Introduction to Discrete Random Process

87203 – Multimedial Signal Processing 1st Module

Politecnico di Milano –

Polo regionale di Como

Agenda

- Ex1: Bernoulli Random Sequence
- Ex2: Gaussian Random Sequence
- Ex3: Jointly distributed random sequences
- Ex4: Stationary vs ergodicity

Exercise 1 (1/6)

- Goal: Generate i.i.d. Bernoulli sequence of 1000 samples with arbitrary p.

$$f_x(x) = \begin{cases} 0 & with\ probability\ p \\ 1 & with\ probability\ 1-p \end{cases}$$

- A DTRS is a sequence of random variables x(n)

$$f_{x(n)}(x,n) = \sum_{k=-\infty}^{\infty} p_{x(n)}(x_k, n)\, \delta(x - x_k)$$

Exercise 1 (2/6)

- A DTRS is independent identically distributed when:
  - the pdf does not depend on the time index

$$f_{x(n)}(x,n) = f_x(x)$$

  - the random variables at different time instant n,m are independent

$$f_{x_j x_k}(x_j, x_k) = f_{x_j}(x_j) \cdot f_{x_k}(x_k)$$

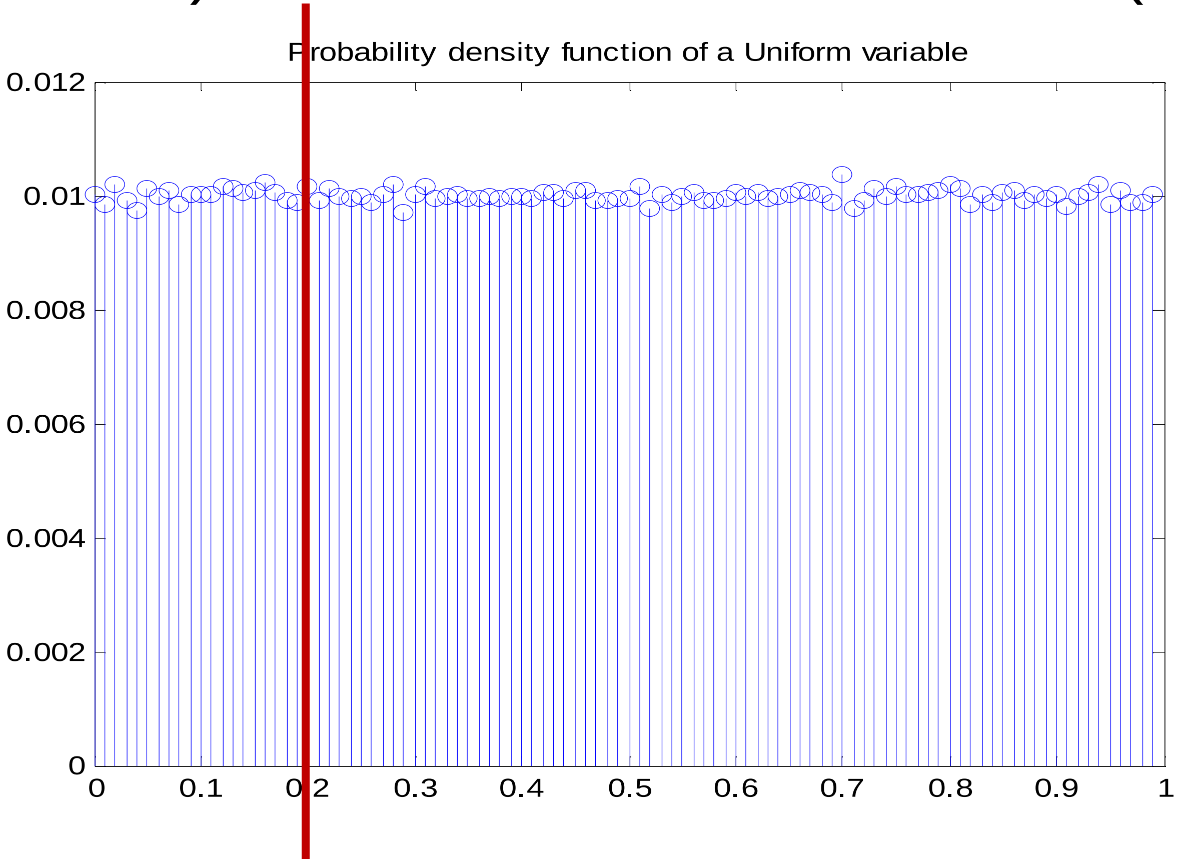
- Hint: A i.i.d. sequence is generated by taking independent sample of the pdf.

Exercise 1 (3/6)

- Pseudocode:
  - generate a iid sequence of 1000 samples with uniform distribution on the interval (0.0 , 1.0).
    - y = rand(1000,1);
  - assign value 0 at the index of each sample that is <p
    - x2 = zeros(length(y),1);
    - p = 0.2;
    - x2(y<p) = 0 ;
  - assign value 1 at the index of each sample that is >=p
    - x2(y>=p) = 1 ;

Exercise 1 (4/6)

- Let's see pdf of a discrete (100 possible set of values) uniform variable in the interval (0 , 1)



Exercise 1 (5/6)

- Compute the expected value and variance of the random variable x of which we have N realizations:
  - mean: $$E[x] = \eta_x = \int_{-\infty}^{\infty} x \; f_x(x) \, \delta(x)$$
    - m=mean(x);
    - ms=(1/N)*sum(x);

$$\eta_x = 1 * p + 0 * (1 - p)$$

  - variance: $$\sigma_x^2 = E\left[\left(x - \eta_x\right)^2\right]$$
    - sigma=var(x);
    - vars=(1/N-1).*sum((x-ms)^2);

$$\sigma_x^2 = (1 - p) * p$$

Exercise 1 (6/6)

- Matlab provides the functions:
- “x=rand(M,N)” that produces a M-by-N matrix with random entries, chosen from a uniform distribution on the interval (0.0,1.0).
- “m=mean(x)” that compute the mean value:
  - For vectors, “m” is the mean value of the elements in “x”
  - For matrices, “m” is a row vector containing the mean value of each column.
- “s=var(x)” that compute the variance:
  - For vectors “s” returns the variance of the values in “x”.
  - For matrices, “s” is a row vector containing the variance of each column of “x”.

Exercise 2 (1/5)

- Goal: Generate i.i.d. Gaussian sequence of 1000 samples with arbitrary mean and variance.

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\eta_x)^2}{2\sigma_x^2}}$$

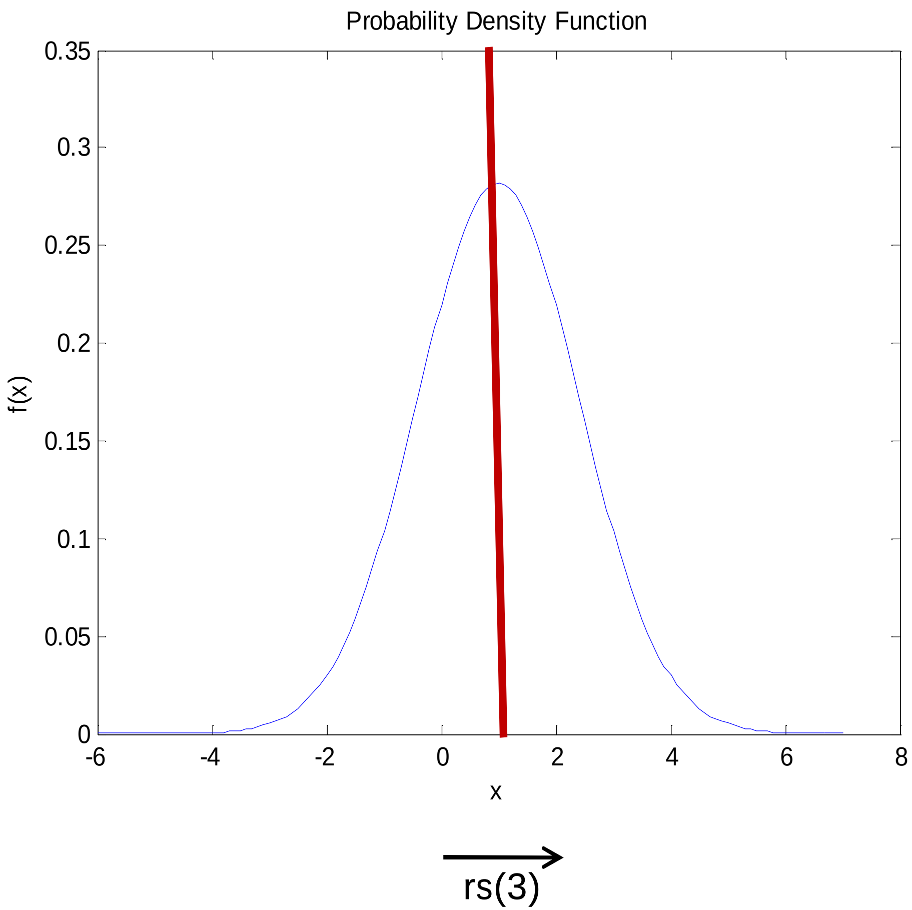
- Hint: A i.i.d. sequence is generated by taking independent sample of the pdf.

Exercise 2 (2/5)

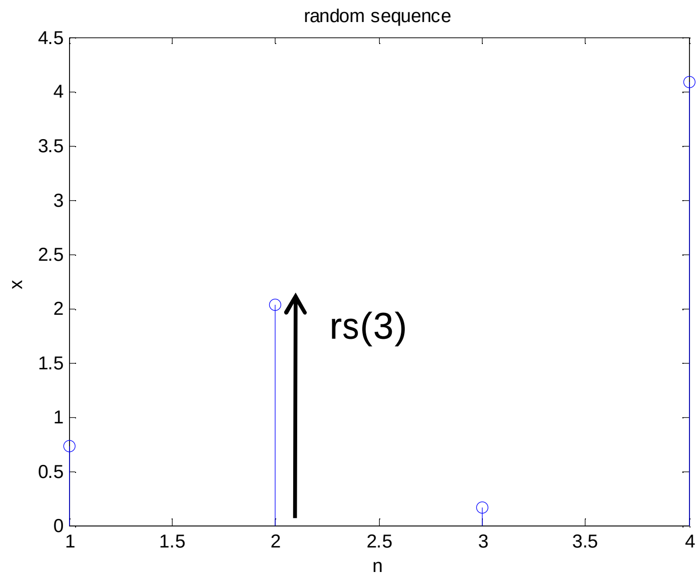
- Pseudocode:
  - generate a iid sequence of 1000 samples with normal distribution
    - x = randn(1000,1);
  - set variance = s
    - x = sqrt(s)*randn(1000,1);
  - set mean = m
    - x = sqrt(s)*randn(1000,1)+m;

- N.B. For a gaussian distribution, 99,7% of possible value are inside the interval m±3s

Exercise 2 (3/5)



rs = [0.7360   2.0264   0.1680   4.0875]

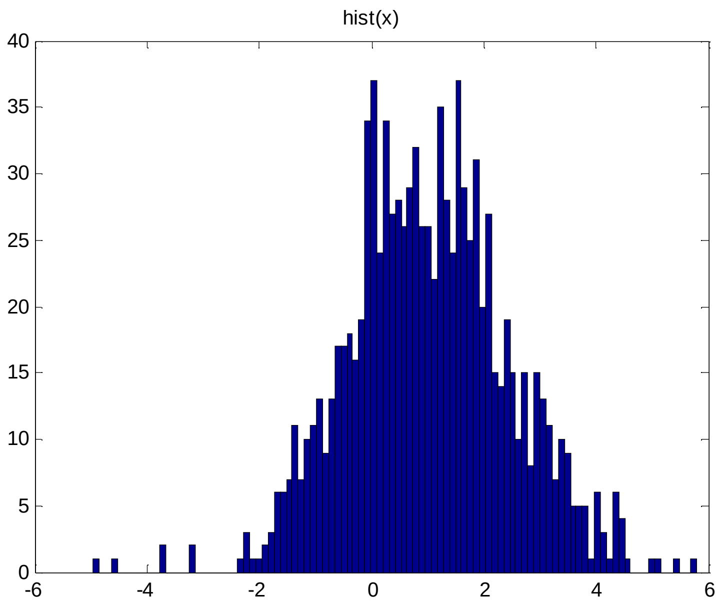


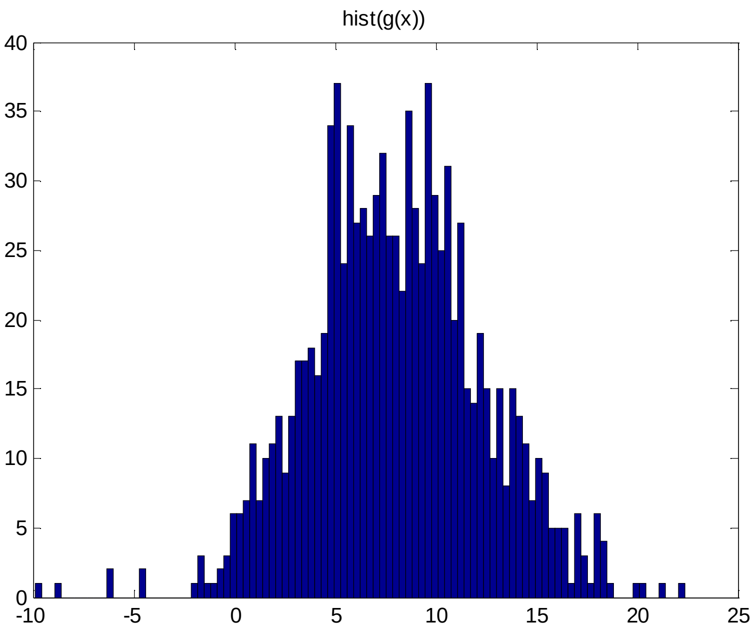
Exercise 2 (4/5)

- Hint: Remember:
  - if x is a gaussian variable $x \approx N\,(\eta_x, \sigma_x^2)$

then g(x)=ax+b is a guassian variable

$$y = g(x) \approx N\,(a\eta_x + b, a^2\sigma_x^2)$$





Exercise 2 (5/5)

- Matlab provides the functions:
- “x=randn(M,N)” that produces a M-by-N matrix with random entries, chosen from a normal distribution with mean zero, variance one and standard deviation one.

- “n=hist(x,M)” that bins the elements of “x” into M equally spaced containers and returns the number of elements in each container “n”. If “x” is a matrix, “hist” works down the columns.

Exercise 3 (1/2)

- Goal: Generate two jointly distributed Gaussian random sequences.
  - Two DTRS x(n) and y(n) are described by the a joint probability function

$$F_{x(n)y(m)}(x,n,y,m) = P\{x(n) \leq x \; and \; y(m) \leq y\}$$

  - Assuming that the joint probability function is indipendent on the time index, it is a function of two variables x and y

Exercise 3 (2/2)

- Pseudocode:
  - generate a iid sequence of 1000 samples x(n)
  - generate a iid sequence of 1000 sample z(n)
  - consider the sequence y(n) defined as:
    - y(n) = x(n) + z(n);

- N.B. {x(n)} and {y(n)} are statistically dependent: the knowledge of $f_x(x)$ and $f_y(y)$ is not enough to describe the joint pdf $f_{x,y}(x,y)$.

Exercise 4 (1/8)

- A DTRS is stationary if its statistical characterization is not affected by a shift in the data sequence:

$$F_{x(n)}(x,n) = F_{x(n+m)}(x,n+m) = F_x(x) \quad \forall n,m$$

- A DTRS is Wide-Sense Stationary if:
  - mean value $\quad E[x(n)] = \eta(n) = \eta \quad \forall n$
  - autocorrelation

$$r_x(m) = E[x(n)x^*(n-m)] = E[x(n+m)x^*(n)]$$

Exercise 4 (2/8)

- A WSS DTRS is also ergodic when time averages are equivalent to ensemble averages:
  - mean value $$< x(n) >= E[x(n)]$$
  - autocorrelation $$< x(n)\, x^*(n-m) >= r_x(m)$$
- N.B.: A WSS sequence is not necessary ergodic

Exercise 4 (3/8)

- Goal: Generate R realization of a WSS random sequence (N samples) that is also ergodic.
  - Compute ensemble averages
  - Compute time averages
  - Demonstrate that the sequence is ergodic

- Hint: Matlab provides the function

  "[r lag]=xcorr(x)" that produces an estimate of the autocorrelation of the stationary sequence "x". "lag" is the vector of lag indices.

Exercise 4 (4/8)

- Pseudocode:
  - Consider an iid sequence of {z(n)} with gaussian zero-mean distribution
    - z = sqrt(varz)*randn(1000,1);
  - Consider the sequence {x(n)} generated by:

$$x(n) = \rho\, x(n-1) + z(n)$$

    - x = filter(1,[1 -rho],z);
  - Compute R sequence {x(n)} of N samples each.
    - z = = sqrt(varz)*randn(N,R);

Exercise 4 (5/8)

- Pseudocode (continued):
  - Compute ensemble averages:
    - mR = mean(x(:,50));
    - vR = var(x(:,50))
    - [rR lagR]=xcorr(x(:,50));
  - Compute time averages
    - mN = mean(x(1,:))
    - vN = var(x(1,:))
    - [rN lagN]=xcorr(x(1,:));
  - Demonstrate that:
    - mN = mR; vN = vR; rN = rR;

Exercise 4 (6/8)

- Goal: Generate R realization of a WSS random sequence (N samples) that is not ergodic.
  - Compute ensemble averages
  - Compute time averages
  - Demonstrate that the sequence is NOT ergodic

Exercise 4 (7/8)

- Pseudocode:
  - Let a be a random number normally distributed:
    - a = randn(R,1);

  - One realization of the process is a constant horizontal time path:
    - x = ones (R,N);
    - for i=1:R

```
    x(i,:) = a(i).*x(i,:);
      end
```

Exercise 4 (8/8)

- Pseudocode (continued):
  - Compute ensemble averages:
    - mR = mean(x(:,50));
    - vR = var(x(:,50))
    - [rR lagR]=xcorr(x(:,50));
  - Compute time averages
    - mN = mean(x(1,:))
    - vN = var(x(1,:))
    - [rN lagN]=xcorr(x(1,:));
  - Demonstrate that:
    - mN ≠ mR;   vN ≠ vR;   rN ≠ rR;