

are discussed in this chapter. An especially important case is when the transformation can be represented on a graph as a straight line. In this case the form of the density does not change; only a shift and scaling occurs.

Distributions may be conditioned on an *event* and the probability of an event can be conditioned on the observance of a random variable. These concepts are developed and applied in the application areas of signal detection and object classification. In these areas, we demonstrate how to form optimal decision rules that minimize the probability of error, and compute the probability of various types of errors that occur when the decision rules are applied.

## References

- [1] Ross L. Finney, Maurice D. Weir, and Frank R. Giordano. *Thomas' Calculus*. Addison-Wesley, Reading, Massachusetts, tenth edition, 2001.
- [2] Alberto Leon-Garcia. *Probability and Random Processes for Electrical Engineering*. Addison-Wesley, Reading, Massachusetts, second edition, 1994.
- [3] Wilbur B. Davenport, Jr. *Probability and Random Processes*. McGraw-Hill, New York, 1970.
- [4] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, New York, second edition, 1990.
- [5] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. John Wiley & Sons, New York, second edition, 2001.

## Problems

### Discrete random variables

- 3.1 In a set of independent trials, a certain discrete random variable  $K$  takes on the values

1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1

- (a) Plot a histogram for the random variable.
- (b) Normalize the histogram to provide an estimate of the PMF of  $K$ .

- 3.2 Repeat Problem 3.1 for the discrete random variable  $J$  that takes on the values

1, 3, 0, 2, 1, 2, 0, 3, 1, 1, 3, 1, 0, 1, 2, 3, 0, 2, 1, 3, 3, 2

- 3.3 A random variable that takes on values of a sequence of 4QAM symbols,  $I$ , is transmitted over a communication channel: 3, 0, 1, 3, 2, 0, 2, 1, 3, 0, 0, 2, 3, 1, 3, 0, 2, 1, 0, 3. The channel introduces additive noise,  $J$ , equivalent to adding the following values to the above symbol sequence: 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1. The addition is performed in modulo form.

- (a) Form the resulting sequence:  $I + J$  modulo 4. Let  $K$  be the random variable that takes on these values.
- (b) Plot histograms for random variables  $I$  and  $K$ .
- (c) Normalize the histograms to provide an estimate of the respective discrete probability density functions.

## PROBLEMS

- 3.4 Consider transmission of ASCII characters, 8 bits long. Define a random variable as the number of 0's in a given ASCII character.

- (a) Determine the range of values that this random variable takes.
- (b) Determine and plot the probability mass function.

### Common discrete probability distributions

- 3.5 Sketch the PMF  $f_K[k]$  for the following random variables and label the values.

- (a) Uniform.  $m = 0, n = 9 \rightarrow$  unif pdf
- (b) Bernoulli.  $p = 1/10 \rightarrow$  binopdf  $\rightarrow n=1$
- (c) Binomial.  $n = 10, p = 1/10$  binopdf
- (d) Geometric.  $p = 1/10$  (plot  $0 \leq k \leq 9$  only)

$k = [0:9]$   
OR  
[1:4]  
sum(y(1:n))

For each of these distributions, compute the probability of the event  $1 \leq k \leq 4$ .

- 3.6 In the transmission of a packet of 128 bytes, byte errors occur independently with probability  $1/9$ . What is the probability that less than five byte errors occur in the packet of 128 bytes?

- 3.7 In transmission of a very long (you may read "infinite") set of characters, byte errors occur independently with probability  $1/9$ . What is the probability that the first error occurs within the first five bytes transmitted?

- 3.8 The Poisson PMF has the form

$$f_K[k] = \frac{\alpha^k}{k!} e^{-\alpha} \quad k = 0, 1, 2, \dots$$

where  $\alpha$  is a parameter.

- (a) Show that this PMF sums to 1.
- (b) What are the restrictions on the parameter  $\alpha$ , that is, what are the allowable values?

- 3.9 (a) Sketch the PMF of the following random variables:

- (i) geometric,  $p = 1/8$
- (ii) uniform,  $m = 0, n = 7 \rightarrow n < 11$
- (iii) Poisson,  $\alpha = 2$

- (b) For each of these, compute the probability that the random variable is greater than 5.

- (c) For each of these, if you desire that  $\Pr[K > 5] \leq 0.5$ , what should be the respective values of the parameters?

$m=1, n=9, k=$

### Continuous random variables

- 3.10 The PDF for a random variable  $X$  is given by

$$f_X(x) = \begin{cases} 2x/9 & 0 \leq x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

- (a) Sketch  $f_X(x)$  versus  $x$ .
- (b) What is the probability of the following events?
  - (i)  $X \leq 1$

$\frac{1}{9} = \frac{1}{9}$   
 $\frac{1}{9} =$

1) Functions  $\rightarrow$  on p14 in formulae.  
2) Format  $\rightarrow$  code  
 $\rightarrow$  solution  
 $\rightarrow$  plot  
3) Be careful of indexing 83

$m \rightarrow n$   
 $1 \rightarrow n-m+1$

Let  $n=1$ , it's  $\frac{1}{n}$   
Something gives us  
MATLAB unif function  
unif  $\rightarrow$  starts at  $m=1$

$P(X > 5) = 1 - P(X \leq 5) \leq 0.5$   
 $0.5 \leq P(X \leq 5)$

Assume  $m=0$  Use plot  
either bar graphs or  
plots  
 $P(X > 5)$

(ii)  $X > 2$

(iii)  $1 < X \leq 2$

(c) Find and sketch the CDF  $F_X(x)$ .

(d) Use the CDF to check your answers to part (b).

**3.11** The PDF of a random variable  $X$  is given by:

$$f_X(x) = \begin{cases} 0.4 + Cx, & 0 \leq x \leq 5 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find  $C$  that makes it a valid PDF.

(b) Determine:  $\Pr[X > 3]$ ,  $\Pr[1 < X \leq 4]$ .

(c) Find the CDF  $F_X(x)$ .

**3.12** A PDF for a continuous random variable  $X$  is defined by

$$f_X(x) = \begin{cases} C & 0 < x \leq 2 \\ 2C & 4 < x \leq 6 \\ C & 7 < x \leq 9 \\ 0 & \text{otherwise} \end{cases}$$

where  $C$  is a constant.

(a) Find the numerical value of the constant  $C$ .

(b) Compute  $\Pr[1 < X \leq 8]$ .

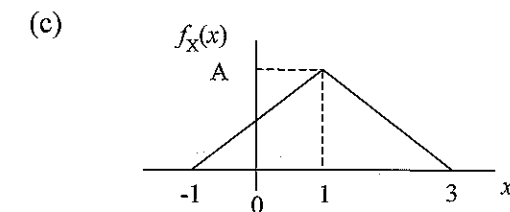
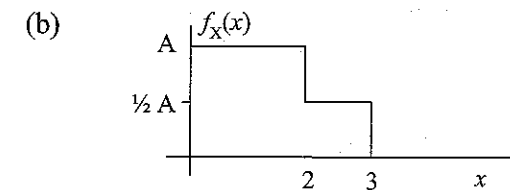
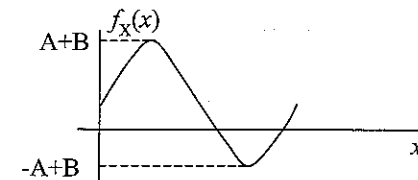
(c) Find the value of  $M$  for which

$$\int_{-\infty}^M f_X(x) dx = \int_M^{\infty} f_X(x) dx = \frac{1}{2}$$

$M$  is known as the *median* of the random variable.

**3.13** For what values or ranges of values of the parameters would the following PDFs be valid? If there are no possible values that would make the PDF valid, state that.

(a) 
$$f_X(x) = \begin{cases} B + A \sin x & 0 \leq x \leq 2\pi \\ 0 & \text{otherwise.} \end{cases}$$



**3.14** The CDF for a random variable  $X$  is given by

$$F_X(x) = \begin{cases} A \arctan(x) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

(a) Find the constant  $A$ .

(b) Find and sketch the PDF.

**3.15** A random variable  $X$  has the CDF specified below.

$$F_X(x) = \begin{cases} 0 & x \leq -\frac{\pi}{2\alpha} \\ \frac{1}{2}(1 + \sin \alpha x) & -\frac{\pi}{2\alpha} < x \leq \frac{\pi}{2\alpha} \\ 1 & x \geq \frac{\pi}{2\alpha} \end{cases}$$

(a) Sketch  $F_X(x)$ .

(b) For what values of the real parameter  $\alpha$  is this expression valid?

(c) What is the PDF  $f_X(x)$ ? Write a formula and sketch it.

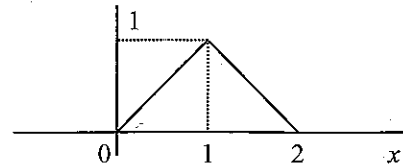
(d) Determine the probability of the following events:

(i)  $X \leq 0$

(ii)  $X \leq -1$

(iii)  $|X| \leq .001$  (approximate)

3.16 The PDF of a random variable  $X$  is shown below:



- (a) Obtain the CDF.  
 (b) Find  $\Pr[0.5 \leq X \leq 1.5]$ ,  $\Pr[1 \leq X \leq 5]$ ,  $\Pr[X < 0.5]$ , and  $\Pr[0.75 \leq X \leq 0.7501]$ .

3.17 Given the CDF,

$$F_X(x) = \begin{cases} 0 & x < 0 \\ Cx^2 & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

- (a) Find  $C$ .  
 (b) Sketch this function.  
 (c) Find the PDF and sketch it.  
 (d) Find  $\Pr[0.5 \leq X \leq 1.5]$ .

3.18 The PDF of a random variable  $X$  is of the form

$$f_X(x) = \begin{cases} Cxe^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Where  $C$  is a normalization constant.

- (a) Find the constant  $C$  and sketch  $f_X(x)$ .  
 (b) Compute the probability of the following events:  
 (i)  $X \leq 1$   
 (ii)  $X > 2$   
 (c) Find the probability of the event  $x_0 < X \leq x_0 + 0.001$  for the following values of  $x_0$ :  
 (i)  $x_0 = 1/2$   
 (ii)  $x_0 = 2$

What value of  $x_0$  is "most likely," i.e., what value of  $x_0$  would give the largest probability for the event  $x_0 < X \leq x_0 + 0.001$ ?

- (d) Find and sketch the CDF.

### Common probability density functions

3.19 Sketch the PDFs of the following random variables:

- (a) Uniform:  
 (i)  $a = 2, b = 6$ ;  
 (ii)  $a = -4, b = -1$ ;  
 (iii)  $a = -7, b = 1$ ;  
 (iv)  $a = 0, b = a + \delta$ , where  $\delta \rightarrow 0$ .

(b) Gaussian:

- (i)  $\mu = 0, \sigma = 1$ ;  
 (ii)  $\mu = 1.5, \sigma = 1$ ;  
 (iii)  $\mu = 0, \sigma = 2$ ;  
 (iv)  $\mu = 1.5, \sigma = 2$ ;  
 (v)  $\mu = -2, \sigma = 1$ ;  
 (vi)  $\mu = -2, \sigma = 2$ ;  
 (vii)  $\mu = 0, \sigma \rightarrow 0$ .

(c) Exponential:

- (i)  $\lambda = 1$ ;  
 (ii)  $\lambda = 2$ ;  
 (iii)  $\lambda \rightarrow 0$ ;  
 (iv)  $\lambda \rightarrow \infty$ ;  
 (v) Can  $\lambda$  be a negative value? Explain.

3.20 The median  $M$  of a random variable  $X$  is defined by the condition

$$\int_{-\infty}^M f_X(u) du = \int_M^{\infty} f_X(u) du$$

Find the median for:

- (a) The uniform PDF.  
 (b) The exponential PDF.  
 (c) The Gaussian PDF.

3.21 A certain random variable  $X$  is described by a Gaussian PDF with parameters  $\mu = -1$  and  $\sigma^2 = 2$ . Use the error function (erf) or the Q function to compute the probability of the following events:

- (a)  $X \leq 0$   
 (b)  $X > +1$   
 (c)  $-2 < X \leq +2$

3.22 Given a Gaussian random variable with mean  $\mu = 1$  and variance  $\sigma^2 = 3$ , find the following probabilities:

- (a)  $\Pr[X > 1]$  and  $\Pr[X > \sqrt{3}]$ .  
 (b)  $\Pr[|X - 1| > \sqrt{12}]$  and  $\Pr[|X - 1| < 6]$ .  
 (c)  $\Pr[X > 1 + \sqrt{3}]$  and  $\Pr[X > 1 + 3\sqrt{3}]$ .

### CDF and PDF for discrete and mixed random variables

3.23 Consider the random variable described in Prob. 3.3.

- (a) Sketch the CDF for  $I$  and  $K$ .  
 (b) Using the CDF, compute the probability that the transmitted signal is between 1 and 2 (inclusive).  
 (c) Using the CDF, compute the probability that the received signal is between 1 and 2 (inclusive).

3.24 Determine and plot the CDF for the random variable described in Prob. 3.4.

3.25 The PDF of a random variable  $X$  is of the form

$$f_X(x) = \begin{cases} Ce^{-x} + \frac{1}{3}\delta(x-2) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Where  $C$  is an unknown constant.

- Find the constant  $C$  and sketch  $f_X(x)$ .
- Find and sketch the CDF.
- What is the probability of the event  $1.5 < X \leq 2.5$ ?

3.26 Given the PDF of a random variable  $X$ :

$$f_X(x) = \begin{cases} Cx + 0.5\delta(x-1), & 0 \leq x \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

- Find  $C$  and sketch the PDF.
- Determine CDF of  $X$ .
- What are the probabilities for the following events?
  - $0.5 \leq X \leq 2$
  - $X > 2$ .

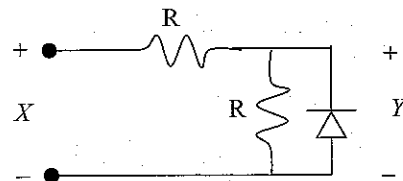
3.27 The CDF of a random variable  $X$  is

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x \leq \frac{1}{2} \\ 1 & \frac{1}{2} \leq x \end{cases}$$

- Sketch  $F_X(x)$ .
- Determine and sketch the PDF of  $X$ .

### Transformation of random variables

3.28 A voltage  $X$  is applied to the circuit shown below and the output voltage  $Y$  is measured.



The diode is assumed to be ideal, acting like a short circuit in the forward direction and an open circuit in the backward direction. If the voltage  $X$  is a Gaussian random variable with parameters  $m = 0$  and  $\sigma^2 = 1$ , find and sketch the PDF of the output voltage  $Y$ .

3.29 A random variable  $X$  is *uniformly distributed* (i.e., it has a uniform PDF) between  $-1$  and  $+1$ . A random variable  $Y$  is defined by  $Y = 3X + 2$ .

- Sketch the density function for  $X$  and the function describing the transformation from  $X$  to  $Y$ .
- Find the PDF and CDF for the random variable  $Y$  and sketch these functions.

3.30 The Laplace PDF is defined by

$$f_X(x) = \frac{\alpha}{2} e^{-\alpha|x-\mu|} \quad -\infty < x < \infty$$

Consider the transformation  $Y = aX + b$ . Demonstrate that if  $X$  is a Laplace random variable, then  $Y$  is also a Laplace random variable. What are the parameters  $\alpha'$  and  $\mu'$  corresponding to the PDF of  $Y$ ?

3.31 Consider a Laplace random variable with PDF

$$f_X(x) = e^{-2|x|}, \quad -\infty < x < \infty.$$

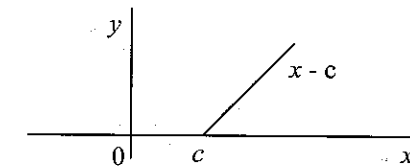
This random variable is passed through a 4-level uniform quantizer with output levels  $\{-1.5d, -0.5d, 0.5d, 1.5d\}$ .

- Determine the PMF of the quantizer output.
- Find  $\Pr[X > 2d]$  and  $\Pr[X < -2d]$ .

3.32 Consider a half-wave rectifier with a Gaussian input having mean 0 and standard deviation 2.

- Find the CDF of the output.
- Find the PDF of output by differentiating the CDF.

3.33 Consider a clipper with the following transfer characteristic:



The random variable  $X$  is exponential with  $\lambda = 2$ .

- Determine the CDF of the output.
- Find the PDF of output by differentiating the CDF.

### Distributions conditioned on an event

3.34 For the PDF given in Example 3.17, find and sketch the conditional density  $f_{X|A}(x|A)$  using the following choices for the event  $A$ :

- $X < 1$
- $\frac{1}{3} < X < \frac{5}{3}$
- $|\sin(X)| < \frac{1}{2}$

3.35 Consider the exponential density function

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find and sketch the conditional density function  $f_{X|A}(x|A)$  where  $A$  is the event  $X \geq 2$ .

3.36 Let  $K$  be a discrete uniform random variable with parameters  $n = -2$ ,  $m = 2$ . Find the conditional PMF  $f_{K|A}[k|A]$  for the following choices of the event  $A$ :

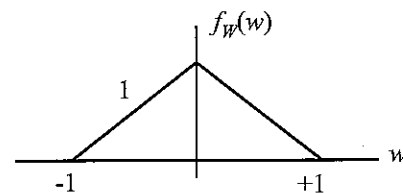
- $K < 0$

(b)  $K \geq 0$ **Applications**

**3.37** In a certain communication problem, the received signal  $X$  is a random variable given by

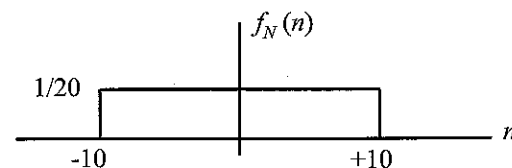
$$X = 2 + W$$

where  $W$  is a random variable representing the noise in the system. The noise  $W$  has a PDF which is sketched below:



- (a) Find and sketch the PDF of  $X$ .  
 (b) Using your answer to Part (a), find  $\Pr[X > 2]$ .

**3.38** A signal is sent over a channel where there is additive noise. The noise  $N$  is uniformly distributed between  $-10$  and  $10$  as shown below:



The receiver wishes to decide between two possible events based on the value of a random variable  $X$ . These two possible events are:

$$\begin{aligned} H_0 : X &= N && \text{(no signal sent, only noise received)} \\ H_1 : X &= S + N && \text{(signal was sent, signal plus noise received)} \end{aligned}$$

Here  $S$  is a fixed constant set equal to 18. It is *not* a random variable.

- (a) Draw the density functions  $f_{X|H_0}(x|H_0)$  and  $f_{X|H_1}(x|H_1)$  in the same picture. Be sure to label your plot carefully.  
 (b) What decision rule, based on the value of  $X$ , would guarantee perfect detection of the signal (i.e.,  $\Pr[\text{Error}|H_1] = 0$ ), but minimize the probability of false alarm (i.e.,  $\Pr[\text{Error}|H_0]$ )? What is this probability of false alarm?  
 (c) What decision rule would make the probability of a false alarm go to 0 but maximize the probability of detection? What is the probability of detection for this decision rule?  
 (d) If we could choose another value of  $S$  rather than  $S=18$ , what is the *minimum* value of  $S$  that would make it possible to have perfect detection and 0 false alarm probability?

**3.39** Carry out the algebraic steps involved to show that the decision rule of (3.42) reduces to the simpler form given by (3.43).

**Computer Projects****Project 3.1**

This project requires writing computer code to generate random variables of various kinds and measure their probability distributions.

1. Generate a sequence of each of the following types of random variables; each sequence should be at least 10,000 points long.
  - (a) A binomial random variable. Let the number of Bernoulli trials be  $n = 12$ . Recall that the binomial random variable is defined as the number of 1's in  $n$  trials for a Bernoulli (binary) random variable. Let the parameter  $p$  in the Bernoulli trials be  $p = 0.5109$ .
  - (b) A Poisson random variable as a limiting case of the binomial random variable with  $p = 0.0125$  or less and  $n = 80$  or more while maintaining  $\alpha = np = 1$ .
  - (c) A type 1 geometric random variable with parameter  $p = 0.09$ .
  - (d) A (continuous) uniform random variable in the range  $[-2, 5]$ .
  - (e) A Gaussian random variable with mean  $\mu = 1.3172$  and variance  $\sigma^2 = 1.9236$ .
  - (f) An exponential random variable with parameter  $\lambda = 1.37$ .

2. Estimate the CDF and PDF or PMF as appropriate and plot these functions next to their theoretical counterparts. Compare and comment on the results obtained.

**MATLAB programming notes**

To generate uniform and Gaussian random variables in Step 1(d) and 1(e), use the MATLAB functions 'rand' and 'randn' respectively with an appropriate transformation.

To generate an exponential random variable, you may use the function 'expon' from the software package for this book. This function uses the "transformation" method to generate the random variable.

In Step 2 you may use the functions 'pmfcdf' and 'pdfcdf' from the software package.

**Project 3.2**

In this project you will explore the problem of detection of a simple signal in Gaussian noise.

1. Consider a binary communication scheme where a logical 1 is represented by a constant voltage of value  $A = 5$  volts, and a logical 0 is represented by zero volts. Additive Gaussian noise is present, so during each bit interval the received signal is a random variable  $X$  described by

$$\begin{aligned} X &= A + N && \text{if a 1 was sent} \\ X &= N && \text{if a 0 was sent} \end{aligned}$$

The noise  $N$  is a Gaussian random variable with mean zero and variance  $\sigma^2$ . The signal-to-noise ratio in decibels is defined by

$$\text{SNR} = 10 \log_{10} \frac{A^2}{\sigma^2} = 20 \log_{10} \frac{A}{\sigma}$$

For this study the SNR will be taken to be 10 dB. This formula allows you to compute a value for  $\sigma^2$ .

- (a) Let  $H_0$  be the event that a logical 0 was sent and  $H_1$  be the event that a 1 was sent. Show that the decision rule for the receiver to minimize the probability of error is of the form

$$x \underset{H_0}{\overset{H_1}{\geq}} \tau$$

Evaluate the threshold for the case where  $P_0 = P_1 = 0.5$  where  $P_0$  and  $P_1$  are the prior probabilities of a 0 and 1, i.e.,  $P_0 = \Pr[H_0]$  and  $P_1 = \Pr[H_1]$ .

- (b) Now consider a decision rule of the above form where  $\tau$  is any arbitrary threshold. Evaluate  $\epsilon_0$  and  $\epsilon_1$  and plot the probability of error (3.45) as a function of  $\tau$  for  $1 \leq \tau \leq 4$ . (Continue to use  $P_0 = P_1 = 0.5$  in this formula.) Find the minimum error, and verify that it occurs for the value of  $\tau$  computed in Step 1(a).
- (c) Repeat Step 1(b) using  $P_0 = 1/3$  and  $P_1 = 2/3$ . Does the minimum error occur where you would expect it to occur? Justify your answer.
2. The detection of targets (by a radar, for example) is an analogous problem. The signal, which is observed in additive noise, will have a mean value of  $A$  if a target is present (event  $H_1$ ), and mean 0 if only noise (no target) is present (event  $H_0$ ). Let  $p_D$  represent the probability of detecting the target and  $p_{FA}$  represent the probability of a false alarm.
- (a) Plot the ROC for this problem for the range of values of  $\tau$  used in Step 1 (above). Indicate the direction of increasing  $\tau$  to show how the probability of detection can be traded off for the probability of false alarm by adjusting the threshold.
- (b) Show where the two threshold values used in Step 1(b) and 1(c) lie on the ROC curve, and find values of  $p_D$  and  $p_{FA}$  corresponding to each threshold value.
3. In this last part of the project you will simulate data to see how experimental results compare to the theory.
- (a) i. Generate 1000 samples of a Gaussian random variable with mean 0 and variance  $\sigma^2$  as determined in Step 1 for the 10 dB SNR. Generate 1000 samples of another random variable with mean 5 and variance  $\sigma^2$ . These two sets of random variables represent the outcomes in 1000 trials of the events  $H_0$  and  $H_1$ . For each of these 2000 random variables make a decision according to the decision rule in Step 1(a) using the values of  $\tau$  that you found in Steps 1(a) and 1(b). Form estimates for the errors  $\epsilon_1$  and  $\epsilon_0$  by computing the relative frequency, that is:

$$\hat{\epsilon}_i = \frac{\text{\#misclassified samples for } H_i}{1000} \quad \text{for } i = 0, 1$$

- ii. Also compute  $p_D$ ,  $p_{FA}$ , and the probability of error (3.45) using these estimates. Plot the values on the graphs generated in Steps 1 and 2 for comparison.

- (b) Repeat Step (a)(i) above using two other threshold values of your choice for  $\tau$ . You can use the *same* 2000 random variables generated in Step (i); just make your decision according to the new thresholds. Plot these points on the ROC of Step 2.

### MATLAB programming notes

The MATLAB function 'erf' is available to help evaluate the integral of the Gaussian density.

To generate uniform and Gaussian random variables in Step 1(d) and 1(e), use the MATLAB functions 'rand' and 'randn' respectively with an appropriate transformation.