MIT 6.02 **DRAFT** *Lecture Notes*
Fall 2010 (Last update: September 26th, 2010)
Comments, questions or bug reports?
Please contact 6.02-staff@mit.edu

LECTURE 4
Noise

There are three kinds of lies: lies, damn lies, and statistics.
—Probably Benjamin Disraeli

There are liars, there are damn liars, and then there are statisticians.
—Possibly Mark Twain

God does not play dice with the universe.
—Albert Einstein, with probability near 1

Any system that measures the physical world and then selects from a finite set of possible outcomes must contend with noise; and communication systems are no exception. In 6.02, we decide whether a transmitted bit was a '1' or a '0' based on comparing a received voltage sample at the "center" of bit period to a threshold voltage. Our bit decision can be affected by random fluctuations in the voltage sample values (known as amplitude noise) or by misindentification of the bit period center caused by random fluctuations in the rate of received samples (known as phase noise). We will be investigating amplitude noise, partly because it is far easier to analyze than phase noise, and partly because amplitude noise dominates in our IR communication system. In this lecture and part of the next, we will be using a model of noise in which each received voltage sample is offset by a small random noise value with a given distribution (typically the **Gaussian** (or **Normal**) distribution), and we will assume that these random offsets are uncorrelated (the random offset at a given sample is independent of the random offset at any other sample). This model of noise is sometimes referred to as **additive white Gaussian noise** or **AWGN**. In this lecture we will be primarily concerned with using the **AWGN** model to estimate the likelyhood of misindentifying bits, or the **bit-error rate**, and will use two important mathematical tools, the **probability density function** (PDF) and its associated **cumulative distribution function** (CDF). We will also use the **variance** of the PDF as a way to define how "big" the noise is.

■ 4.1 When is noise noise?

Imagine you are in a room with other people, where each person is playing music streamed from a different web site, and each is using a set of speakers (no earbuds!). Then, since you will be nearest your own speakers, you will hear your music loudest, but will also hear the music of others in the background. If there are only a few people in the room, you could probably pick out each of the other songs being played, and ignore them. Some people are surprisingly good at that, others are not. If there are thousands of people in the room, you will probably be able to hear your music, but those thousands of other songs will probably combine together to sound to you like background noise. But there is nothing random going on, you could presumably get the playlists of all the web streams and know exactly what the background was, but it would hardly be worth your time. Describing what you hear as background noise is good enough. Now, if those thousands of other people switched at random times to randomly selected web sites, the background noise truely would be random, though it is unlikely you would hear much of a difference.

In communication links, we have the same three cases. Sometimes there are only a few sources of interference, and if their effects can be determined easily, the effects can be eliminated. Inter-symbol interference is an example of this case. Sometimes there are so many sources of interference that even if it were possible to determine the effect of each source, acquiring the data to eliminate all the interfering sources becomes impractical. A more tractable approach is to approximate the effect of the combination of interferering sources as the result of noise of the appropriate amplitude. Finally, sometimes the sources of interference really are random, with an unknown distribution, though the Gaussian or Normal distribution described below is usually a good appoximation.

As you will see when we begin examining error detection and correction codes, there are many alternatives for dealing with bit errors that inevitably occur in any communication system. We hope that by understanding noise, you will be better able to select the right strategy from among these alternatives.

■ 4.2 Origins of noise

In a communication link, noise is an undesirable perturbation of the signal being sent over the communication channel (e.g. electrical signals on a wire, optical signals on a fiber, or electromagnetic signals through the air). The physical mechanism that is the dominant noise source in a channel varies enormously with the channel, but is rarely associated with a fundamentally random process. For example, electric current in an integrated circuit is generated by electrons moving through wires and across transistors. The electrons must navigate a sea of obstacles (atomic nuclei), and behave much like marbles traveling through a Pachinko machine. They collide randomly with nuclei and have transit times that vary randomly. The result is that electric currents have random noise, but the amplitude of the noise is typically five to six orders of magnitude smaller than the nominal current. Even in the interior of an integrated circuit, where digital information is transported on micron-wide wires, the impact of electron transit time fluctuations is still negligible.

If the communication channel is a wire on an integrated circuit, the primary source of noise is capacitive coupling between signals on neighboring wires. If the channel is a

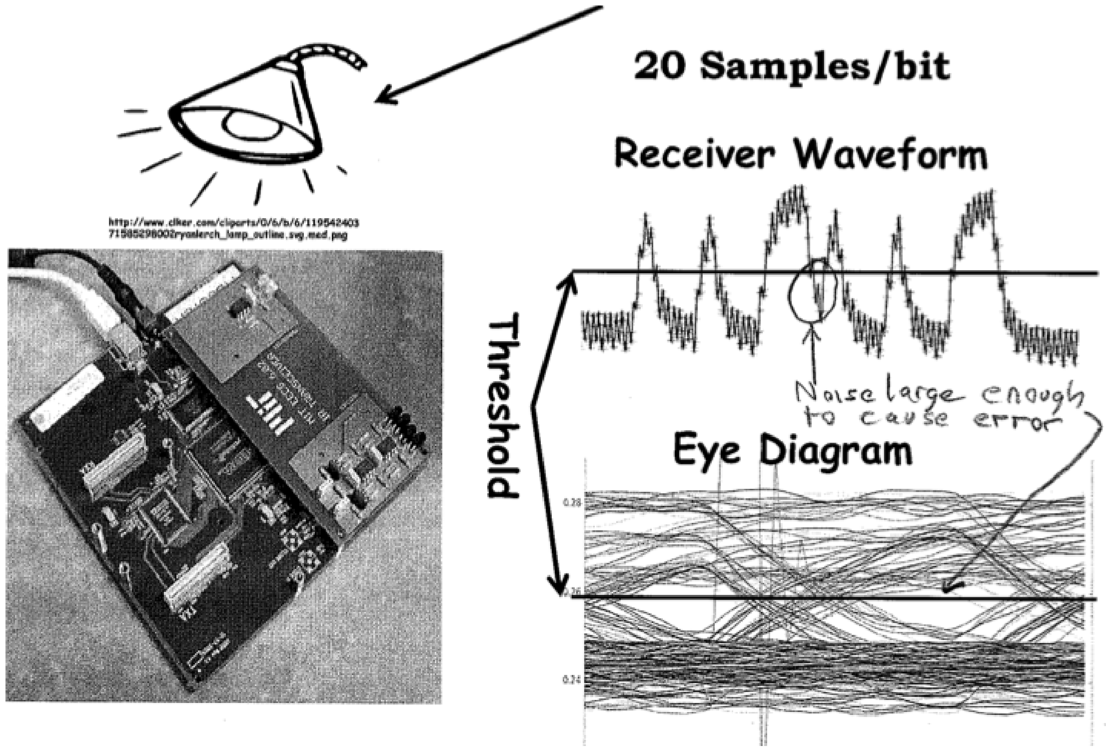

Figure 4-1: Bright ambient lighting is noise for the 6.02 IR communication link.

wire on a printed circuit board, signal coupling is still the primary source of noise, but the coupling between wires is due to unintended electromagnetic radiation. In both these wire examples, one might argue that the noise is not random, as the signals generating the noise are under the designer's control. However, signals on a wire in an integrated circuit or on a printed circuit board will frequently be affected by signals on thousands of nearby wires, so approximating the interference using a random noise model seems eminently practical. In wireless communication networks, like cell phones or Wi-Fi, noise can be generated by concurrent users, or by signals reflecting off walls and arriving late(multi-path), or by background signals generated from appliances like microwave ovens and cordless telephones. Are these noise sources really random? Well, some of the concurrent users might be pretty random.

Optical channels are one of the few cases where fluctuations in electron transit times is a dominant source of noise. Though, the noise is not generated by any mechanism in the optical fiber, but rather by circuits used to convert between optical to electronic signals at the ends of the fiber.

For the IR communication channel in the Athena cluster, the dominant noise source is the fluorescent lighting (see Figure 4-1). Though, as we will see in subsequent weeks, the noise in the IR channel has a structure we can exploit.

Although there are a wide variety of mechanisms that can be the source of noise, the bottom line is that *it is physically impossible to construct a noise-free channel*. But, by understanding noise, we can develop approaches that reduce the probably that noise will lead to bit errors. Though, it will never be possible to entirely eliminate errors. In fact, *there is a* *fundamental trade-off between how fast we send bits and how low we make the probability of error*. That is, you can reduce the probably of making a bit error to zero, but only if you use an infinite interval of time to send a bit. And if you are using a finite interval of time to send a bit, then the probability of a bit error *must* be greater than zero.

■ 4.3 Additive Noise

Given the variety of mechanisms that could be responsible for noise, and how little detailed information we are likely to have about those mechanisms, it might seem prudent to pick a model for noise in a channel that is easy to analyze. So, consider dividing the result of transmitting samples through a channel in to a two step process. First, the input sample sequence, X, is processed by a noise-free channel to produce a noise-free sample sequence, Y_{nf}. Then, noise is added to Y_{nf} to produce the actual received samples, Y. Diagrammatically,

$$X \rightarrow CHANNEL \rightarrow Y_{nf} \rightarrow Add\ NOISE \rightarrow Y. \tag{4.1}$$

If we assume the noise-free channel is LTI and described by a unit sample response, H, we can write a more detailed description,

$$y_{nf}[n] = \sum_{m=0}^{m=n} h[m]x[n-m] \tag{4.2}$$

and

$$y[n] = y_{nf}[n] + noise[n] = \sum_{m=0}^{m=n} h[m]x[n-m] + noise[n], \tag{4.3}$$

where $y_{nf}[n]$ is the output at the n^{th} sample of the noise-free channel, and $noise[n]$ is noise voltage offset generated at the n^{th} sample.

Formally, we will model $noise[n]$ as the n^{th} sample of a random process, and a simple way of understanding what that means is to consider the following thought experiment. Start with a coin with -1 on the head side and $+1$ on the tail side. Flip the coin 1000 times, sum the 1000 values, divide by a 1000, and record the result as $noise[0]$. Then forget that result, and again flip the coin 1000 times, sum the 1000 values, divide by a 1000, and record the result as $noise[1]$. And continue. What you will generate are values for $noise[0], noise[1], \ldots, noise[n], \ldots$ that are independent and identically distributed.

By identically distributed, we mean, for example, that

$$P(noise[j] > 0.5) = P(noise[k] > 0.5) \tag{4.4}$$

for any j and k, where we used $P(expression)$ to denote the probability that *expression* is true. By independent, we mean, for example, that knowing $noise[j] = 0.5$ tells you nothing about the values of $noise[k], k \neq j$.

4.4 Analyzing Bit Errors

Noise disrupts the quality of communication between sender and receiver because the received noisy voltage samples can cause the receiver to incorrectly identify the transmitted bit, thereby generating a **bit error**. If we transmit a long stream of bits and count the fraction of received bits that are in error, we obtain a quantity called the **bit error rate**. This quantity is equivalent to the *probability that any given bit is in error*.

Communication links exhibit a wide range of bit error rates. At one end, high-speed (multiple gigabits per second) fiber-optic links implement various mechanisms that reduce the bit error rates to be as low as 1 part in 10^{12}, or $P(bit\ \ error) = 10^{-12}$.[1] Wireless communication links usually have errors anywhere between one part in 10^4, or $P(bit\ \ error) = 10^{-4}$, for a relatively noisy environments, down to to one part in 10^7, or $P(bit\ \ error) = 10^{-7}$. Very noisy links can still be useful even if they have bit error rates as high as one part in 10^2 or 10^3.

The eye diagram can be to used to gain some intuition about the relationship between bit error rate and the amount of noise in the received samples. Recall that we have been converting samples to bits by selecting a bit detection sample from each bit period, and then comparing the bit detection sample to a threshold. The bit detection sample should correspond to the sample in the eye diagram associated with widest open part of the eye. If the bit detection sample has be selected correctly, then a channel with a wide open eye, will generate fewer bit errors for a given amount of noise than a channel with a more narrowly open eye. For reasons we will make clearer below, we refer to one-half the width of the widest open part of a channel's *noise-free* eye diagram as the channel's **noise margin**. For a given amount of noise, the larger the channel's noise margin, the lower the bit error rate of the channel.

4.4.1 Bit Error Probabilities

If we make the strong assumption the that bit period of the transmitter and receiver are equal and do not drift apart, as is the case in the IR communication channel, then we can greatly simplify the analysis of bit errors in the channel. The relation between the sequence of received voltage samples, Y, and the sequence of received bits, B, can then be written as

$$bit[k] = 1 \;\; if \;\; y[i+sk] > v_{th} \tag{4.5}$$

$$bit[k] = 0 \;\; otherwise \tag{4.6}$$

where s is the number of samples in a bit period, v_{th} is the threshold used to digitize the bit detection sample, $bit[k]$ is the k^{th} received bit, and i is the index of the bit detection sample for the zeroth received bit.

Note that the problem of selecting the bit detection sample and the threshold voltage in the presence of noise is not trivial. The bit detection sample should correspond to the sample associated with the most open part of the eye in the noise-free eye diagram, but there is no way to generate the noise-free eye diagram using only the noisy received samples. If, for the moment, we assume that some strategy has identified the best bit detection sample

[1] This error rate looks exceptionally low, but a link that can send data at 10 gigabits per second with such an error rate will encounter a bit error every 100 seconds of continuous activity, so it does need ways of masking errors that occur.

index, i, and the best threshold voltage, v_{th}, for a given channel, then in the *noise-free* case, the correct value for k^{th} received bit should be '1' if $y_{nf}[i+sk] > v_{th}$ and '0' otherwise. We can also specify the noise margin in terms of the noise-free received samples as

$$noise\ \ margin \equiv \min_k \| y_{nf}[i+sk] - v_{th} \| \tag{4.7}$$

where the above equation just says the noise margin is equal to the smallest distance between the bit detection sample voltage and the threshold voltage. Again note that the noise margin is defined based on channel behavior in the absence of noise.

Assuming no period drift also simplifies the expression for the probability of that an error is made when receiving the k^{th} transmitted bit,

$$P(bit[k]\ \ error) = P(y_{nf}[i+sk] > v_{th}\ \&\ y[i+sk] \le v_{th}) + P(y_{nf}[i+sk] \le v_{th}\ \&\ y[i+sk] > v_{th}). \tag{4.8}$$

or

$$P(bit[k]\ \ error) = P(xbit[k] = '1'\ \&\ y[i+sk] \le v_{th}) + P(xbit[k] = '0'\ \&\ y[i+sk] > v_{th}) \tag{4.9}$$

where $xbit[k]$ is the k^{th} transmitted bit.

Note that we can not yet estimate the probability of a bit error (or equivalently the bit error rate). We will need to invoke the additive noise model to go any further.

■ 4.4.2 Additive Noise and No ISI

If we assume additive noise, as in (4.3), then (4.9) can be simplified to

$$P(bit[k]\ \ error) = P(xbit[k] = '1'\ and\ noise[i+sk] \le -(y_{nf}[i+sk] - v_{th})) \tag{4.10}$$
$$+P(xbit[k] = '0'\ and\ noise[i+sk] > (v_{th} - y_{nf}[i+sk])) \tag{4.11}$$

The quantity in (4.11), $-(y_{nf}[i+sk] - v_{th})$, indicates how negative the noise must be to cause a bit error when receiving a transmitted '1' bit, and the quantity $(v_{th} - y_{nf}[i+sk])$ indicates how positive the noise must be to cause an error when receiving a transmitted '0' bit.

If there is little or no intersymbol interference in the channel, then $y_{nf}[i+sk]$ will be equal to maximum voltage at the receiver when a transmitted '1' is being received, and will be equal to the minimum receiver voltage when a transmitted '0' is being received. To make the best use of this observation to simplify (4.11), it will be helpful to separate the probability that the transmitted bit is a '1' or a '0' (or equivalently in this ISI-free case, that the noise-free received voltage is at its maximum or minimum), from the probability that a received bit is in error *given* a particular value for the transmitted bit. The latter probability statement is referred to as a **conditional probability**. Conditional probability statements are of the form *the probability that a is true* **given** *b is true*, which we denote as $P(a|b)$.

Using our definition of noise margin in (4.7), we can rewrite part of (4.11) using conditional probabilities as

$$p_{01}[k] \equiv P(bit[k]\ \ error|xbit[k] = 0) = P(noise[i+sk] > noise\ margin) \tag{4.12}$$
$$p_{10}[k] \equiv P(bit[k]\ \ error|xbit[k] = 1) = P(noise[i+sk] \le -noise\ margin) \tag{4.13}$$

where we have introduced the notation $p_{01}[k]$ to mean the probability that the k^{th} bit is received as a '1' given the k^{th} transmitted bit was a '0', and $p_{10}[k]$ to mean the probability that the k^{th} bit is received as a '0' given the k^{th} transmitted bit was a '1'.

Finally, we can use the assumption that the noise samples are identically distributed to eliminate the dependence on i and k in (4.13), yeilding

$$P(bit\ \ error) = p_{01} \cdot p_0 + p_{10} \cdot p_1 \tag{4.14}$$

where p_0 is the probability that the transmitted bit is a '0', p_1 is the probability that the transmitted bit is a '1', and we have used the standard identity $P(a\ and\ b) = P(a|b)P(b)$.

At this point we have an estimate of the bit error rate (recall BER is equivalent to the probability of a bit error), provided we can evaluate the noise probabilities. We turn to that problem in the next section.

4.5 Noise Statistics

We are modeling noise in the received samples as the result of sampling a **random process**, where the description of the values generated by sampling this random process are in terms of probabilities. For example, if we transmit a sequence of zero volt samples and observe the received samples, we can process these received samples to determine some statistics of the noise process. If the observed samples are $noise[0], noise[1], \ldots, noise[N-1]$, then the **sample mean** is given by

$$\mu = \frac{\sum_{n=0}^{N-1} noise[n]}{N}. \tag{4.15}$$

For our model of additive noise, the noise samples should have zero mean ($\mu = 0$), so the sample mean does not provide much information about the noise. A quantity that is more indicative of the amount of noise in the received samples is given by the sample variance, defined as

$$\sigma^2 = \frac{\sum_{n=0}^{N-1} (noise[n] - \mu)^2}{N}. \tag{4.16}$$

The sample standard deviation, σ, is in some sense, the amplitude of the noise. To ensure that noise does not corrupt the digitization of a bit detection sample, the distance between the noise-free value of the bit detection sample and the digitizing threshold should be much larger than the amplitude of the noise. As explained above, one-half of the width of the eye is defined as the noise margin, because any noise that is larger than the noise margin will always lead to an incorrect digitization; if the standard deviation of the noise process is not much smaller than the noise margin, a huge number of bits will be received incorrectly.

4.5.1 Probability density functions

A convenient way to model noise is using a probability density function, or PDF. To understand what a PDF is, let us imagine that we generate 100 or 1000 independent noise samples and plot each one on a histogram. We might see pictures that look like the ones shown in Figure 4-2 (the top two pictures), where the horizontal axis is the value of the

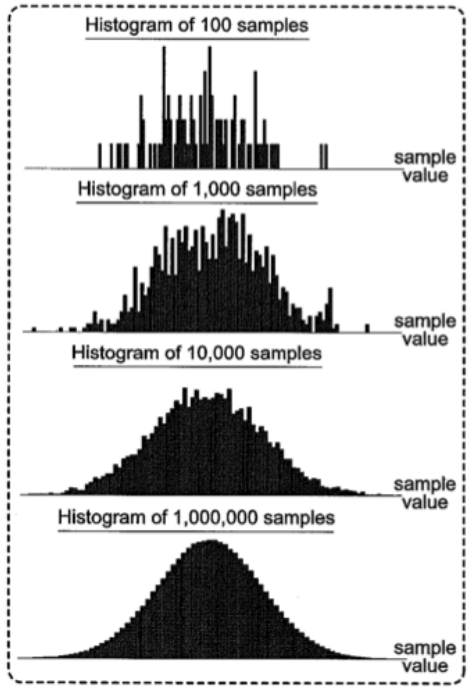

Figure 4-2: Histograms become smoother and more continuous when they are made from an increasing number of samples. In the limit when the number of samples approaches infinity and the bin width approaches zero, the resulting curve is a probability density function.

noise sample (binned) and the vertical axis is the frequency with which values showed up in each noise bin. As we increase the number of noise samples, we might see pictures as in the middle and bottom of Figure 4-2. The histogram is becoming increasingly smooth and continuous. In the limit when the number of noise samples approaches infinity, and the width of the bins approaches zero, the resulting histogram is called a probability density function (PDF).

Formally, let X be the random variable of interest, and suppose x can take on any value in the range $(-\infty, \infty)$. The PDF of the random variable is denoted $f_X(x)$. What $f_X(x)$ means is that the probability that the random variable X takes on a value between x and $x + dx$, where dx is a vanishingly small increment about x, is given by the product $f_X(x)\, dx$. Example PDF's are shown in Figure 4-3 and Figure 4-4.

The PDF is *not* a probability; but the*area* under the $f_X(x)$ curve, for any interval of x values, is a probability (see Figure 4-4). Note that therefore $f_X(x)$ may exceed one, but $f_X(x)\, dx$, the area under a tiny sliver, is a probability, and can never exceed one.

Any legitimate PDF must satisfy a *normalization condition* because the area under $f_X(x)$ for $x \in (-\infty, \infty)$ is the probability of all possible outcomes and must be exactly one. That is, $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.

One can use the definition of the PDF to calculate the probability that a random variable x lies in the range $[x_1, x_2]$:

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f_X(x)\, dx. \tag{4.17}$$

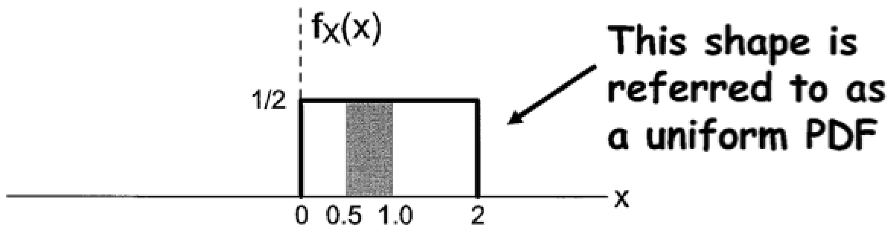

Figure 4-3: PDF of a uniform distribution.

Mean and variance. The mean (or average value) of a random variable X, denoted μ_X, can be computed from its PDF as follows:

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x)\, dx. \tag{4.18}$$

This definition of the mean directly follows from the definition of the mean of a discrete random process, defined in 4.15, and taking the limit when $N \to \infty$ in that equation. Strictly speaking, some assumptions must be made regarding the existence of the mean, but under these typical satisfied assumptions, the following fact holds. If $noise[n]$ is generated by a discrete random process with underlying probability density $f_X(x)$, then the sample mean approaches the mean as the number of samples approaches ∞,

$$\lim^{N\to\infty} \sum_{n=0}^{N} noise[n] = \mu_X. \tag{4.19}$$

Similarly, one defines the variance,

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx. \tag{4.20}$$

and the standard deviation, (σ_X), is the square root of the variance.

To summarize: *If the noise (or any random variable) is described by a PDF $f_X(x)$, then the sample mean and the sample variance converge to the mean and variance of the PDF as the number of samples goes to ∞.*

■ 4.5.2 Examples

Some simple examples may help illustrate the idea of a PDF better, especially for those who haven't see this notion before.

Uniform distribution. Suppose that a random variable X can take on any value between 0 and 2 with equal probability, and always lies in that range. What is the corresponding PDF?

Because the probability of X being in the range $(x, x + dx)$ is independent of x as long as x is in $[0, 2]$, it must be the case that the PDF $f_X(x)$ is some constant, k, for $x \in [0, 2]$. Moreover, it must be 0 for any x outside this range. We need to determine k. To do so,

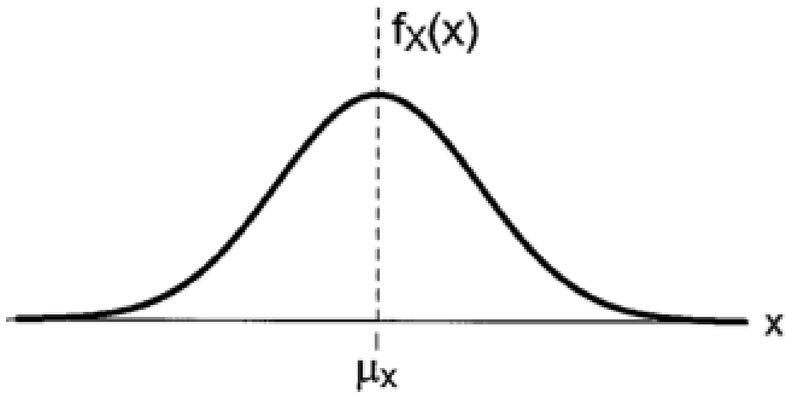

Figure 4-4: PDF of a Gaussian distribution, aka a "bell curve".

observe that the PDF must be normalized, so

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^2 k\,dx = 1, \tag{4.21}$$

which implies that $k = 0.5$. Hence, $f_X(x) = 0.5$ when $0 \leq x \leq 2$ and 0 otherwise. Figure 4-3 shows this *uniform PDF*.

One can easily calculate the probability that an x chosen from this distribution lies in the range $(0.3, 0.7)$. It is equal to $\int_{0.3}^{0.7}(0.5)\,dx = 0.2$.

A uniform PDF also provides a simple example that shows how the PDF, $f_X(x)$, could easily exceed 1. A uniform distribution whose values are always between 0 and δ, for some $\delta < 1$, has $f_X(x) = 1/\delta$, which is always larger than 1. To reiterate a point made before: the PDF $f_X(x)$ is *not a probability*, it is a probability *density*, and as such, could take on any non-negative value. The only constraint on it is that the total area under its curve (the integral over the possible values it can take) is 1.

Gaussian distribution. The Gaussian, or "normal", distribution is of particular interest to us because it turns out to be an accurate model for noise in many communication (and other) systems. The reason for the accuracy of this model will become clear later in this lecture—it follows from the *central limit theorem*—but let us first understand it mathematically.

The PDF of a Gaussian distribution is

$$f_X(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}. \tag{4.22}$$

This formula captures a *bell shape* (Figure 4-4), and because of that, is colloquially referred to as a "bell curve". It is symmetric about the mean, μ and tapers off to 0 quite rapidly because of the e^{-x^2} dependence. A noteworthy property of the Gaussian distribution is that it is **completely characterized** by the mean and the variance, σ^2. If you tell me the mean and variance and tell me that the random process is Gaussian, then you have told me *everything* about the distribution.

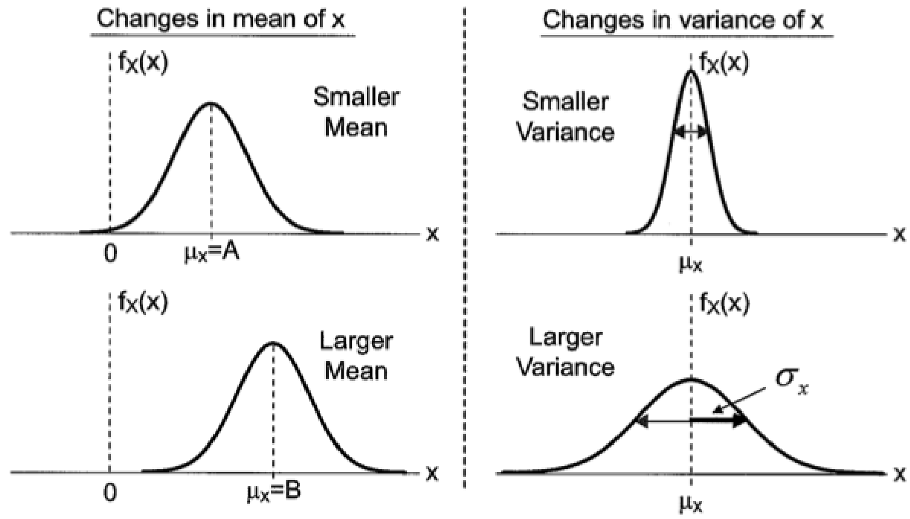

Figure 4-5: Changing the mean of a Gaussian distribution merely shifts the center of mass of the distribution because it just shifts the location of the peak. Changing the variance widens the curve.

Changing the mean simply shifts the distribution to the left or right on the horizontal axis, as shown in the pictures on the left of Figure 4-5. Increasing the variance is more interesting from a physical standpoint; it widens (or fattens) the distribution and makes it more likely for values further from the mean to be selected, compared to a Gaussian with a smaller variance.

■ 4.5.3 Calculating the bit error rate

Given the PDF of the noise random process, we can calculate the bit error rate by observing that each received sample is in fact a noise-free value plus a value drawn from a probability distribution with zero mean, $\mu = 0$, and a given variance σ^2 (or equivalently, a given standard deviation, σ). The standard deviation is often referred to informally as the "amplitude" of the noise. So, if the noise-free received voltage is zero volts, the noisy received voltage will be $0 + noise$, where *noise* is drawn from a noise distribution determined by the noise PDF, denoted $f_{NOISE}(noise)$. For brevity, let us write the noisy received voltage as $0 + x$, and replace $f_{NOISE}(noise)$ with $f_X(x)$. In this briefer notation, if the noise-free received voltage is one volt, then noisy received voltage will be $1 + x$, where x is again drawn from the distribution determined by $f_X(x)$.

Suppose the digitizing threshold is 0.5, and suppose that in the noise-free case, the received bit detection sample is zero volts when receiving a transmitted '0' bit, and one volt when receiving a transmitted '1' bit. If the probability of transmitting a '0' bit is p_0 and the probability of transmitting a '1' bit is p_1, the probability of a bit error in the noisy case is given by

$$\begin{aligned} P(\text{bit error}) &= p_0 \cdot P(x > 0.5) + p_1 \cdot P(1 + x < 0.5) \\ &= p_0 \cdot P(x > 0.5) + p_1 \cdot P(x < -0.5) \\ &= p_0 \cdot \int_{0.5}^{\infty} f_X(x)\,dx + p_1 \cdot \int_{-\infty}^{-0.5} f_X(x)\,dx. \end{aligned} \tag{4.23}$$

Since we are considering ISI-free case at the moment, we are only concerned with two cases: a transmitted '0' or a transmitted '1'. Since there are only two cases, $p_0 = 1 - p_1$, but since the *a priori* probabilities of transmitting a '0' or a '1' are usually equal, typically $p_0 = p_1 = 1/2$. If in addition to having values for p_0 and p_1, we know the noise PDF, $f_X(x)$, then we we can evaluate the integrals in (4.23) and determine the bit error rate. In fact, if the noise process is symmetric about the mean (as is the case for a Gaussian or a uniform distribution), then the two integrals in (4.23) are identical, and only one integral need be evaluated. So, if '0''s and '1''s are transmitted equally often, and the noise PDF is symmetric about the mean, then the formula for the bit error rate simplifies to

$$P(\mathit{bit\ error}) = \int_{0.5}^{\infty} f_X(x)\,dx = 1 - \int_{-\infty}^{0.5} f_X(x)\,dx. \tag{4.24}$$

The integral of any PDF from $-\infty$ to x, $\int_{-\infty}^{x} f_X(x')\,dx'$, has a special name, it is called the **cumulative distribution function (CDF)**, because it represents the cumulative probability that the random variable X takes on any value $\leq x$. From the definition of the CDF, it follows that the value of the CDF approaches one as $x \to \infty$.

To summarize: the probabilty of a bit error, also called the bit error rate, requires the evaluation of a single CDF when the noise random process is symmetric about the mean. The Gaussian noise distribution has this property, in addition to being completely characterized in this case by the variance alone (as the mean is assumed to be zero as part of the additive noise assumption). In the next lecture, we will discuss some salient features of Gaussian noise, why it is a good model for noise over a communication channel, and how to recover signals when *both* ISI and noise occur together.