself (the same point). similar...
- $R(0, 1) = 0.0151$: if shifted to neighboring sample, still very similar to time $n_1 = 0$.
- $R(0, 5) = 0.0030$: times $n_1 = 0$ and $n_2 = 5$ are not similar at all.
- $R(0, 11) = -0.0133$: in times $n_1 = 0$ and $n_2 = 11$ the process is similar to itself, but with inverse sign ! It is probable, that if the value $\xi[n_1]$ is positive, $\xi[n_2]$ will be negative and vice versa.

Correlation function for

$n_1 = 0$, $n_2 = n_1 + k$ for $k = 0 \ldots 40$, comparison with $n_1 = 100$, $n_2 = n_1 + k$ for $k = 0 \ldots 40$:



STATIONARITY OF RANDOM PROCESS

simply said, the behavior of random process is not changing with the time. The values do not depend on time t or n. Correlation function does not depend on the precise values of t_1, t_2 or n_1, n_2, but only on their difference: $\tau = t_2 - t_1$, $k = n_2 - n_1$. For stationary continuous-time signal:

$$F(x,t) \rightarrow F(x) \quad p(x,t) \rightarrow p(x)$$

$$a(t) \rightarrow a \quad D(t) \rightarrow D \quad \sigma(t) \rightarrow \sigma$$

$$p(x_1, x_2, t_1, t_2) \rightarrow p(x_1, x_2, \tau) \quad R(t_1, t_2) \rightarrow R(\tau)$$

Similarly for discrete time:

$$F(x,n) \rightarrow F(x) \quad p(x,n) \rightarrow p(x)$$

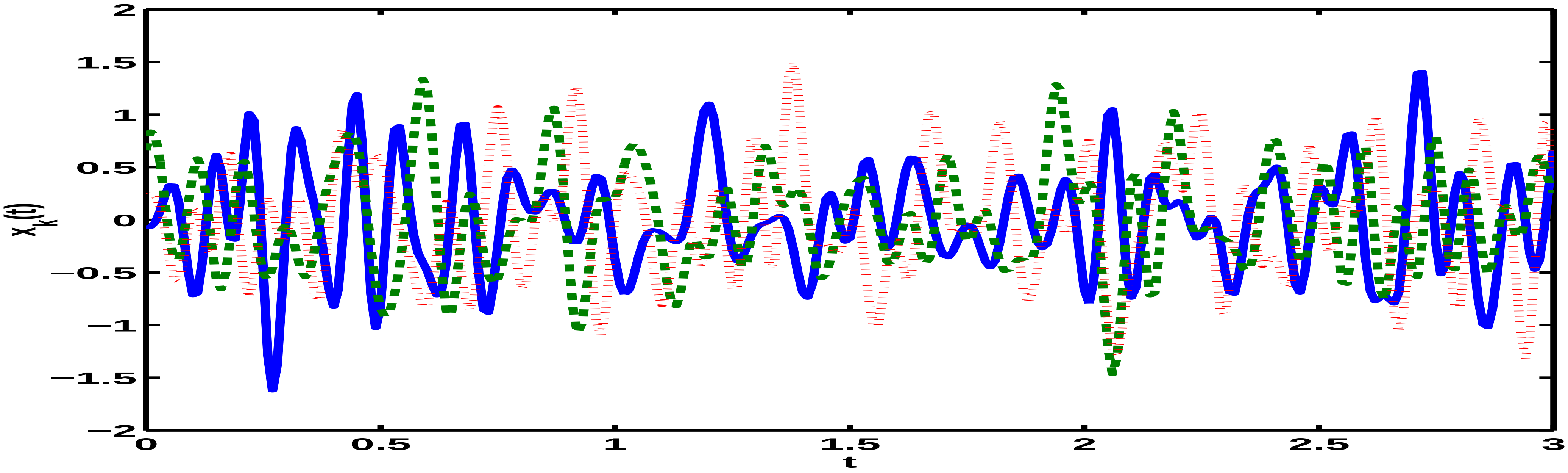
$$a[n] \rightarrow a \quad D[n] \rightarrow D \quad \sigma[n] \rightarrow \sigma$$

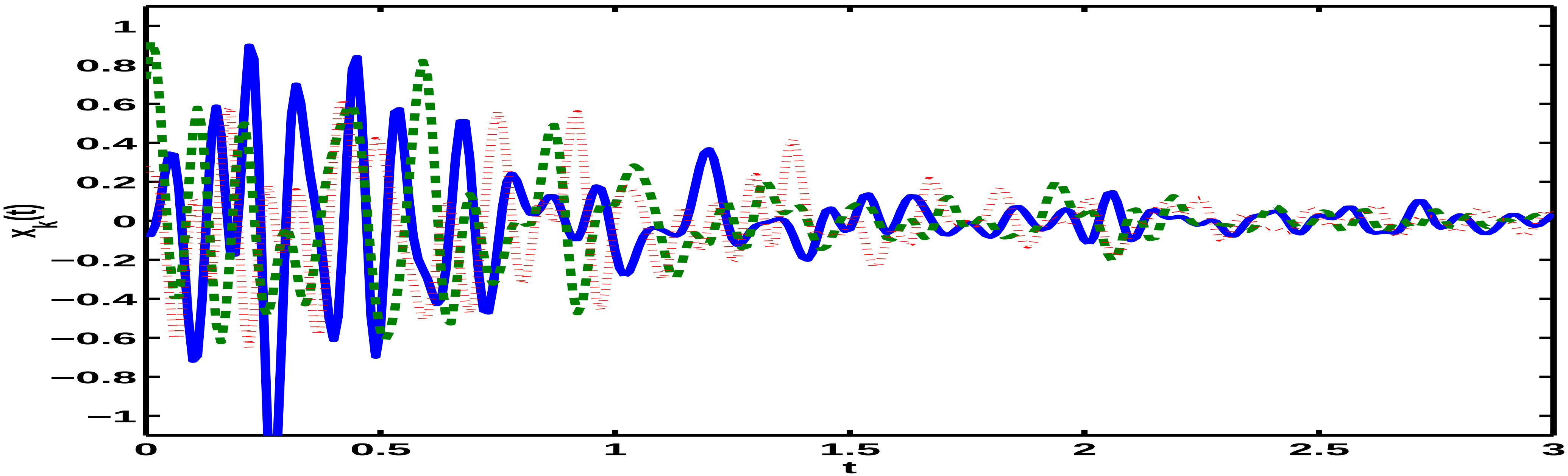
$$p(x_1, x_2, n_1, n_2) \rightarrow p(x_1, x_2, k) \quad R(n_1, n_2) \rightarrow R(k)$$

The “water”-example signal was obviously stationary, as:

- the mean value was similar for all times. In case we had more realizations, it would be even more similar.
- dtto for standard deviation.
- Correlation function for $n_1 = 0, n_2 = n_1 + k$ and for $n_1 = 100, n_2 = n_1 + k$ was similar.

Stationary vs. non-stationary signal:

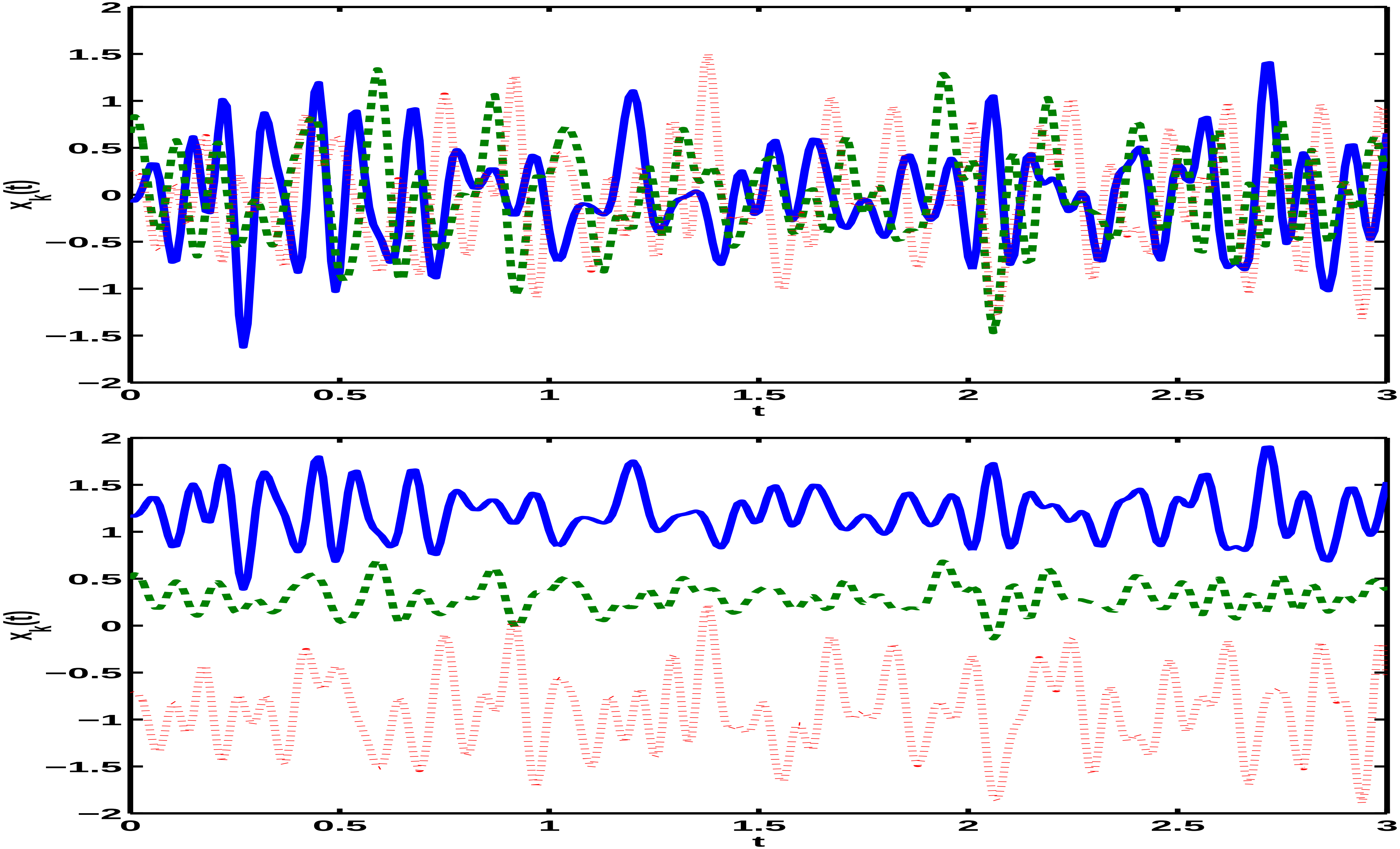




ERGODICITY OF RANDOM PROCESS

all parameters can be estimated from 1 realization:

Example of stationary and ergodic and of stationary but non-ergodic random process.



All **ensemble estimates** can be replaced by **temporal estimates**, we dispose of interval of lenght T (continuous time) or of N samples (discrete time). The only realization we have will be simply denoted $x(t)$, resp. $x[n]$:

- histograms can be applied to estimate distribution function and PDF:
- mean value, variance, std:

$$\hat{a} = \frac{1}{T}\int_0^T x(t)dt \qquad \hat{D} = \frac{1}{T}\int_0^T [x(t) - \hat{a}]^2 dt \qquad \hat{\sigma} = \sqrt{\hat{D}}$$

$$\hat{a} = \frac{1}{N}\sum_{n=0}^{N-1} x[n] \qquad \hat{D} = \frac{1}{N}\sum_{n=0}^{N-1} [x[n] - \hat{a}]^2 \qquad \hat{\sigma} = \sqrt{\hat{D}}$$

- correlation function:

$$\hat{R}(\tau) = \frac{1}{T}\int_0^T x(t)x(t+\tau)dt$$

$$\hat{R}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x[n]x[n+k]$$