EE353 Lecture 20: Intro To Random Processes

Chapter 9: 9.1: Definition of Random Processes

In certain random experiments, the outcome is a function of time and space. In the example we used last time, speech recognition systems make decisions based on the voltage waveform corresponding to a speech utterance. If we recorded everyone in the class saying the exact same phrase (e.g., “Go Navy” or better yet, “Go Hokies”), we would find that no two recordings were identical. How then does Siri, or Google Voice, or the NSA wiretaps recognize certain keywords?

Voice Recording and Corresponding Spectrum
https://commons.wikimedia.org/wiki/File:Voice_waveform_and_spectrum.png
Taking it one step further, what about those wonderful biometric identification systems (such as the fingerprint scanner, hand geometry scanner, iris recognition, or voice recognition over in the ECE Biometrics Lab)? What if your phone is locked via your voice and if you get a cold? Does that mean you can no longer access your phone? What if your classified iPad is biometrically locked to your specific EKG pattern (don’t laugh; Dr. Anderson was one time part of a panel discussion on this very topic). What if you’re taking heavy fire or are otherwise in a stressful situation, does that mean you can’t access the SIPRNET?

Clearly something else must be going on here to support those applications.

The functions in the above examples can be viewed as numerical quantities that evolve randomly as a function of time or space. What we really have then is a family of random variables indexed by the time or space variable. For the final part of the course, we will investigate these random processes.

**Important Distinction**

A sine-wave is a deterministic signal. Given the frequency, amplitude and initial phase offset, we can predict/determine what its value will be at any time $t$.

The alternative is a signal that we cannot predict where the value of the signal is random. However, if we can say that it is “probably” going to be within a certain range, then we can say that we have a stochastic signal.

In communications, we often describe many real signals as a random signals as value being transmitted in the signal as unknown before it is received (known in advance to transmitter but not to the receiver).

**Definition:**

A stochastic signal is one which we may describe future values only in terms of probability of the signal being within a certain range.
**Definition of a Random Process (9.1)**

Consider a random experiment specified by the outcomes $\zeta$ from some sample space $S$, by the events defined on $S$, and by the probabilities on those events. Suppose that for every outcome $\zeta \in S$, we assign a function of time according to some rule: $X(t, \zeta)$, $t \in \mathbb{R}$

**Definition:** The graph of the function $X(t, \zeta)$ for a fixed $\zeta$ is called a **realization**, **sample path**, or **sample function** of the random process. Thus, we can view the outcome of the random process as producing an entire function of time, as illustrated below.

![Image showing the definition of a random process](image)

On the other hand, if we fix the time index to some time $t_k$, then our function $X(t_k, \zeta)$ will be a **random variable**, since we are mapping $\zeta$ into a real number. As a result, we have created a family (or **ensemble**) of random variables **indexed by time**. This family of random variables is known as a **random (or stochastic)** process.

For notational convenience, we usually suppress the $\zeta$ and use $X(t)$ to denote a random process.
**Double Important Distinction!**

- **Random variables** model unknown events.
- **Random processes** model unknown signals.
- A *random process* is just a collection of *random variables*.

A **random process** is an indexed set of functions of some parameter (usually time) that has certain statistical properties.

If \( X(t) \) is a random process then \( X(1) \), \( X(1.5) \), and \( X(37.25) \) are all **random variables** for each specific time \( t \).

Samples of \( X(t) \) are **Joint Random Variables**: 
\[
P\left[ X(t_0) \leq x_0 \right] = P\left[ \bigcup_{t} x_i(t_0) \leq x_0 \right]
\]

A **Stochastic Process** is said to be **Discrete Time** if the index is a countable set (i.e., the set of integers or the set of nonnegative integers). For notational convenience, we usually use the variable \( n \) to denote the time index and \( X_n \) to denote the random process.

A **Stochastic Process** is said to be **Continuous Time** if the index is continuous (i.e., the real line or the nonnegative real line).

In theory, we could calculate the probabilities of events involving \( X(t, \zeta) \) in terms of the underlying events \( \zeta \), using the random variable transformation techniques we taught previously. However, these functions become extremely complex extremely quickly in practice; thus we generally have to work with an alternative technique.
Specifying a Random Process (9.2)

Many electrical and computer engineering problems cannot be answered with knowledge of the distribution of a random variable at a single instant in time. For example, we may be interested in the temperature of a location at two different times, which requires the following information:

\[ P\left[ x_1 < X(t_1) \leq x_2, \ x_1 < X(t_2) \leq x_2 \right] \]

For another example, the Voice Encoders (known as Vocoders) that perform speech compression in mobile phones use a technique that predicts the amplitude of the speech signal at the next sampling time based on the previous \( k \) samples. Thus, we may be interested in the following probability:

\[ P\left[ a < X(t_{k+1}) \leq b \mid X(t_1) = x_1, \ X(t_i) = x_i, \ldots X(t_k) = x_k \right] \]

In both of these cases, we are interested in the probability of a vector of samples of the random process.

It bears repeating that the Random Process is a function that assigns every outcome \( \zeta \) in our samples space \( S \) to a real valued function. Hence \( X(t, \zeta) \) is a mapping of the sample space to a family of functions.
Example

A random experiment consists of rolling 1D4 (one four-sided dice) and noting the result. Thus, the sample space is $S = \{1,2,3,4\}$. Suppose we define the following mapping:

$$X(n, \zeta) = \sin \left( \frac{\pi n}{2\zeta} \right) = \begin{cases} 
\sin \left( \frac{\pi n}{2} \right) & \zeta = 1 \\
\sin \left( \frac{\pi n}{\zeta} \right) & \zeta = 2 \\
\sin \left( \frac{\pi n}{\zeta} \right) & \zeta = 3 \\
\sin \left( \frac{\pi n}{\zeta} \right) & \zeta = 4 
\end{cases}$$

Then the Discrete Random Process looks like the following:
As an illustrative example, thermal noise (the noise generated by any component or device which is not at a
temperature of absolute zero) is generally modeled as voltage which is a **Gaussian Stochastic Process**. What
does that mean?

Let’s presume we have a noisy resistor, and we connect that resistor to four different oscilloscopes and
measure the voltage generated (note that the same effect could be achieved if we measured four similar
resistors simultaneously on four oscilloscopes or four channels of the same oscilloscope). The experimental
setup and resulting plots are shown below.

What we observe is that the sample functions are **different**. In fact, if we repeated this test over and over
again, they would still be **different**. How could we possibly model this signal, you ask? Well, because it is a
Gaussian Stochastic Process, the amplitude at any given time (e.g., \( t = 65 \) as shown in the figure) is given by
the underlying **Gaussian Distribution** (with its specific \( \mu, \sigma \)). If we repeated this experiment hundreds and
hundreds of times, and each time recorded the amplitude of the noise voltage at \( t = 65 \), we would see that the
amplitudes fit a Gaussian PDF.
Example (9.2)

Let $\zeta$ be selected at random (uniformly) from the interval $[-1,1]$. Define two continuous-time random processes $X(t, \zeta)$ by:

$$X(t, \zeta_1) = \zeta_1 \cos(2\pi f_1 t + 0)$$
$$X(t, \zeta_2) = \cos(2\pi f_1 t + \zeta_2 [\text{rad}])$$

Plot the realizations of the sinusoids generated from this random process.

Matlab is our friend here:

```matlab
zeta = -1:0.1:1;
t=0:0.01:1;
fc = 1;

figure(1)
X_1(1,:) = zeta(1).*cos(2.*pi.*fc.*t);
plot(t,X_1(1,:),'k-','linewidth',2)
hold on

figure(2)
X_2(1,:) = cos(2.*pi.*fc.*t + zeta(1));
plot(t,X_2(1,:),'k-','linewidth',2)
hold on

for i = 2:length(zeta)
    X_1(i,:) = zeta(i).*cos(2.*pi.*fc.*t);
    X_2(i,:) = cos(2.*pi.*fc.*t + zeta(i));
    figure(1)
    plot(t,X_1(i,:),'k-','linewidth',2)
    figure(2)
    plot(t,X_2(i,:),'k-','linewidth',2)
end

figure(1)
xlabel('Time (s)')
ylabel('Amplitude')
title('Random Sinusoids with Amplitude Uniformly Distributed on [-1,1]')

figure(2)
xlabel('Time (s)')
ylabel('Amplitude')
title('Random Sinusoids with Phase Uniformly Distributed on [-1,1]')
```
Which Results in:

**Note That:** The randomness in \( \zeta \) induces randomness in the observed function \( X(t, \zeta) \). In principle, we can deduce the probability of events involving a stochastic process at various instants of time from probabilities involving \( \zeta \).

At first glance, it does not appear that we’ve made much progress in specifying a random process, as we are now confronted with the task of specifying a vast, multi-dimensional CDF and PDF (and we thought two dimensions was hard work!). However, the vast majority of stochastic processes that Electrical or Computer Engineers are interested in can be obtained by elaborating on a few simple models. Moreover, in many cases, and *in lieu of* completely specifying the random process, we can use a few statistical techniques to tell us virtually everything we need to know about the process. These techniques are similar to using the mean and variance to specify most of what we need to know about a Random Variable.
Bernoulli Random Process (Example 9.5)

Let \( X_n \) be a sequence of independent, identically distributed Bernoulli random variables with \( p = \frac{1}{2} \). We can think of these as flipping a coin multiple times, or perhaps flipping a bit multiple times. It can be shown, using the techniques in Chapter 5, that the Joint PMF for any \( k \) time samples is then:

\[
P[ X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k ] = P[ X_1 = x_1 ] \cdots P[ X_k = x_k ] = \left( \frac{1}{2} \right)^k
\]

Where \( x_i \in \{0, 1\} \)

**Note That:** This gives us the probability of a particular binary sequence being transmitted. If we think about what this means for a moment:

- **Probability of a Binary Sequence of Length = 1:** \( P = \left( \frac{1}{2} \right)^1 = \frac{1}{2} \)
- **Probability of a Binary Sequence of Length = 2:** \( P = \left( \frac{1}{2} \right)^2 = \frac{1}{4} \)
- **Probability of a Binary Sequence of Length = 3:** \( P = \left( \frac{1}{2} \right)^3 = \frac{1}{8} \)
- **Probability of a Binary Sequence of Length = N:** \( P = \left( \frac{1}{2} \right)^N = \frac{1}{2^N} \)

**What this means:** If we toss a fair coin \( N \) times, there are \( 2^N \) different sequences of heads and tails possible, all of them equally likely. So the probability of getting the one sequence among them that contains exactly \( N \) heads is \( \frac{1}{2^N} \). We could extend this analysis to an experiment that has \( d \) equally probable outcomes for a single trial (e.g., rolling a \( d \)-sided dice). If we carry out \( N \) trials, there would be \( d^N \) different sequences of results possible, and the probability of getting the one sequence where a particular outcome – say, rolling a string of 1’s, is \( \frac{1}{d^N} \).
**Gaussian Random Process (Example 9.6/9.8)**

Let $X_n$ be a sequence of independent, identically distributed Gaussian random variables with zero mean and variance $\sigma^2_X$. The joint PDF for any $k$ time samples is given by:

$$f_{x_1, x_2, ..., x_k}(x_1, x_2, ..., x_k) = \frac{1}{(2\pi\sigma^2)^{k/2}} e^{-\frac{1}{2\sigma^2}(x_1^2 + x_2^2 + \cdots + x_k^2)}$$

How is that useful? Consider observing a signal embedded in noise (something all of you will do throughout your career). Let $X_j$ be a sequence of $J$ IID observations of a signal voltage $\mu$ corrupted by zero-mean Gaussian noise $N_j$ with variance $\sigma^2$:

$$X_j = \mu + N_j \quad \text{for} \quad j = 1, 2, ..., J$$

Consider what would happen if we average the sequence of observations:

$$S_J = \frac{X_1 + X_2 + \cdots + X_J}{J}$$

From our discussion on the *Central Limit Theorem*, we know that $S_J$ is the sample mean of an IID sequence of Gaussian random variables. We also know that $S_J$ is itself a Gaussian random variable with mean $\mu$ and variance $\text{VAR}[S_J] = \frac{\sigma^2}{J}$.

What that means: As $J$ increases, the variance of $S_J$ decreases, and the calculated mean will tend towards the value of the true mean (i.e., the amplitude $\mu$ of our signal). The more observations we can make of a noisy signal, the more accurate our estimate of its true value will be. Of course, this presumes that our signal is periodic, but this is a powerful result for the world of communications! Think about it for a moment; our signal could be completely obliterated by noise, *and yet*, we could still recover its value simply by *averaging out the noise*!