EN 257: Applied Stochastic Processes
Problem Set 2

Douglas Lanman
dlanman@brown.edu
9 February 2007

Problem 3.16

The objective is to generate numbers from the pdf shown in Figure P3.16 on page 164 in [3]. All that is available is a random number generator that generates numbers uniformly in $(0, 1)$. Explain what procedure you would use to meet the objective.

If we assume that $f(x)$ is symmetric about $x = 0$ and is composed of two linear segments on $-1 \leq x < 0$ and $0 \leq x < 1$, then the resulting triangle must have unit area (by Equation 2.4-3 on page 66 in [3]). It directly follows that the pdf $f_X(x)$ must have the following form.

$$f_X(x) = \begin{cases} 0, & \text{for } x < -1 \\ 1 - x, & \text{for } -1 \leq x < 0 \\ x - 1, & \text{for } 0 \leq x < 1 \\ 0, & \text{for } x \geq 1 \end{cases} \tag{1}$$

Now recall that the PDF $F_X(x)$ is defined in terms of $f_X(x)$ as follows.

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi) d\xi \tag{2}$$

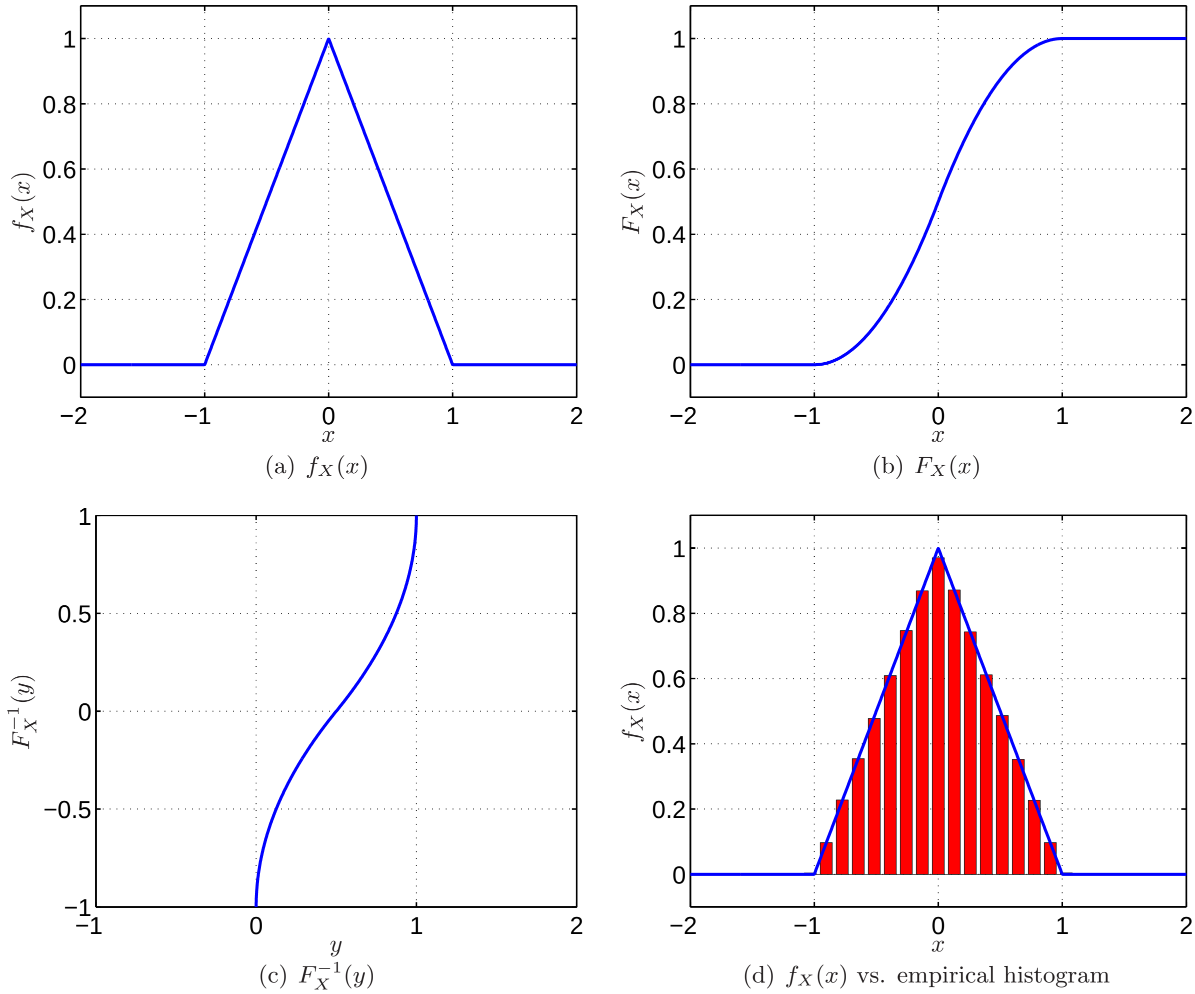
Applying Equation 2 to Equation 1 (and solving the resulting integrals) yields in the following expression for the PDF $F_X(x)$.

$$F_X(x) = \begin{cases} 0, & \text{for } x < -1 \\ \frac{x^2}{2} + x + \frac{1}{2}, & \text{for } -1 \leq x < 0 \\ -\frac{x^2}{2} + x + \frac{1}{2}, & \text{for } 0 \leq x < 1 \\ 1, & \text{for } x \geq 1 \end{cases} \tag{3}$$

At this point we recall the following procedure for generating a r.v. X with PDF $F_X(x)$ from a uniform r.v. Y. As stated on page 125 in [3], "...given a uniform r.v. Y, the transformation $X = F_X^{-1}(Y)$ will generate a r.v. with PDF $F_X(x)$". Note that $F_X^{-1}(y)$ denotes the inverse function, such that $F_X^{-1}(F_X(x)) = x$ [4]. From the plot of $F_X(x)$ shown in Figure 1(b), it is apparent that $F_X(x)$ maps $x \in \mathbb{R}$ onto the open interval $(0, 1)$. As a result we only require a closed-form expression for the inverse PDF $F_X^{-1}(y)$ for $y \in (0, 1)$. Application of the quadratic formula gives the following inverse functions.

$$y = \frac{x^2}{2} + x + \frac{1}{2}, \text{ for } -1 \leq x < 0 \quad \Rightarrow \quad \sqrt{2y} - 1, \text{ for } 0 \leq y < \frac{1}{2} \tag{4}$$

$$y = -\frac{x^2}{2} + x + \frac{1}{2}, \text{ for } 0 \leq x < 1 \quad \Rightarrow \quad 1 - \sqrt{2 - 2y}, \text{ for } \frac{1}{2} \leq y < 1 \tag{5}$$



(a) $f_X(x)$ (b) $F_X(x)$ (c) $F_X^{-1}(y)$ (d) $f_X(x)$ vs. empirical histogram

Figure 1: Generating a r.v. X with pdf $f_X(x)$ from a uniform r.v. Y. (a) The desired pdf $f_X(x)$. (b) The corresponding PDF $F_X(x)$. (c) The inverse PDF $F_X^{-1}(y)$. (d) The empirical pdf using 10^6 samples (i.e., the normalized histogram shown in red) and the desired pdf $f_X(x)$ (shown in blue).

Combining Equation 4 and Equation 5 gives the following solution for the inverse PDF $F_X^{-1}(y)$.

$$F_X^{-1}(y) = \begin{cases} \sqrt{2y} - 1, & \text{for } 0 < y < \frac{1}{2} \\ 1 - \sqrt{2 - 2y}, & \text{for } \frac{1}{2} \leq y < 1 \end{cases} \tag{6}$$

In conclusion, we propose the following algorithm for generating a r.v. X with pdf $f_X(x)$, as defined by Equation 1, from a uniform r.v. Y.

1. Generate N random samples y_i, for $i = 0, \ldots, N$, of the r.v. Y which is uniform on $(0, 1)$.
2. Transform each sample such that $x_i = F_X^{-1}(y_i)$, for $F_X^{-1}(y)$ as defined by Equation 6.

This procedure was implemented using the MATLAB function `prob1.m` (attached at the end of this write-up). As shown in Figure 1(d), the histogram of 10^6 transformed samples closely approximates the desired pdf $f_X(x)$.

Problem 3.19

Let X and Y be independent, continuous r.v.'s. Let $Z = \min(X, Y)$. (a) Compute $F_Z(z)$ and $f_Z(z)$. (b) Sketch the result if X and Y are uniform r.v.'s in $(0, 1)$. (c) Sketch the result for $f_X(x) = f_Y(x) = \alpha \exp(-\alpha x) \cdot u(x)$.

Part (a)

Note that the probability distribution function $F_Z(z)$ of $Z = \min(X, Y)$ can be expressed as follows.

$$F_Z(z) = P[\min(X,Y) \le z] = 1 - P[X > z, Y > z]$$

In other words, the region of interest is $\mathbb{R}^2 \setminus \{X > z, Y > z\}$ (i.e., the entire real plane except where X and Y are greater than z). Since X and Y are independent, we can write

$$\begin{aligned} F_Z(z) &= 1 - P[X > z]P[Y > z] = 1 - (1 - P[X \le z])\,(1 - P[Y \le z]) \\ &= 1 - (1 - F_X(z))(1 - F_Y(z)) = F_X(z) + F_Y(z) - F_X(z)F_Y(z), \end{aligned}$$

were $F_X(x) = P[X \le x]$. Finally, we differentiate $F_Z(z)$ with respect to z to obtain the probability density function $f_Z(z)$.

$$\boxed{\begin{aligned} F_Z(z) &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) \\ f_Z(z) &= f_X(z) + f_Y(z) - f_X(z)F_Y(z) - F_X(z)f_Y(z) \end{aligned}} \tag{7}$$

Part (b)

First, we recall that the PDF $F_X(x)$ of a uniform r.v. X in $(0, 1)$ can be expressed using Equation 2.3-3 on page 64 in [3].

$$F_X(x) = \begin{cases} 0, & \text{for } x \le 0 \\ x, & \text{for } 0 < x \le 1 \\ 1, & \text{for } x > 1 \end{cases} \tag{8}$$

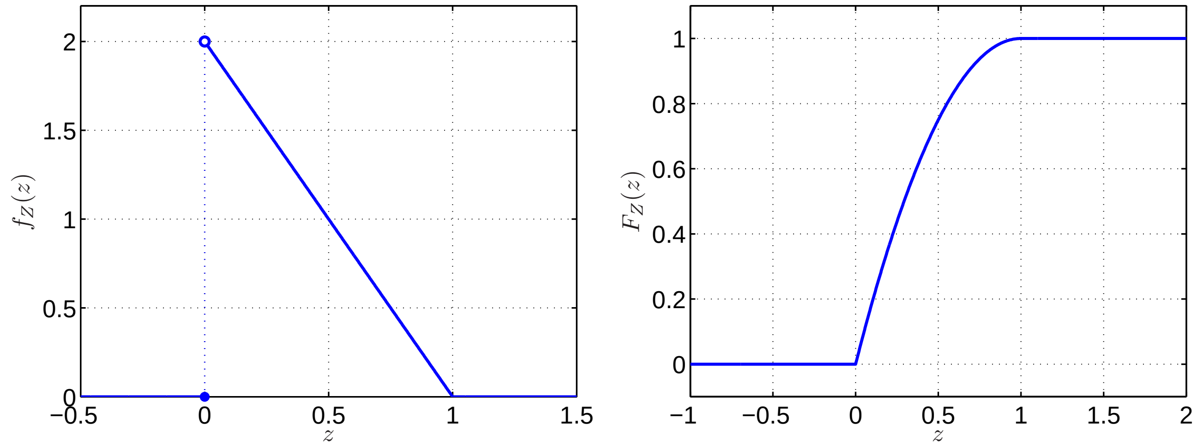
Similarly, the pdf $f_X(x)$ is given by Equation 2.4-17 on page 72.

$$f_X(x) = \begin{cases} 1, & \text{for } 0 < x \le 1 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Substituting Equations 8 and 9 into Equation 7 yields the desired expressions.

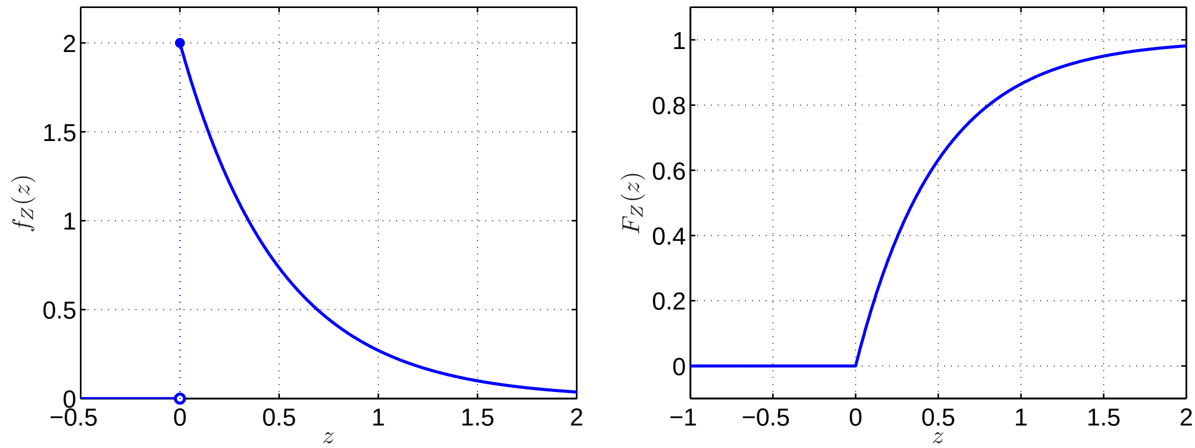
$$\boxed{f_Z(z) = \begin{cases} -2z + 2, & \text{for } 0 < z \le 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad F_Z(z) = \begin{cases} 0, & \text{for } z \le 0 \\ -z^2 + 2z, & \text{for } 0 < z \le 1 \\ 1, & \text{for } z > 1 \end{cases}}$$

Plots of $f_Z(z)$ and $F_Z(z)$ are shown in Figure 2.



(a) $f_Z(z)$ (b) $F_Z(z)$

Figure 2: Plots of $f_Z(z)$ and $F_Z(z)$ derived in part (b) of Problem 3.19.



(a) $f_Z(z)$, for $\alpha = 1$ (b) $F_Z(z)$, for $\alpha = 1$

Figure 3: Plots of $f_Z(z)$ and $F_Z(z)$ derived in part (c) of Problem 3.19.

Part (c)

To begin our analysis we note that, for $f_X(x) = \alpha \exp(-\alpha x) \cdot u(x)$, X is an exponential random variable. Given the pdf $f_X(x)$, the PDF $F_X(x)$ can be obtained by integration.

$$F_X(x) = \int_{-\infty}^{x} f(\xi) d\xi = \begin{cases} \alpha \int_0^x e^{-\alpha \xi} d\xi, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases} = \begin{cases} 1 - e^{-\alpha x}, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases}$$

Substituting for $f_X(x)$ and $F_X(x)$ in Equation 7 yields the desired expressions.

$$\boxed{f_Z(z) = \begin{cases} 2\alpha e^{-2\alpha z}, & \text{for } z \geq 0 \\ 0, & \text{for } z < 0 \end{cases} \quad \text{and} \quad F_Z(z) = \begin{cases} 1 - e^{-2\alpha z}, & \text{for } z \geq 0 \\ 0, & \text{for } z < 0 \end{cases}}$$

Plots of $f_Z(z)$ and $F_Z(z)$ are shown in Figure 3.

Problem 3.28

(a) Compute the joint pdf of

$$Z \triangleq g(X,Y) = X^2 + Y^2$$
$$W \triangleq h(X,Y) = X$$

when

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2} e^{-[(x^2+y^2)/2\sigma^2]}. \tag{10}$$

(b) Compute $f_Z(z)$ from your results.

Part (a)

First, we note that this problem is primarily instructive since direct methods exist (e.g., Example 3.3-8 on page 149 in [3]). As described, this problem demonstrates the use of an *auxiliary random variable* $W \triangleq X$ to determine $f_Z(z)$ from the joint pdf $f_{ZW}(z,w)$. We begin our analysis by observing that the equations

$$z - g(x,y) = 0$$
$$w - h(x,y) = 0$$

have two real roots, for $|w| \leq \sqrt{z}$ and $z \geq 0$, given by

$$\begin{array}{ll} x_1 = \phi_1(z,w) = w & x_2 = \phi_2(z,w) = w \\ y_1 = \varphi_1(z,w) = \sqrt{z - w^2} & y_2 = \varphi_2(z,w) = -\sqrt{z-w^2}. \end{array} \tag{11}$$

At this point we recall that $f_{ZW}(z,w)$ can be obtained directly from $f_{XY}(x,y)$ using the methods outlined in Section 3.4 in [3]. From that section we note that the joint pdf can be expressed as

$$f_{ZW}(z,w) = \sum_{i=1}^{n} f_{XY}(x_i, y_i)|\tilde{J}_i|, \tag{12}$$

where $|\tilde{J}_i|$ is the magnitude of the Jacobian transformation such that

$$|\tilde{J}_i| = \left| \det \begin{pmatrix} \partial\phi_i/\partial z & \partial\phi_i/\partial w \\ \partial\varphi_i/\partial z & \partial\varphi_i/\partial w \end{pmatrix} \right| \tag{13}$$

and n is the number of solutions to the equations $z = g(x,y)$ and $w = h(x,y)$. Substituting Equation 11 into Equation 13 gives the following Jacobian magnitutdes.

$$|\tilde{J}_1| = \left| \det \begin{pmatrix} 0 & 1 \\ \frac{1}{2\sqrt{z-w^2}} & \frac{-w}{\sqrt{z-w^2}} \end{pmatrix} \right| = \frac{1}{2\sqrt{z-w^2}} \tag{14}$$

$$|\tilde{J}_2| = \left| \det \begin{pmatrix} 0 & 1 \\ \frac{-1}{2\sqrt{z-w^2}} & \frac{w}{\sqrt{z-w^2}} \end{pmatrix} \right| = \frac{1}{2\sqrt{z-w^2}} \tag{15}$$

Before we proceed, we observe that Equation 10 can be expressed as a function of z, such that

$$f_{XY}(z) = \frac{1}{2\pi\sigma^2} e^{-z/2\sigma^2}. \tag{16}$$

Substituting Equations 14-16 into Equation 12 gives

$$f_{ZW}(z,w) = \begin{cases} f_{XY}(z)|\tilde{J}_1| + f_{XY}(z)|\tilde{J}_2|, & \text{for } |w| \le \sqrt{z} \text{ and } z \ge 0 \\ 0, & \text{otherwise,} \end{cases}$$

which yields the desired expression for the joint pdf $f_{ZW}(z,w)$.

$$\boxed{f_{ZW}(z,w) = \begin{cases} \frac{1}{2\pi\sigma^2\sqrt{z-w^2}}\, e^{-z/2\sigma^2}, & \text{for } |w| \le \sqrt{z} \text{ and } z \ge 0 \\ 0, & \text{otherwise} \end{cases}}$$

Part (b)

Given the joint pdf $f_{ZW}(z,w)$, we can obtain the marginal pdf $f_Z(z)$ as follows.

$$f_Z(z) = \int_{-\infty}^{\infty} f_{ZW}(z,w)dw.$$

Substituting the result from part (b), we find

$$f_Z(z) = \frac{1}{2\pi\sigma^2}\, e^{-z/2\sigma^2} \left[\int_{-\sqrt{z}}^{\sqrt{z}} \frac{dw}{\sqrt{z-w^2}} \right] u(z).$$

To complete our derivation, we note that the remaining integral can be solved by a change of variables. If we let $w \triangleq \sqrt{z}\sin\theta$, then $dw = \sqrt{z}\cos\theta d\theta$ and $\sqrt{z-w^2} = \sqrt{z}\cos\theta$. As a result, we find

$$\int_{-\sqrt{z}}^{\sqrt{z}} \frac{dw}{\sqrt{z-w^2}} = \int_{-\pi/2}^{\pi/2} d\theta = \pi.$$

In conclusion, we obtain the following solution for the marginal pdf $f_Z(z)$.

$$\boxed{f_Z(z) = \frac{1}{2\sigma^2}\, e^{-z/2\sigma^2} u(z)}$$

Note that this result is identical to that obtained using direct methods in Example 3.3-8.

Problem 4.44

Let X_i for $i = 1, \ldots, 4$ be four zero-mean Gaussian random variables. Use the joint characteristic function to show that

$$E[X_1X_2X_3X_4] = E[X_1X_2]E[X_3X_4] + E[X_1X_3]E[X_2X_4] + E[X_2X_3]E[X_1X_4]. \tag{17}$$

Recall from Equation 4.7-12 on page 222 in [3] that the joint characteristic function is given by

$$\Phi_{X_1 \ldots X_N}(\omega_1, \ldots, \omega_N) = E\left[\exp\left(j\sum_{i=1}^{N}\omega_i X_i\right)\right],$$

for N random variables $X_1, \ldots, X_N$. From Equation 4.7-14 and 5.7-5, we also recall that the joint characteristic function can be used to obtain the joint moments as follows.

$$E[X_1^{k_1} \ldots X_N^{k_N}] = (-j)^{k_1+\ldots+k_N} \left.\frac{\partial^{k_1+\ldots+k_N}\Phi_{X_1\ldots X_N}(\omega_1, \ldots, \omega_N)}{\partial\omega_1^{k_1}\ldots\partial\omega_N^{k_N}}\right|_{\omega_1=\ldots=\omega_N=0} \tag{18}$$

For this problem we have $N = 4$ and $k_1 = \ldots = k_4 = 1$. As a result, substituting Equation 18 into Equation 17 gives the following equality in terms of the joint characteristic functions.

$$\begin{aligned}\Phi^{(1,1,1,1)}_{X_1X_2X_3X_4}(0,0,0,0) = \Phi^{(1,1)}_{X_1X_2}(0,0)\Phi^{(1,1)}_{X_3X_4}(0,0)+ \\ \Phi^{(1,1)}_{X_1X_3}(0,0)\Phi^{(1,1)}_{X_2X_4}(0,0) + \Phi^{(1,1)}_{X_2X_3}(0,0)\Phi^{(1,1)}_{X_1X_4}(0,0)\end{aligned} \tag{19}$$

Notice that we have simplified the previous expression by using the following shorthand notation.

$$\Phi^{(k_1,\ldots,k_N)}_{X_1\ldots X_N}(0,\ldots,0) \triangleq \left.\frac{\partial^{k_1+\ldots+k_N}\Phi_{X_1\ldots X_N}(\omega_1,\ldots,\omega_N)}{\partial\omega_1^{k_1}\ldots\partial\omega_N^{k_N}}\right|_{\omega_1=\ldots=\omega_N=0}$$

At this point, we require a closed-form expression for the joint characteristic function of two or more zero-mean Gaussian random variables. Conveniently, this has already been derived in Section 5.7 in [3]. Following the derivation of Equation 5.7-20 we obtain the following joint characteristic function for two zero-mean Gaussian random variables X_i and X_j.

$$\Phi_{X_iX_j}(\omega_i, \omega_j) = e^{-\frac{1}{2}\omega^T\mathbf{K}\omega}, \text{ for } \mathbf{K} = \begin{pmatrix} K_{ii} & K_{ij} \\ K_{ji} & K_{jj}\end{pmatrix} \text{ and } \omega = \begin{pmatrix}\omega_i \\ \omega_j\end{pmatrix} \tag{20}$$

Similarly, for four zero-mean Gaussian random variables $X_1, \ldots, X_4$, we have

$$\Phi_{X_1X_2X_3X_4}(\omega_1,\omega_2,\omega_3,\omega_4) = e^{-\frac{1}{2}\omega^T\mathbf{K}\omega}, \text{ for } \mathbf{K} = \begin{pmatrix} K_{11} & \ldots & K_{14} \\ \vdots & \ddots & \vdots \\ K_{41} & \ldots & K_{44}\end{pmatrix} \text{ and } \omega = \begin{pmatrix}\omega_1 \\ \vdots \\ \omega_4\end{pmatrix}. \tag{21}$$

Substituting Equation 20 into Equation 18 gives the following expression for the joint moment $E[X_iX_j]$.

$$E[X_iX_j] = -\left.\frac{\partial^2\Phi_{X_iX_j}(\omega_i,\omega_j)}{\partial\omega_i\partial\omega_j}\right|_{\omega_i=\omega_j=0} = \frac{1}{2}(K_{ij}+K_{ji}) = K_{ij} \tag{22}$$

Note that in the previous expression we have applied the result that the covariance matrix is symmetric (i.e., $K_{ij} = K_{ji}$). As a result, we conclude that $E[X_i X_j] = K_{ij}$; this result is expected since, from the definition of the joint moment in Equation 5.3-2, the covariance and correlation are identical for $\mu = 0$. Substituting the result from Equation 22 into Equation 17 gives the following expression for the joint moment of four zero-mean Gaussian random variables.

$$E[X_1X_2X_3X_4] \stackrel{?}{=} K_{12}K_{34} + K_{13}K_{24} + K_{23}K_{14} \tag{23}$$

All that remains is to demonstrate that the left-hand and right-hand sides of Equation 23 are equivalent. To proceed, we substitute Equation 21 into Equation 18 to obtain the following result.

$$E[X_1X_2X_3X_4] = \left. \frac{\partial^4 e^{-\frac{1}{2}\omega^T \mathbf{K} \omega}}{\partial\omega_1 \partial\omega_2 \partial\omega_3 \partial\omega_4} \right|_{\omega_1 = \ldots = \omega_4 = 0}, \text{ for } \mathbf{K} = \begin{pmatrix} K_{11} & \ldots & K_{14} \\ \vdots & \ddots & \vdots \\ K_{41} & \ldots & K_{44} \end{pmatrix} \text{ and } \omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_4 \end{pmatrix}$$

Evaluating the first partial derivative with respect to ω_1 gives the following result.

$$\begin{aligned} E[X_1X_2X_3X_4] &= \frac{\partial^3}{\partial\omega_2\partial\omega_3\partial\omega_4} \left\{ \left. \frac{\partial\, e^{-\frac{1}{2}\omega^T \mathbf{K}\omega}}{\partial\omega_1} \right|_{\omega_1=0} \right\} \Bigg|_{\omega_2=\ldots=\omega_4=0} \\ &= \frac{\partial^3}{\partial\omega_2\partial\omega_3\partial\omega_4} \left\{ -(K_{12}\omega_2 + K_{13}\omega_3 + K_{14}\omega_4) e^{-\frac{1}{2}\omega^T\mathbf{K}\omega} \right\} \Bigg|_{\omega_2=\ldots=\omega_4=0} \end{aligned}$$

Note that in the previous expression we have assumed that ω is now given by $\omega = (0, \omega_2, \omega_3, \omega_4)^T$. Continuing, we can now evaluate the partial derivative with respect to ω_2.

$$\begin{aligned} E[X_1X_2X_3X_4] &= -\frac{\partial^2}{\partial\omega_3\partial\omega_4} \left\{ \left. \frac{\partial\,(K_{12}\omega_2 + K_{13}\omega_3 + K_{14}\omega_4) e^{-\frac{1}{2}\omega^T\mathbf{K}\omega}}{\partial\omega_2} \right|_{\omega_2=0} \right\} \Bigg|_{\omega_3=\omega_4=0} \\ &= \frac{\partial^2}{\partial\omega_3\partial\omega_4} \left\{ \left[(K_{13}\omega_3 + K_{14}\omega_4)(K_{23}\omega_3 + K_{24}\omega_4) - K_{12} \right] e^{-\frac{1}{2}\omega^T\mathbf{K}\omega} \right\} \Bigg|_{\omega_3=\omega_4=0} \end{aligned}$$

As before, we have reduced ω to be $\omega = (0, 0, \omega_3, \omega_4)^T$. Next, we evaluate the partial derivative with respect to ω_3.

$$\begin{aligned} E[X_1X_2X_3X_4] &= \frac{\partial}{\partial\omega_4} \left\{ \left. \frac{\partial\, \left[(K_{13}\omega_3 + K_{14}\omega_4)(K_{23}\omega_3 + K_{24}\omega_4) - K_{12} \right] e^{-\frac{1}{2}\omega^T\mathbf{K}\omega}}{\partial\omega_3} \right|_{\omega_3=0} \right\} \Bigg|_{\omega_4=0} \\ &= \frac{\partial}{\partial\omega_4} \left\{ \left[(K_{12}K_{34} + K_{13}K_{24} + K_{23}K_{14})\omega_4 - K_{14}K_{24}K_{34}\omega_4^3 \right] e^{-\frac{1}{2}K_{44}\omega_4^2} \right\} \Bigg|_{\omega_4=0} \end{aligned}$$

To complete our derivation, we evaluate the partial derivative in ω_4 to yield the desired result.

$$E[X_1X_2X_3X_4] = K_{12}K_{34} + K_{13}K_{24} + K_{23}K_{14} \tag{24}$$

Since Equation 23 and Equation 24 agree, we conclude that the following expression will hold for any four zero-mean Gaussian random variables.

$$\boxed{E[X_1X_2X_3X_4] = E[X_1X_2]E[X_3X_4] + E[X_1X_3]E[X_2X_4] + E[X_2X_3]E[X_1X_4]}$$

(QED)

Problem 4.48

Show that

$$W_n \triangleq \sum_{i=1}^{n} \left[\frac{1}{\sigma} \left(X_i - \frac{1}{n} \sum_{j=1}^{n} X_j \right) \right]^2 \tag{25}$$

is Chi-square with $n-1$ degrees of freedom.

To begin our derivation we first recall that the sample mean estimator $\hat{\mu}_n$ is defined as

$$\hat{\mu}_n \triangleq \frac{1}{n} \sum_{i=1}^{n} X_i, \tag{26}$$

where $X_1, \ldots, X_n$ are n independent observations of a Normal random variable with unknown mean μ and variance σ^2. Substituting this expression into Equation 25 gives the following result.

$$W_n = \sum_{i=1}^{n} \left(\frac{X_i - \hat{\mu}_n}{\sigma} \right)^2$$

At this point we are free to add and subtract the true mean μ as follows.

$$\begin{aligned} W_n &= \sum_{i=1}^{n} \left[\frac{(X_i - \mu) + (\mu - \hat{\mu}_n)}{\sigma} \right]^2 \\ &= \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 + \frac{2}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)(\mu - \hat{\mu}_n) + \frac{1}{\sigma^2} \sum_{i=1}^{n} (\mu - \hat{\mu}_n)^2 \end{aligned}$$

Note that the quantity $(\mu - \hat{\mu}_n)$ is a constant. As a result, we can further reduce the previous expression.

$$W_n = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 + \frac{2(\mu - \hat{\mu}_n)}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu) + n \left(\frac{\mu - \hat{\mu}_n}{\sigma} \right)^2$$

As an aside we also note that $\sum_{i=1}^{n} (X_i - \mu) = n(\hat{\mu}_n - \mu)$. Substituting this result into the previous expression yields

$$W_n = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 - n \left(\frac{\mu - \hat{\mu}_n}{\sigma} \right)^2 = \sum_{i=1}^{n} \left[\left(\frac{X_i - \mu}{\sigma} \right)^2 - \left(\frac{\mu - \hat{\mu}_n}{\sigma} \right)^2 \right].$$

Finally, by applying Equation 26, we can reduce W_n to obtain the following simple form for Equation 25.

$$W_n = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 - \frac{1}{n} \left[\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) \right]^2 \tag{27}$$

As described on page 234 in [3], if we make n observations of a Normal random variable with variance σ^2 and mean μ, then the random variable $U_i \triangleq (X_i - \mu)/\sigma$ is $N(0,1)$. In addition, we also know that $Z_n \triangleq \sum_{i=1}^{n} U_i^2$ is Chi-square with n degrees of freedom. Examining Equation 27, we notice that the first term corresponds identically with this situation. The second term, however, is simply a sum of standard Normal random variables which is multiplied by $1/n$. By the fundamental properties of Normal random variables, we can conclude that the right-hand side is also a Normal random variable $N(0,1)$. In conclusion, W_n is composed of a summation over $n-1$ *independent* Normal random variables. As a result we can conclude that

$$W_n \triangleq \sum_{i=1}^{n} \left[\frac{1}{\sigma} \left(X_i - \frac{1}{n} \sum_{j=1}^{n} X_j \right) \right]^2$$

is Chi-square with $n-1$ degrees of freedom. (QED)

References

[1] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes (Third Edition)*. Oxford University Press, 2001.

[2] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[3] Henry Stark and John W. Woods. *Probability and Random Processes with Applications to Signal Processing (Third Edition)*. Prentice-Hall, 2002.

[4] Eric W. Weisstein. Inverse function. `http://mathworld.wolfram.com/InverseFunction.html`.