Carleton University

SYSC 5504 **Department of Systems and Computer Engineering**

ELG 6154 **Principles of Digital Communications** **Fall 2017/18**

Assignment #1

Due on Tuesday, November 7th, 2017

1. Let X be a random variable with a mean of zero and a variance of σ_X^2. For $k \in \{1, 2, \ldots, N\}$, let

$$Y_k = \alpha_k X + W_k \ ,$$

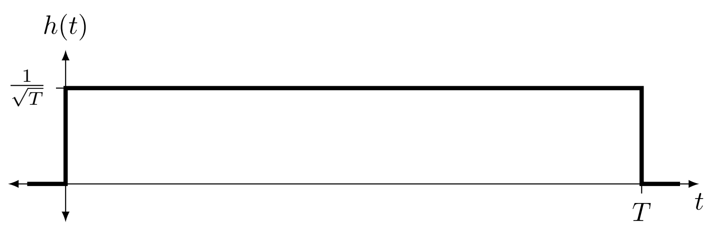
be noisy observations of X, where $\{\alpha_k\}$ are known constants, and $\{W_k\}$ are independent Gaussian random variables, each with zero mean and a variance of σ_W^2. Suppose that you want to estimate X, given the observations, $\{Y_k\}$. A *linear predictor* for X has the form

$$Z = b + \sum_{k=1}^{N} p_k Y_k$$

where b and $\{p_k\}$ are the *predictor coefficients*. Find the predictor coefficients that minimize the mean squared prediction error between the predictor, Z, and X. That is, find b and $\{p_k\}$ that minimizes

$$J = \mathbf{E}\left[(Z - X)^2\right] \ .$$

2. Suppose $x(t)$ is a zero-mean stationary Gaussian random process with power spectral density $\Phi_x(f)$ is the input to a linear filter whose impulse response show below. A sample, Y, is taken of the output of the filter at time T.



   (a) Calculate the mean and variance of Y in terms of $\Phi_x(f)$ and T.
   (b) Upper bound the variance under the condition $\Phi_x(f) \leq S$ for all f.

3. Sketch the spectrum

$$\Phi_x(f) = 2T \left(\frac{\sin(\pi[f - f_c]T)}{\pi[f - f_c]T} \right)^2 \sin^2(\pi[f - f_c]T)$$

with $f_c = 1$ GHz and $T = 10^{-7}$ seconds, then find the

   (a) null-to-null bandwidth,
   (b) the 3 dB bandwidth, and
   (c) the 99% power bandwidth (use numerical integration).

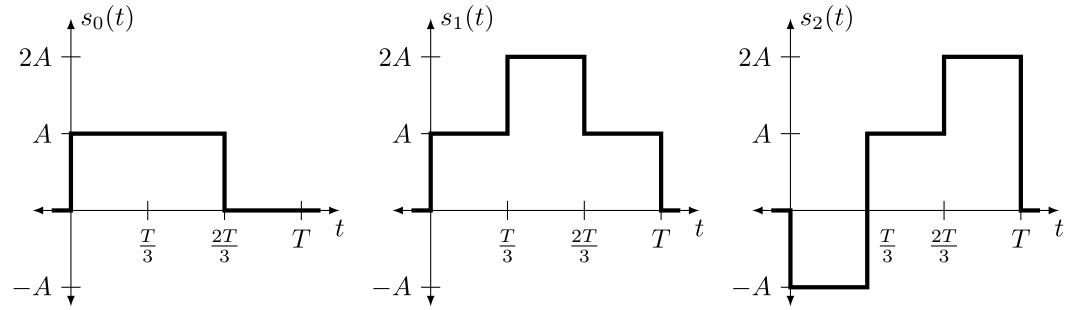
4. For the signals $s_0(t)$, $s_1(t)$, and $s_2(t)$ shown below:

   (a) Use the Gram-Schmidt orthogonalization procedure to find a set of basis signals.

   (b) Express $s_0(t)$, $s_1(t)$, and $s_2(t)$ as linear combinations of the basis signals.

   (c) Show $s_0(t)$, $s_1(t)$, and $s_2(t)$ on a signal space diagram.

   (d) On the signal space diagram indicate the optimal decision regions and carefully draw the boundaries between the regions. Assume that all three signals are equally likely to be transmitted.



5. One reason for running a computer simulation of a digital communication system is to estimate the system relability in terms for the probability of error. Although computer simulation does not replace the need for field trials with actual hardware, it is nonetheless a very useful tool. For example, the trouble and expense of implementing a system in hardware can be avoided if it can be shown through simulation that the system will not achieve acceptable performance. In addition, computer simulation can be used to compare the performance of different systems under operating conditions that may be difficult to create in the field.

   For a given communication system operating in a specific environment, a computer simulation can provide an estimate of the probability of a bit error. Suppose you simulate the transmission of N_a message bits, and discover that N_ε of those bits are received incorrectly. The probability of a bit error can be estimated as

$$\widehat{p} = \frac{N_\varepsilon}{N_a} \; .$$

   It is important to have some knowledge of the accuracy of the estimate. That is, we need some measure of how close is $\widehat{p}$ likely to be to the actual probability of a bit error for the system, p.

   One useful measure of the accuracy of the estimate is the confidence interval $[d_1, d_2]$, such that

$$\Pr\{d_1 \leq p \leq d_2\} = \beta \; ,$$

   where d_1 and d_2 are functions of $\widehat{p}$, and β is the confidence level. Typical values of β are 90%, 95%, or 99%. For example, suppose a value of $\widehat{p} = 5 \times 10^{-5}$ was found by simulation and used to construct a confidence interval of

$$\Pr\left\{4 \times 10^{-5} \leq p \leq 6 \times 10^{-5}\right\} = 0.95 \; .$$

   This means that there is a 95% chance that the true value of p falls within $[4 \times 10^{-5}, 6 \times 10^{-5}]$. In this assignment a method for constructing the confidence interval will be found.

   (a) Consider a digital communication system in which transmission bit errors occur independently with a probability of p. Suppose a block of N_a bits are transmitted and received. Let ε_n be a random variable indicating the error status of each bit, so

$$\varepsilon_n = \begin{cases} 0, & \text{if the } n^{th} \text{ bit is received correctly} \\ 1, & \text{if the } n^{th} \text{ bit is received incorrectly} \end{cases} ,$$

   for $n \in \{0, 1, 2, \ldots, N_a - 1\}$. Each ε_n is a Bernoulli random variable with $\Pr\{\varepsilon_n = 0\} = 1 - p$ and $\Pr\{\varepsilon_n = 1\} = p$, and ε_n is independent of ε_m for all $m \neq n$. <u>Find the mean and variance of ε_n</u>.

(b) Let N_ε be a random variable representing the total number of bit errors in the block, so

$$N_\varepsilon = \sum_{n=0}^{N_a-1} \varepsilon_n \ .$$

Find the mean and variance of N_ε. What type of probability distribution does N_ε have?

(c) Let

$$\widehat{p} = \frac{N_\varepsilon}{N_a}$$

be an estimate of p. Find the mean and variance of $\widehat{p}$.

(d) If $N_a p(1-p) \geq 10$ then $\widehat{p}$ can be approximated as a Gaussian random variable with a probability density function of

$$f_{\widehat{p}}(\widehat{p}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\widehat{p}-p)^2}{2\sigma^2}\right\} \ ,$$

where $\sigma^2 = \frac{p(1-p)}{N_a}$. Using this Gaussian approximation for $\widehat{p}$, find an expression for the probability that $\widehat{p}$ will be in the interval $[(1-\delta)p, (1+\delta)p]$ for some constant, δ. That is, find an expression for

$$\beta = \Pr\left\{(1-\delta)p \leq \widehat{p} \leq (1+\delta)p\right\} \tag{1}$$

in terms of the error function,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2}\, dz \ .$$

To verify your answer, with $p = 0.1$, $N_a = 1000$, and $\delta = 0.2$, then

$$\beta = \Pr\left\{0.08 \leq \widehat{p} \leq 0.12\right\} \cong 0.965 \ .$$

In this case there is roughly a 97% chance that your simulation will produce a value of $\widehat{p}$ within the range [0.08,0.12].

(e) Using the expression for β found above, derive expressions for δ and N_a in terms of the inverse error function, $\mathrm{erf}^{-1}(\cdot)$.

(f) Suppose that you know p, and you want to be reasonably certain that $\widehat{p}$ will fall within the range $[0.8p, 1.2p]$. It is worthwhile to determine in advance how many bits you will need to process in your simulation. If $p = 10^{-4}$, what is the minimum value of N_a that will ensure that

$$\Pr\left\{0.8 \times 10^{-4} \leq \widehat{p} \leq 1.2 \times 10^{-4}\right\} \geq 0.95 \ ?$$

You can use the `matlab erfinv` function to evaluate the inverse of the error function. [Answer: 960,269 bits]

(g) In practice, p is not generally known (and hence the reason for running the simulation). As a result, it is not possible to determine in advance how many bits must be processed. Instead, it is necessary to first perform a fixed number of tests, and then determine the accuracy of the result. If it is not accurate enough, further testing must be performed. As mentioned above, accuracy can be described in terms of the confidence interval. By modifying Eq. (1), and using the expression for δ found in part (e), show that the confidence interval can be written as:

$$\Pr\left\{d_1 \leq p \leq d_2\right\} = \beta \ ,$$

where

$$d_1 = \frac{N_a}{N_a+\gamma^2}\left[\widehat{p} + \frac{\gamma^2}{2N_a} - \gamma\sqrt{\frac{\widehat{p}(1-\widehat{p})}{N_a} + \frac{\gamma^2}{4N_a^2}}\right]$$

and

$$d_2 = \frac{N_a}{N_a+\gamma^2}\left[\widehat{p} + \frac{\gamma^2}{2N_a} + \gamma\sqrt{\frac{\widehat{p}(1-\widehat{p})}{N_a} + \frac{\gamma^2}{4N_a^2}}\right] \ ,$$

and $\gamma = \sqrt{2}\mathrm{erf}^{-1}(\beta)$. [HINT: Make use of the fact that

$$\begin{aligned}\Pr\left\{-c \leq (X-a) \leq c\right\} &= \Pr\left\{|X-a| \leq c\right\} \\ &= \Pr\left\{(X-a)^2 \leq c^2\right\} \ .\end{aligned}$$

]

(h) Suppose you had run a simulation of $N_a = 1,000,000$ bits, and obtained an estimate of the bit error probability of $\widehat{p} = 10^{-4}$. Find numerical values for d_1 and d_2 for the 95% confidence interval ($\beta = 0.95$).