

Digital Communication Using GNU Radio

Prof Kumar Appiah

Department of Electrical Engineering

Indian Institute of Technology Bombay

Week-02

Lecture-08

Welcome to this lecture on Digital Communication Using GNU Radio. My name is Kumar Appiah. In this lecture, we will be going through a quick recap of random variables and random processes. While you may be familiar with this material from the prerequisites, we will still do a quick run through so that you have a recap and we will also go through the material with an eye on the applications specific to digital communication. Random variables and random processes. As you may recall, a random variable is a mapping from the probability space to real or complex numbers.

To be more specific, if you remember, a probability space consists of elementary events and mapping these events or subsets of these events into real numbers or complex numbers is what yields random variables. A common example will be the role of a die. If you roll a die, you end up getting a number which is 1, 2, 3 or some number up to 6. This is an example of a discrete random variable and assuming that you have a fair die, all the numbers are equally likely.

An example of a continuous random variable is the standard Gaussian. In the case of a Gaussian, you know that the probability distribution function has a bell shape and generating a random number from the standard Gaussian essentially yields a realization of a random variable that has a probability distribution function corresponding to the Gaussian. Implicitly, there are concepts such as probability mass function and probability density function. Probability mass function of course is for discrete and probability density function is for continuous random variables. Cumulative distribution functions and so on, these are all characteristics that define the random variable.

In other words, the probability mass function or probability density function and other such functions essentially determine the statistics of the random variable, what values it can take, with what probability and so on. Even though the random variable is essentially random, this gives you a hint of statistics as to what values are likely, unlikely and so on. An extension of these random variables to the case where you have a collection of these that can be indexed with a variable such as time is essentially a random process. A

simple example would be to roll a die multiple times, we can say let us say once per second and then tabulate those values and this becomes a nice discrete time random process. Therefore, a collection of random variables that are indexed by time in case time is measured using a continuous approach

$$t \in \mathbb{R},$$

but if you want to measure it in a discrete approach then you can choose

$$n$$

as n an integer.

Both of these are random processes of course, the former one is a continuous time random process while the latter one is a discrete time random process. Since random processes deal with collections of random variables, we need ways to understand the relationship across two random variables that are part of the random process. As we just discussed, random processes can either be continuous time that is

$$X(t)$$

or discrete time that is

$$X[n].$$

To keep consistent with the common literature whenever we have continuous time processes we will be using the parenthesis notation and generally the variable will be

$$t$$

while whenever we have discrete time processes we will use the square bracket notation and stick to

$$n$$

as our time index. Now just like random variables whenever we deal with random processes we need some tools or ways to analyze these.

So, we need to know the statistics such as mean, variance or the relationship between successive random variables that are part of the random process so that we can understand what this random process is and how we can design around it. For example, if you are talking about something like a noise random variable we need to know how much noise there is so that we will be able to signal in a way so as to peak that noise. Therefore, in practice the statistics of the process needs to be known or you must be able to measure it. The ability to measure the statistics of the random process is very

important for practical systems for example if you are making a call on your phone your phone has to learn the environment and learn the statistics of the environment so that you can successfully conduct a phone call. Some examples of random processes are your data streams themselves because let us say that you are dealing with a voice communication system.

The system does not know in advance the words that you are going to speak. Therefore, all these systems model the input that is actually to be communicated as a random process. Noise as we just discussed is a random process. You may recall from your circuits experiments if you connect just a noise source or a source which has no voltage to the oscilloscope and then view it in the millivolt or microvolt range you will start seeing some oscillations that are unpredictable. This is called thermal noise or circuit noise and while it is unpredictable thankfully its amplitude is generally limited or low which is why we are able to communicate despite the presence of this noise.

Yet the statistics of this noise is extremely important for us as the designers of digital communication systems. Finally, channel measurements are also random processes. Practically speaking let us say that you are conducting a call on your phone. If you walk the environment around you changes the strength of the signal from the base station to your phone changes and therefore the impulse response of the system that connects the base station to your phone essentially changes and the characteristics of this particular channel or the parameters are also examples of random processes. Therefore, practically speaking random processes occur extremely frequently in digital communication systems.

When dealing with random processes like I just mentioned it is essential to know the statistics and it is also essential to know that the statistics do not change with time. Continuing with the same example of your phone, your phone can learn some characteristics of the random processes it is dealing with such as your voice, such as the environment around it, such as the noise. But if these parameters themselves undergo radical changes within short durations of time, then you cannot successfully communicate because by the time you have essentially designed the waveform to communicate your voice to the base station, if the parameters of the system have changed then your signal will not be able to reach the base station. Therefore, we must operate in a regime where some statistics of these random processes remain constant with time. This practical aspect translates mathematically to the case of wide sense stationarity.

That is if your random process is said to be wide sense stationary, the second order statistics, of course first order statistics implicitly and second order statistics do not vary with time. First order statistics are generally, is generally mean. Second order statistics

are like variance and correlations. The reason why they are called second order is because it involves multiplication of a signal by a signal. So it is like square of signal.

If the second order statistics do not vary with time, then you can design your system to account for this and successfully communicate despite all the obstacles that it faces in the form of unknown realizations of random processes. Let us take a very simple example of a random process that takes plus or minus one, it is a discrete time random process that takes the values plus or minus one with equal probabilities and for every instant of time it takes the value plus one or minus one and these are independent across time. That is

$$X[0] = \pm 1$$

with probability half each,

$$X[1] = \pm 1$$

with probability half each independent of what

$$X[0]$$

was and similarly

$$X[2]$$

and so on. It is very easy for you to check that for the above process the mean is zero. In other words

$$E[X[n]] = 0$$

independent of n.

That is its first order statistics do not vary with time. But more importantly if we define the autocorrelation function as

,

it is very easy to show that this does not depend on n. It depends only on k. Why is this? It is very easy. If you look at

$$E[X[n] X[n-k]],$$

let us choose

$$k=1.$$

You have

$$E[X[n]X[n-1]].$$

As we just discussed

$$X[n] \text{ and } X[n-1]$$

are independent. Therefore, you can separate these and you get

$$E[X[n]]E[X[n-1]]=0$$

and does not depend on n. Similarly, if you choose

$$k=2, 3, -1, -2$$

and so on, you will still get

$$E[X[n]X[n-k]]=0$$

and the only exception is if you take

$$E[X[n]X[n-k]]$$

for the case where

$$k=0.$$

In this case, you are essentially finding the

$$E[(X[n])^2]$$

which is the second order statistic and since mean is zero, this will also be the variance.

So this is going to be

$$1(1/2) + (-1)(1/2) = 1$$

and this is still not dependent on n. Therefore, this process is wide sense stationary because its second order statistics do not vary with time. There is a stronger condition called strict sense stationarity. In the case of strict sense stationarity, the distribution or rather the joint distribution of the variables in the random process themselves depend only on k, the lag and not n, but that is a stronger condition which we do not need for our design processes in this particular application. Nevertheless, it is a useful characteristic and strict sense stationarity implies wide sense stationarity, but the converse is not true.

Finally, as we emphasize, wide sense stationary processes are very convenient for our design because once you have an idea of the mean, the variance and the correlations that a process exhibits, you can then tune your design in order to handle these and yet be successful in communicating your signal. As we discussed previously, the auto correlation function for a wide sense stationary process can be defined in this manner. For a random process

$$s(t),$$

the auto correlation function

In this particular case, it depends only on

$$\tau$$

and not

$$t,$$

that is the

$$t$$

dependence is not there and this

$$\tau$$

is called the lag. The reason it is called lag is because this corresponds to the difference between the current time instant and the next time instant across which you are performing the correlation.

In other words,

$$s(0)$$

and

$$s(3)$$

have the same auto correlation or same correlation as

$$s(1)$$

and

$s(4)$

and

$s(10)$

and

$s(13)$

because all of them have the same gap of 3 or lag of 3 or -3 depending on the definition. For a discrete time random process similarly

$x[n]$,

the definition can be written as

$$E[x[n]x^*[n-k]].$$

In this case again, there is no n dependence, the dependence is only on the lag k . Therefore,

$x[0]$ and $x[10]$

are correlated in the same way as

$x[1]$ and $x[11]$, and $x[2]$ and $x[12]$

because they have a gap of 10 samples each. Intuitively the auto correlation function specifies the temporal correlation of the random process, that is you get an idea of how much correlation there is between the realization of the random variable, the current time instant with some other time instant let us say at some number of seconds apart.

Let us look at a practical example. If we take

where

$z[n]$

is an iid random variable independent of

$x[n]$.

In fact, we can also say it is Gaussian with mean 0 and variance 1. Then, we can write

$$R_x[k] = \alpha^{|k|}.$$

This is something which you can refer to some text books and find out.

This is called an AR(1) random process. The reason is because the random process depends on itself, that is

$$x[n]$$

depends only on

$$x[n-1]$$

and it depends on only one past time instant. So, if you have dependence on

$$\text{and } x[n-2]$$

you will call it an AR(2) process and so on. Let us assume that

$$\alpha \in [0,1].$$

Then, the auto correlation function is

$$\alpha^{|k|}$$

will look like this.

Let us say I draw it continuously it will look like this. That is it will have some amount of decay as you keep going further and further. Of course, if you substitute k equal to 0 you will get 1. Therefore, the variance of the process

$$x[n]$$

is going to be 1. In fact, for this particular model it is going to be Gaussian as well.

However, unlike

$$z[n]$$

which is an iid Gaussian

$$x[n]$$

is going to be correlated and the correlation is

$$\alpha^{|k|}$$

is going to be reasonably higher for smaller values. As you can see

$$x[n] = \alpha x[n-1]$$

which means

$$x[n]$$

is going to be correlated with

$$x[n-1]$$

to the tune of

$$\alpha.$$

Similarly,

$$x[n]$$

is going to be correlated with

$$x[n-2]$$

to the tune of

$$\alpha^2$$

and so on. But as

$$\alpha \in [0,1]$$

as k becomes larger let us say 10, 12, 13 the correlation goes to a very small value. In fact, the higher the value of alpha the more

$$x[n] \text{ and } x[n-1]$$

are

correlated.

But if you choose

$$\alpha$$

to be a number much closer to 0 like 0.01 then

$$x[n]$$

is also close to an iid process. It is not specifically an iid process because of the correlation. In the limit where

$$\alpha=0$$

then

$$x[n]=z[n]$$

it will be an iid process. Therefore,

$$R_x[k]$$

gives you an idea of the correlation properties of

$$x[n].$$

In a similar manner we can define cross correlation. I am implementing the correlation implicitly defining what a jointly wide sense stationary process is. Jointly wide sense stationary processes or pairs of processes are those where the joint second order statistics do not vary with time. In other words, if you define

$$E[s_1(t)s_2^*(t-\tau)]$$

this should depend only on

$$\tau$$

and not on t then you can call such pairs of random processes jointly wide sense stationary. In the case of discrete processes as well if their joint statistics depends only on the lag then they are said to be jointly wide sense stationary.

This tracks the temporal correlation across a pair of random processes. Here this becomes significant because if you have joint statistics across let us say your data and the channel that are both random processes then it is very easy to design your communication to handle that. If however there are some properties that get affected and your channel statistics vary differently from the actual data and they are not jointly wide sense stationary the design process becomes more complex. We will restrict our consideration to the scenarios where the processes are indeed jointly wide sense stationary. The next thing we must be aware of is power spectral density.

So, how can we have a Fourier transform for random processes? If you have a realization of a random process and take its Fourier transform that does not make much sense directly because a realization has a Fourier transform a different realization may

have a different Fourier transform and this analysis may not make much sense especially because it is a random process. A better approach would be for us to have a notion of second order statistics or expectations of power over frequency content. So, how does this translate? The power spectral density can be considered like the energy within the random process at a particular frequency. In other words if we assume that the random process is constituted by addition of several sinusoids the power spectral density gives you the statistics of how much power a particular frequency contributes to the random process. In other words you can consider it to be the amount of power that a particular frequency contributes if you assume that the frequency is part of the random process.

A more direct definition is that the power spectral density is the Fourier transform of the auto correlation function. We will give a quick justification of why that is the case. So as I try to give you an intuition if you say

$$S(f)$$

is the power spectral density of a random process in this case we are looking at a continuous time random process you can get it by taking the auto correlation of sorry the Fourier transform of the auto correlation of a random process. An intuition of how to interpret this is this take a very narrow window of Δf of course we will take it at both sides because the auto correlation will always have a symmetric power spectral density. There is a spectral content in a small frequency range if you now measure the power or let us say the exact contribution of this particular frequency range by placing a filter a narrow filter which is centered around this frequency and has a narrow band that power essentially corresponds to the power of that frequency in your random process.

How does this translate to as the Fourier transform of the auto correlation function? If in the auto correlation function you have a repeated pattern of statistics where a particular random variable is correlated highly with something else at another gap and something else at the similar gap and so on. This repeated correlation statistics essentially means that you can detect a particular sinusoid at that frequency that is present in the random process with a reasonably high power. The exact proof of the fact that the auto correlation functions Fourier transform is the power spectral density is out of the scope of this but you can easily find it in the references. The final thing which we will deal with is Ergodicity. Ergodicity essentially implies that time averages are the same as ensemble averages.

These are practically very useful for probabilistic systems. Let us take an example of two scenarios. In scenario one if we toss a coin and then we are just going to output a string of zeros if we get heads and a string of ones if we get tails and let us define this to be our random process. In this particular case you are only going to have two

realizations. The first realization is where the random process essentially looks like 0 0 0 0 and so on.

The alternate realization is where it just looks like 1 1 1 1 1 and so on. Let us suppose that you wanted to find whether the coin is fair or not. If I give you this realization of 0 0 0 and so on and average it you will end up getting 0 and you will say that the coin is heads all the time. Similarly for the other case you will end up saying that the coin is tail all the time. So in this particular example the random process defined in this way will not be ergodic because if I take the average of the realizations that does not correspond to the true average even if the coin were a fair coin.

But if you now toss the coin multiple times and each time you toss you get a head you write 0 you get a tail you write 1 you get a head you write 0 you get another head you write 0 you get a tail you write 1 and so on. In this particular scenario you have reasonably fair realizations and if you take an average over a large enough realization string you will end up getting close to half that reflects the true fairness of the coin in that the coin gives you a head or tail with equal probability. In practical situations in digital communication we need ergodicity because only then can we measure a particular statistic in time and then take its average and conclude that that average is also the ensemble average. If you do not have ergodicity practical design becomes complicated. So how do we use these in digital communication systems? So the bits or symbols that you communicate are essentially discrete time random processes.

Then as we discussed uncertainty in the environment or continuous time random processes because as you move the environment changes impulse response can change and this is also modeled as a random process. There is of course noise and other forms of degradation that is also not known a priori and that also changes in unexpected ways and several other scenarios that can occur in your design which are all treated as random processes. For the purposes of our discussion we will be handling many of these and how to overcome them using our practical design. To summarize joint wide sense stationary processes are those where the second order statistics remain constant that is auto and cross correlations depend only on lag. The power spectral density is the Fourier transform of the auto correlation function.

It is a measure of the power per frequency of a random process and this also gives you an idea of what frequency content is present significantly in a random process. Finally ergodicity is where time average of a random process statistics correspond to ensemble averages and this is extremely useful in several practical scenarios where you can use samples of a random process in order to estimate statistics by taking time averages. In future lectures we will be using these tools to aid our design and analysis. Thank you.