Class Exercise 7 – Random Processes, Stationarity

Theoretical Background:

General Random Processes:

Random process $X(\omega,t)$ is a function of two parameters: the time parameter – $t \in \mathbf{R}$ in continuous time ($n \in \mathbf{Z}$ in discrete time), and the "luck" parameter – $\omega \in \Omega$ (mapping of the sample space to the space of all the functions).

- When the time parameter is constant, we get the random variable $X(\omega,t_0)$ with some distribution. Generally speaking, for different times we get a different distribution.
- When the "luck" parameter is constant, we get a function of time – sample function $X(\omega_0,t)$. Generally speaking, for each $\omega_i \in \Omega$ we get a different sample function.

<u>Shortened symbol[1]:</u> from now on we will represent random processes by uppercase letters with time specified, when the values that they take (namely, the sample functions) will be represented by lowercase letters: for example, $X(t)$ is a random process in continuous time and $Y[n]$ [2] is a random process in discrete time, having sample functions $x(t)$ and $y[n]$ [3], respectively.

Highlights:

- A random process is a mapping of the sample space to the functions space (or the series in the case of a random process in discrete time).
- A random process is overall an (infinite[4]) reserve of random variables – the time parameter gives names to the random variables in the reserve. In this way, random processes are a called for generalization of random vectors.
- In most cases, there will not be any statistical differences when dealing with processes in continuous time to those in discrete time. Therefore, only the definitions and formulas for continuous time will be given in all cases in which a trivial replacing of the time parameter can be done to get the appropriate definitions or formulas in the discrete time.

[1] This is one of the accepted conventions for representing random processes; there are others, too.

[2] It is accepted to mark discrete time random processes as: Y_n, and their sample series as: y_n.

[3] This symbol is the same symbol of random variables (which are, as you may recall, a mapping from the sample space to the real numbers axis, and they, too, may be marked as $X(\omega)$, where $\omega \in \Omega$ is the "luck" parameter).

[4] But take notice: in continuous time random processes, this is an infinity that is not countable (continuous), whereas in discrete time it is a countable infinity.

First Order Marginal Distribution

When the time parameter is constant $(t = t_0)$, we get a unique random variable $X(t_0)$. This random variable has a PDF , $f_X(x;t_0)$. With this density function, the expectation and variance can be found, which will generally be a function of the time parameter:

$$E\{X(t)\} = \eta_X(t)$$
$$VAR\{X(t)\} = \sigma_X{}^2(t)$$

Second Order Marginal Distribution

Let us sample the process at two different times t_1 and t_2, so as to get two random variables $X(t_1)$ and $X(t_2)$ with some statistical dependence between them. For any two times, a joint PDF exists $f_{X_1X_2}(x_1, x_2; t_1, t_2)$. From this function, we may find the **second order statistics** (namely, all the second order moments) of the process:

Auto-correlation[5] function:

$$R_X(t_1,t_2) = E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{X_1X_2}(x_1, x_2; t_1, t_2) dx_1 dx_2$$

Auto-covariance function:

$$C_X(t_1,t_2) = E\{(X(t_1) - \eta_X(t_1))(X(t_2) - \eta_X(t_2))\} = R_X(t_1,t_2) - \eta_X(t_1)\eta_X(t_2)$$

Auto-correlation coefficient function:

$$\rho_X(t_1,t_2) = \frac{C_X(t_1,t_2)}{\sqrt{C_X(t_1,t_1)C_X(t_2,t_2)}}$$

Marginal Distribution of Order n:

Let us take a look at the samples of the random process in n different times, $t_1, t_2, ..., t_n$. These samples are random variables with some interrelationship $X(t_1), ..., X(t_n)$ defined by the n-dimensioned PDF:

$$f_{X_1 \dots X_n}(x_1, ..., x_n; t_1 ... t_n)$$

The complete statistical information of the process is defined by the joint PDF of any n samples (for any value of n).

Gaussian Random Process

Definition: $X(\bullet)$ is a Gaussian random process iff any vector of its samples is a Gaussian random vector.

Claim: the complete statistics of a Gaussian random process are characterized by the auto-correlation function and the expectation function.

[5] It is also accepted to represent the auto-covariance and auto-correlation functions as $C_{XX,}(\bullet), R_{XX}(\bullet)$, respectively.

Explanation:

Let us assume $X[n]$ is a Gaussian random process in discrete time, with expected value $\mu_X[n]$ and auto-correlation $R_X[l,m]$. Let us denote a vector of size k of its samples:

$$\underline{Y} = [X[n_1]\ X[n_2] \ldots X[n_k]]^T$$

By definition, this is a Gaussian random vector; thus, to characterize its statistics, only the expectation $\underline{\mu}_{\underline{Y}} = E\{\underline{Y}\}$ and covariance matrix $C_{\underline{Y}} = E\{\underline{Y}\underline{Y}^T\} - \underline{\mu}_{\underline{Y}} \cdot \underline{\mu}_{\underline{Y}}^T$ are needed.

However:

$$\underline{\mu}_{\underline{Y}} = E\{[X[n_1]\ X[n_2] \ldots X[n_k]]^T\} = [E\{X[n_1]\}\ E\{X[n_2]\} \ldots E\{X[n_k]\}]^T = [\mu_X[n_1]\ \mu_X[n_2] \ldots \mu_X[n_k]]^T$$

Similarly: $[C_{\underline{Y}}]_{l,m} = R_X[n_l, n_m] - \mu_X[n_l]\mu_X[n_m] \qquad 1 \le m, l \le k$

Stationarity

A random process is called strict sense stationary (S.S.S.) if for all values of n, the PDF of order n of the process does not change due to a shifting in time, namely:

Continuous time random process $X(t)$:

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;t_1,\ldots,t_n) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;t_1+\tau,\ldots,t_n+\tau) \quad \forall n \in N \qquad \forall t_i, \tau \in R \quad i \in [1,n]$$

Discrete time random process $X[k]$:

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;k_1,\ldots,k_n) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;k_1+m,\ldots,k_n+m) \quad \forall n \in N \quad \forall k_i, m \in Z \quad i \in [1,n]$$

In other words, the PDF is dependent only on time difference, and invariant to shifting in the time axis.

A random process is said to be wide sense stationary (W.S.S.) if the following two conditions are met:

In continuous time:

- $E\{X(t)\} = \eta_X$ - the expected value does not depend on time.
- $R_X(t_1,t_2) = R_X(t_1 - t_2)$ - the auto-correlation function is dependent only on the time difference.

In discrete time the conditions are: $E\{X[k]\} = \eta_X$ and $R_X[k_1,k_2] = R_X[k_1 - k_2]$.

[!!!] A random process that is S.S.S. is necessarily W.S.S. The opposite is not necessarily true. If the process is Gaussian, then the opposite is true.

Properties of the auto-correlation function for a real W.S.S. random process:

- $R_X(0) = E\{X^2(t)\} \geq 0$ - positive
- $R_X(0) \geq |R_X(\tau)|$
- $R_X(\tau) = R_X(-\tau)$ - even function
- $\sum_{i=1}^{M}\sum_{j=1}^{N} a_i a_j R_X(t_i, t_j) \geq 0$ - the correlation matrix is positive semi-definite. **This does not mean that the elements of the matrix are non-negative!!!**

<u>Stationarity and Processes Independent of Time and Memoryless</u>

In the problem set you will prove the following claim that deals with passing random processes through a time-invariant and memoryless system:

Let $X(t)$ be a random process and $g(\bullet)$ be an arbitrary real function:

$$g : R \longrightarrow R$$

Let us define the random process $Z(t)$ as follows:

$$Z(t) = g(X(t)) \quad \forall t$$

It is said that the random process $Z(t)$ is obtained by passing the random process $X(t)$ through a time-invariant and memoryless system.

then:

- if $X(t)$ is S.S.S. then necessarily $Z(t)$ is S.S.S.
- if $X(t)$ is W.S.S. then $Z(t)$ is not necessarily W.S.S.

<u>Generalizations and highlights (we will expanded on this later in the course):</u>

[1] It is easy to generalize the claim for systems with memory and time-invariant. Namely, if $g : R^m \longrightarrow R$ and $Z(t) = g\left(\left\{X(t-\tau_i)\right\}_{i=1}^{m}\right)$, $\forall t$, such that $\{\tau_i\}_{i=1}^{m}$ is a series of constants not dependent of the time t, then if $X(t)$ is S.S.S. then necessarily $Z(t)$ is S.S.S.

[2] If $g(\bullet)$ is a system independent of time but linear to its parameters, i.e.:

$$Z(t) = g\left(\{X(t-\tau_i)\}_{i=1}^{m}\right) = \sum_{i=1}^{m} a_i X(t-\tau_i) + b$$

where $\{a_i\}_{i=1}^{i=m}, b$ are constants not dependent of time, then if $X(t)$ is W.S.S. then $Z(t)$ is necessarily W.S.S.

Problem 1:

The following random processes are given:

1. $X(t) = B\sin(2\pi t)$, $\quad B = \begin{cases} 0 & w.p. 0.5 \\ 1 & w.p. 0.5 \end{cases}$
2. $Y(t) = A\sin(2\pi t)$, $\quad A \sim N(0,1)$
3. Let $\{G_i\}_{-\infty}^{\infty}$ be a series of i.i.d random variables, $G_i \sim U[-1,1]$.

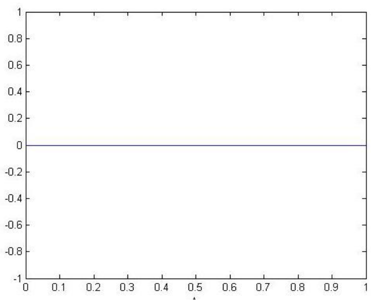
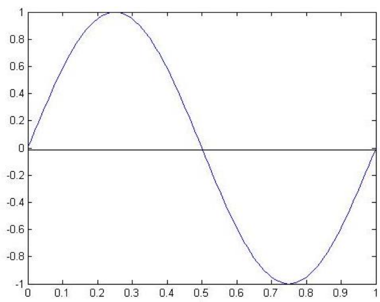
   Let us denote $Z(t) = \sum_{i=-\infty}^{\infty} G_i \cdot r(t-i)$ where $r(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & \text{else} \end{cases}$.

For the above random processes:

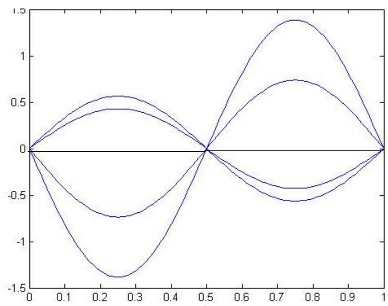
a. Draw a number of sample functions of each process.
b. For each process, find the density function (in the continuous case) or the probability function (in the discrete case) of first order.

Solution:

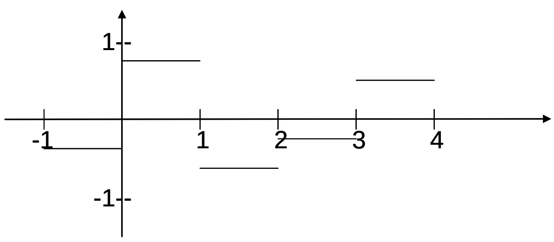
1. The sample function of the process:
   a. $X(t)$ - there are only two possible sample functions:

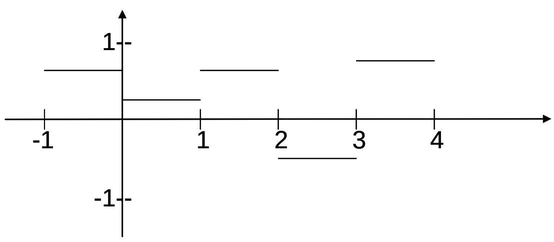



   b. $Y(t)$ - there is an infinite amount of sample functions possible:



   c. $Z(t)$ - there is an infinite amount of sample functions possible:





2. Calculating the density/probability function:
   a. $X(t)$ - for all t, $X(t)$ can take one of two values with probability 0.5, thus:

$$P(X(t)=\alpha)=\begin{cases} 0.5 \ , & \alpha=0 \\ 0.5 \ , & \alpha=\sin(2\pi t) \end{cases}$$

   b. $Y(t)$ - for all t, $Y(t)$ is a Gaussian random variable: $Y(t)=\underbrace{\sin(2\pi t)}_{const\ given\ t}A$

   Therefore:

$$Y(t)\sim N(\sin(2\pi t)E(A),\sin^2(2\pi t)\sigma_A^2)=N(0,\sin^2(2\pi t))$$

   c. $Z(t)$ - for all t, $Z(t)$ is a random variable G_i, uniform on the interval $[-1,1]$, and, therefore:

$$f_{Z(t)}(\alpha)=\begin{cases} 0.5 \ , & \alpha\in[-1,1] \\ 0 & else \end{cases} \Rightarrow Z(t)\sim U[-1,1]$$

<u>Auto-Regressive (A.R.) Random Process</u>

An A.R. random process in discrete time $X[n]$ is defined as:

$$X[n]=g(X[n-1],W[n]) \qquad \forall n>n_0$$

$g(\bullet)$ is a deterministic function, and $W[n]$ is the innovations process, which is assumed to be i.i.d. n_0 is the starting point of the process, and can even be $-\infty$. If n_0 is finite, the initial conditions must also be specified (i.e. the distribution of the sample $X[n_0]$) in order to characterize $X[n]$. $X[n_0]$ is independent of $\{W[n]\}_{n=0}^{\infty}$.

<u>Problem 2:</u>

Given is the following A.R.[6] process:

$$X_n=aX_{n-1}+W_n \qquad |a|<1$$

where W_n is an i.i.d process that satisfies $E\{W_n\}=\eta_W$ and $Var\{W_n\}=\sigma_W^2$. Furthermore, it is given that X_0 is independent of W_n and satisfies $E\{X_0\}=\eta_{X_0}$ and $Var\{X_0\}=\sigma_{X_0}^2$.

1. Find the expectation and the auto-covariance function of the process X_n.
2. Find conditions on $\eta_{X_0},\sigma_{X_0}^2,\eta_W,\sigma_W^2,a$ such that the above process will be W.S.S.

<u>Solution:</u>

[6] More specifically, it is a first order A.R. process.

1. Let us solve the following difference equation for later use:

$$z[n] = az[n-1] + b \quad n \geq 1 \quad a \neq 1$$

where $z[n]$ is some deterministic function with initial condition $z[0]$,

$$z[n] = az[n-1] + b = a(az[n-2] + b) + b = \ldots = a^n z[0] + b\sum_{i=0}^{n-1} a^i =$$

$$= a^n z[0] + b\frac{1-a^n}{1-a} = a^n\left(z[0] - \frac{b}{1-a}\right) + \frac{b}{1-a}$$

Therefore:

$$\boxed{z[n] = a^n\left(z[0] - \frac{b}{1-a}\right) + \frac{b}{1-a} \quad n \geq 0}$$

Let us start by finding the expectation:

$$\eta_X[n] = E\{X_n\} = E\{aX_{n-1} + W_n\} = a\eta_X[n-1] + \eta_W \qquad n \geq 1$$

$$\Rightarrow \eta_X[n] = a^n\left(\eta_{X_0} - \frac{\eta_W}{1-a}\right) + \frac{\eta_W}{1-a} \qquad n \geq 0$$

In order to find the auto-covariance function, we will first find the variance:

$$C_X[n,n] = E\left\{\left(X_n - \eta_X[n]\right)^2\right\} = E\left\{\left(\underbrace{aX_{n-1} + W_n}_{X_n} - \underbrace{\left(a\eta_X[n-1] + \eta_W\right)}_{\eta_X[n]}\right)^2\right\} =$$

$$= E\left\{\left(a\left(X_{n-1} - \eta_X[n-1]\right) + W_n - \eta_W\right)^2\right\} =$$

$$= E\left\{a^2\left(X_{n-1} - \eta_X[n-1]\right)^2\right\} + 2aE\left\{\left(X_{n-1} - \eta_X[n-1]\right)\left(W_n - \eta_W\right)\right\} + E\left\{\left(W_n - z\eta_W\right)^2\right\} =$$

$$= a^2 C_X[n-1,n-1] + 2a \cdot Cov\left(X_{n-1}, W_n\right) + \sigma_W^2 =$$

$$\underset{(1)}{=} a^2 C_X[n-1,n-1] + \sigma_W^2 \qquad n \geq 1$$

Notice that $C_X[n,n]$, in this case, is a function of only one variable (both its inputs take the same index), therefore we can use here, too, the result we got above, and we get:

$$\Rightarrow C_X[n,n] = \left(a^2\right)^n\left(C_X[0,0] - \frac{\sigma_W^2}{1-a^2}\right) + \frac{\sigma_W^2}{1-a^2} = a^{2n}\left(\sigma_{X_0}^2 - \frac{\sigma_W^2}{1-a^2}\right) + \frac{\sigma_W^2}{1-a^2} \qquad n \geq 0$$

Now, the auto-covariance function is:

$$
\begin{aligned}
&C_X[n,n-k] = E\left\{\left(X_n - \eta_X[n]\right)\left(X_{n-k} - \eta_X[n-k]\right)\right\} = \\
&= E\left\{\left(a\left(X_{n-1} - \eta_X[n-1]\right) + W_n - \eta_W\right)\left(X_{n-k} - \eta_X[n-k]\right)\right\} = \\
&= aE\left\{\left(X_{n-1} - \eta_X[n-1]\right)\left(X_{n-k} - \eta_X[n-k]\right)\right\} + E\left\{\left(W_n - \eta_W\right)\left(X_{n-k} - \eta_X[n-k]\right)\right\} = \\
&= aC_X(n-1,n-k) + Cov\left(W_n, X_{n-k}\right) = \\
&\underset{(1)}{=} aC_X(n-1,n-k) = \ldots = a^k C_X(n-k,n-k) = a^k\left\{a^{2(n-k)}\left(\sigma_{X_0}^2 - \frac{\sigma_W^2}{1-a^2}\right) + \frac{\sigma_W^2}{1-a^2}\right\} = \\
&= a^{2n-k}\left(\sigma_{X_0}^2 - \frac{\sigma_W^2}{1-a^2}\right) + a^k \frac{\sigma_W^2}{1-a^2} \qquad\qquad n \ge k \ge 1
\end{aligned}
$$

Justification for (1):
It holds that:

$$Cov\left(W_n, X_{n-k}\right) = 0 \qquad \forall k > 0$$

because:

$$X_{n-k} = func\left(X_0, W_1, W_2, \ldots, W_{n-k}\right)$$

and, therefore, X_{n-k} is independent of W_n (being a function of a random vector that is independent of W_n).

2. In order for the process to be W.S.S., the expectation must be constant in time:

$$
\begin{aligned}
&\eta_X[n] = a^n\left(\eta_{X_0} - \frac{\eta_W}{1-a}\right) + \frac{\eta_W}{1-a} \\
&\Rightarrow \left(\eta_{X_0} - \frac{\eta_W}{1-a}\right) = 0 \\
&\Rightarrow (1-a)\eta_{X_0} = \eta_W
\end{aligned}
$$

Moreover, the auto-covariance function must be dependent of time differences only. We will show that, specifically, the variance of the process must be constant in time:

$$
\begin{aligned}
&Var\left(X[n]\right) = C_X[n,n] = a^{2n}\left(\sigma_{X_0}^2 - \frac{\sigma_W^2}{1-a^2}\right) + \frac{\sigma_W^2}{1-a^2} \\
&\Rightarrow \left(\sigma_{X_0}^2 - \frac{\sigma_W^2}{1-a^2}\right) = 0 \\
&\Rightarrow \left(1-a^2\right)\sigma_{X_0}^2 = \sigma_W^2
\end{aligned}
$$

Notice that here we must also require that $|a| < 1$.

Let us check that this condition is indeed enough (**in this case**) for the auto-correlation function to be dependent of time differences only:

$$C_X[n,n-k] = a^k C_X(n-k,n-k) =$$
$$= a^k Var(X[n]) = a^k \frac{\sigma_W^2}{1-a^2} \qquad n \geq k, k \geq 0$$

and, in general:

$$C_X[n,n-k] = a^{|k|} \frac{\sigma_W^2}{1-a^2} \qquad n \geq k$$

<u>Summary of Necessary and Sufficient Conditions for W.S.S.:</u>

$$(I)\ |a|<1, \quad (II) \quad \eta_{X_0} = \frac{\eta_W}{1-a},\ (III)\ \sigma_{X_0}^2 = \frac{\sigma_W^2}{1-a^2}$$

Notice that if condition (I) holds, then it is guaranteed that:

$$\eta_X[n] \xrightarrow{n\to\infty} \frac{\eta_W}{1-a}$$
$$C_X[n,n-k] \xrightarrow{n\to\infty} a^{|k|} \frac{\sigma_W^2}{1-a^2}$$

regardless if whether the other two conditions hold. In this case, we say that the process is asymptotically W.S.S.

3. Now it is given that W_n is an i.i.d Gaussian series. Are the conditions we found in the previous section enough to say that the process $X[n]$ is asymptotically S.S.S.?

   <u>Answer:</u> yes. For the conditions we found in the previous section, the initial conditions of the process decay, namely for a big enough n, $X[n]$ is not affected by $X[0]$. For such a value of n, any sample of the process is a linear function of the samples $\{W_i\}_{i=1}^n$, which constitute a Gaussian random vector since W_n is a Gaussian random process. Therefore, $X[n]$ is also a Gaussian random process, asymptotically. Under the conditions of the previous section, $X[n]$ is also asymptotically W.S.S., and from here that it is asymptotically S.S.S.

Extra Questions

Problem 3:

Determine whether the following claim is correct:

1. If $X[n]$ is a Gaussian random process, then $X[-n]$ is also a Gaussian random process.

In each one of the following cases, determine if $X[n]$ is **necessarily** an i.i.d random process. If yes, prove! Otherwise, give an explicit counter-example and explain.

2. Given: $X[n] = X_1[n] + X_2[n]$, where the processes $X_1[n], X_2[n]$ are independent of one another and each one of them is an i.i.d random process.
3. Given: $X[n] = \begin{cases} X_1[n] & Y = 1 \\ X_2[n] & Y = 0 \end{cases}$, where the processes $X_1[n], X_2[n]$ are independent of one another and each one of them is an i.i.d random process, and Y is a Bernoulli random variable, independent of the processes $X_1[n], X_2[n]$.
4. The **optimal** MMSE estimator of a single sample $X[n]$ from the sample $X[m]$ is 0 for all $m \neq n$.

Solution:

1. Correct.
   Any set of samples of $X[-n]$ is also a set of samples of $X[n]$ and, thus, necessarily the samples constitute a Gaussian random vector. A random process for which **all** of its samples constitute a Gaussian random vector, is a Gaussian random process.

2. The process $X[n]$ is necessarily i.i.d.

   Let us prove that all of its samples are identically distributed:
   $$\forall n \ X[n] \sim X_1 + X_2$$
   where X_1, X_2 are two independent random variables having the distribution of the samples $X_1[\ ], X_2[\ ]$, respectively.

   Let us prove that all the samples are independent:
   Any vector of the samples of the process:
   $$\begin{bmatrix} X[k_1] & X[k_2] & \cdots & X[k_n] \end{bmatrix}^T$$

is a vector of independent samples, since all of its elements are functions of samples that are independent of all the samples that the rest of the elements of the vector are (deterministically) dependent of.

3. The process $X[n]$ is not necessarily i.i.d, let us provide a counter-example:
Let us denote the following processes:

$$X_1[n] = \begin{cases} 2 & w.p.\ 0.5 \\ -2 & w.p.\ 0.5 \end{cases}$$

$$X_2[n] = \begin{cases} 4 & w.p.\ 0.5 \\ -4 & w.p.\ 0.5 \end{cases}$$

where $X_1[n], X_2[n]$ are independent, and each one of the processes is i.i.d.
Since the process $X[n]$ takes $X_1[n]$ with probability half and $X_2[n]$ with probability half, the marginal distribution of $X[n]$ is:

$$X[n] = \begin{cases} -4 & w.p. 0.25 \\ -2 & w.p. 0.25 \\ 2 & w.p. 0.25 \\ 4 & w.p. 0.25 \end{cases}$$

However, it is clear that, given $X[n-1]$, we can know deterministically if we will choose the process $X_1[n]$ or $X_2[n]$, and, therefore, $X[n]$ is dependent of $X[n-1]$.

4. The process $X[n]$ is not necessarily i.i.d, it is enough to show an example of a process that is not identically distributed yet does satisfy the conditions. Let us take a look at the process:

$$X[n] = \begin{cases} X_1[n] & n \text{ is odd} \\ X_2[n] & n \text{ is even} \end{cases}$$

where the processes $X_1[n], X_2[n]$ are independent of one another and each one of them is i.i.d with distribution:

$$X_1[n] \sim Unif(-b, b)$$
$$X_2[n] \sim Unif(-c, c) \qquad 0 < b < c$$

Since for all $m \neq n$ it holds that $X[n]$ is independent of $X[m]$, the optimal estimator would be $E\{X[n] | X[m]\} = E\{X[n]\} = 0$ and the conditions of the problem are satisfied, even though $X[n]$ is not identically distributed.