ECE438 - Laboratory 7:
Discrete-Time Random Processes (Week 1)
October 6, 2010

1   Introduction

Many of the phenomena that occur in nature have uncertainty and are best characterized statistically as random processes. For example, the thermal noise in electronic circuits, radar detection, and games of chance are best modeled and analyzed in terms of statistical averages.

This lab will cover some basic methods of analyzing random processes. Section 2 reviews some basic definitions and terminology associated with random variables, observations, and estimation. Section 3 investigates a common estimate of the cumulative distribution function. Section 4 discusses the problem of transforming a random variable so that it has a given distribution, and lastly, Section 5 illustrates how the *histogram* may be used to estimate the probability density function.

Note that this lab assumes an introductory background in probability theory. Some review is provided, but it is unfeasible to develop the theory in detail. A secondary reference such as [1] is strongly encouraged.

---

2 Random Variables

The following section contains an abbreviated review of some of the basic definitions associated with random variables. Then we will discuss the concept of an *observation* of a random event, and introduce the notion of an *estimator*.

2.1 Basic Definitions

A *random variable* is a function that maps a set of possible outcomes of a random experiment into a set of real numbers. The probability of an event can then be interpreted as the probability that the random variable will take on a value in a corresponding subset of the real line. This allows a fully numerical approach to modeling probabilistic behavior.

A very important function used to characterize a random variable is the *cumulative distribution function (CDF)*, defined as

$$F_X(x) = P(X \leq x) \quad x \in (-\infty, \infty) . \tag{1}$$

Here, X is the random variable, and $F_X(x)$ is the probability that X will take on a value in the interval $(-\infty, x]$. It is important to realize that x is simply a dummy variable for the function $F_X(x)$, and is therefore not random at all.

The derivative of the cumulative distribution function, if it exists, is known as the *probability density* function, denoted as $f_X(x)$. By the fundamental theorem of calculus, the probability density has the following property:

$$\begin{aligned}
\int_{t_0}^{t_1} f_X(x)dx &= F_X(t_1) - F_X(t_0) \\
&= P(t_0 < X \leq t_1) .
\end{aligned} \tag{2}$$

Since the probability that X lies in the interval $(-\infty, \infty)$ equals one, the entire area under the density function must also equal one.

Expectations are fundamental quantities associated with random variables. The expected value of some function of X, call it $g(X)$, is defined by the following.

$$\begin{aligned}
E[g(X)] &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \quad \text{(for X continuous)} \\
E[g(X)] &= \sum_{x=-\infty}^{\infty} g(x)P(X = x) \quad \text{(for X discrete)}
\end{aligned}$$

Note that expected value of g(X) is a deterministic number. Note also that due to the properties of integration, expectation is a linear operator.

The two most common expectations are the mean μ_X and variance σ_X^2, defined by

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{3}$$

$$\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx . \tag{4}$$

A very important type of random variable is the *Gaussian* or *normal* random variable. A Gaussian random variable has a density function of the following form:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2\sigma_X^2}(x - \mu_X)^2\right) \ . \tag{5}$$

Note that a Gaussian random variable is completely characterized by its mean and variance. This is not necessarily the case for other types of distributions. Sometimes, the notation $X \sim N(\mu, \sigma^2)$ is used to identify X as being Gaussian with mean μ and variance σ^2.

2.2 Samples of a Random Variable

Suppose some random experiment may be characterized by a random variable X whose distribution is unknown. For example, suppose we are measuring a deterministic quantity v, but our measurement is subject to a random measurement error ε. We can then characterize the observed value, X, as a random variable, $X = v + \varepsilon$.

If the distribution of X does not change over time, we may gain further insight into X by making several independent observations $\{X_1, X_2, \ldots, X_N\}$. These observations X_i, also known as *samples*, will be independent random variables and have the same distribution $F_X(x)$. In this situation, the X_i's are referred to as *i.i.d.*, for *independent* and *identically distributed*. We also sometimes refer to $\{X_1, X_2, \ldots, X_N\}$ collectively as a sample, or observation, of size N.

Suppose we want to use our observation $\{X_1, X_2, \ldots, X_N\}$ to estimate the mean and variance of X. Two estimators which should already be familiar to you are the *sample mean* and *sample variance* defined by

$$\hat{\mu}_X = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{6}$$

$$\hat{\sigma}_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{\mu}_X)^2 \ . \tag{7}$$

It is important to realize that these sample estimates are functions of random variables, and are therefore themselves random variables. Therefore we can also talk about the statistical properties of the estimators. For example, we can compute the mean and variance of the sample mean $\hat{\mu}_X$.

$$E\left[\hat{\mu}_X\right] = E\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N}\sum_{i=1}^{N} E\left[X_i\right] = \mu_X \tag{8}$$

$$
\begin{aligned}
Var\left[\hat{\mu}_X\right] &= Var\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N^2}Var\left[\sum_{i=1}^{N} X_i\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N} Var\left[X_i\right] = \frac{\sigma_X^2}{N}
\end{aligned}
\tag{9}
$$

In both (8) and (9) we have used the i.i.d. assumption. We can also show that $E[\hat{\sigma}_X^2] = \sigma_X^2$.

An estimate $\hat{a}$ for some parameter a which has the property $E[\hat{a}] = a$ is said to be an *unbiased* estimate. An estimator such that $Var[\hat{a}] \to 0$ as $N \to \infty$ is said to be *consistent*. These two properties are highly desirable because they imply that if a large number of samples are used the estimate will be close to the true parameter.

Suppose X is a Gaussian random variable with mean 0 and variance 1. Use the Matlab function *random* or *randn* to generate 1000 samples of X, denoted as X_1, X_2, ..., X_{1000}. See the online help for the **random** function. Plot them using the Matlab function *plot*. We will assume our generated samples are i.i.d.

Write Matlab functions to compute the sample mean and sample variance of equations (6) and (7) **without** using the predefined *mean* and *var* functions. Use these functions to compute the sample mean and sample variance of the samples you just generated.

INLAB REPORT:

1. Submit the plot of samples of X.

2. Submit the sample mean and the sample variance that you calculated. Why are they not equal to the true mean and true variance?

2.3   Linear Transformation of a Random Variable

A linear transformation of a random variable X has the following form

$$Y = aX + b \tag{10}$$

where a and b are real numbers, and $a \neq 0$. A very important property of linear transformations is that they are *distribution-preserving*, meaning that Y will be random variable with a distribution of the same form as X. For example, in (10), if X is Gaussian then Y will also be Gaussian, but not necessarily with the same mean and variance.

Using the linearity property of expectation, find the mean μ_Y and variance σ_Y^2 of Y in terms of a, b, μ_X, and σ_X^2. Show your derivation in detail.
Hint: First find the mean, then substitute the result when finding the variance.

Consider a linear transformation of a Gaussian random variable X with mean 0 and variance 1. Calculate the constants a and b which make the mean and the variance of Y 3 and 9, respectively. Using equation (5), find the probability density function for Y.

Generate 1000 samples of X, and then calculate 1000 samples of Y by applying the linear transformation in equation (10), using the a and b that you just determined. Plot the resulting samples of Y, and use your functions to calculate the sample mean and sample variance of the samples of Y.

INLAB REPORT:

1. Submit your derivation of the mean and variance of Y.

2. Submit the transformation you used, and the probability density function for Y.

3. Submit the plot of samples of Y and the Matlab code used to generate Y. Include the calculated sample mean and sample variance for Y.

3 Estimating the Cumulative Distribution Function

Suppose we want to model some phenomenon as a random variable X with distribution $F_X(x)$. How can we assess whether or not this is an accurate model? One method would be to make many observations and estimate the distribution function based on the observed values. If the distribution estimate is "close" to our proposed model $F_X(x)$, we have evidence that our model is a good characterization of the phenomenon. This section will introduce a common estimate of the cumulative distribution function.

Given a set of i.i.d. random variables $\{X_1, X_2, ..., X_N\}$ with CDF $F_X(x)$, the *empirical cumulative distribution function* $\hat{F}_X(x)$ is defined as the following.

$$\hat{F}_X(x) \;=\; \frac{1}{N} \sum_{i=1}^{N} I_{\{X_i \le x\}} \tag{11}$$

$$I_{\{X_i \le x\}} \;=\; \begin{cases} 1, & \text{if } X_i \le x \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

In words, $\hat{F}_X(x)$ is the fraction of the X_i's which are less than or equal to x.

To get insight into the estimate $\hat{F}_X(x)$, let's compute its mean and variance. To do so, it is easiest to first define N_x as the number of X_i's which are less than or equal to x.

$$N_x = \sum_{i=1}^{N} I_{\{X_i \le x\}} = N\hat{F}_X(x) \tag{13}$$

Notice that $P(X_i \le x) = F_X(x)$, so

$$\begin{aligned} P(I_{\{X_i \le x\}} = 1) &= F_X(x) \\ P(I_{\{X_i \le x\}} = 0) &= 1 - F_X(x) \end{aligned}$$

Now we can compute the mean of $\hat{F}_X(x)$ as follows,

$$\begin{aligned} E\left[\hat{F}_X(x)\right] &= \frac{1}{N} E[N_x] \\ &= \frac{1}{N} \sum_{i=1}^{N} E\left[I_{\{X_i \le x\}}\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{N} N E \left[I_{\{X_i \leq x\}} \right] \\
&= 0 \cdot P \left(I_{\{X_i \leq x\}} = 0 \right) + 1 \cdot P \left(I_{\{X_i \leq x\}} = 1 \right) \\
&= F_X(x) \ .
\end{aligned}
$$

This shows that $\hat{F}_X(x)$ is an unbiased estimate of $F_X(x)$. By a similar approach, we can show that

$$
Var \left[\hat{F}_X(x) \right] = \frac{1}{N} F_X(x)(1 - F_X(x)) \ .
$$

Therefore the empirical CDF $\hat{F}_X(x)$ is both an unbiased and consistent estimate of the true CDF.

3.1 Exercise

Write a function `F=empcdf(X,t)` to compute the empirical CDF $\hat{F}_X(t)$ from the sample vector X at the points specified in the vector t. (**Hint:** The expression `sum(X<=s)` will return the number of elements in the vector X which are less than or equal to s.)

To test your function, generate a sample of Uniform[0,1] random variables using the function $X = rand(1, N)$. Plot two CDF estimates: one using a sample size $N = 20$, and one using $N = 200$. Plot these functions in the range `t=[-1:0.001:2]`, and on each plot superimpose the true distribution for a Uniform[0,1] random variable.

INLAB REPORT:
Hand in your *empcdf* function and the two plots.

---

4 Generating Samples from a Given Distribution

It is oftentimes necessary to generate samples from a particular distribution. For example, we might want to run simulations to test how an algorithm performs on noisy inputs. In this section we will address the problem of generating random numbers from a given distribution $F_X(x)$.

Suppose we have a continuous random variable X with distribution $F_X(x)$, and we form the new random variable $Y = F_X(X)$. In other words Y is a function of X, and the particular function is the CDF of the random variable X.

$$
X \longrightarrow \boxed{F_X(\cdot)} \longrightarrow Y \tag{14}
$$

How is Y distributed? First notice that $F_X(\cdot)$ is a probability, so that Y can only take values in the interval $[0, 1]$.

$$
P(Y \leq y) = \begin{cases} 0, & \text{for } y < 0 \\ 1, & \text{for } y > 1 \end{cases} \tag{15}
$$

Since $F_X(x)$ is a monotonically increasing function of x, the event $\{Y \leq y\}$ is equivalent to $\{X \leq x\}$ if we define $y = F_X(x)$. This implies that for $0 \leq y \leq 1$,

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(F_X(X) \leq F_X(x)) \\
&= P(X \leq x) \qquad \text{(monotonicity)} \\
&= F_X(x) \\
&= y \ .
\end{aligned}
$$

Therefore Y is uniformly distributed on the interval [0,1].

Conversely, if $F_X(\cdot)$ is a one-to-one function, we may use the inverse transformation $F_X^{-1}(U)$ to transform a Uniform[0,1] random variable U to a random variable with distribution $F_X(\cdot)$.

$$
U \longrightarrow \boxed{F_X^{-1}(\cdot)} \longrightarrow X \tag{16}
$$

Note that combining these results allows us to transform any continuous random variable $X \sim F_X(x)$ to any other continuous random variable $Z \sim F_Z(z)$, provided that $F_Z(\cdot)$ is a one-to-one function.

$$
X \longrightarrow \boxed{F_X(\cdot)} \xrightarrow{U} \boxed{F_Z^{-1}(\cdot)} \longrightarrow Z \tag{17}
$$

4.1 Exercise

Your task is to use i.i.d. Uniform[0,1] random variables to generate a set of i.i.d. exponentially distributed random variables with CDF

$$
F_X(x) = (1 - e^{-3x})u(x) \ . \tag{18}
$$

Derive the required transformation.

Generate the Uniform[0,1] random variables using the function $rand(1, N)$. Use your *empcdf* function to plot two CDF estimates for the exponentially distributed random variables: one using a sample size $N = 20$, and one using $N = 200$. Plot these functions in the range `x=[-1:0.001:2]`, and on each plot superimpose the true exponential distribution of equation (18).

INLAB REPORT:

- Hand in the derivation of the required transformation, and your Matlab code.

- Hand in the two plots.

---

5 Estimating the Probability Density Function

The statistical properties of a random variable are completely described by its probability density function (assuming it exists, of course). Therefore, it is oftentimes useful to estimate the PDF, given an observation of a random variable. For example, similar to the empirical CDF, probability density estimates may be used to test a proposed model. They may also be used in non-parametric classification problems, where we need to classify data as belonging to a particular group but without any knowledge of the true underlying class distributions.

Notice that we cannot form a density estimate by simply differentiating the empirical CDF, since this function contains discontinuities at the sample locations X_i. Rather, we need to estimate the probability that a random variable will fall within a particular interval of the real axis. In this section, we will describe a common method known as the *histogram*.

5.1 The Histogram

Our goal is to estimate an arbitrary probability density function, $f_X(x)$, within a finite region of the x-axis. We will do this by partitioning the region into L equally spaced subintervals, or "bins", and forming an approximation for $f_X(x)$ within each bin. Let our region of support start at the value x_0, and end at x_L. Our L subintervals of this region will be $[x_0, x_1], (x_1, x_2], ..., (x_{L-1}, x_L]$. To simplify our notation we will define $bin(k)$ to represent the interval $(x_{k-1}, x_k]$, $k = 1, 2, \ldots, L$, and define the quantity Δ to be the length of each subinterval.

$$
\begin{aligned}
bin(k) &= (x_{k-1}, x_k] \quad k = 1, 2, \ldots, L \\
\Delta &= \frac{x_L - x_0}{L}
\end{aligned}
$$

We will also define $\tilde{f}(k)$ to be the probability that X falls into $bin(k)$.

$$
\begin{aligned}
\tilde{f}(k) &= P(X \in bin(k)) \\
&= \int_{x_{k-1}}^{x_k} f_X(x)dx \quad (19) \\
&\approx f_X(x)\Delta \quad \text{for } x \in bin(k) \quad (20)
\end{aligned}
$$

The approximation in (20) only holds for an appropriately small *bin width* Δ.

Next we introduce the concept of a *histogram* of a collection of i.i.d. random variables $\{X_1, X_2, \ldots, X_N\}$. Let us start by defining a function that will indicate whether or not the random variable X_n falls within $bin(k)$.

$$
I_n(k) = \begin{cases} 1, & \text{if } X_n \in bin(k) \\ 0, & \text{if } X_n \notin bin(k) \end{cases} \quad (21)
$$

The *histogram* of X_n at $bin(k)$, denoted as $H(k)$, is simply the number of random variables that fall within $bin(k)$. This can be written as

$$
H(k) = \sum_{n=1}^{N} I_n(k) . \quad (22)
$$

We can show that the *normalized* histogram, $H(k)/N$, is an unbiased estimate of the probability of X falling in $bin(k)$. Let us compute the expected value of the normalized histogram.

$$
\begin{aligned}
E\left[\frac{H(k)}{N}\right] &= \frac{1}{N}\sum_{n=1}^{N} E[I_n(k)] \\
&= \frac{1}{N}\sum_{n=1}^{N}\{1\cdot P(X_n \in bin(k)) + 0\cdot P(X_n \notin bin(k))\} \\
&= \tilde{f}(k)
\end{aligned}
$$

The last equality results from the definition of $\tilde{f}(k)$, and from the assumption that the X_n's have the same distribution. A similar argument may be used to show that the variance of $H(k)$ is given by

$$
Var\left[\frac{H(k)}{N}\right] = \frac{1}{N}\tilde{f}(k)(1-\tilde{f}(k)) .
$$

Therefore, as N grows large, the bin probabilities $\tilde{f}(k)$ can be approximated by the normalized histogram $H(k)/N$.

$$
\tilde{f}(k) \approx \frac{H(k)}{N} \tag{23}
$$

Using (20), we may then approximate the density function $f_X(x)$ within $bin(k)$ by

$$
f_X(x) \approx \frac{H(k)}{N\Delta} \quad \text{for } x \in bin(k) . \tag{24}
$$

Notice this estimate is a staircase function of x which is constant over each interval $bin(k)$. It can also easily be verified that this density estimate integrates to 1.

5.2 Exercise

Let U be a uniformly distributed random variable on the interval $[0,1]$ with the following cumulative probability distribution, $F_U(u)$:

$$
F_U(u) = \begin{cases} 0, & \text{if } u < 0 \\ u, & \text{if } 0 \le u \le 1 \\ 1, & \text{if } u > 1 \end{cases}
$$

We can calculate the cumulative probability distribution for the new random variable $X = U^{\frac{1}{3}}$.

$$
\begin{aligned}
F_X(x) &= P(X \le x) \\
&= P(U^{\frac{1}{3}} \le x) \\
&= P(U \le x^3) \\
&= F_U(u)|_{u=x^3} \\
&= \begin{cases} 0, & \text{if } x < 0 \\ x^3, & \text{if } 0 \le x \le 1 \\ 1, & \text{if } x > 1 \end{cases}
\end{aligned}
$$

Plot $F_X(x)$ for $x \in [0, 1]$. Also, analytically calculate the probability density $f_X(x)$, and plot it for $x \in [0, 1]$.

Using $L = 20$, $x_0 = 0$ and $x_L = 1$, use Matlab to compute $\tilde{f}(k)$, the probability of X falling into $bin(k)$. (Hint: Use the fact that $\tilde{f}(k) = F_X(x_k) - F_X(x_{k-1})$.) Plot $\tilde{f}(k)$ for $k = 1, \ldots, L$ using the *stem* function.

INLAB REPORT:

1. Submit your plots of $F_X(x)$, $f_X(x)$ and $\tilde{f}(k)$. Use *stem* to plot $\tilde{f}(k)$, and put all three plots on a single figure using *subplot*.

2. Show (mathematically) how $f_X(x)$ and $\tilde{f}(k)$ are related.

Generate 1000 samples of a random variable U that is uniformly distributed between 0 and 1 (using the *rand* command). Then form the random vector X by computing $X = U^{\frac{1}{3}}$.

Use the Matlab function *hist* to plot a **normalized** histogram for your samples of X, using 20 bins uniformly spaced on the interval $[0, 1]$. (Hint: Use the Matlab command H=hist(X,(0.5:19.5)/20) to obtain the histogram, and then normalize H.) Use the *stem* command to plot the normalized histogram $H(k)/N$ and $\tilde{f}(k)$ together on the same figure using *subplot*.

INLAB REPORT:

1. Submit your two stem plots of $H(k)/N$ and $\tilde{f}(k)$. How do these plots compare?

2. Discuss the tradeoffs (advantages and the disadvantages) between selecting a very large or very small bin-width.

References

[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, 3rd ed.*, McGraw-Hill, New York, 1991.