Expected Values of Sums

Theorem 9.1

For any set of random variables $X_1, \ldots, X_n$, the sum $W_n = X_1 + \cdots + X_n$ has expected value

$$\mathrm{E}[W_n] = \mathrm{E}[X_1] + \mathrm{E}[X_2] + \cdots + \mathrm{E}[X_n].$$

Proof: Theorem 9.1

We prove this theorem by induction on n. In Theorem 5.11, we proved $E[W_2] = E[X_1] + E[X_2]$. Now we assume $E[W_{n-1}] = E[X_1] + \cdots + E[X_{n-1}]$. Notice that $W_n = W_{n-1} + X_n$. Since W_n is a sum of the two random variables W_{n-1} and X_n, we know that $E[W_n] = E[W_{n-1}] + E[X_n] = E[X_1] + \cdots + E[X_{n-1}] + E[X_n]$.

Theorem 9.2

The variance of $W_n = X_1 + \cdots + X_n$ is

$$\text{Var}[W_n] = \sum_{i=1}^{n} \text{Var}[X_i] + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{Cov}\left[X_i, X_j\right].$$

Proof: Theorem 9.2

From the definition of the variance, we can write $\text{Var}[W_n] = \text{E}[(W_n - \text{E}[W_n])^2]$. For convenience, let μ_i denote $\text{E}[X_i]$. Since $W_n = \sum_{i=1}^n X_n$ and $\text{E}[W_n] = \sum_{i=1}^n \mu_i$, we can write

$$\begin{aligned} \text{Var}[W_n] = \text{E}\left[\left(\sum_{i=1}^n (X_i - \mu_i)\right)^2\right] &= \text{E}\left[\sum_{i=1}^n (X_i - \mu_i) \sum_{j=1}^n (X_j - \mu_j)\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}\left[X_i, X_j\right]. \end{aligned} \tag{9.2}$$

In terms of the random vector $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}'$, we see that $\text{Var}[W_n]$ is the sum of all the elements of the covariance matrix $\mathbf{C_X}$. Recognizing that $\text{Cov}[X_i, X_i] = \text{Var}[X]$ and $\text{Cov}[X_i, X_j] = \text{Cov}[X_j, X_i]$, we place the diagonal terms of $\mathbf{C_X}$ in one sum and the off-diagonal terms (which occur in pairs) in another sum to arrive at the formula in the theorem.

Theorem 9.3

When $X_1, \ldots, X_n$ are uncorrelated,

$$\text{Var}[W_n] = \text{Var}[X_1] + \cdots + \text{Var}[X_n].$$

Example 9.1 Problem

$X_0, X_1, X_2, \ldots$ is a sequence of random variables with expected values $\mathrm{E}[X_i] = 0$ and covariances, $\mathrm{Cov}[X_i, X_j] = 0.8^{|i-j|}$. Find the expected value and variance of a random variable Y_i defined as the sum of three consecutive values of the random sequence

$$Y_i = X_i + X_{i-1} + X_{i-2}. \tag{9.3}$$

Example 9.1 Solution

Theorem 9.1 implies that

$$\mathrm{E}[Y_i] = \mathrm{E}[X_i] + \mathrm{E}[X_{i-1}] + \mathrm{E}[X_{i-2}] = 0. \tag{9.4}$$

Applying Theorem 9.2, we obtain for each i,

$$\begin{aligned}\mathrm{Var}[Y_i] = \mathrm{Var}[X_i] + \mathrm{Var}[X_{i-1}] + \mathrm{Var}[X_{i-2}] \\ + 2\,\mathrm{Cov}[X_i, X_{i-1}] + 2\,\mathrm{Cov}[X_i, X_{i-2}] + 2\,\mathrm{Cov}[X_{i-1}, X_{i-2}].\end{aligned} \tag{9.5}$$

We next note that $\mathrm{Var}[X_i] = \mathrm{Cov}[X_i, X_i] = 0.8^{i-i} = 1$ and that

$$\mathrm{Cov}[X_i, X_{i-1}] = \mathrm{Cov}[X_{i-1}, X_{i-2}] = 0.8^1, \qquad \mathrm{Cov}[X_i, X_{i-2}] = 0.8^2. \tag{9.6}$$

Therefore,

$$\mathrm{Var}[Y_i] = 3 \times 0.8^0 + 4 \times 0.8^1 + 2 \times 0.8^2 = 7.48. \tag{9.7}$$

Example 9.2 Problem

At a party of $n \geq 2$ people, each person throws a hat in a common box. The box is shaken and each person blindly draws a hat from the box without replacement. We say a match occurs if a person draws his own hat. What are the expected value and variance of V_n, the number of matches?

Example 9.2 Solution

Let X_i denote an indicator random variable such that

$$X_i = \begin{cases} 1 & \text{person } i \text{ draws his hat,} \\ 0 & \text{otherwise.} \end{cases} \tag{9.8}$$

The number of matches is $V_n = X_1 + \cdots + X_n$. Note that the X_i are generally not independent. For example, with $n = 2$ people, if the first person draws his own hat, then the second person must also draw her own hat. Note that the ith person is equally likely to draw any of the n hats, thus $P_{X_i}(1) = 1/n$ and $\mathrm{E}[X_i] = P_{X_i}(1) = 1/n$. Since the expected value of the sum always equals the sum of the expected values,

$$\mathrm{E}[V_n] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_n] = n(1/n) = 1. \tag{9.9}$$

To find the variance of V_n, we will use Theorem 9.2. The variance of X_i is

$$\mathrm{Var}[X_i] = \mathrm{E}\left[X_i^2\right] - (\mathrm{E}[X_i])^2 = \frac{1}{n} - \frac{1}{n^2}. \tag{9.10}$$

To find $\mathrm{Cov}[X_i, X_j]$, we observe that

$$\mathrm{Cov}[X_i, X_j] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\,\mathrm{E}[X_j]. \tag{9.11}$$

[Continued]

Example 9.2 Solution (Continued 2)

Note that $X_iX_j = 1$ if and only if $X_i = 1$ and $X_j = 1$, and $X_iX_j = 0$ otherwise. Thus

$$\mathrm{E}[X_iX_j] = P_{X_i,X_j}(1,1) = P_{X_i|X_j}(1|1)\, P_{X_j}(1)\,. \tag{9.12}$$

Given $X_j = 1$, that is, the jth person drew his own hat, then $X_i = 1$ if and only if the ith person draws his own hat from the $n-1$ other hats. Hence $P_{X_i|X_j}(1|1) = 1/(n-1)$ and

$$\mathrm{E}[X_iX_j] = \frac{1}{n(n-1)}, \qquad \mathrm{Cov}[X_i, X_j] = \frac{1}{n(n-1)} - \frac{1}{n^2}. \tag{9.13}$$

Finally, we can use Theorem 9.2 to calculate

$$\mathrm{Var}[V_n] = n\,\mathrm{Var}[X_i] + n(n-1)\,\mathrm{Cov}[X_i, X_j] = 1. \tag{9.14}$$

That is, both the expected value and variance of V_n are 1, no matter how large n is!

Example 9.3 Problem

Continuing Example 9.2, suppose each person immediately returns to the box the hat that he or she drew. What is the expected value and variance of V_n, the number of matches?

Example 9.3 Solution

In this case the indicator random variables X_i are independent and identically distributed (iid) because each person draws from the same bin containing all n hats. The number of matches $V_n = X_1 + \cdots + X_n$ is the sum of n iid random variables. As before, the expected value of V_n is

$$E[V_n] = n\,E[X_i] = 1. \tag{9.15}$$

In this case, the variance of V_n equals the sum of the variances,

$$\text{Var}[V_n] = n\,\text{Var}[X_i] = n\left(\frac{1}{n} - \frac{1}{n^2}\right) = 1 - \frac{1}{n}. \tag{9.16}$$

Quiz 9.1

Let W_n denote the sum of n independent throws of a fair four-sided die. Find the expected value and variance of W_n.

Quiz 9.1 Solution

Let $K_1, \ldots, K_n$ denote a sequence of iid random variables each with PMF

$$P_K(k) = \begin{cases} 1/4 & k = 1, \ldots, 4, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

We can write $W_n = K_1 + \cdots + K_n$. First, we note that the first two moments of K_i are

$$\mathrm{E}[K_i] = \frac{1+2+3+4}{4} = 2.5, \tag{2}$$

$$\mathrm{E}\left[K_i^2\right] = \frac{1^2+2^2+3^2+4^2}{4} = 7.5. \tag{3}$$

Thus the variance of K_i is

$$\begin{aligned} \mathrm{Var}[K_i] &= \mathrm{E}\left[K_i^2\right] - (\mathrm{E}[K_i])^2 \\ &= 7.5 - (2.5)^2 = 1.25. \end{aligned} \tag{4}$$

Since $\mathrm{E}[K_i] = 2.5$, the expected value of W_n is

$$\mathrm{E}[W_n] = \mathrm{E}[K_1] + \cdots + \mathrm{E}[K_n] = 2.5n. \tag{5}$$

Since the rolls are independent, the random variables $K_1, \ldots, K_n$ are independent. Hence, by Theorem 9.3, the variance of the sum equals the sum of the variances. That is,

$$\mathrm{Var}[W_n] = \mathrm{Var}[K_1] + \cdots + \mathrm{Var}[K_n] = 1.25n. \tag{6}$$

Section 9.2

Moment Generating Functions

Definition 9.1 (MGF)

For a random variable X, *the* moment generating function (MGF) *of* X *is*

$$\phi_X(s) = \mathsf{E}\left[e^{sX}\right].$$

Theorem 9.4

A random variable X with MGF $\phi_X(s)$ has nth moment

$$\mathrm{E}\left[X^n\right] = \left.\frac{d^n \phi_X(s)}{ds^n}\right|_{s=0}.$$

Proof: Theorem 9.4

The first derivative of $\phi_X(s)$ is

$$\frac{d\phi_X(s)}{ds} = \frac{d}{ds}\left(\int_{-\infty}^{\infty} e^{sx} f_X(x)\ dx\right) = \int_{-\infty}^{\infty} x e^{sx} f_X(x)\ dx. \tag{9.19}$$

Evaluating this derivative at $s = 0$ proves the theorem for $n = 1$.

$$\left.\frac{d\phi_X(s)}{ds}\right|_{s=0} = \int_{-\infty}^{\infty} x f_X(x)\ dx = \mathrm{E}\left[X\right]. \tag{9.20}$$

Similarly, the nth derivative of $\phi_X(s)$ is

$$\frac{d^n\phi_X(s)}{ds^n} = \int_{-\infty}^{\infty} x^n e^{sx} f_X(x)\ dx. \tag{9.21}$$

The integral evaluated at $s = 0$ is the formula in the theorem statement.

Example 9.4 Problem

X is an exponential random variable with MGF $\phi_X(s) = \lambda/(\lambda - s)$. What are the first and second moments of X? Write a general expression for the nth moment.

Example 9.4 Solution

The first moment is the expected value:

$$E[X] = \left.\frac{d\phi_X(s)}{ds}\right|_{s=0} = \left.\frac{\lambda}{(\lambda - s)^2}\right|_{s=0} = \frac{1}{\lambda}. \tag{9.22}$$

The second moment of X is the mean square value:

$$E\left[X^2\right] = \left.\frac{d^2\phi_X(s)}{ds^2}\right|_{s=0} = \left.\frac{2\lambda}{(\lambda - s)^3}\right|_{s=0} = \frac{2}{\lambda^2}. \tag{9.23}$$

Proceeding in this way, it should become apparent that the nth moment of X is

$$E[X^n] = \left.\frac{d^n\phi_X(s)}{ds^n}\right|_{s=0} = \left.\frac{n!\lambda}{(\lambda - s)^{n+1}}\right|_{s=0} = \frac{n!}{\lambda^n}. \tag{9.24}$$

Theorem 9.5

The MGF of $Y = aX + b$ is $\phi_Y(s) = e^{sb}\phi_X(as)$.

Proof: Theorem 9.5

From the definition of the MGF,

$$\phi_Y(s) = \mathsf{E}\left[e^{s(aX+b)}\right] = e^{sb}\,\mathsf{E}\left[e^{(as)X}\right] = e^{sb}\phi_X(as). \tag{9.25}$$

Quiz 9.2

Random variable K has PMF

$$P_K(k) = \begin{cases} 0.2 & k = 0, \ldots, 4, \\ 0 & \text{otherwise.} \end{cases} \tag{9.26}$$

Use $\phi_K(s)$ to find the first, second, third, and fourth moments of K.

Quiz 9.2 Solution

The MGF of K is

$$\phi_K(s) = \mathsf{E}\left[e^{sK}\right] = \sum_{k=0}^{4} \frac{1}{5} e^{sk} = \frac{1 + e^s + e^{2s} + e^{3s} + e^{4s}}{5}. \qquad (1)$$

We find the moments by taking derivatives. The first derivative of $\phi_K(s)$ is

$$\frac{d\phi_K(s)}{ds} = \frac{e^s + 2e^{2s} + 3e^{3s} + 4e^{4s}}{5}. \qquad (2)$$

Evaluating the derivative at $s = 0$ yields

$$\mathsf{E}[K] = \left.\frac{d\phi_K(s)}{ds}\right|_{s=0} = \frac{1 + 2 + 3 + 4}{5} = 2. \qquad (3)$$

[Continued]

Quiz 9.2 Solution (Continued 2)

To find higher-order moments, we continue to take derivatives:

$$\begin{aligned} \mathrm{E}\left[K^2\right] &= \left.\frac{d^2\phi_K(s)}{ds^2}\right|_{s=0} \\ &= \left.\frac{e^s + 4e^{2s} + 9e^{3s} + 16e^{4s}}{5}\right|_{s=0} = 6. \end{aligned} \tag{4}$$

$$\begin{aligned} \mathrm{E}\left[K^3\right] &= \left.\frac{d^3\phi_K(s)}{ds^3}\right|_{s=0} \\ &= \left.\frac{e^s + 8e^{2s} + 27e^{3s} + 64e^{4s})}{5}\right|_{s=0} = 20. \end{aligned} \tag{5}$$

$$\begin{aligned} \mathrm{E}\left[K^4\right] &= \left.\frac{d^4\phi_K(s)}{ds^4}\right|_{s=0} \\ &= \left.\frac{e^s + 16e^{2s} + 81e^{3s} + 256e^{4s}}{5}\right|_{s=0} = 70.8. \end{aligned} \tag{6}$$

Section 9.3

MGF of the Sum of Independent Random Variables

Theorem 9.6

For a set of independent random variables $X_1, \ldots, X_n$, the moment generating function of $W = X_1 + \cdots + X_n$ is

$$\phi_W(s) = \phi_{X_1}(s)\phi_{X_2}(s) \cdots \phi_{X_n}(s).$$

When $X_1, \ldots, X_n$ are iid, each with MGF $\phi_{X_i}(s) = \phi_X(s)$,

$$\phi_W(s) = [\phi_X(s)]^n.$$

Proof: Theorem 9.6

From the definition of the MGF,

$$\phi_W(s) = \mathsf{E}\left[e^{s(X_1+\cdots+X_n)}\right] = \mathsf{E}\left[e^{sX_1}e^{sX_2}\cdots e^{sX_n}\right]. \tag{9.28}$$

Here, we have the expected value of a product of functions of independent random variables. Theorem 8.4 states that this expected value is the product of the individual expected values:

$$\mathsf{E}\left[g_1(X_1)g_2(X_2)\cdots g_n(X_n)\right] = \mathsf{E}\left[g_1(X_1)\right]\mathsf{E}\left[g_2(X_2)\right]\cdots\mathsf{E}\left[g_n(X_n)\right]. \tag{9.29}$$

By Equation (9.29) with $g_i(X_i) = e^{sX_i}$, the expected value of the product is

$$\phi_W(s) = \mathsf{E}\left[e^{sX_1}\right]\mathsf{E}\left[e^{sX_2}\right]\cdots\mathsf{E}\left[e^{sX_n}\right] = \phi_{X_1}(s)\phi_{X_2}(s)\cdots\phi_{X_n}(s). \tag{9.30}$$

When $X_1, \ldots, X_n$ are iid, $\phi_{X_i}(s) = \phi_X(s)$ and thus $\phi_W(s) = (\phi_W(s))^n$.

Example 9.5 Problem

J and K are independent random variables with probability mass functions

$$\begin{array}{c|ccc} j & 1 & 2 & 3 \\ \hline P_J(j) & 0.2 & 0.6 & 0.2 \end{array}, \qquad \begin{array}{c|cc} k & -1 & 1 \\ \hline P_K(k) & 0.5 & 0.5 \end{array}. \tag{9.31}$$

Find the MGF of $M = J + K$. What are $P_M(m)$ and $E[M^3]$?

Example 9.5 Solution

J and K have have moment generating functions

$$\phi_J(s) = 0.2e^s + 0.6e^{2s} + 0.2e^{3s}, \qquad \phi_K(s) = 0.5e^{-s} + 0.5e^s. \tag{9.32}$$

Therefore, by Theorem 9.6, $M = J + K$ has MGF

$$\phi_M(s) = \phi_J(s)\phi_K(s) = 0.1 + 0.3e^s + 0.2e^{2s} + 0.3e^{3s} + 0.1e^{4s}. \tag{9.33}$$

The value of $P_M(m)$ at any value of m is the coefficient of e^{ms} in $\phi_M(s)$:

$$\phi_M(s) = \mathrm{E}\left[e^{sM}\right] = \underbrace{0.1}_{P_M(0)} + \underbrace{0.3}_{P_M(1)} e^s + \underbrace{0.2}_{P_M(2)} e^{2s} + \underbrace{0.3}_{P_M(3)} e^{3s} + \underbrace{0.1}_{P_M(4)} e^{4s}.$$

From the coefficients of $\phi_M(s)$, we construct the table for the PMF of M:

m	0	1	2	3	4
$P_M(m)$	0.1	0.3	0.2	0.3	0.1

To find the third moment of M, we differentiate $\phi_M(s)$ three times:

$$\begin{aligned} \mathrm{E}\left[M^3\right] &= \left.\frac{d^3\phi_M(s)}{ds^3}\right|_{s=0} \\ &= 0.3e^s + 0.2(2^3)e^{2s} + 0.3(3^3)e^{3s} + 0.1(4^3)e^{4s}\Big|_{s=0} = 16.4. \end{aligned} \tag{9.34}$$

Theorem 9.7

If $K_1, \ldots, K_n$ are independent Poisson random variables, $W = K_1 + \cdots + K_n$ is a Poisson random variable.

Proof: Theorem 9.7

We adopt the notation $\text{E}[K_i] = \alpha_i$ and note in Table 9.1 that K_i has MGF

$$\phi_{K_i}(s) = e^{\alpha_i(e^s-1)}. \tag{9.35}$$

By Theorem 9.6,

$$\begin{aligned} \phi_W(s) &= e^{\alpha_1(e^s-1)} e^{\alpha_2(e^s-1)} \cdots e^{\alpha_n(e^s-1)} \\ &= e^{(\alpha_1+\cdots+\alpha_n)(e^s-1)} \\ &= e^{(\alpha_T)(e^s-1)} \end{aligned} \tag{9.36}$$

where $\alpha_T = \alpha_1 + \cdots + \alpha_n$. Examining Table 9.1, we observe that $\phi_W(s)$ is the moment generating function of the Poisson (α_T) random variable. Therefore,

$$P_W(w) = \begin{cases} \alpha_T^w e^{-\alpha}/w! & w = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases} \tag{9.37}$$

Theorem 9.8

The sum of n independent Gaussian random variables $W = X_1 + \cdots + X_n$ is a Gaussian random variable.

Proof: Theorem 9.8

For convenience, let $\mu_i = \mathsf{E}[X_i]$ and $\sigma_i^2 = \mathsf{Var}[X_i]$. Since the X_i are independent, we know that

$$\begin{aligned}\phi_W(s) &= \phi_{X_1}(s)\phi_{X_2}(s)\cdots\phi_{X_n}(s) \\ &= e^{s\mu_1+\sigma_1^2 s^2/2}e^{s\mu_2+\sigma_2^2 s^2/2}\cdots e^{s\mu_n+\sigma_n^2 s^2/2} \\ &= e^{s(\mu_1+\cdots+\mu_n)+(\sigma_1^2+\cdots+\sigma_n^2)s^2/2}. \qquad (9.38)\end{aligned}$$

From Equation (9.38), we observe that $\phi_W(s)$ is the moment generating function of a Gaussian random variable with expected value $\mu_1 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$.

Theorem 9.9

If $X_1, \ldots, X_n$ are iid exponential (λ) random variables, then $W = X_1 + \cdots + X_n$ has the Erlang PDF

$$f_W(w) = \begin{cases} \frac{\lambda^n w^{n-1} e^{-\lambda w}}{(n-1)!} & w \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: Theorem 9.9

In Table 9.1 we observe that each X_i has MGF $\phi_X(s) = \lambda/(\lambda - s)$. By Theorem 9.6, W has MGF

$$\phi_W(s) = \left(\frac{\lambda}{\lambda - s}\right)^n. \tag{9.39}$$

Returning to Table 9.1, we see that W has the MGF of an Erlang (n, λ) random variable.

Quiz 9.3(A)

Let $K_1, K_2, \ldots, K_m$ be iid discrete uniform random variables with PMF

$$P_K(k) = \begin{cases} 1/n & k = 1, 2, \ldots, n, \\ 0 & \text{otherwise.} \end{cases} \tag{9.40}$$

Find the MGF of $J = K_1 + \cdots + K_m$.

Quiz 9.3(A) Solution

Each K_i has MGF

$$\phi_K(s) = \mathsf{E}\left[e^{sK_i}\right] = \frac{e^s + e^{2s} + \cdots + e^{ns}}{n} = \frac{e^s(1 - e^{ns}}{n(1 - e^s)}. \tag{1}$$

Since the sequence of K_i is independent, Theorem 9.6 says the MGF of J is

$$\phi_J(s) = (\phi_K(s))^m = \frac{e^{ms}(1 - e^{ns})^m}{n^m(1 - e^s)^m}. \tag{2}$$

Quiz 9.3(B)

Let $X_1, \ldots, X_n$ be independent Gaussian random variables with E$[X_i = 0]$ and Var$[X_i] = i$. Find the PDF of

$$W = \alpha X_1 + \alpha^2 X_2 + \cdots + \alpha^n X_n. \tag{9.41}$$

Quiz 9.3(B) Solution

Since the set of $\alpha^j X_j$ are independent Gaussian random variables, Theorem 9.8 says that W is a Gaussian random variable. Thus to find the PDF of W, we need only find the expected value and variance. Since the expectation of the sum equals the sum of the expectations:

$$E[W] = \alpha E[X_1] + \alpha^2 E[X_2] + \cdots + \alpha^n E[X_n] = 0. \tag{1}$$

Since the $\alpha^j X_j$ are independent, the variance of the sum equals the sum of the variances:

$$\begin{aligned} \text{Var}[W] &= \alpha^2 \text{Var}[X_1] + \alpha^4 \text{Var}[X_2] + \cdots + \alpha^{2n} \text{Var}[X_n] \\ &= \alpha^2 + 2(\alpha^2)^2 + \cdots + n(\alpha^2)^n. \end{aligned} \tag{2}$$

Defining $q = \alpha^2$, we can use Math Fact B.6 to write

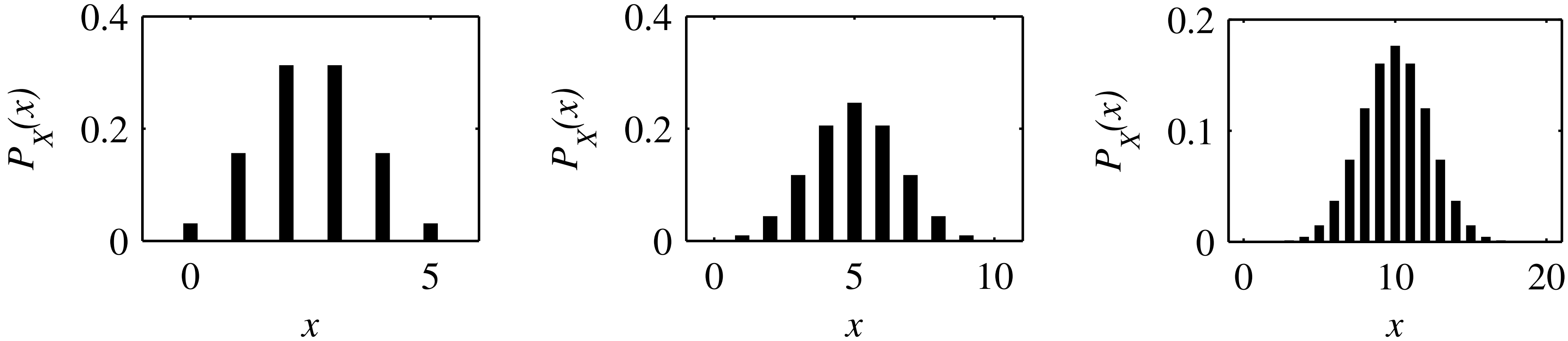
$$\text{Var}[W] = \frac{\alpha^2 - \alpha^{2n+2}[1 + n(1 - \alpha^2)]}{(1 - \alpha^2)^2}. \tag{3}$$

With $E[W] = 0$ and $\sigma_W^2 = \text{Var}[W]$, we can write the PDF of W as

$$f_W(w) = \frac{1}{\sqrt{2\pi\sigma_W^2}} e^{-w^2/2\sigma_W^2}. \tag{4}$$

Central Limit Theorem

Figure 9.1



$n = 5$ $n = 10$ $n = 20$

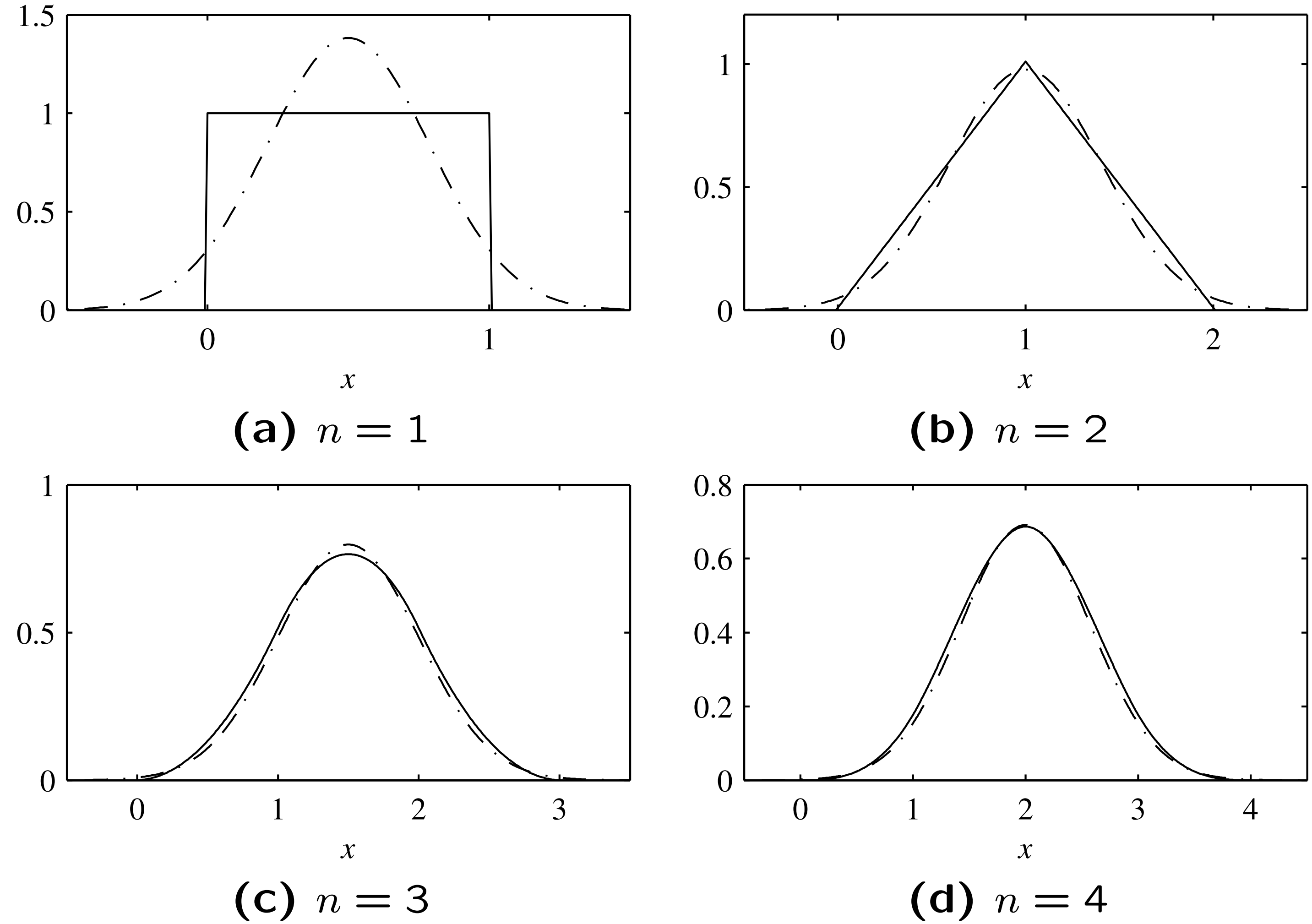
The PMF of the X, the number of heads in n coin flips for $n = 5, 10, 20$. As n increases, the PMF more closely resembles a bell-shaped curve.

Theorem 9.10 Central Limit Theorem

Given $X_1, X_2, \ldots$, a sequence of iid random variables with expected value μ_X and variance σ_X^2, the CDF of $Z_n = (\sum_{i=1}^{n} X_i - n\mu_X)/\sqrt{n\sigma_X^2}$ has the property

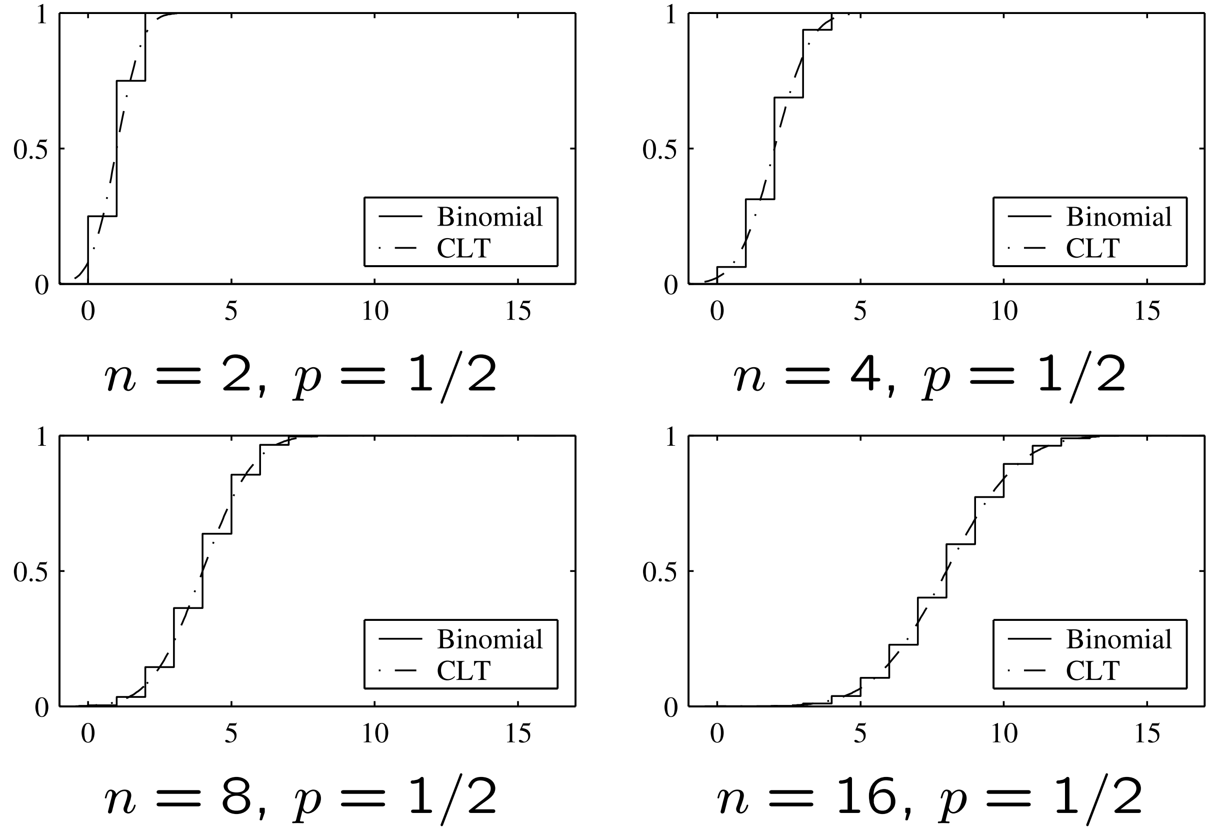
$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z).$$

Figure 9.2



The PDF of W_n, the sum of n uniform $(0, 1)$ random variables, and the corresponding central limit theorem approximation for $n = 1, 2, 3, 4$. The solid — line denotes the PDF $f_{W_n}(w)$, and the broken $- \cdot -$ line denotes the Gaussian approximation.

Figure 9.3



The binomial (n, p) CDF and the corresponding central limit theorem approximation for $n = 4, 8, 16, 32$, and $p = 1/2$.

Definition 9.2 Approximation

Let $W_n = X_1 + \cdots + X_n$ be the sum of n iid random variables, each with $\mathsf{E}[X] = \mu_X$ and $\mathsf{Var}[X] = \sigma_X^2$. The central limit theorem approximation to the CDF of W_n is

$$F_{W_n}(w) \approx \Phi\left(\frac{w - n\mu_X}{\sqrt{n\sigma_X^2}}\right).$$

Example 9.6

To gain some intuition into the central limit theorem, consider a sequence of iid continuous random variables X_i, where each random variable is uniform (0,1). Let

$$W_n = X_1 + \cdots + X_n. \tag{9.46}$$

Recall that E$[X] = 0.5$ and Var$[X] = 1/12$. Therefore, W_n has expected value E$[W_n] = n/2$ and variance $n/12$. The central limit theorem says that *the CDF* of W_n should approach a Gaussian CDF with the same expected value and variance. Moreover, since W_n is a continuous random variable, we would also expect that the PDF of W_n would converge to a Gaussian PDF. In Figure 9.2, we compare the PDF of W_n to the PDF of a Gaussian random variable with the same expected value and variance. First, W_1 is a uniform random variable with the rectangular PDF shown in Figure 9.2(a). This figure also shows the PDF of W_1, a Gaussian random variable with expected value $\mu = 0.5$ and variance $\sigma^2 = 1/12$. Here the PDFs are very dissimilar. When we consider $n = 2$, we have the situation in Figure 9.2(b). The PDF of W_2 is a triangle with expected value 1 and variance 2/12. The figure shows the corresponding Gaussian PDF. The following figures show the PDFs of $W_3, \ldots, W_6$. The convergence to a bell shape is apparent.

Example 9.7

Now suppose $W_n = X_1 + \cdots + X_n$ is a sum of independent Bernoulli (p) random variables. We know that W_n has the binomial PMF

$$P_{W_n}(w) = \binom{n}{w} p^w (1-p)^{n-w}. \tag{9.47}$$

No matter how large n becomes, W_n is always a discrete random variable and would have a PDF consisting of impulses. However, the central limit theorem says that the CDF of W_n converges to a Gaussian CDF. Figure 9.3 demonstrates the convergence of the sequence of binomial CDFs to a Gaussian CDF for $p = 1/2$ and four values of n, the number of Bernoulli random variables that are added to produce a binomial random variable. For $n \geq 32$, Figure 9.3 suggests that approximations based on the Gaussian distribution are very accurate.

Example 9.8 Problem

A compact disc (CD) contains digitized samples of an acoustic waveform. In a CD player with a "one bit digital to analog converter," each digital sample is represented to an accuracy of ± 0.5 mV. The CD player oversamples the waveform by making eight independent measurements corresponding to each sample. The CD player obtains a waveform sample by calculating the average (sample mean) of the eight measurements. What is the probability that the error in the waveform sample is greater than 0.1 mV?

Example 9.8 Solution

The measurements $X_1, X_2, \ldots, X_8$ all have a uniform distribution between $v - 0.5$ mV and $v + 0.5$ mV, where v mV is the exact value of the waveform sample. The compact disk player produces the output $U = W_8/8$, where

$$W_8 = \sum_{i=1}^{8} X_i. \tag{9.48}$$

To find $\mathrm{P}[|U - v| > 0.1]$ exactly, we would have to find an exact probability model for W_8, either by computing an eightfold convolution of the uniform PDF of X_i or by using the moment generating function. Either way, the process is extremely complex. Alternatively, we can use the central limit theorem to model W_8 as a Gaussian random variable with $\mathrm{E}[W_8] = 8\mu_X = 8v$ mV and variance $\mathrm{Var}[W_8] = 8\,\mathrm{Var}[X] = 8/12$. Therefore, U is approximately Gaussian with $\mathrm{E}[U] = \mathrm{E}[W_8]/8 = v$ and variance $\mathrm{Var}[W_8]/64 = 1/96$. Finally, the error, $U - v$ in the output waveform sample is approximately Gaussian with expected value 0 and variance 1/96. It follows that

$$\mathrm{P}\left[|U - v| > 0.1\right] = 2\left[1 - \Phi\left(0.1/\sqrt{1/96}\right)\right] = 0.3272. \tag{9.49}$$

Example 9.9 Problem

Transmit one million bits. Let A denote the event that there are at least 499,000 ones but no more than 501,000 ones. What is $\mathrm{P}[A]$?

Example 9.9 Solution

Let X_i be the value of bit i (either 0 or 1). The number of ones in one million bits is $W = \sum_{i=1}^{10^6} X_i$. Because X_i is a Bernoulli (0.5) random variable, $\mathsf{E}[X_i] = 0.5$ and $\mathsf{Var}[X_i] = 0.25$ for all i. Note that $\mathsf{E}[W] = 10^6\,\mathsf{E}[X_i] = 500{,}000$, and $\mathsf{Var}[W] = 10^6\,\mathsf{Var}[X_i] = 250{,}000$. Therefore, $\sigma_W = 500$. By the central limit theorem approximation,

$$\begin{aligned} \mathsf{P}[A] &= \mathsf{P}[W \leq 501{,}000] - \mathsf{P}[W < 499{,}000] \\ &\approx \Phi\left(\frac{501{,}000 - 500{,}000}{500}\right) - \Phi\left(\frac{499{,}000 - 500{,}000}{500}\right) \\ &= \Phi(2) - \Phi(-2) = 0.9544. \end{aligned} \tag{9.50}$$

Quiz 9.4

X milliseconds, the total access time (waiting time + read time) to get one block of information from a computer disk, is the continuous (0,12) random variable. Before performing a certain task, the computer must access 12 different blocks of information from the disk. (Access times for different blocks are independent of one another.) The total access time for all the information is a random variable A milliseconds.

(a) Find the expected value and variance of the access time X.

(b) Find the expected value and standard deviation of the total access time A.

(c) Use the central limit theorem to estimate P[$A > 75$ ms].

(d) Use the central limit theorem to estimate P[$A < 48$ ms].

Quiz 9.4 Solution

(a) The expected access time is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^{12} \frac{x}{12}\, dx = 6 \text{ ms.} \tag{1}$$

(b) The second moment of the access time is

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^{12} \frac{x^2}{12}\, dx = 48. \tag{2}$$

The variance of the access time is $\text{Var}[X] = E[X^2] - (E[X])^2 = 12$.

(c) Using X_i to denote the access time of block i, we can write

$$A = X_1 + X_2 + \cdots + X_{12} \tag{3}$$

Since the expectation of the sum equals the sum of the expectations,

$$E[A] = E[X_1] + \cdots + E[X_{12}] = 12\, E[X] = 72 \text{ ms.} \tag{4}$$

[Continued]

(d) Since the X_i are independent,

$$\text{Var}[A] = \text{Var}[X_1] + \cdots + \text{Var}[X_{12}] = 12\,\text{Var}[X] = 144. \tag{5}$$

Thus A has standard deviation $\sigma_A = 12$.

(e) To use the central limit theorem, we use Table 4.1 to evaluate

$$\begin{aligned} P[A \leq 75] &= P\left[\frac{A - E[A]}{\sigma_A} \leq \frac{75 - E[A]}{\sigma_A}\right] \\ &\approx \Phi\left(\frac{75 - 72}{12}\right) = 0.5987. \end{aligned} \tag{6}$$

Then $P[A > 75] = 1 - P[A \leq 75] = 0.4013$.

(f) Once again, we use the central limit theorem and Table 4.1 to estimate

$$\begin{aligned} P[A < 48] &= P\left[\frac{A - E[A]}{\sigma_A} < \frac{48 - E[A]}{\sigma_A}\right] \\ &\approx \Phi\left(\frac{48 - 72}{12}\right) = 0.0227. \end{aligned} \tag{7}$$

Section 9.5

Matlab

Example 9.10 Problem

X_1 and X_2 are independent discrete random variables with PMFs

$$P_{X_1}(x) = \begin{cases} 0.04 & x = 1, \ldots, 25, \\ 0 & \text{otherwise}, \end{cases} \quad P_{X_2}(x) = \begin{cases} \frac{x}{550} & x = 10, 20, \ldots, 100, \\ 0 & \text{otherwise}. \end{cases}$$

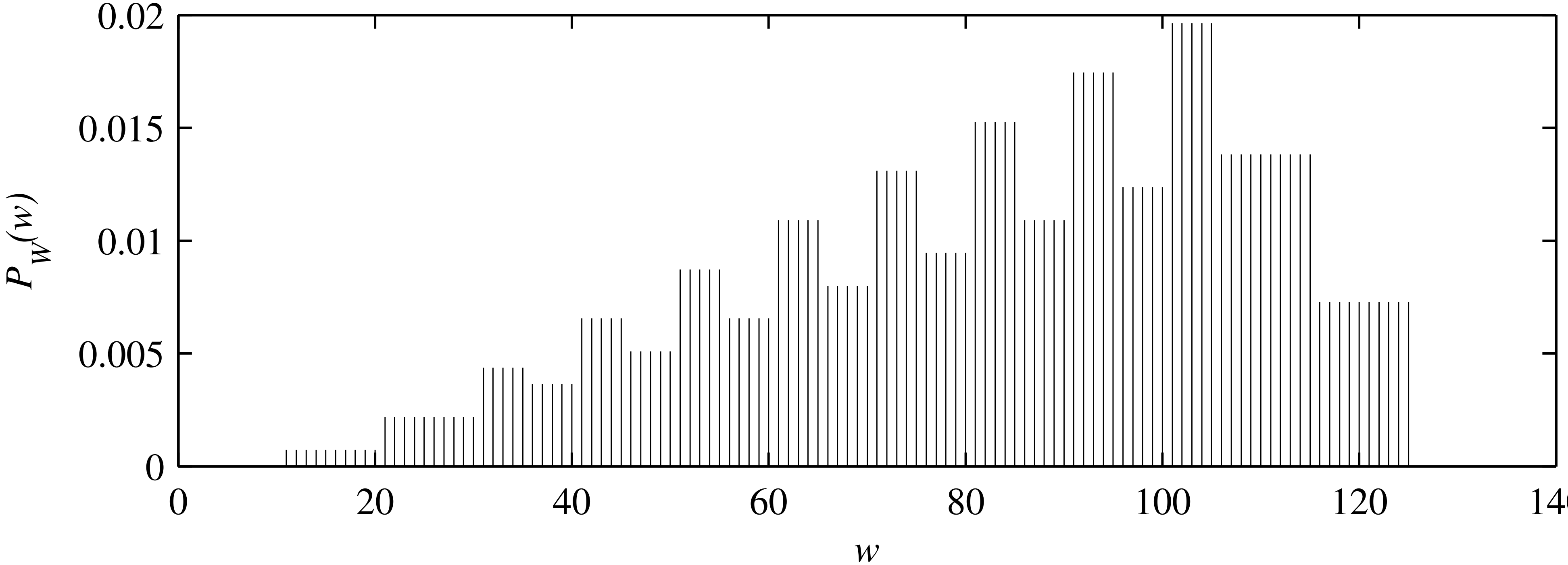
What is the PMF of $W = X_1 + X_2$?

Example 9.10 Solution

```
%sumx1x2.m
sx1=(1:25);px1=0.04*ones(1,25);
sx2=10*(1:10);px2=sx2/550;
[SX1,SX2]=ndgrid(sx1,sx2);
[PX1,PX2]=ndgrid(px1,px2);
SW=SX1+SX2;PW=PX1.*PX2;
sw=unique(SW);
pw=finitepmf(SW,PW,sw);
pmfplot(sw,pw);
```

As in Example 5.25, `sumx1x2.m` uses `ndgrid` to generate a grid for all possible pairs of X_1 and X_2. The matrix `SW` holds the sum $x_1 + x_2$ for each possible pair x_1, x_2. The probability $P_{X_1,X_2}(x_1, x_2)$ of each such pair is in the matrix `PW`. For each unique w generated by pairs x_1+x_2, `finitepmf` finds the probability $P_W(w)$. The graph of $P_W(w)$ appears in Figure 9.4.

Figure 9.4



The PMF $P_W(w)$ for Example 9.10.

Figure 9.5

```
>> uniform12(10000);
ans =
 -3.0000 -2.0000 -1.0000       0  1.0000  2.0000  3.0000
  0.0013  0.0228  0.1587  0.5000  0.8413  0.9772  0.9987
  0.0005  0.0203  0.1605  0.5027  0.8393  0.9781  0.9986
>> uniform12(10000);
ans =
 -3.0000 -2.0000 -1.0000       0  1.0000  2.0000  3.0000
  0.0013  0.0228  0.1587  0.5000  0.8413  0.9772  0.9987
  0.0015  0.0237  0.1697  0.5064  0.8400  0.9778  0.9993
```

Two sample runs of `uniform12.m`.

Example 9.11 Problem

Write a Matlab program to generate $m = 10{,}000$ samples of the random variable $X = \sum_{i=1}^{12} U_i - 6$. Use the data to find the relative frequencies of the following events $\{X \leq T\}$ for $T = -3, -2 \ldots, 3$. Calculate the probabilities of these events when X is a Gaussian $(0, 1)$ random variable.

Example 9.11 Solution

```
function FX=uniform12(m);
x=sum(rand(12,m))-6;
T=(-3:3);FX=(count(x,T)/m)';
[T;phi(T);FX]
```

In `uniform12(m)`, `x` holds the m samples of X. The function `n=count(x,T)` returns `n(i)` as the number of elements of `x` less than or equal to `T(i)`. The output is a three-row table: T on the first row, the true probabilities $P[X \leq T] = \Phi(T)$ second, and the relative frequencies third. Two sample runs of `uniform12` are shown in Figure 9.5. We see that the relative frequencies and the probabilities diverge as T moves farther from zero. In fact this program will never produce a value of $|X| > 6$, no matter how many times it runs. By contrast, $Q(6) = 9.9 \times 10^{-10}$. This suggests that in a set of one billion independent samples of the Gaussian (0, 1) random variable, we can expect two samples with $|X| > 6$, one sample with $X < -6$, and one sample with $X > 6$.

Quiz 9.5

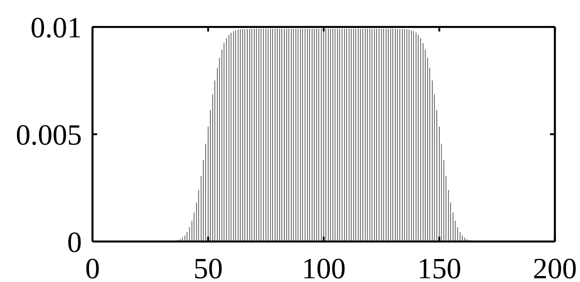
X is the binomial $(100, 0.5)$ random variable and Y is the discrete uniform $(0, 100)$ random variable. Calculate and graph the PMF of $W = X + Y$.

Quiz 9.5 Solution

One solution to this problem is to follow the approach of Example 9.10:

```
%unifbinom100.m
sx=0:100;sy=0:100;
px=binomialpmf(100,0.5,sx);
py=duniformpmf(0,100,sy);
[SX,SY]=ndgrid(sx,sy);
[PX,PY]=ndgrid(px,py);
SW=SX+SY; PW=PX.*PY;
sw=unique(SW);
pw=finitepmf(SW,PW,sw);
pmfplot(sw,pw);
```

Here is a graph of the PMF $P_W(w)$:



With some thought, it should be apparent that the `finitepmf` function is implementing the convolution of the two PMFs.