

CHAPTER 10

Orthogonal Polynomials

The subject of orthogonal polynomials can be traced back to the work of the French mathematician Adrien-Marie Legendre (1752–1833) on planetary motion. These polynomials have important applications in physics, quantum mechanics, mathematical statistics, and other areas in mathematics.

This chapter provides an exposition of the properties of orthogonal polynomials. Emphasis will be placed on Legendre, Chebyshev, Jacobi, Laguerre, and Hermite polynomials. In addition, applications of these polynomials in statistics will be discussed in Section 10.9.

10.1. INTRODUCTION

Suppose that $f(x)$ and $g(x)$ are two continuous functions on $[a, b]$. Let $w(x)$ be a positive function that is Riemann integrable on $[a, b]$. The dot product of $f(x)$ and $g(x)$ with respect to $w(x)$, which is denoted by $(f \cdot g)_\omega$, is defined as

$$(f \cdot g)_\omega = \int_a^b f(x)g(x)w(x)dx.$$

The norm of $f(x)$ with respect to $w(x)$, denoted by $\|f\|_\omega$, is defined as $\|f\|_\omega = [\int_a^b f^2(x)w(x)dx]^{1/2}$. The functions $f(x)$ and $g(x)$ are said to be orthogonal [with respect to $w(x)$] if $(f \cdot g)_\omega = 0$. Furthermore, a sequence $\{f_n(x)\}_{n=0}^\infty$ of continuous functions defined on $[a, b]$ are said to be orthogonal with respect to $w(x)$ if $(f_m \cdot f_n)_\omega = 0$ for $m \neq n$. If, in addition, $\|f_n\|_\omega = 1$ for all n , then the functions $f_n(x)$, $n = 0, 1, 2, \dots$, are called orthonormal. In particular, if $S = \{p_n(x)\}_{n=0}^\infty$ is a sequence of polynomials such that $(p_n \cdot p_m)_\omega = 0$ for all $m \neq n$, then S forms a sequence of polynomials orthogonal with respect to $w(x)$.

A sequence of orthogonal polynomials can be constructed on the basis of the following theorem:

Theorem 10.1.1. The polynomials $\{p_n(x)\}_{n=0}^{\infty}$ which are defined according to the following recurrence relation are orthogonal:

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= x - \frac{(xp_0 \cdot p_0)_{\omega}}{\|p_0\|_{\omega}^2} = x - \frac{(x \cdot 1)_{\omega}}{\|1\|_{\omega}^2}, \\ &\vdots \\ p_n(x) &= (x - a_n)p_{n-1}(x) - b_n p_{n-2}(x), \quad n = 2, 3, \dots, \end{aligned} \tag{10.1}$$

where

$$a_n = \frac{(xp_{n-1} \cdot p_{n-1})_{\omega}}{\|p_{n-1}\|_{\omega}^2} \tag{10.2}$$

$$b_n = \frac{(xp_{n-1} \cdot p_{n-2})_{\omega}}{\|p_{n-2}\|_{\omega}^2} \tag{10.3}$$

Proof. We show by mathematical induction on n that $(p_n \cdot p_i)_{\omega} = 0$ for $i < n$. For $n = 1$,

$$\begin{aligned} (p_1 \cdot p_0)_{\omega} &= \int_a^b \left[x - \frac{(x \cdot 1)_{\omega}}{\|1\|_{\omega}^2} \right] w(x) dx \\ &= (x \cdot 1)_{\omega} - (x \cdot 1)_{\omega} \frac{\|1\|_{\omega}^2}{\|1\|_{\omega}^2} \\ &= 0. \end{aligned}$$

Now, suppose that the assertion is true for $n - 1$ ($n \geq 2$). To show that it is true for n . We have that

$$\begin{aligned} (p_n \cdot p_i)_{\omega} &= \int_a^b [(x - a_n)p_{n-1}(x) - b_n p_{n-2}(x)] p_i(x) w(x) dx \\ &= (xp_{n-1} \cdot p_i)_{\omega} - a_n (p_{n-1} \cdot p_i)_{\omega} - b_n (p_{n-2} \cdot p_i)_{\omega}. \end{aligned}$$

Thus, for $i = n - 1$,

$$\begin{aligned} (p_n \cdot p_i)_{\omega} &= (xp_{n-1} \cdot p_{n-1})_{\omega} - a_n \|p_{n-1}\|_{\omega}^2 - b_n (p_{n-2} \cdot p_{n-1})_{\omega} \\ &= 0, \end{aligned}$$

by the definition of a_n in (10.2) and the fact that $(p_{n-2} \cdot p_{n-1})_\omega = (p_{n-1} \cdot p_{n-2})_\omega = 0$. Similarly, for $i = n - 2$,

$$\begin{aligned} (p_n \cdot p_i)_\omega &= (xp_{n-1} \cdot p_{n-2})_\omega - a_n(p_{n-1} \cdot p_{n-2})_\omega - b_n(p_{n-2} \cdot p_{n-2})_\omega \\ &= (xp_{n-1} \cdot p_{n-2})_\omega - b_n \|p_{n-2}\|_\omega^2 \\ &= 0, \quad \text{by (10.3).} \end{aligned}$$

Finally, for $i < n - 2$, we have

$$\begin{aligned} (p_n \cdot p_i)_\omega &= (xp_{n-1} \cdot p_i)_\omega - a_n(p_{n-1} \cdot p_i)_\omega - b_n(p_{n-2} \cdot p_i)_\omega \\ &= \int_a^b xp_{n-1}(x)p_i(x)w(x)dx. \end{aligned} \quad (10.4)$$

But, from the recurrence relation,

$$p_{i+1}(x) = (x - a_{i+1})p_i(x) - b_{i+1}p_{i-1}(x),$$

that is,

$$xp_i(x) = p_{i+1}(x) + a_{i+1}p_i(x) + b_{i+1}p_{i-1}(x).$$

It follows that

$$\begin{aligned} &\int_a^b xp_{n-1}(x)p_i(x)w(x)dx \\ &= \int_a^b p_{n-1}(x)[p_{i+1}(x) + a_{i+1}p_i(x) + b_{i+1}p_{i-1}(x)]w(x)dx \\ &= (p_{n-1} \cdot p_{i+1})_\omega + a_{i+1}(p_{n-1} \cdot p_i)_\omega + b_{i+1}(p_{n-1} \cdot p_{i-1})_\omega \\ &= 0. \end{aligned}$$

Hence, by (10.4), $(p_n \cdot p_i)_\omega = 0$. $\square$

It is easy to see from the recurrence relation (10.1) that $p_n(x)$ is of degree n , and the coefficient of x^n is equal to one. Furthermore, we have the following corollaries:

Corollary 10.1.1. An arbitrary polynomial of degree $\leq n$ is uniquely expressible as a linear combination of $p_0(x), p_1(x), \dots, p_n(x)$.

Corollary 10.1.2. The coefficient of x^{n-1} in $p_n(x)$ is $-\sum_{i=1}^n a_i$ ($n \geq 1$).

Proof. If d_n denotes the coefficient of x^{n-1} in $p_n(x)$ ($n \geq 2$), then by comparing the coefficients of x^{n-1} on both sides of the recurrence relation

(10.1), we obtain

$$d_n = d_{n-1} - a_n, \quad n = 2, 3, \dots \quad (10.5)$$

The result follows from (10.5) and by noting that $d_1 = -a_1$ $\square$

Another property of orthogonal polynomials is given by the following theorem:

Theorem 10.1.2. If $\{p_n(x)\}_{n=0}^\infty$ is a sequence of orthogonal polynomials with respect to $w(x)$ on $[a, b]$, then the zeros of $p_n(x)$ ($n \geq 1$) are all real, distinct, and located in the interior of $[a, b]$.

Proof. Since $(p_n \cdot p_0)_\omega = 0$ for $n \geq 1$, then $\int_a^b p_n(x)w(x)dx = 0$. This indicates that $p_n(x)$ must change sign at least once in (a, b) [recall that $w(x)$ is positive]. Suppose that $p_n(x)$ changes sign between a and b at just k points, denoted by $x_1, x_2, \dots, x_k$. Let $g(x) = (x - x_1)(x - x_2) \cdots (x - x_k)$. Then, $p_n(x)g(x)$ is a polynomial with no zeros of odd multiplicity in (a, b) . Hence, $\int_a^b p_n(x)g(x)w(x)dx \neq 0$, that is, $(p_n \cdot g)_\omega \neq 0$. If $k < n$, then we have a contradiction by the fact that p_n is orthogonal to $g(x)$ [$g(x)$, being a polynomial of degree k , can be expressed as a linear combination of $p_0(x), p_1(x), \dots, p_k(x)$ by Corollary 10.1.1]. Consequently, $k = n$, and $p_n(x)$ has n distinct zeros in the interior of $[a, b]$. $\square$

Particular orthogonal polynomials can be derived depending on the choice of the interval $[a, b]$, and the weight function $w(x)$. For example, the well-known orthogonal polynomials listed below are obtained by the following selections of $[a, b]$ and $w(x)$:

Orthogonal Polynomial	a	b	$w(x)$
Legendre	-1	1	1
Jacobi	-1	1	$(1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1$
Chebyshev of the first kind	-1	1	$(1-x^2)^{-1/2}$
Chebyshev of the second kind	-1	1	$(1-x^2)^{1/2}$
Hermite	$-\infty$	∞	$e^{-x^2/2}$
Laguerre	0	∞	$e^{-x}x^\alpha$, $\alpha > -1$

These polynomials are called classical orthogonal polynomials. We shall study their properties and methods of derivation.

10.2. LEGENDRE POLYNOMIALS

These polynomials are derived by applying the so-called *Rodrigues formula*

$$p_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n}, \quad n = 0, 1, 2, \dots$$

Thus, for $n = 0, 1, 2, 3, 4$, for example, we have

$$\begin{aligned} p_0(x) &= 1, \\ p_1(x) &= x, \\ p_2(x) &= \frac{3}{2}x^2 - \frac{1}{2}, \\ p_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x, \\ p_4(x) &= \frac{35}{8}x^4 - \frac{30}{8}x^2 + \frac{3}{8}. \end{aligned}$$

From the Rodrigues formula it follows that $p_n(x)$ is a polynomial of degree n and the coefficient of x^n is $\binom{2n}{n}/2^n$. We can multiply $p_n(x)$ by $2^n/\binom{2n}{n}$ to make the coefficient of x^n equal to one ($n = 1, 2, \dots$).

Another definition of Legendre polynomials is obtained by means of the generating function,

$$g(x, r) = \frac{1}{(1 - 2rx + r^2)^{1/2}},$$

by expanding it as a power series in r for sufficiently small values of r . The coefficient of r^n in this expansion is $p_n(x)$, $n = 0, 1, \dots$, that is,

$$g(x, r) = \sum_{n=0}^{\infty} p_n(x) r^n.$$

To demonstrate this, let us consider expanding $(1 - z)^{-1/2}$ in a neighborhood of zero, where $z = 2xr - r^2$:

$$\begin{aligned} \frac{1}{(1 - z)^{1/2}} &= 1 + \frac{1}{2}z + \frac{3}{8}z^2 + \frac{5}{16}z^3 + \dots, \quad |z| < 1, \\ &= 1 + \frac{1}{2}(2xr - r^2) + \frac{3}{8}(2xr - r^2)^2 + \frac{5}{16}(2xr - r^2)^3 + \dots \\ &= 1 + xr + \left(\frac{3}{2}x^2 - \frac{1}{2}\right)r^2 + \left(\frac{5}{2}x^3 - \frac{3}{2}x\right)r^3 + \dots. \end{aligned}$$

We note that the coefficients of $1, r, r^2$, and r^3 are the same as $p_0(x), p_1(x), p_2(x), p_3(x)$, as was seen earlier. In general, it is easy to see that the coefficient of r^n is $p_n(x)$ ($n = 0, 1, 2, \dots$).

By differentiating $g(x, r)$ with respect to r , it can be seen that

$$(1 - 2rx + r^2) \frac{\partial g(x, r)}{\partial r} - (x - r)g(x, r) = 0.$$

By substituting $g(x, r) = \sum_{n=0}^{\infty} p_n(x)r^n$ in this equation, we obtain

$$(1 - 2rx + r^2) \sum_{n=1}^{\infty} np_n(x)r^{n-1} - (x - r) \sum_{n=0}^{\infty} p_n(x)r^n = 0.$$

The coefficient of r^n must be zero for each n and for all values of x ($n = 1, 2, \dots$). We thus have the following identity:

$$(n+1)p_{n+1}(x) - (2n+1)xp_n(x) + np_{n-1}(x) = 0, \quad n = 1, 2, \dots \quad (10.6)$$

This is a recurrence relation that connects any three successive Legendre polynomials. For example, for $p_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$, $p_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$, we find from (10.6) that

$$\begin{aligned} p_4(x) &= \frac{1}{4}[7xp_3(x) - 3p_2(x)] \\ &= \frac{35}{8}x^4 - \frac{30}{8}x^2 + \frac{3}{8}. \end{aligned}$$

10.2.1. Expansion of a Function Using Legendre Polynomials

Suppose that $f(x)$ is a function defined on $[-1, 1]$ such that $\int_{-1}^1 f(x)p_n(x) dx$ exists for $n = 0, 1, 2, \dots$. Consider the series expansion

$$f(x) = \sum_{i=0}^{\infty} a_i p_i(x). \quad (10.7)$$

Multiplying both sides of (10.7) by $p_n(x)$ and then integrating from -1 to 1 , we obtain, by the orthogonality of Legendre polynomials,

$$a_n = \left[\int_{-1}^1 p_n^2(x) dx \right]^{-1} \int_{-1}^1 f(x)p_n(x) dx, \quad n = 0, 1, \dots$$

It can be shown that (see Jackson, 1941, page 52)

$$\int_{-1}^1 p_n^2(x) dx = \frac{2}{2n+1}, \quad n = 0, 1, 2, \dots$$

Hence, the coefficient of $p_n(x)$ in (10.7) is given by

$$a_n = \frac{2n+1}{2} \int_{-1}^1 f(x)p_n(x) dx. \quad (10.8)$$

If $s_n(x)$ denotes the partial sum $\sum_{i=0}^n a_i p_i(x)$ of the series in (10.7), then

$$s_n(x) = \frac{n+1}{2} \int_{-1}^1 \frac{f(t)}{t-x} [p_{n+1}(t)p_n(x) - p_n(t)p_{n+1}(x)] dt, \\ n = 0, 1, 2, \dots \quad (10.9)$$

This is known as *Christoffel's identity* (see Jackson, 1941, page 55). If $f(x)$ is continuous on $[-1, 1]$ and has a derivative at $x = x_0$, then $\lim_{n \rightarrow \infty} s_n(x_0) = f(x_0)$, and hence the series in (10.7) converges at x_0 to the value $f(x_0)$ (see Jackson, 1941, pages 64–65).

10.3. JACOBI POLYNOMIALS

Jacobi polynomials, named after the German mathematician Karl Gustav Jacobi (1804–1851), are orthogonal on $[-1, 1]$ with respect to the weight function $w(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha > -1$, $\beta > -1$. The restrictions on α and β are needed to guarantee integrability of $w(x)$ over the interval $[-1, 1]$. These polynomials, which we denote by $p_n^{(\alpha, \beta)}(x)$, can be derived by applying the Rodrigues formula:

$$p_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n [(1-x)^{\alpha+n} (1+x)^{\beta+n}]}{dx^n}, \\ n = 0, 1, 2, \dots \quad (10.10)$$

This formula reduces to the one for Legendre polynomials when $\alpha = \beta = 0$. Thus, Legendre polynomials represent a special class of Jacobi polynomials.

Applying the so-called Leibniz formula (see Exercise 4.2 in Chapter 4) concerning the n th derivative of a product of two functions, namely, $f_n(x) = (1-x)^{\alpha+n}$ and $g_n(x) = (1+x)^{\beta+n}$ in (10.10), we obtain

$$\frac{d^n [(1-x)^{\alpha+n} (1+x)^{\beta+n}]}{dx^n} = \sum_{i=0}^n \binom{n}{i} f_n^{(i)}(x) g_n^{(n-i)}(x), \quad n = 0, 1, \dots, \\ (10.11)$$

where for $i = 0, 1, \dots, n$, $f_n^{(i)}(x)$ is a constant multiple of $(1-x)^{\alpha+n-i} = (1-x)^\alpha(1-x)^{n-i}$, and $g_n^{(n-i)}(x)$ is a constant multiple of $(1+x)^{\beta+i} = (1+x)^\beta(1+x)^i$. Thus, the n th derivative in (10.11) has $(1-x)^\alpha(1+x)^\beta$ as a factor. Using formula (10.10), it can be shown that $p_n^{(\alpha, \beta)}(x)$ is a polynomial of degree n with the leading coefficient (that is, the coefficient of x^n) equal to $(1/2^n n!) \Gamma(2n + \alpha + \beta + 1) / \Gamma(n + \alpha + \beta + 1)$.

10.4. CHEBYSHEV POLYNOMIALS

These polynomials were named after the Russian mathematician Pafnuty Lvovich Chebyshev (1821–1894). In this section, two kinds of Chebyshev polynomials will be studied, called, Chebyshev polynomials of the first kind and of the second kind.

10.4.1. Chebyshev Polynomials of the First Kind

These polynomials are denoted by $T_n(x)$ and defined as

$$T_n(x) = \cos(n \operatorname{Arccos} x), \quad n = 0, 1, \dots, \quad (10.12)$$

where $0 \leq \operatorname{Arccos} x \leq \pi$. Note that $T_n(x)$ can be expressed as

$$T_n(x) = x^n + \binom{n}{2} x^{n-2} (x^2 - 1) + \binom{n}{4} x^{n-4} (x^2 - 1)^2 + \dots, \quad n = 0, 1, \dots, \quad (10.13)$$

where $-1 \leq x \leq 1$. Historically, the polynomials defined by (10.13) were originally called Chebyshev polynomials without any qualifying expression. Using (10.13), it is easy to obtain the first few of these polynomials:

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1, \\ T_5(x) &= 16x^5 - 20x^3 + 5x, \\ &\vdots \end{aligned}$$

The following are some properties of $T_n(x)$:

1. $-1 \leq T_n(x) \leq 1$ for $-1 \leq x \leq 1$.
2. $T_n(-x) = (-1)^n T_n(x)$.
3. $T_n(x)$ has simple zeros at the following n points:

$$\xi_i = \cos \left[\frac{(2i-1)\pi}{2n} \right], \quad i = 1, 2, \dots, n.$$

We may recall that these zeros, also referred to as Chebyshev points, were instrumental in minimizing the error of Lagrange interpolation in Chapter 9 (see Theorem 9.2.3).

4. The weight function for $T_n(x)$ is $w(x) = (1 - x^2)^{-1/2}$. To show this, we have that for two nonnegative integers, m, n ,

$$\int_0^\pi \cos m\theta \cos n\theta \, d\theta = 0, \quad m \neq n, \quad (10.14)$$

and

$$\int_0^\pi \cos^2 n\theta \, d\theta = \begin{cases} \pi/2, & n \neq 0, \\ \pi, & n = 0 \end{cases}. \quad (10.15)$$

Making the change of variables $x = \cos \theta$ in (10.14) and (10.15), we obtain

$$\begin{aligned} \int_{-1}^1 \frac{T_m(x)T_n(x)}{(1-x^2)^{1/2}} \, dx &= 0, \quad m \neq n, \\ \int_{-1}^1 \frac{T_n^2(x)}{(1-x^2)^{1/2}} \, dx &= \begin{cases} \pi/2, & n \neq 0, \\ \pi, & n = 0. \end{cases} \end{aligned}$$

This shows that $\{T_n(x)\}_{n=0}^\infty$ forms a sequence of orthogonal polynomials on $[-1, 1]$ with respect to $w(x) = (1 - x^2)^{-1/2}$.

5. We have

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \dots \quad (10.16)$$

To show this recurrence relation, we use the following trigonometric identities:

$$\begin{aligned} \cos[(n+1)\theta] &= \cos n\theta \cos \theta - \sin n\theta \sin \theta, \\ \cos[(n-1)\theta] &= \cos n\theta \cos \theta + \sin n\theta \sin \theta. \end{aligned}$$

Adding these identities, we obtain

$$\cos[(n+1)\theta] = 2\cos n\theta \cos \theta - \cos[(n-1)\theta]. \quad (10.17)$$

If we set $x = \cos \theta$ and $\cos n\theta = T_n(x)$ in (10.17), we obtain (10.16). Recall that $T_0(x) = 1$ and $T_1(x) = x$.

10.4.2. Chebyshev Polynomials of the Second Kind

These polynomials are defined in terms of Chebyshev polynomials of the first kind as follows: Differentiating $T_n(x) = \cos n\theta$ with respect to $x = \cos \theta$, we

obtain,

$$\begin{aligned}\frac{dT_n(x)}{dx} &= -n \sin n\theta \frac{d\theta}{dx} \\ &= n \frac{\sin n\theta}{\sin \theta}.\end{aligned}$$

Let $U_n(x)$ be defined as

$$\begin{aligned}U_n(x) &= \frac{1}{n+1} \frac{dT_{n+1}(x)}{dx} \\ &= \frac{\sin[(n+1)\theta]}{\sin \theta}, \quad n = 0, 1, \dots\end{aligned}\quad (10.18)$$

This polynomial, which is of degree n , is called a Chebyshev polynomial of the second kind. Note that

$$\begin{aligned}U_n(x) &= \frac{\sin n\theta \cos \theta + \cos n\theta \sin \theta}{\sin \theta} \\ &= xU_{n-1}(x) + T_n(x), \quad n = 1, 2, \dots,\end{aligned}\quad (10.19)$$

where $U_0(x) = 1$. Formula (10.19) provides a recurrence relation for $U_n(x)$. Another recurrence relation that is free of $T_n(x)$ can be obtained from the following identity:

$$\sin[(n+1)\theta] = 2 \sin n\theta \cos \theta - \sin[(n-1)\theta].$$

Hence,

$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x), \quad n = 2, 3, \dots\quad (10.20)$$

Using the fact that $U_0(x) = 1$, $U_1(x) = 2x$, formula (10.20) can be used to derive expressions for $U_n(x)$, $n = 2, 3, \dots$. It is easy to see that the leading coefficient of x^n in $U_n(x)$ is 2^n .

We now show that $\{U_n(x)\}_{n=0}^{\infty}$ forms a sequence of orthogonal polynomials with respect to the weight function, $w(x) = (1-x^2)^{1/2}$, over $[-1, 1]$. From the formula

$$\int_0^\pi \sin[(m+1)\theta] \sin[(n+1)\theta] d\theta = 0, \quad m \neq n,$$

we get, after making the change of variables $x = \cos \theta$,

$$\int_{-1}^1 U_m(x) U_n(x) (1-x^2)^{1/2} dx = 0, \quad m \neq n.$$

This shows that $w(x) = (1-x^2)^{1/2}$ is a weight function for the sequence $\{U_n(x)\}_{n=0}^{\infty}$. Note that $\int_{-1}^1 U_n^2(x) (1-x^2)^{1/2} dx = \pi/2$.

10.5. HERMITE POLYNOMIALS

Hermite polynomials, denoted by $\{H_n(x)\}_{n=0}^{\infty}$, were named after the French mathematician Charles Hermite (1822–1901). They are defined by the Rodrigues formula,

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n(e^{-x^2/2})}{dx^n}, \quad n = 0, 1, 2, \dots \quad (10.21)$$

From (10.21), we have

$$\frac{d^n(e^{-x^2/2})}{dx^n} = (-1)^n e^{-x^2/2} H_n(x). \quad (10.22)$$

By differentiating the two sides in (10.22), we obtain

$$\frac{d^{n+1}(e^{-x^2/2})}{dx^{n+1}} = (-1)^n \left[-xe^{-x^2/2} H_n(x) + e^{-x^2/2} \frac{dH_n(x)}{dx} \right]. \quad (10.23)$$

But, from (10.21),

$$\frac{d^{n+1}(e^{-x^2/2})}{dx^{n+1}} = (-1)^{n+1} e^{-x^2/2} H_{n+1}(x). \quad (10.24)$$

From (10.23) and (10.24) we then have

$$H_{n+1}(x) = xH_n(x) - \frac{dH_n(x)}{dx}, \quad n = 0, 1, 2, \dots,$$

which defines a recurrence relation for the sequence $\{H_n(x)\}_{n=0}^{\infty}$. Since $H_0(x) = 1$, it follows by induction, using this relation, that $H_n(x)$ is a polynomial of degree n . Its leading coefficient is equal to one.

Note that if $w(x) = e^{-x^2/2}$, then

$$\begin{aligned} w(x-t) &= \exp\left(-\frac{x^2}{2} + tx - \frac{t^2}{2}\right) \\ &= w(x) \exp\left(tx - \frac{t^2}{2}\right). \end{aligned}$$

Applying Taylor's expansion to $w(x-t)$, we obtain

$$\begin{aligned} w(x-t) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^n \frac{d^n[w(x)]}{dx^n} \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x) w(x). \end{aligned}$$

Consequently, $H_n(x)$ is the coefficient of $t^n/n!$ in the expansion of $\exp(tx - t^2/2)$. It follows that

$$H_n(x) = x^n - \frac{n^{[2]}}{2 \cdot 1!} x^{n-2} + \frac{n^{[4]}}{2^2 \cdot 2!} x^{n-4} - \frac{n^{[6]}}{2^3 \cdot 3!} x^{n-6} + \dots,$$

where $n^{[r]} = n(n-1)(n-2)\cdots(n-r+1)$. This particular representation of $H_n(x)$ is given in Kendall and Stuart (1977, page 167). For example, the first seven Hermite polynomials are

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= x, \\ H_2(x) &= x^2 - 1, \\ H_3(x) &= x^3 - 3x, \\ H_4(x) &= x^4 - 6x^2 + 3, \\ H_5(x) &= x^5 - 10x^3 + 15x, \\ H_6(x) &= x^6 - 15x^4 + 45x^2 - 15, \\ &\vdots \end{aligned}$$

Another recurrence relation that does not use the derivative of $H_n(x)$ is given by

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x), \quad n = 1, 2, \dots, \quad (10.25)$$

with $H_0(x) = 1$ and $H_1(x) = x$. To show this, we use (10.21) in (10.25):

$$\begin{aligned} & (-1)^{n+1} e^{x^2/2} \frac{d^{n+1}(e^{-x^2/2})}{dx^{n+1}} \\ &= x(-1)^n e^{x^2/2} \frac{d^n(e^{-x^2/2})}{dx^n} - n(-1)^{n-1} e^{x^2/2} \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}} \end{aligned}$$

or equivalently,

$$-\frac{d^{n+1}(e^{-x^2/2})}{dx^{n+1}} = x \frac{d^n(e^{-x^2/2})}{dx^n} + n \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}}.$$

This is true given the fact that

$$\frac{d(e^{-x^2/2})}{dx} = -xe^{-x^2/2}.$$

Hence,

$$\begin{aligned} -\frac{d^{n+1}(e^{-x^2/2})}{dx^{n+1}} &= \frac{d^n(xe^{-x^2/2})}{dx^n} \\ &= n \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}} + x \frac{d^n(e^{-x^2/2})}{dx^n}, \end{aligned} \quad (10.26)$$

which results from applying Leibniz's formula to the right-hand side of (10.26).

We now show that $\{H_n(x)\}_{n=0}^\infty$ forms a sequence of orthogonal polynomials with respect to the weight function $w(x) = e^{-x^2/2}$ over $(-\infty, \infty)$. For this purpose, let m, n be nonnegative integers, and let c be defined as

$$c = \int_{-\infty}^{\infty} e^{-x^2/2} H_m(x) H_n(x) dx. \quad (10.27)$$

Then, from (10.21) and (10.27), we have

$$c = (-1)^n \int_{-\infty}^{\infty} H_m(x) \frac{d^n(e^{-x^2/2})}{dx^n} dx.$$

Integrating by parts gives

$$\begin{aligned} c &= (-1)^n \left\{ \left[H_m(x) \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{dH_m(x)}{dx} \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}} dx \right\} \\ &= (-1)^{n+1} \int_{-\infty}^{\infty} \frac{dH_m(x)}{dx} \frac{d^{n-1}(e^{-x^2/2})}{dx^{n-1}} dx. \end{aligned} \quad (10.28)$$

Formula (10.28) is true because $H_m(x) d^{n-1}(e^{-x^2/2})/dx^{n-1}$, which is a polynomial multiplied by $e^{-x^2/2}$, has a limit equal to zero as $x \rightarrow \mp\infty$. By repeating the process of integration by parts $m-1$ more times, we obtain for $n > m$

$$c = (-1)^{m+n} \int_{-\infty}^{\infty} \frac{d^m[H_m(x)]}{dx^m} \frac{d^{n-m}(e^{-x^2/2})}{dx^{n-m}} dx. \quad (10.29)$$

Note that since $H_m(x)$ is a polynomial of degree m with a leading coefficient equal to one, $d^m[H_m(x)]/dx^m$ is a constant equal to $m!$. Furthermore, since $n > m$, then

$$\int_{-\infty}^{\infty} \frac{d^{n-m}(e^{-x^2/2})}{dx^{n-m}} dx = 0.$$

It follows that $c = 0$. We can also arrive at the same conclusion if $n < m$. Hence,

$$\int_{-\infty}^{\infty} e^{-x^2/2} H_m(x) H_n(x) dx = 0, \quad m \neq n.$$

This shows that $\{H_n(x)\}_{n=0}^{\infty}$ is a sequence of orthogonal polynomials with respect to $w(x) = e^{-x^2/2}$ over $(-\infty, \infty)$.

Note that if $m = n$ in (10.29), then

$$\begin{aligned} c &= \int_{-\infty}^{\infty} \frac{d^n[H_n(x)]}{dx^n} e^{-x^2/2} dx \\ &= n! \int_{-\infty}^{\infty} e^{-x^2/2} dx \\ &= 2n! \int_0^{\infty} e^{-x^2/2} dx \\ &= n! \sqrt{2\pi} \end{aligned}$$

By comparison with (10.27), we conclude that

$$\int_{-\infty}^{\infty} e^{-x^2/2} H_n^2(x) dx = n! \sqrt{2\pi}. \quad (10.30)$$

Hermite polynomials can be used to provide the following series expansion of a function $f(x)$:

$$f(x) = \sum_{n=0}^{\infty} c_n H_n(x), \quad (10.31)$$

where

$$c_n = \frac{1}{n! \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} f(x) H_n(x) dx, \quad n = 0, 1, \dots \quad (10.32)$$

Formula (10.32) follows from multiplying both sides of (10.31) with $e^{-x^2/2} H_n(x)$, integrating over $(-\infty, \infty)$, and noting formula (10.30) and the orthogonality of the sequence $\{H_n(x)\}_{n=0}^{\infty}$.

10.6. LAGUERRE POLYNOMIALS

Laguerre polynomials were named after the French mathematician Edmond Laguerre (1834–1886). They are denoted by $L_n^{(\alpha)}(x)$ and are defined over the interval $(0, \infty)$, $n = 0, 1, 2, \dots$, where $\alpha > -1$.

The development of these polynomials is based on an application of Leibniz formula to finding the n th derivative of the function

$$\phi_n(x) = x^{\alpha+n} e^{-x}.$$

More specifically, for $\alpha > -1$, $L_n^{(\alpha)}(x)$ is defined by a Rodrigues-type formula, namely,

$$L_n^{(\alpha)}(x) = (-1)^n e^x x^{-\alpha} \frac{d^n (x^{n+\alpha} e^{-x})}{dx^n}, \quad n = 0, 1, 2, \dots$$

We shall henceforth use $L_n(x)$ instead of $L_n^{(\alpha)}(x)$.

From this definition, we conclude that $L_n(x)$ is a polynomial of degree n with a leading coefficient equal to one. It can also be shown that Laguerre polynomials are orthogonal with respect to the weight function $w(x) = e^{-x} x^{\alpha}$ over $(0, \infty)$, that is,

$$\int_0^{\infty} e^{-x} x^{\alpha} L_m(x) L_n(x) dx = 0, \quad m \neq n$$

(see Jackson, 1941, page 185). Furthermore, if $m = n$, then

$$\int_0^{\infty} e^{-x} x^{\alpha} [L_n(x)]^2 dx = n! \Gamma(\alpha + n + 1), \quad n = 0, 1, \dots$$

A function $f(x)$ can be expressed as an infinite series of Laguerre polynomials of the form

$$f(x) = \sum_{n=0}^{\infty} c_n L_n(x),$$

where

$$c_n = \frac{1}{n! \Gamma(\alpha + n + 1)} \int_0^\infty e^{-x} x^\alpha L_n(x) f(x) dx, \quad n = 0, 1, 2, \dots$$

A recurrence relation for $L_n(x)$ is developed as follows: From the definition of $L_n(x)$, we have

$$(-1)^n x^\alpha e^{-x} L_n(x) = \frac{d^n(x^{n+\alpha} e^{-x})}{dx^n}, \quad n = 0, 1, 2, \dots \quad (10.33)$$

Replacing n by $n + 1$ in (10.33) gives

$$(-1)^{n+1} x^\alpha e^{-x} L_{n+1}(x) = \frac{d^{n+1}(x^{n+\alpha+1} e^{-x})}{dx^{n+1}}. \quad (10.34)$$

Now,

$$x^{n+\alpha+1} e^{-x} = x(x^{n+\alpha} e^{-x}). \quad (10.35)$$

Applying the Leibniz formula for the $(n + 1)$ st derivative of the product on the right-hand side of (10.35) and noting that the n th derivative of x is zero for $n = 2, 3, 4, \dots$, we obtain

$$\begin{aligned} \frac{d^{n+1}(x^{n+\alpha+1} e^{-x})}{dx^{n+1}} &= x \frac{d^{n+1}(x^{n+\alpha} e^{-x})}{dx^{n+1}} + (n+1) \frac{d^n(x^{n+\alpha} e^{-x})}{dx^n} \\ &= x \frac{d}{dx} \left[\frac{d^n(x^{n+\alpha} e^{-x})}{dx^n} \right] + (n+1) \frac{d^n(x^{n+\alpha} e^{-x})}{dx^n}, \end{aligned} \quad (10.36)$$

Using (10.33) and (10.34) in (10.36) gives

$$\begin{aligned} &(-1)^{n+1} x^\alpha e^{-x} L_{n+1}(x) \\ &= x \frac{d}{dx} \left[(-1)^n x^\alpha e^{-x} L_n(x) \right] + (-1)^n (n+1) x^\alpha e^{-x} L_n(x) \\ &= (-1)^n x^\alpha e^{-x} \left[(\alpha + n + 1 - x) L_n(x) + x \frac{dL_n(x)}{dx} \right]. \end{aligned} \quad (10.37)$$

Multiplying the two sides of (10.37) by $(-1)^{n+1} e^x x^{-\alpha}$, we obtain

$$L_{n+1}(x) = (x - \alpha - n - 1) L_n(x) - x \frac{dL_n(x)}{dx}, \quad n = 0, 1, 2, \dots$$

This recurrence relation gives $L_{n+1}(x)$ in terms of $L_n(x)$ and its derivative. Note that $L_0(x) = 1$.

Another recurrence relation that does not require using the derivative of $L_n(x)$ is given by (see Jackson, 1941, page 186)

$$L_{n+1}(x) = (x - \alpha - 2n - 1)L_n(x) - n(\alpha + n)L_{n-1}(x), \quad n = 1, 2, \dots$$

10.7. LEAST-SQUARES APPROXIMATION WITH ORTHOGONAL POLYNOMIALS

In this section, we consider an approximation problem concerning a continuous function. Suppose that we have a set of polynomials, $\{p_i(x)\}_{i=0}^n$, orthogonal with respect to a weight function $w(x)$ over the interval $[a, b]$. Let $f(x)$ be a continuous function on $[a, b]$. We wish to approximate $f(x)$ by the sum $\sum_{i=0}^n c_i p_i(x)$, where $c_0, c_1, \dots, c_n$ are constants to be determined by minimizing the function

$$\gamma(c_0, c_1, \dots, c_n) = \int_a^b \left[\sum_{i=0}^n c_i p_i(x) - f(x) \right]^2 w(x) dx,$$

that is, γ is the square of the norm, $\|\sum_{i=0}^n c_i p_i - f\|_\omega$. If we differentiate γ with respect to $c_0, c_1, \dots, c_n$ and equate the partial derivatives to zero, we obtain

$$\frac{\partial \gamma}{\partial c_i} = 2 \int_a^b \left[\sum_{j=0}^n c_j p_j(x) - f(x) \right] p_i(x) w(x) dx = 0, \quad i = 0, 1, \dots, n.$$

Hence,

$$\sum_{j=0}^n \left[\int_a^b p_i(x) p_j(x) w(x) dx \right] c_j = \int_a^b f(x) p_i(x) w(x) dx, \quad i = 0, 1, \dots, n. \quad (10.38)$$

Equations (10.38) can be written in vector form as

$$\mathbf{S}\mathbf{c} = \mathbf{u}, \quad (10.39)$$

where $\mathbf{c} = (c_0, c_1, \dots, c_n)'$, $\mathbf{u} = (u_0, u_1, \dots, u_n)'$ with $u_i = \int_a^b f(x) p_i(x) w(x) dx$, and $\mathbf{S}$ is an $(n+1) \times (n+1)$ matrix whose (i, j) th element, s_{ij} , is given by

$$s_{ij} = \int_a^b p_i(x) p_j(x) w(x) dx, \quad i, j = 0, 1, \dots, n.$$

Since $p_0(x), p_1(x), \dots, p_n(x)$ are orthogonal, then $\mathbf{S}$ must be a diagonal matrix with diagonal elements given by

$$s_{ii} = \int_a^b p_i^2(x) w(x) dx = \|p_i\|_\omega^2, \quad i = 0, 1, \dots, n.$$

From equation (10.39) we get the solution $\mathbf{c} = \mathbf{S}^{-1}\mathbf{u}$. The i th element of $\mathbf{c}$ is therefore of the form

$$\begin{aligned} c_i &= \frac{u_i}{s_{ii}} \\ &= \frac{(f \cdot p_i)_\omega}{\|p_i\|_\omega^2}, \quad i = 0, 1, \dots, n. \end{aligned}$$

For such a value of $\mathbf{c}$, γ has an absolute minimum, since $\mathbf{S}$ is positive definite. It follows that the linear combination

$$\begin{aligned} p_n^*(x) &= \sum_{i=0}^n c_i p_i(x) \\ &= \sum_{i=0}^n \frac{(f \cdot p_i)_\omega}{\|p_i\|_\omega^2} p_i(x) \end{aligned} \quad (10.40)$$

minimizes γ . We refer to $p_n^*(x)$ as the least-squares polynomial approximation of $f(x)$ with respect to $p_0(x), p_1(x), \dots, p_n(x)$.

If $\{p_n(x)\}_{n=0}^\infty$ is a sequence of orthogonal polynomials, then $p_n^*(x)$ in (10.40) represents a partial sum of the infinite series $\sum_{n=0}^\infty [(f \cdot p_n)_\omega / \|p_n\|_\omega^2] p_n(x)$. This series may fail to converge point by point to $f(x)$. It converges, however, to $f(x)$ in the norm $\|\cdot\|_\omega$. This is shown in the next theorem.

Theorem 10.7.1. If $f: [a, b] \rightarrow R$ is continuous, then

$$\int_a^b [f(x) - p_n^*(x)]^2 w(x) dx \rightarrow 0$$

as $n \rightarrow \infty$, where $p_n^*(x)$ is defined by formula (10.40).

Proof. By the Weierstrass theorem (Theorem 9.1.1), there exists a polynomial $b_n(x)$ of degree n that converges uniformly to $f(x)$ on $[a, b]$, that is,

$$\sup_{a \leq x \leq b} |f(x) - b_n(x)| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Hence,

$$\int_a^b |f(x) - b_n(x)|^2 w(x) dx \rightarrow 0$$

as $n \rightarrow \infty$, since

$$\int_a^b |f(x) - b_n(x)|^2 w(x) dx \leq \sup_{a \leq x \leq b} |f(x) - b_n(x)|^2 \int_a^b w(x) dx.$$

Furthermore,

$$\int_a^b |f(x) - p_n^*(x)|^2 w(x) dx \leq \int_a^b |f(x) - b_n(x)|^2 w(x) dx, \quad (10.41)$$

since, by definition, $p_n^*(x)$ is the least-squares polynomial approximation of $f(x)$. From inequality (10.41) we conclude that $\|f - p_n^*\|_\omega \rightarrow 0$ as $n \rightarrow \infty$. $\square$

10.8. ORTHOGONAL POLYNOMIALS DEFINED ON A FINITE SET

In this section, we consider polynomials, $p_0(x), p_1(x), \dots, p_n(x)$, defined on a finite set $D = \{x_0, x_2, \dots, x_n\}$ such that $a \leq x_i \leq b$, $i = 0, 1, \dots, n$. These polynomials are orthogonal with respect to a weight function $w^*(x)$ over D if

$$\sum_{i=0}^n w^*(x_i) p_m(x_i) p_\nu(x_i) = 0, \quad m \neq \nu; \quad m, \nu = 0, 1, \dots, n.$$

Such polynomials are said to be orthogonal of the discrete type. For example, the set of discrete Chebyshev polynomials, $\{t_i(j, n)\}_{i=0}^n$, which are defined over the set of integers $j = 0, 1, \dots, n$, are orthogonal with respect to $w^*(j) = 1$, $j = 0, 1, 2, \dots, n$, and are given by the following formula (see Abramowitz and Stegun, 1964, page 791):

$$t_i(j, n) = \sum_{k=0}^i (-1)^k \binom{i}{k} \binom{i+k}{k} \frac{j!(n-k)!}{(j-k)!n!},$$

$$i = 0, 1, \dots, n; \quad j = 0, 1, \dots, n. \quad (10.42)$$

For example, for $i = 0, 1, 2$, we have

$$\begin{aligned} t_0(j, n) &= 1, & j &= 0, 1, 2, \dots, n, \\ t_1(j, n) &= 1 - \frac{2}{n}j, & j &= 0, 1, \dots, n, \\ t_2(j, n) &= 1 - \frac{6}{n} \left[j - \frac{j(j-1)}{n-1} \right] \\ &= 1 - \frac{6j}{n} \left(1 - \frac{j-1}{n-1} \right) \\ &= 1 - \frac{6j}{n} \left(\frac{n-j}{n-1} \right), & j &= 0, 1, 2, \dots, n. \end{aligned}$$

A recurrence relation for the discrete Chebyshev polynomials is of the form

$$(i+1)(n-i)t_{i+1}(j, n) = (2i+1)(n-2j)t_i(j, n) - i(n+i+1)t_{i-1}(j, n), \\ i = 1, 2, \dots, n. \quad (10.43)$$

10.9. APPLICATIONS IN STATISTICS

Orthogonal polynomials play an important role in approximating distribution functions of certain random variables. In this section, we consider only univariate distributions.

10.9.1. Applications of Hermite Polynomials

Hermite polynomials provide a convenient tool for approximating density functions and quantiles of distributions using convergent series. They are associated with the normal distribution, and it is therefore not surprising that they come up in various investigations in statistics and probability theory. Here are some examples.

10.9.1.1. Approximation of Density Functions and Quantiles of Distributions

Let $\phi(x)$ denote the density function of the standard normal distribution, that is,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty. \quad (10.44)$$

We recall from Section 10.5 that the sequence $\{H_n(x)\}_{n=0}^{\infty}$ of Hermite polynomials is orthogonal with respect to $w(x) = e^{-x^2/2}$, and that

$$H_n(x) = x^n - \frac{n^{[2]}}{2 \cdot 1!} x^{n-2} + \frac{n^{[4]}}{2^2 \cdot 2!} x^{n-4} - \frac{n^{[6]}}{2^3 \cdot 3!} x^{n-6} + \dots, \quad (10.45)$$

where $n^{[r]} = n(n-1)(n-2)\cdots(n-r+1)$. Suppose now that $g(x)$ is a density function for some continuous distribution. We can represent $g(x)$ as a series of the form

$$g(x) = \sum_{n=0}^{\infty} b_n H_n(x) \phi(x), \quad (10.46)$$

where, as in formula (10.32),

$$b_n = \frac{1}{n!} \int_{-\infty}^{\infty} g(x) H_n(x) dx. \quad (10.47)$$

By substituting $H_n(x)$, as given by formula (10.45), in formula (10.47), we obtain an expression for b_n in terms of the central moments, $\mu_0, \mu_1, \dots, \mu_n, \dots$, of the distribution whose density function is $g(x)$. These moments are defined as

$$\mu_n = \int_{-\infty}^{\infty} (x - \mu)^n g(x) dx, \quad n = 0, 1, 2, \dots,$$

where μ is the mean of the distribution. Note that $\mu_0 = 1$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$, the variance of the distribution. In particular, if $\mu = 0$, then

$$\begin{aligned} b_0 &= 1, \\ b_1 &= 0, \\ b_2 &= \frac{1}{2}(\mu_2 - 1), \\ b_3 &= \frac{1}{6}\mu_3, \\ b_4 &= \frac{1}{24}(\mu_4 - 6\mu_2 + 3), \\ b_5 &= \frac{1}{120}(\mu_5 - 10\mu_3), \\ b_6 &= \frac{1}{720}(\mu_6 - 15\mu_4 + 45\mu_2 - 15), \\ &\vdots \end{aligned}$$

The expression for $g(x)$ in formula (10.46) can then be written as

$$\begin{aligned} g(x) = \phi(x) \Big[&1 + \frac{1}{2}(\mu_2 - 1)H_2(x) + \frac{1}{6}\mu_3H_3(x) \\ &+ \frac{1}{24}(\mu_4 - 6\mu_2 + 3)H_4(x) + \dots \Big]. \end{aligned} \quad (10.48)$$

This expression is known as the *Gram-Charlier series of type A*.

Thus the Gram-Charlier series provides an expansion of $g(x)$ in terms of its central moments, the standard normal density, and Hermite polynomials. Using formulas (10.21) and (10.46), we note that $g(x)$ can be expressed as a series of derivatives of $\phi(x)$ of the form

$$g(x) = \sum_{n=0}^{\infty} \frac{c_n}{n!} \left(\frac{d}{dx} \right)^n \phi(x), \quad (10.49)$$

where

$$c_n = (-1)^n \int_{-\infty}^{\infty} g(x) H_n(x) dx, \quad n = 0, 1, \dots$$

Cramér (1946, page 223) gave conditions for the convergence of the series on the right-hand side of formula (10.49), namely, if $g(x)$ is continuous and of bounded variation on $(-\infty, \infty)$, and if the integral $\int_{-\infty}^{\infty} g(x) \exp(x^2/4) dx$ is convergent, then the series in formula (10.49) will converge for every x to $g(x)$.

We can utilize Gram–Charlier series to find the upper α -quantile, x_α , of the distribution with the density function $g(x)$. This point is defined as

$$\int_{-\infty}^{x_\alpha} g(x) dx = 1 - \alpha.$$

From (10.46) we have that

$$g(x) = \phi(x) + \sum_{n=2}^{\infty} b_n H_n(x) \phi(x).$$

Then

$$\int_{-\infty}^{x_\alpha} g(x) dx = \int_{-\infty}^{x_\alpha} \phi(x) dx + \sum_{n=2}^{\infty} b_n \int_{-\infty}^{x_\alpha} H_n(x) \phi(x) dx. \quad (10.50)$$

However,

$$\int_{-\infty}^{x_\alpha} H_n(x) \phi(x) dx = -H_{n-1}(x_\alpha) \phi(x_\alpha).$$

To prove this equality we note that by formula (10.21)

$$\begin{aligned} \int_{-\infty}^{x_\alpha} H_n(x) \phi(x) dx &= (-1)^n \int_{-\infty}^{x_\alpha} \left(\frac{d}{dx} \right)^n \phi(x) dx \\ &= (-1)^n \left(\frac{d}{dx} \right)^{n-1} \phi(x_\alpha), \end{aligned}$$

where $(d/dx)^{n-1} \phi(x_\alpha)$ denotes the value of the $(n-1)$ st derivative of $\phi(x)$ at x_α . By applying formula (10.21) again we obtain

$$\begin{aligned} \int_{-\infty}^{x_\alpha} H_n(x) \phi(x) dx &= (-1)^n (-1)^{n-1} H_{n-1}(x_\alpha) \phi(x_\alpha) \\ &= -H_{n-1}(x_\alpha) \phi(x_\alpha). \end{aligned}$$

By making the substitution in formula (10.50), we get

$$\int_{-\infty}^{x_\alpha} g(x) dx = \int_{-\infty}^{x_\alpha} \phi(x) dx - \sum_{n=2}^{\infty} b_n H_{n-1}(x_\alpha) \phi(x_\alpha). \quad (10.51)$$

Now, suppose that z_α is the upper α -quantile of the standard normal distribution. Then

$$\int_{-\infty}^{x_\alpha} g(x) dx = 1 - \alpha = \int_{-\infty}^{z_\alpha} \phi(x) dx.$$

Using the expansion (10.51), we obtain

$$\int_{-\infty}^{x_\alpha} \phi(x) dx - \sum_{n=2}^{\infty} b_n H_{n-1}(x_\alpha) \phi(x_\alpha) = \int_{-\infty}^{z_\alpha} \phi(x) dx. \quad (10.52)$$

If we expand the right-hand side of equation (10.52) using Taylor's series in a neighborhood of x_α , we get

$$\begin{aligned} \int_{-\infty}^{z_\alpha} \phi(x) dx &= \int_{-\infty}^{x_\alpha} \phi(x) dx + \sum_{j=1}^{\infty} \frac{(z_\alpha - x_\alpha)^j}{j!} \left(\frac{d}{dx} \right)^{j-1} \phi(x_\alpha) \\ &= \int_{-\infty}^{x_\alpha} \phi(x) dx + \sum_{j=1}^{\infty} \frac{(z_\alpha - x_\alpha)^j}{j!} (-1)^{j-1} H_{j-1}(x_\alpha) \phi(x_\alpha), \\ &\quad \text{using formula (10.21)} \\ &= \int_{-\infty}^{x_\alpha} \phi(x) dx - \sum_{j=1}^{\infty} \frac{(x_\alpha - z_\alpha)^j}{j!} H_{j-1}(x_\alpha) \phi(x_\alpha). \end{aligned} \quad (10.53)$$

From formulas (10.52) and (10.53) we conclude that

$$\sum_{n=2}^{\infty} b_n H_{n-1}(x_\alpha) \phi(x_\alpha) = \sum_{j=1}^{\infty} \frac{(x_\alpha - z_\alpha)^j}{j!} H_{j-1}(x_\alpha) \phi(x_\alpha).$$

By dividing both sides by $\phi(x_\alpha)$ we obtain

$$\sum_{n=2}^{\infty} b_n H_{n-1}(x_\alpha) = \sum_{j=1}^{\infty} \frac{(x_\alpha - z_\alpha)^j}{j!} H_{j-1}(x_\alpha). \quad (10.54)$$

This provides a relationship between x_α , the α -quantile of the distribution with the density function $g(x)$, and z_α , the corresponding quantile for the standard normal. Since the b_n 's are functions of the moments associated with $g(x)$, then it is possible to use (10.54) to express x_α in terms of z_α and the moments of $g(x)$. This was carried out by Cornish and Fisher (1937). They provided an expansion for x_α in terms of z_α and the cumulants (instead of the moments) associated with $g(x)$. (See Section 5.6.2 for a definition of cumulants. Note that there is a one-to-one correspondence between moments and cumulants.) Such an expansion became known as the

Cornish–Fisher expansion. It is reported in Johnson and Kotz (1970, page 34) (see Exercise 10.11). See also Kendall and Stuart (1977, pages 175–178).

10.9.1.2. Approximation of a Normal Integral

A convergent series representing the integral

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$$

was derived by Kerridge and Cook (1976). Their method is based on the fact that

$$\int_0^x f(t) dt = 2 \sum_{n=0}^{\infty} \frac{(x/2)^{2n+1}}{(2n+1)!} f^{(2n)}\left(\frac{x}{2}\right) \quad (10.55)$$

for any function $f(t)$ with a suitably convergent Taylor's expansion in a neighborhood of $x/2$, namely,

$$f(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(t - \frac{x}{2}\right)^n f^{(n)}\left(\frac{x}{2}\right). \quad (10.56)$$

Formula (10.55) results from integrating (10.56) with respect to t from 0 to x and noting that the even terms vanish. Taking $f(t) = e^{-t^2/2}$, we obtain

$$\int_0^x e^{-t^2/2} dt = 2 \sum_{n=0}^{\infty} \frac{(x/2)^{2n+1}}{(2n+1)!} \frac{d^{2n}(e^{-t^2/2})}{dt^{2n}} \bigg|_{t=x/2}. \quad (10.57)$$

Using the Rodrigues formula for Hermite polynomials [formula (10.21)], we get

$$\frac{d^n(e^{-x^2/2})}{dx^n} = (-1)^n e^{-x^2/2} H_n(x), \quad n = 0, 1, \dots$$

By making the substitution in (10.57), we find

$$\begin{aligned} \int_0^x e^{-t^2/2} dt &= 2 \sum_{n=0}^{\infty} \frac{(x/2)^{2n+1}}{(2n+1)!} e^{-x^2/2} H_{2n}\left(\frac{x}{2}\right) \\ &= 2e^{-x^2/8} \sum_{n=0}^{\infty} \frac{(x/2)^{2n+1}}{(2n+1)!} H_{2n}\left(\frac{x}{2}\right). \end{aligned} \quad (10.58)$$

This expression can be simplified by letting $\Theta_n(x) = x^n H_n(x)/n!$ in (10.58), which gives

$$\int_0^x e^{-t^2/2} dt = xe^{-x^2/8} \sum_{n=0}^{\infty} \frac{\Theta_{2n}(x/2)}{2n+1}.$$

Hence,

$$\psi(x) = \frac{1}{\sqrt{2\pi}} x e^{-x^2/8} \sum_{n=0}^{\infty} \frac{\Theta_{2n}(x/2)}{2n+1}. \quad (10.59)$$

Note that on the basis of formula (10.25), the recurrence relation for $\Theta_n(x)$ is given by

$$\Theta_{n+1} = \frac{x^2(\Theta_n - \Theta_{n-1})}{n+1}, \quad n = 1, 2, \dots$$

The $\Theta_n(x)$'s are easier to handle numerically than the Hermite polynomials, as they remain relatively small, even for large n . Kerridge and Cook (1976) report that the series in (10.59) is accurate over a wide range of x . Divgi (1979), however, states that the convergence of the series becomes slower as x increases.

10.9.1.3. Estimation of Unknown Densities

Let $X_1, X_2, \dots, X_n$ represent a sequence of independent random variables with a common, but unknown, density function $f(x)$ assumed to be square integrable. From (10.31) we have the representation

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x),$$

or equivalently,

$$f(x) = \sum_{j=0}^{\infty} a_j h_j(x), \quad (10.60)$$

where $h_j(x)$ is the so-called normalized Hermite polynomial of degree j , namely,

$$h_j(x) = \frac{1}{(\sqrt{2\pi} j!)^{1/2}} e^{-x^2/4} H_j(x), \quad j = 0, 1, \dots,$$

and $a_j = \int_{-\infty}^{\infty} f(x) h_j(x) dx$, since $\int_{-\infty}^{\infty} h_j^2(x) dx = 1$ by virtue of (10.30).

Schwartz (1967) considered an estimate of $f(x)$ of the form

$$\hat{f}_n(x) = \sum_{j=0}^{q(n)} \hat{a}_{jn} h_j(x),$$

where

$$\hat{a}_{jn} = \frac{1}{n} \sum_{k=1}^n h_j(X_k),$$

and $q(n)$ is a suitably chosen integer dependent on n such that $q(n) = o(n)$ as $n \rightarrow \infty$. Under these conditions, Schwartz (1967, Theorem 1) showed that $\hat{f}_n(x)$ is a consistent estimator of $f(x)$ in the mean integrated squared error sense, that is,

$$\lim_{n \rightarrow \infty} E \int [\hat{f}_n(x) - f(x)]^2 dx = 0.$$

Under additional conditions on $f(x)$, $\hat{f}_n(x)$ is also consistent in the mean squared error sense, that is,

$$\lim_{n \rightarrow \infty} E [f(x) - \hat{f}_n(x)]^2 = 0$$

uniformly in x .

10.9.2. Applications of Jacobi and Laguerre Polynomials

Dasgupta (1968) presented an approximation to the distribution function of $X = \frac{1}{2}(r + 1)$, where r is the sample correlation coefficient, in terms of a beta density and Jacobi polynomials. Similar methods were used by Durbin and Watson (1951) in deriving an approximation of the distribution of a statistic used for testing serial correlation in least-squares regression.

Quadratic forms in random variables, which can often be regarded as having joint multivariate normal distributions, play an important role in analysis of variance and in estimation of variance components for a random or a mixed model. Approximation of the distributions of such quadratic forms can be carried out using Laguerre polynomials (see, for example, Gurland, 1953, and Johnson and Kotz, 1968). Tiku (1964a) developed Laguerre series expansions of the distribution functions of the nonnormal variance ratios used for testing the homogeneity of treatment means in the case of one-way classification for analysis of variance with nonidentical group-to-group error distributions that are not assumed to be normal. Tiku (1964b) also used Laguerre polynomials to obtain an approximation to the first negative moment of a Poisson random variable, that is, the value of $E(X^{-1})$, where X has the Poisson distribution.

More recently, Schöne and Schmid (2000) made use of Laguerre polynomials to develop a series representation of the joint density and the joint distribution of a quadratic form and a linear form in normal variables. Such a representation can be used to calculate, for example, the joint density and the joint distribution function of the sample mean and sample variance. Note that for autocorrelated variables, the sample mean and sample variance are, in general, not independent.

10.9.3. Calculation of Hypergeometric Probabilities Using Discrete Chebyshev Polynomials

The hypergeometric distribution is a discrete distribution, somewhat related to the binomial distribution. Suppose, for example, we have a lot of M items,

r of which are defective and $M - r$ of which are nondefective. Suppose that we choose at random m items without replacement from the lot ($m \leq M$). Let X be the number of defectives found. Then, the probability that $X = k$ is given by

$$P(X = k) = \frac{\binom{r}{k} \binom{M-r}{m-k}}{\binom{M}{m}}, \quad (10.61)$$

where $\max(0, m - M + r) \leq k \leq \min(m, r)$. A random variable with the probability mass function (10.61) is said to have a hypergeometric distribution. We denote such a probability function by $h(k; m, r, M)$.

There are tables for computing the probability value in (10.61) (see, for example, the tables given by Lieberman and Owen, 1961). There are also several algorithms for computing this probability. Recently, Alvo and Cabilio (2000) proposed to represent the hypergeometric distribution in terms of discrete Chebyshev polynomials, as was seen in Section 10.8. The following is a summary of this work: Consider the sequence $\{t_n(k, m)\}_{n=0}^m$ of discrete Chebyshev polynomials defined over the set of integers $k = 0, 1, 2, \dots, m$ [see formula (10.42)], which is given by

$$t_n(k, m) = \sum_{i=0}^n (-1)^i \binom{n}{i} \binom{n+i}{i} \frac{k!(m-i)!}{(k-i)!m!},$$

$$n = 0, 1, \dots, m, \quad k = 0, 1, \dots, m. \quad (10.62)$$

Let X have the hypergeometric distribution as in (10.61). Then according to Theorem 1 in Alvo and Cabilio (2000),

$$\sum_{k=0}^m t_n(k, m) h(k; m, r, M) = t_n(r, M) \quad (10.63)$$

for all $n = 0, 1, \dots, m$ and $r = 0, 1, \dots, M$. Let $\mathbf{t}_n = [t_n(0, m), t_n(1, m), \dots, t_n(m, m)]$, $n = 0, 1, \dots, m$, be the base vectors in an $(m+1)$ -dimensional Euclidean space determined from the Chebyshev polynomials. Let $g(k)$ be any function defined over the set of integers, $k = 0, 1, \dots, m$. Then $g(k)$ can be expressed as

$$g(k) = \sum_{n=0}^m g_n t_n(k, m), \quad k = 0, 1, \dots, m, \quad (10.64)$$

where $g_n = \mathbf{g} \cdot \mathbf{t}_n / \|\mathbf{t}_n\|^2$, and $\mathbf{g} = [g(0), g(1), \dots, g(m)]$. Now, using the result

in (10.63), the expected value of $g(X)$ is given by

$$\begin{aligned}
 E[g(X)] &= \sum_{k=0}^m g(k)h(k; m, r, M) \\
 &= \sum_{k=0}^m \sum_{n=0}^m g_n t_n(k, m)h(k; m, r, M) \\
 &= \sum_{n=0}^m g_n \sum_{k=0}^m t_n(k, m)h(k; m, r, M) \\
 &= \sum_{n=0}^m g_n t_n(r, M). \tag{10.65}
 \end{aligned}$$

This shows that the expected value of $g(X)$ can be computed from knowledge of the coefficients g_n and the discrete Chebyshev polynomials up to order m , evaluated at r and M .

In particular, if $g(x)$ is an indicator function taking the value one at $x = k$ and the value zero elsewhere, then

$$\begin{aligned}
 E[g(X)] &= P(X = k) \\
 &= h(k; m, r, M).
 \end{aligned}$$

Applying the result in (10.65), we then obtain

$$h(k; m, r, M) = \sum_{n=0}^m \frac{t_n(k, m)}{\|t_n\|^2} t_n(r, M). \tag{10.66}$$

Because of the recurrence relation (10.43) for discrete Chebyshev polynomials, calculating the hypergeometric probability using (10.66) can be done simply on a computer.

FURTHER READING AND ANNOTATED BIBLIOGRAPHY

- Abramowitz, M., and I. A. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Wiley, New York. (This useful volume was prepared by the National Bureau of Standards. It was edited by Milton Abramowitz and Irene A. Stegun.)
- Alvo, M., and P. Cabilio (2000). "Calculation of hypergeometric probabilities using Chebyshev polynomials." *Amer. Statist.*, **54**, 141–144.
- Cheney, E. W. (1982). *Introduction to Approximation Theory*, 2nd ed. Chelsea, New York. (Least-squares polynomial approximation is discussed in Chap. 4.)

- Chihara, T. S. (1978). *An Introduction to Orthogonal Polynomials*. Gordon and Breach, New York. (This text deals with the general theory of orthogonal polynomials, including recurrence relations, and some particular systems of orthogonal polynomials.)
- Cornish, E. A., and R. A. Fisher (1937). "Moments and cumulants in the specification of distributions." *Rev. Internat. Statist. Inst.*, **5**, 307–320.
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton. (This classic book provides the mathematical foundation of statistics. Chap. 17 is a good source for approximation of density functions.)
- Dasgupta, P. (1968). "An approximation to the distribution of sample correlation coefficient, when the population is non-normal." *Sankhyā, Ser. B.*, **30**, 425–428.
- Davis, P. J. (1975). *Interpolation and Approximation*. Dover, New York. (Chaps. 8 and 10 discuss least-squares approximation and orthogonal polynomials.)
- Divgi, D. R. (1979). "Calculation of univariate and bivariate normal probability functions." *Ann. Statist.*, **7**, 903–910.
- Durbin, J., and G. S. Watson (1951). "Testing for serial correlation in least-squares regression II." *Biometrika*, **38**, 159–178.
- Freud, G. (1971). *Orthogonal Polynomials*. Pergamon Press, Oxford. (This book deals with fundamental properties of orthogonal polynomials, including Legendre, Chebyshev, and Jacobi polynomials. Convergence theory of series of orthogonal polynomials is discussed in Chap. 4.)
- Gurland, J. (1953). "Distributions of quadratic forms and ratios of quadratic forms." *Ann. Math. Statist.*, **24**, 416–427.
- Jackson, D. (1941). *Fourier Series and Orthogonal Polynomials*. Mathematical Association of America. (This classic monograph provides a good coverage of orthogonal polynomials, including Legendre, Jacobi, Hermite, and Laguerre polynomials. The presentation is informative and easy to follow.)
- Johnson, N. L., and S. Kotz (1968). "Tables of distributions of positive definite quadratic forms in central normal variables." *Sankhyā, Ser. B.*, **30**, 303–314.
- Johnson, N. L., and S. Kotz (1970). *Continuous Univariate Distributions—1*. Houghton Mifflin, Boston. (Chap. 12 contains a good discussion concerning the Cornish–Fisher expansion of quantiles.)
- Kendall, M. G., and A. Stuart (1977). *The Advanced Theory of Statistics*, Vol. 1, 4th ed. Macmillan, New York. (This classic book provides a good source for learning about the Gram–Charlier series of Type A and the Cornish–Fisher expansion.)
- Kerridge, D. F., and G. W. Cook (1976). "Yet another series for the normal integral." *Biometrika*, **63**, 401–403.
- Lieberman, G. J., and Owen, D. B. (1961). *Tables of the Hypergeometric Probability Distribution*. Stanford University Press, Palo Alto, California.
- Ralston, A., and P. Rabinowitz (1978). *A First Course in Numerical Analysis*. McGraw-Hill, New York. (Chap. 7 discusses Chebyshev polynomials of the first kind).
- Rivlin, T. (1990). *Chebyshev Polynomials*, 2nd ed. Wiley, New York. (This book gives a survey of the most important properties of Chebyshev polynomials.)
- Schöne, A., and W. Schmid (2000). "On the joint distribution of a quadratic and a linear form in normal variables." *J. Mult. Analysis*, **72**, 163–182.

- Schwartz, S. C. (1967). "Estimation of probability density by an orthogonal series," *Ann. Math. Statist.*, **38**, 1261–1265.
- Subrahmaniam, K. (1966). "Some contributions to the theory of non-normality—I (univariate case)." *Sankhyā, Ser. A*, **28**, 389–406.
- Szegő, G. (1975). *Orthogonal Polynomials*, 4th ed. Amer. Math. Soc., Providence, Rhode Island. (This much-referenced book provides a thorough coverage of orthogonal polynomials.)
- Tiku, M. L. (1964a). "Approximating the general non-normal variance ratio sampling distributions." *Biometrika*, **51**, 83–95.
- Tiku, M. L. (1964b). "A note on the negative moments of a truncated Poisson variate." *J. Amer. Statist. Assoc.*, **59**, 1220–1224.
- Viskov, O. V. (1992). "Some remarks on Hermite polynomials." *Theory of Probability and its Applications*, **36**, 633–637.

EXERCISES

In Mathematics

- 10.1.** Show that the sequence $\{1/\sqrt{2}, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx, \dots\}$ is orthonormal with respect to $w(x) = 1/\pi$ over $[-\pi, \pi]$.
- 10.2.** Let $\{p_n(x)\}_{n=0}^\infty$ be a sequence of Legendre polynomials
- (a) Use the Rodrigues formula to show that
 - (i) $\int_{-1}^1 x^m p_n(x) dx = 0$ for $m = 0, 1, \dots, n-1$,
 - (ii) $\int_{-1}^1 x^n p_n(x) dx = \frac{2^{n+1}}{2n+1} \binom{2n}{n}^{-1}$, $n = 0, 1, 2, \dots$
 - (b) Deduce from (a) that $\int_{-1}^1 p_n(x) \pi_{n-1}(x) dx = 0$, where $\pi_{n-1}(x)$ denotes an arbitrary polynomial of degree at most equal to $n-1$.
 - (c) Make use of (a) and (b) to show that $\int_{-1}^1 p_n^2(x) dx = 2/(2n+1)$, $n = 0, 1, \dots$.
- 10.3.** Let $\{T_n(x)\}_{n=0}^\infty$ be a sequence of Chebyshev polynomials of the first kind. Let $\zeta_i = \cos[(2i-1)\pi/2n]$, $i = 1, 2, \dots, n$.
- (a) Verify that $\zeta_1, \zeta_2, \dots, \zeta_n$ are zeros of $T_n(x)$, that is, $T_n(\zeta_i) = 0$, $i = 1, 2, \dots, n$.
 - (b) Show that $\zeta_1, \zeta_2, \dots, \zeta_n$ are simple zeros of $T_n(x)$.
[Hint: show that $T'_n(\zeta_i) \neq 0$ for $i = 1, 2, \dots, n$.]
- 10.4.** Let $\{H_n(x)\}_{n=0}^\infty$ be a sequence of Hermite polynomials. Show that
- (a) $\frac{dH_n(x)}{dx} = nH_{n-1}(x)$,
 - (b) $\frac{d^2H_n(x)}{dx^2} - x \frac{dH_n(x)}{dx} + nH_n(x) = 0$.

10.5. Let $\{T_n(x)\}_{n=0}^\infty$ and $\{U_n(x)\}_{n=0}^\infty$ be sequences of Chebyshev polynomials of the first and second kinds, respectively,

(a) Show that $|U_n(x)| \leq n+1$ for $-1 \leq x \leq 1$.

[Hint: Use the representation (10.18) and mathematical induction on n .]

(b) Show that $|dT_n(x)/dx| \leq n^2$ for all $-1 \leq x \leq 1$, with equality holding only if $x = \mp 1$ ($n \geq 2$).

(c) Show that

$$\int_{-1}^x \frac{T_n(t)}{\sqrt{1-t^2}} dt = -\frac{U_{n-1}(x)}{n} \sqrt{1-x^2}$$

for $-1 \leq x \leq 1$ and $n \neq 0$.

10.6. Show that the Laguerre polynomial $L_n(x)$ of degree n satisfies the differential equation

$$x \frac{d^2 L_n(x)}{dx^2} + (\alpha + 1 - x) \frac{dL_n(x)}{dx} + nL_n(x) = 0.$$

10.7. Consider the function

$$H(x, t) = \frac{1}{(1-t)^{\alpha+1}} e^{-xt/(1-t)}, \quad \alpha > -1.$$

Expand $H(x, t)$ as a power series in t and let the coefficient of t^n be denoted by $(-1)^n g_n(x)/n!$ so that

$$H(x, t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} g_n(x) t^n.$$

Show that $g_n(x)$ is identical to $L_n(x)$ for all n , where $L_n(x)$ is the Laguerre polynomial of degree n .

10.8. Find the least-squares polynomial approximation of the function $f(x) = e^x$ over the interval $[-1, 1]$ by using a Legendre polynomial of degree not exceeding 4.

10.9. A function $f(x)$ defined on $-1 \leq x \leq 1$ can be represented using an infinite series of Chebyshev polynomials of the first kind, namely,

$$f(x) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n T_n(x),$$

where

$$c_n = \frac{2}{\pi} \int_{-1}^1 \frac{f(x)T_n(x)}{\sqrt{1-x^2}} dx, \quad n = 0, 1, \dots$$

This series converges uniformly whenever $f(x)$ is continuous and of bounded variation on $[-1, 1]$. Approximate the function $f(x) = e^x$ using the first five terms of the above series.

In Statistics

- 10.10.** Suppose that from a certain distribution with a mean equal to zero we have knowledge of the following central moments: $\mu_2 = 1.0$, $\mu_3 = -0.91$, $\mu_4 = 4.86$, $\mu_5 = -12.57$, $\mu_6 = 53.22$. Obtain an approximation for the density function of the distribution using Gram–Charlier series of type A.
- 10.11.** The Cornish–Fisher expansion for x_α , the upper α -quantile of a certain distribution, standardized so that its mean and variance are equal to zero and one, respectively, is of the following form (see Johnson and Kotz, 1970, page 34):

$$\begin{aligned} x_\alpha = & z_\alpha + \frac{1}{6}(z_\alpha^2 - 1)\kappa_3 + \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\kappa_4 \\ & - \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\kappa_3^2 + \frac{1}{120}(z_\alpha^4 - 6z_\alpha^2 + 3)\kappa_5 \\ & - \frac{1}{24}(z_\alpha^4 - 5z_\alpha^2 + 2)\kappa_3\kappa_4 + \frac{1}{324}(12z_\alpha^4 - 53z_\alpha^2 + 17)\kappa_3^3 \\ & + \frac{1}{720}(z_\alpha^5 - 10z_\alpha^3 + 15z_\alpha)\kappa_6 \\ & - \frac{1}{180}(2z_\alpha^5 - 17z_\alpha^3 + 21z_\alpha)\kappa_3\kappa_5 \\ & - \frac{1}{384}(3z_\alpha^5 - 24z_\alpha^3 + 29z_\alpha)\kappa_4^2 \\ & + \frac{1}{288}(14z_\alpha^5 - 103z_\alpha^3 + 107z_\alpha)\kappa_3^2\kappa_4 \\ & - \frac{1}{7776}(252z_\alpha^5 - 1688z_\alpha^3 + 1511z_\alpha)\kappa_3^4 + \dots, \end{aligned}$$

where z_α is the upper α -quantile of the standard normal distribution, and κ_r is the r th cumulant of the distribution ($r = 3, 4, \dots$). Apply this expansion to finding the upper 0.05-quantile of the central chi-squared distribution with $n = 5$ degrees of freedom.

[Note: The mean and variance of a central chi-squared distribution with n degrees of freedom are n and $2n$, respectively. Its r th cumu-

lant, denoted by κ'_r , is

$$\kappa'_r = n(r-1)! 2^{r-1}, \quad r = 1, 2, \dots$$

Hence, the r th cumulant, κ_r , of the standardized chi-squared distribution is $\kappa_r = (2n)^{-r/2} \kappa'_r$ ($r = 2, 3, \dots$).]

10.12. The normal integral $\int_0^x e^{-t^2/2} dt$ can be calculated from the series

$$\int_0^x e^{-t^2/2} dt = x - \frac{x^3}{2 \cdot 3 \cdot 1!} + \frac{x^5}{2^2 \cdot 5 \cdot 2!} - \frac{x^7}{2^3 \cdot 7 \cdot 3!} + \dots$$

(a) Use this series to obtain an approximate value for $\int_0^1 e^{-t^2/2} dt$.

(b) Redo part (a) using the series given by formula (10.59), that is,

$$\int_0^x e^{-t^2/2} dt = x e^{-x^2/8} \sum_{n=0}^{\infty} \frac{\Theta_{2n}(x/2)}{2n+1}.$$

(c) Compare the results from (a) and (b) with regard to the number of terms in each series needed to achieve an answer correct to five decimal places.

10.13. Show that the expansion given by formula (10.46) is equivalent to representing the density function $g(x)$ as a series of the form

$$g(x) = \sum_{n=0}^{\infty} \frac{c_n}{n!} \frac{d^n \phi(x)}{dx^n},$$

where $\phi(x)$ is the standard normal density function, and the c_n 's are constant coefficients.

10.14. Consider the random variable

$$W = \sum_{i=1}^n X_i^2,$$

where $X_1, X_2, \dots, X_n$ are independent random variables from a distribution with the density function

$$f(x) = \phi(x) - \frac{\lambda_3}{6} \frac{d^3 \phi(x)}{dx^3} + \frac{\lambda_4}{24} \frac{d^4 \phi(x)}{dx^4} + \frac{\lambda_3^2}{72} \frac{d^6 \phi(x)}{dx^6},$$

where $\phi(x)$ is the standard normal density function and the quantities λ_3 and λ_4 are, respectively, the standard measures of skewness and kurtosis for the distribution. Obtain the moment generating function of W , and compare it with the moment generating function of a chi-squared distribution with n degrees of freedom. (See Example 6.9.8 in Section 6.9.3.)

[*Hint*: Use Hermite polynomials.]

- 10.15.** A lot of $M = 10$ articles contains $r = 3$ defectives and 7 good articles. Suppose that a sample of $m = 4$ articles is drawn from the lot without replacement. Let X denote the number of defective articles in the sample. Find the expected value of $g(X) = X^3$ using formula (10.65).