A NEW ASYMPTOTIC EXPANSION FOR DISTRIBUTION OF SUMS OF RANDOM VARIABLES

By

JAMES CHERNESKY JR.

A Thesis Submitted to the Graduate Faculty of

WAKE FOREST UNIVERSITY

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF ARTS

in the Department of Mathematics

May 2010

Winston-Salem, North Carolina

Approved By:

Dr. Kenneth S. Berenhaut, Ph.D., Advisor ____________________

Examining Committee:

Dr. James Norris, Ph.D., Chairperson ____________________

Dr. Miaohua Jiang, Ph.D. ____________________

Table of Contents

Acknowledgments

There are many people who helped to make this thesis possible. First I would like to thank my committee members. Their insightful questions and flexibility helped to make this whole process run smoothly. Many thanks to the fellow graduate students in the department. They were always around to act as a sound board, or provide support. I would like to thank my professors for providing me with the knowledge and guidance to become the student that I am today. I would like to thank my family for their never-ending supply of love and support. No matter what else was going on, they always helped keep a head on my shoulders, and always keep me looking towards the finish line. Without their encouragement, I would not be where I am today. Finally, I would like to thank my advisor, Dr. Berenhaut. He has been an advisor to me in ways that far surpass this thesis, or even this program. Whether it was a random pep talk, or one of the countless meetings until all hours of the night, his top priority was always to keep my spirits up, and to help me do my best. Without that I don't think I would have been able to finish this thesis, let alone this program.

Abstract

James Chernesky Jr.

In this paper we look to improve upon local Edgeworth expansions for probability distributions of sums of independent identically distributed random variables. Let X be a random variable with finite variance, and $X_1, X_2, ...,$ a sequence of i.i.d. random variables each with the same distribution as X. In addition, suppose that X has probability function p, where p is either a density or a probability mass function. Define the partial sum $S_i = X_1 + X_2 + ... + X_i$, with probability function $p^{(i)}$.

In the standard Edgeworth case, the local expansion for $p^{(n)}$ is in terms of powers of $1/\sqrt{n}$. This expansion involves the cumulants of X, as well as Hermite polynomials. The expansion defined in this thesis uses a rival sequence of coefficients dependant on the probability function p, which does not require the calculation of cumulants, or the use of Hermite polynomials. The advantages of this expansion are most noticeable when X is a discrete random variable on a small support set. The approach produces better asymptotic results in many cases, especially for symmetric random variables X. The main error term is often simpler (in terms of the value of interest in the support set), and as a result, the asymptotic crossing points of the estimate with the target can be more easily computed (given information on the location of zeros of Hermite polynomials).

Chapter 1: Introduction

We begin with a random variable X with finite variance. For the purpose of this thesis, we will mainly be dealing with discrete random variables on a relatively small support set. Suppose $X_1, X_2, \ldots$ forms a sequence of independent identically distributed random variables each with the same distribution as X. Now define the partial sum $S_n = X_1 + \cdots + X_n$ with probability function $p^{(n)}$. We are interested in legitimate expansions of the form

$$p^{(n)}(M) = \sum_{i=1}^{\infty} C_i(n, p, M) \frac{1}{n^{i/2}}$$

with concrete error terms, i.e. for $k \geq 1$

$$p^{(n)}(M) = \sum_{i=1}^{k} C_i(n, p, M) \frac{1}{n^{i/2}} + o\left(\frac{1}{n^{k/2}}\right),$$

where p is the distribution of X. The standard Edgeworth expansion is of this form. We have been able to obtain a rival expansion for $p^{(n)}$ in terms of powers of $1/\sqrt{n}$. This new expansion has many advantages over the Edgeworth which will be discussed throughout this thesis.

In order to properly define these expansions, there are a number of preliminary topics that will be introduced. Chapter 1 will cover some basics statistical concepts, as well as Stirling numbers of the second kind. These will be needed in the proof of our new expansion. Following these topics, Chapter 2 will comprise a discussion of Hermite polynomials; this chapter will be concerned with not only defining these polynomials, but obtaining information about their zeros. To date there is no closed form equation for finding the zeros to these polynomials, so Chapter 2 will discuss various properties of the zeros, as well as some of the types of bound that may be put on them.

Chapter 3 will include definition and discussion of the well known Edgeworth expansion. Chapter 4 will first define our new expansion, then discuss the main theorem of this thesis, which can be summarized by the following. For any integer $h \in \{1, \ldots, n\}$, we can define the independent sum $\Psi_{h,n} = X_1 + X_2 + \cdots + X_h + N_{h+1} + \cdots + N_n$, where X_i are our random variables from above with mean μ and variance σ^2, and N_i are normal $N(\mu, \sigma^2)$ random variables (see Section 1.1). Then, let $\eta_h = \eta_h^{(n)}$ be the probability function for $\Psi_{h,n}$. Note that for relatively small h and small support set for X_i, η_h is quite simple to compute via conditioning. Now define

$$\eta_h^*(M) = \frac{\eta_h(M) - \phi_{\sigma,n}(x)}{\phi_{\sigma,n}(x)}.$$

Combining these with the delta difference operator (see Chapter 5), we can employ a result of Petrov on Edgeworth-type expansions for general sums of non-identically distributed random variables to obtain

$$f_n(M) = \eta_1(M) + (n-1)(\eta_2(M) - \eta_1(M)) + \frac{(n-1)(n-2)}{2}(\eta_3(M) - 2\eta_2(M) + \eta_1(M)) + \ldots$$

i.e.

$$f_n(M) = \sum_{j=1}^{s} \binom{n-1}{j-1} \Delta^{j-1} \eta_1^*(M) + o\left(\frac{1}{n^{s/2}}\right).$$

This is the formal definition of our new expansion. It will be compared to the Edgeworth expansion in various scenarios. In Chapter 6, we will show the advantages of the new expansion over the Edgeworth in two key scenarios. The first is the simple lattice case with a small support set; the second is the symmetric random variable case.

The last chapter will discuss an additional advantage of the new expansion. This deals with the asymptotic crossing points of the estimate with the target. After discussing how to explicitly solve for these crossing points, it will be apparent that these are often easier to compute for the new expansion as compared with the Edgeworth.

Finally, we will show that certain properties will always hold for the crossing points of the new expansion that are not guaranteed for the Edgeworth.

Chapter 2: Preliminaries and Notation

2.1 Introductory Concepts

Let us start by introducing some elementary probabilistic topics that will be used throughout this thesis.

Definition 1 *The* <u>*support set*</u> $\mathcal{M}$ *of a discrete random variable* X *is defined via*

$$\mathcal{M} = \mathcal{M}_x(X) = \{a : P(X = a) > 0\}$$

Definition 2 *Let* X *be a discrete random variable with support set*

$$\mathcal{X} = \mathcal{X}_h = \{x_1, x_2, ...\}$$

Then X *is said to have a* <u>*lattice*</u> *distribution if there exists a constant* b *such that, for any* j *and* k, $x_j - x_k = mb$ *for some* $m \in \mathbb{Z}$. *Notice this can only happen if there exists an* a *such that*

$$\mathcal{X} \subseteq \{a + bj, j = 0, \pm 1, \pm 2, ...\}.$$

Definition 3 *Suppose the random variable* X *has a distribution such that* $P(X \in B) = 0$ *for any set* B *consisting of a finite or countable number of real values. Then this distribution is said to be* <u>*continuous*</u>*, and* Y *is a* <u>*continuous*</u> *random variable.*

The distribution function $F = F_X$ can be defined via

$$F_X(x) = P(X \leq x),$$

and is absolutely continuous if and only if

$$F(x) = \int_{-\infty}^{x} p(t)\, dt$$

for all x, where p is a non-negative function that is integrable on the real line. This new function p is the probability density function (pdf) of our distribution, or the probability density. The pdf will only be associated with continuous distributions because by our above definition, any lattice distribution F will be purely discontinuous, since $P(X = x) > 0$ for all $x \in \mathcal{X}$.

Definition 4 *Suppose μ is a real number and σ^2 is a positive real number. Then the random variable X is a normal random variable if it has the pdf*

$$\phi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}. \tag{2.1}$$

Given these conditions, we say X has a normal (μ, σ^2) $(N(\mu, \sigma^2))$ distribution.

Definition 5 *The N(0, 1) pdf of the random variable X is denoted ϕ and is written*

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

This is also referred to as the standardized normal pdf.

Definition 6 *Consider the real-valued random variable X distribution function F. Then the expected value, or the mean of X, denoted $\mathbf{E}[X] = \mu$, is given in two ways. If X is continuous*

$$\mu = \mathbf{E}[X] = \int_{-\infty}^{\infty} x\, dF(x). \tag{2.2}$$

If X has a lattice distribution, then

$$\mu = \sum_{x \in \mathcal{M}} xP(X = x). \tag{2.3}$$

Notice that equation (2.2) is satisfied for the μ in equation (2.1). The value σ^2 used above is known as the variance of X.

Definition 7 *The* <u>*variance*</u>*,* σ^2*, of a random variable* X *is the mean of the square of the deviation of* X *from* μ*. In the discrete and continuous cases, the functions are,*

$$\sigma^2 = \sum_{i \in \mathcal{M}} (i-\mu)^2 P(X=i),$$

and

$$\sigma^2 = \int_{\infty}^{\infty} (x-\mu)^2 \, dF(x).$$

respectively.

Let k be a positive integer. Then the expected value of X^k, if it exists, is denoted α_k. This is called the k^{th} order moment, or the k^{th} raw moment of X about the origin. The formula for α_k in the discrete and continuous case are

$$\alpha_k = \mathbf{E}[X^k] = \sum_{\mathcal{M}} x^k P(X=x)$$

and

$$\alpha_k = \mathbf{E}[X^k] = \int_{-\infty}^{\infty} x^k \, dF(x).$$

respectively.

Notice that the existence of α_m for $m \geq 1$ guarantees the existance of α_i for all $i \leq m$. The raw moments of X can be used to find the central moments, which will be needed later. The central moments are the moments, of order k about the mean. They are denoted μ_k and are written as

$$\mu_k \quad = \quad \mathbf{E}[(X - \mathbf{E}[X])^k].$$

Definition 8 *The joint* <u>*cumulative distribution function*</u> F *in both discrete and continuous cases is defined via*

$$F(x_1, ..., x_n) = P(X_1 \leq x_1, ..., X_n \leq x_n).$$

Definition 9 *Let* $X_1, ..., X_n$ *be real random variables. Then they are said to be* *independent random variables, if and only if*

$$F(x_1, ..., x_n) = \prod_{k=1}^{n} F_k(x_k),$$

for every real $x_1, ..., x_n$*, where* F_k *is the distribution function corresponding to* X_k*. In the discrete case,*

$$P(X_1 = x_1, ..., X_n = x_n) = \prod_{m=1}^{n} P(X_m = x_m)$$

for real numbers $x_1, ..., x_n$.

Definition 10 *Let* $X_1, ..., X_n$ *be a sequence of random variables. Then the sequence is said to be independent and identically distributed variables if the probability distributions for all* X_i *are the same, and for any* $i, j \in \{1, 2, ..., n\}$*,* $i \neq j$*,* X_i *and* X_j *are independent random variables.*

The distribution function encapsulates fully the behavior of the random variable X. Another useful tool is the characteristic function of X. For the purpose of this thesis, we will be dealing with a sequence of independent identically distributed random variables such that the characteristic function of the standardized sum converges to that of a Normal $N(0, 1)$.

Definition 11 *The characteristic function,* χ_X*, of a random variable* X *can be written as the expected value of the function* e^{itx} *with parameter* $t \in \mathbf{R}$ *and* $i = \sqrt{-1}$*, i.e.*

$$\begin{aligned} \chi_X(t) &= \mathbf{E}[e^{itX}] \\ &= \int_{-\infty}^{\infty} e^{itx} dF(x). \end{aligned}$$

The characteristic function gives us a simple way of calculating the cumulants of a random variable X.

2.2 Cumulants

The cumulants represent the raw moments in a different way and will be used to define the expansions in this thesis. Let χ_X be the characteristic function of our distribution, then the cumulant generating function, g, is the power series expansion for $\log(\chi_X)$ (see [1]). The jth cumulants can then be found by taking the coefficient of $\frac{1}{j!}(it)^j$ in our expansion

$$\log(\chi_X(t)) = \gamma_1 it + \frac{1}{2}\gamma_2(it)^2 + ... + \frac{1}{j!}\gamma_j(it)^j + \dots .$$

However, our original expansion of the characteristic function gives us

$$\chi_X(t) = 1 + \mathbf{E}[X]it + \frac{1}{2}\mathbf{E}[X^2](it)^2 + ... + \frac{1}{j!}\mathbf{E}[X^j](it)^j + \dots ,$$

which after taking the log can give us a formal identity

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{1}{j!}\gamma_j(it)^j &= \log\left(1 + \sum_{j=1}^{\infty} \frac{1}{j!}\mathbf{E}[Y^j](it)^j\right) \\ &= \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}\left(\sum_{j=1}^{\infty} \frac{1}{j!}\mathbf{E}[Y^j](it)^j\right)^k . \end{aligned}$$

From this, we can equate the coefficients of $(it)^j$ to give us our cumulants. However, as an alternate means of calculating the γ_j we may take

$$g^{(j)}(0) = \left(\frac{d}{dt}\right)^j \log(\chi_X(0)) = \gamma_j .$$

Now we find that the first few cumulants in terms of the moments are

$$
\begin{aligned}
\gamma_1 &= \alpha \\
\gamma_2 &= \alpha_2 - \alpha \\
\gamma_3 &= 2\alpha^3 - 3\alpha\alpha_2 + \alpha_3 \\
\gamma_4 &= -6\alpha^4 + 12\alpha^2\alpha_2 - 3\alpha_2^2 - 4\alpha\alpha_3 + \alpha_4 \\
\gamma_5 &= 24\alpha^5 - 60\alpha^3\alpha_2 + 20\alpha^2\alpha_3 - 10\alpha_2\alpha_3 + 5\alpha(6\alpha_2^2 - \alpha_4) + \alpha_5.
\end{aligned}
$$

Once calculated, we can write the cumulants in terms of the central moments. This representation will be used for the remainder of this paper. The first few cumulants in terms of the central moments are

$$
\begin{aligned}
\gamma_1 &= 0 \\
\gamma_2 &= \mu_2 \\
\gamma_3 &= \mu_3 \\
\gamma_4 &= \mu_4 - 3\mu_2^2 \\
\gamma_5 &= \mu_5 - 10\mu_2\mu_3 \\
\gamma_6 &= \mu_6 - 15\mu_2\mu_4 + 30\mu_2^3.
\end{aligned}
$$

2.3 Stirling Numbers of the Second Kind

Before defining the new expansion, we will introduce the Stirling numbers of the second kind. These will become very important when proving the validity of the new expansion.

Definition 12 *The Stirling numbers of the second kind $S(n,k)$ represent the number of ways to partition a set of n elements into k nonempty subsets. It is defined by* *the explicit formula*

$$S(n,k) = \left\{ {n \atop k} \right\} = \frac{1}{k!}\sum_{j=0}^{k}(-1)^{k-(j-1)}\frac{(j-1)^n}{(j-1)!(k-j)!}, \tag{2.4}$$

where $S(n,k) = 0$, for all $k > n$.

The following recurrence relation [6] holds for Stirling numbers of the second kind.

Proposition 1 *For all n, and all $k \leq n$*

$$S(n,k) = kS(n-1,k) + S(n-1,k-1), \tag{2.5}$$

given the initial condition that

$$S(n,n) = S(n,1) = 1. \tag{2.6}$$

Proof. By definition, we know that $S(n,k)$ is the number of ways to partition a set of n elements into k nonempty subsets. Clearly, given a set of n elements, there is only one way to partition them into one nonempty subset. Similarly, there is also only one way to partition these n elements into n nonempty subsets; each subset would then consist of a single element. Therefore, the initial condition holds.

Now let us consider what it means to partition a set of n elements into k nonempty subsets. This process can be broken down into two parts. The first being partitioning the first $n-1$ elements into k nonempty subsets, then adding the final element to a subset. By definition, there are $S(n-1,k)$ ways to partition the first $n-1$ elements, and k groups to which the final element could be added. Therefore, there are $kS(n-1,k)$ ways to do this. Now the second approach is to partition the first $n-1$ elements into $k-1$ nonempty subsets, leaving the final elements as its own subset. By definition, there are $S(n-1,k-1)$ ways to do this. This gives us exactly

$$S(n,k) = kS(n-1,k) + S(n-1,k-1).$$

□

As an example, let $n = 3$ and $k = 2$. By definition, $S(3,2) = 3$. By the above lemma,

$$S(3,2) = 2S(3,1) + S(2,1) = 2((-1)(-1) + 0) + 1 = 3.$$

Definition 13 *For $x \in \mathbb{R}$, let the falling number $(x)_n$ be defined as*

$$(x)_n = \prod_{i=0}^{n-1} (x - n), \tag{2.7}$$

for all $n \in \mathbb{N}$ such that $n \geq 1$, and $(x)_0 = 1$.

Now we can introduce the following well known theorem.

Theorem 2.1 *For all Stirling numbers of the second kind,*

$$\sum_{k=1}^{n} S(n,k)(x)_k = x^n. \tag{2.8}$$

Proof. This proof follows from information given in [6]. Clearly this holds for $x = 0$. Now suppose the result holds up to $n-1$ for $n \geq 1$. Then

$$\begin{aligned}
x^n &= xx^{n-1} \\
&= x\sum_{k=1}^{n-1} S(n-1,k)(x)_k \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_k(x-k+k) \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_{k+1} + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k \\
&= \sum_{j=2}^{n-1} S(n-1,j-1)(x)_j + \sum_{k=1}^{n-1}[S(n,k) - S(n-1,k-1)](x)_k \\
&= \sum_{k=2}^{n-1} S(n-1,k-1)(x)_k + \sum_{k=1}^{n-1} S(n,k)(x)_k - \sum_{k=1}^{n-1} S(n-1,k-1)(x)_k \\
&= \sum_{k=1}^{n-1} S(n,k)(x)_k + [S(n-1,n-1)(x)_k - S(n-1,0)(x)_k] \\
&= \sum_{k=1}^{n-1} S(n,k)(x)_k + S(n,n)(x)_k \\
&= \sum_{k=1}^{n} S(n,k)(x)_k
\end{aligned}$$

Therefore, by induction, the above equality holds. □

As an example, let $n = 5$. Then we have

$$\begin{aligned}
\sum_{k=1}^{5} S(5,k)(x)_k &= S(5,1)(x)_1 + S(5,2)(x)_2 + S(5,3)(x)_3 + S(5,4)(x)_4 + S(5,5)(x)_5 \\
&= (x)_1 + 15(x)_2 + 25(x)_3 + 10(x)_4 + (x)_5 \\
&= x + 15(x)(x-1) + 25(x)(x-1)(x-2) + ... \\
&= x^5 + x^4(10-1-2-3-4) + x^3(25-10-20-30+35) + \\
&\quad +x^2(15-75+110-50) + x(1-15+50-60+24) \\
&= x^5.
\end{aligned}$$

Chapter 3: Hermite Polynomials

3.1 Definition and Elementary Properties

In order to properly define the upcoming expansions, we must first look at what are called the Hermite polynomials. These polynomial sequences appear in different applications ranging from probability to quantum mechanics. They also have the property of being orthogonal with respect to the standard normal probability density function (see [6]). It was the French mathematician Charles Hermite (1822-1901), for whom they are named, who studied these extensively in the mid 1800's.

Definition 14 *The Hermite polynomials are defined in one of two ways. The first form is often employed by physicists and is defined via*

$$H_n^{phys}(x) = 2^{n/2}(-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}(\sqrt{2}x). \tag{3.1}$$

The second form, used by probabilists, is defined via

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

Notice these two definitions are slightly different. They have the following relationship

$$H_n^{phys}(x) = 2^{n/2} H_n(\sqrt{2}x).$$

For the purpose of this paper, we are only concerned with the probabilistic form of the Hermite polynomial. This can be rewritten by using the power series expansion of e (see [5]), giving us

$$H_n(x) = n! \sum_{k=0}^{n/2} \frac{(-1)^k x^{n-2k}}{k!(n-2k)!2^k} \tag{3.2}$$

for all $n \in \mathbb{N}_0$. These polynomials are well known for being solutions to the differential equations

$$H_n^{''}(x) - xH_n^{'}(x) + nH_n(x) = 0 \tag{3.3}$$

$$y'' + (2n+1-x^2)y = 0 \tag{3.4}$$

for $y = e^{-x^2/2}H_n$.

Proposition 2 *The probabilistic form of the Hermite polynomial satisfies the recursive relation:*

$$\begin{aligned} H_0 &= 1 \\ H_1 &= x \\ H_{n+1}(x) &= xH_n(x) - nH_{n-1}(x) \end{aligned}$$

Proof. The proof of this is modeled after that in [7]. Using the expanded definition of our probabilistic Hermite polynomial (3.2), we see

$$H_n(x) = x^n - \frac{n(n-1)x^{n-2}}{2} + \frac{n(n-1)(n-2)(n-3)x^{n-4}}{2\cdot 4} - \ldots.$$

If we differentiate this expansion once, we get

$$\begin{aligned} \frac{dH_n(x)}{dx} &= n\left[(x^{n-1} - \frac{(n-1)(n-2)x^{n-2}}{2} + \frac{(n-1)(n-2)(n-3)x^{n-4}}{2\cdot 4} - \ldots\right] \\ &= nH_{n-1}(x). \end{aligned}$$

Then, if we differentiate H_n a second time, we get

$$\begin{aligned} \frac{d^2H_n(x)}{dx^2} &= n(n-1)\left[(x^{n-2} - \frac{(n-2)(n-3)x^{n-4}}{2} + \frac{(n-2)(n-3)(n-4)x^{n-5}}{2\cdot 4} - \ldots\right] \\ &= n\frac{dH_{n-1}(x)}{dx} \\ &= n(n-1)H_{n-2}(x). \end{aligned}$$

Finally, if we combine these two derivatives with the equation (3.3), we get

$$n(n-1)H_{n-2}(x) - xnH_{n-1} + nH_n = 0,$$

which simplifies to our recurrence relation

$$H_{n+1} = xH_n - nH_{n-1}.$$

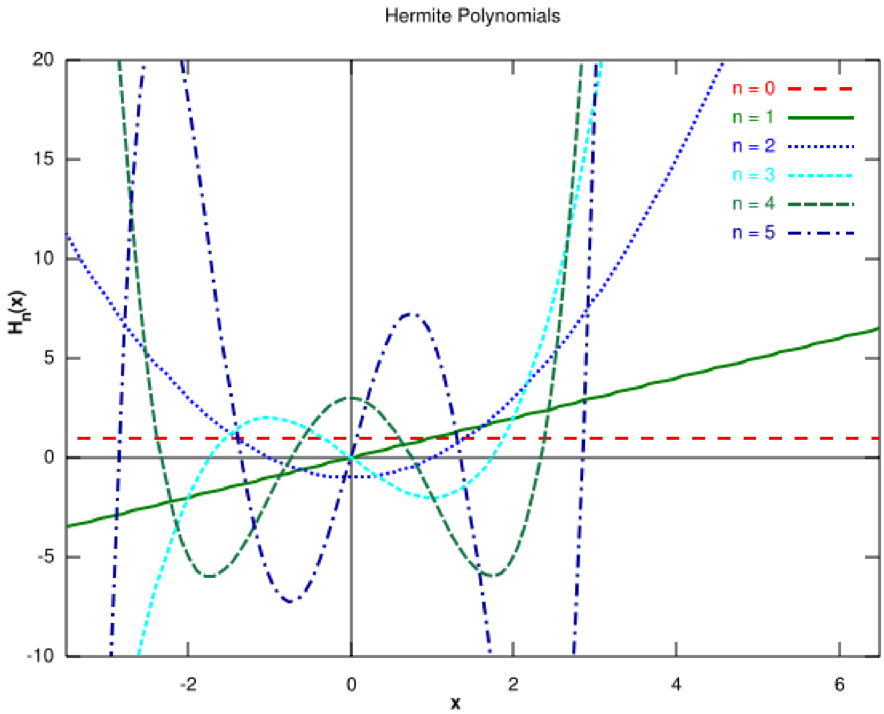
□

The first ten probabilistic Hermite polynomials are as follows

$$\begin{aligned}
H_0 &= 1 \\
H_1 &= x \\
H_2 &= x^2 - 1 \\
H_3 &= x^3 - 3x \\
H_4 &= x^4 - 6x^2 + 3 \\
H_5 &= x^5 - 10x^3 + 15x \\
H_6 &= x^6 - 15x^4 + 45x^2 - 15 \\
H_7 &= x^7 - 21x^5 + 105x^3 - 105x \\
H_8 &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105 \\
H_9 &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x \\
H_{10} &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945.
\end{aligned}$$

3.2 Locating the Zeros of Hermite Polynomials

We can see from the above that the Hermite polynomials become very complicated very quickly. Determining the zeros to these polynomials has been studied in several different ways. To date there is no closed form solution for finding these zeros; however, there have been various properties proven about them. Note that many of the ideas mentioned in this section are based on work in [3].



From the graph above [1], we see trends in the first few H_n. For instance, notice that all of the zeros appear to be are simple and real. This begs the question whether the zeros of H_n are always simple and real. Also notice that for all odd n, $x = 0$ will be a zero of H_n. This is clear from equation (3.2).

If we examine the polynomials themselves, we see that for even n, the power of x is even for every non-zero term of H_n. Therefore, for even n, H_n is an even function. Likewise, for all odd n, the power of x is odd in every non-zero term. Thus, for odd n, H_n is an odd function. This tells us an important fact about the zeros. Since H_n is either an even or odd function for all n, for all zeros x_0 of H_n,

$$H_n(x_0) = H_n(-x_0) = 0$$

(assuming all real zeros).

Now we have that when locating the zeros of H_n, it suffices to consider all distinct

[1] $http://www.absoluteastronomy.com/topics/Hermite\_polynomials$

zeros up to sign.

Definition 15 *Let f be a polynomial of degree n. Also let $\{x_1, ..., x_n\}$ be the zeros of f (including multiplicity). Then define x_i to be a simple zero of f if for all $j \neq i$, $x_j \neq x_i$.*

Definition 16 *Let x_0 be zero of a polynomial f of degree n. We can find a function g such that*

$$f(x) = (x - x_0)^k g(x),$$

where g is a polynomial of degree $n - k$ satisfying $g(x_0) \neq 0$. We refer to such a k as the multiplicity of the zero x_0.

Proposition 3 *Let f be a polynomial of degree n. Assume x_0 is a non-simple zero of f, if multiplicity equals k (where $1 < k \leq n$). Then*

$$f^{(i)}(x_0) = 0$$

for all $i \leq k - 1$.

Proof. To prove this, let $f(x) = (x - x_0)^k g(x)$, where g is a polynomial of degree $n - k$. The proposition follows by the product rule since each term in the expression for $f^{(i)}$ for $i \leq k - 1$ will contain a factor of $(x - x_0)$, and therefore $f^{(i)}(x_0) = 0$ for $i \leq k - 1$. □

Now we can relate this back to our Hermite polynomials with the following theorem.

Theorem 3.1 *Every zero of H_n is simple for all n*

Proof. For this proof, we let x_0 be a zero with multiplicity k of H_n. Then using the differential equation used to define H_n, we get

$$\begin{aligned} H_n^{''}(x_0) - x_0 H_n^{'}(x_0) + nH_n(x_0) &= 0 \\ H_n^{''}(x_0) &= x_0 H_n^{'}(x_0). \end{aligned}$$

Once we have this relation, we can take the derivative again to get

$$H_n^{'''}(x_0) = x_0 H_n^{''}(x_0) = x_0^2 H_n^{'}(x_0).$$

Inductively we obtain

$$H_n^{(k)}(x_0) = x_0^{k-1} H_n^{'}(x_0).$$

So by Proposition 3, if x_0 is not a simple zero, then $H_n^{(n)}(x_0) = 0$. Thus, x_0 must be the only zero of H_n. Also note that by definition, the coefficients of every other term in the Hermite polynomial is zero. Hence

$$H_n(x) \neq (x - x_0)^n.$$

Therefore, x_0 must be a simple zero of H_n, and all zeros must be simple. □

Lemma 1 *Let x_0 be a simple zero of a monic polynomial f. Then any circle through the points x_0 and*

$$x_0^{'} = x_0 - 2(n-1)\frac{f^{'}(x_0)}{f^{''}(x_0)}$$

contains some zeros of $f(x)$ in both domains bounded by it, unless all the zeros lie on the circumference of the circle.

Proof. (see Section 6.2 in [2]) □

Implying lemma 3.2, we have the following theorem.

Theorem 3.2 *Every zero of H_n is real.*

Proof. First let x_0 be a zero of H_n. Then define

$$x_0' = x_0 - 2(n-1)\frac{H_n''(x_0)}{H_n'(x_0)}$$
$$= x_0 - \frac{2(n-1)}{x_0}.$$

Now assume x_0 is a non-real zero of H_n with the greatest imaginary part $\mathcal{I}(x_0)$. Then $\mathcal{I}(x_0')>\mathcal{I}(x_0)$ if $\mathcal{I}(x_0) > 0$ since for $z = a + ib$,

$$\frac{1}{a+ib} = \frac{a-ib}{a^2+b^2}.$$

A circle can now be drawn through x_0' and x_0 which contains no zeros of H_n. By the theorem above, the zeros must all lie on this circle. However, if they did they would all have to coincide with x_0, which is impossible. Therefore, x_0 must be real, and H_n must have all real zeros. □

Proposition 4 *Let x_0 be the largest zero of H_n, then*

$$x_0 > \sqrt{n-1}.$$

Proof. If x_0 is the largest zero of H_n, then given the symmetry of these polynomials, $-x_0$ would be the smallest zero. Then from the theorem above, we obtain the following inequality

$$-x_0 < x_0' = x_0 - \frac{2(n-1)}{x_0} < x_0. \tag{3.5}$$

From equation (3.5) it follows that

$$x_0 - \frac{2(n-1)}{x_0} > x_0$$
$$x_0 > \sqrt{n-1}$$

for $n > 2$. In the case of $n = 2$, we have an equality. □

This gives us a lower bound for the greatest zero of H_n. This turns out to be a somewhat weak bound on which one can improve.

Proposition 5 *Let x_0 be the largest zero of H_n, then*

$$|x_0| \leq \frac{2(n-1)}{\sqrt{5n+2}}.$$

Proof. First, from Equation (3.3), and by the product rule we have

$$H_n^{''}(x_0) = x_0 H_n^{'}(x_0)$$

$$H_n^{'''}(x_0) = (x_0^2 + 1 - n) H_n^{'}(x_0).$$

Now, let $H_n(x) = (x - x_0)h(x)$. Recall that by definition, H_n is a monic polynomial for all n. Notice again by the product rule, $h(x_0) = \frac{1}{3} H_n(x_0)$. We can rewrite h in terms of its factors as

$$h(x) = \prod_{i=1}^{n-1} (x - x_i).$$

Now, if we follow a similar process to that above, we get

$$\begin{aligned}
\log(h(x)) &= \sum_{i=1}^{n-1} \log(x - x_i) \\
\frac{d^2}{dx_0^2}(\log(h(x_0))) &= \sum_{i=1}^{n-1} \frac{1}{(x_0 - x_i)^2} = \frac{d}{dx_0}\left(\frac{h^{'}(x_0)}{h(x_0)}\right) \\
&= \frac{h^{'}(x_0) - h(x_0)h^{''}(x_0)}{h(x_0)^2} = \frac{3H_n^{''}(x_0) - 4H_n^{'}(x_0)H_n^{'''}(x_0)}{12H_n^{'}(x_0)}.
\end{aligned}$$

According to the Cauchy-Schwartz inequality,

$$\begin{aligned}
\left((n-1)\sum_{i=1}^{n-1} \frac{1}{x_0 - x_i}\right)^2 &= \left(\frac{H_n^{''}(x_0)}{2H_n(x_0)}\right)^2 \leq (n-1)\sum_{i=1}^{n-1} \frac{1}{(x_0 - x_i)^2} \\
0 &\leq (n-1)\frac{3H_n^{''}(x_0) - 4H_n^{'}(x_0)H_n^{'''}(x_0)}{12H_n^{'}(x_0)} - \frac{H_n^{''}(x_0)}{4H_n(x_0)} \\
0 &\leq 3(n-2)(H_n^{''}(x_0))^2 - 4(n-1)H_n^{'}(x_0)H_n^{'''}(x_0).
\end{aligned}$$

Finally, if we apply the equations from above, we get

$$H_n^{'} x_0^2(3(n-2)x_0^2 - 4(n-1)(2x_0^2 - n + 1)) \geq 0.$$

Solving for x_0 then yields

$$|x_0| \geq \frac{2(n-1)}{\sqrt{5n-2}}.$$

□

We close this section with a few notes regarding the constant term of H_n. Notice that from the expansion of H_n, we can calculate the last term of H_n specifically as

$$\frac{n!}{\left(\frac{n}{2}\right)!2^{n/2}} = (1 \cdot 3 \cdot 5 \cdots (n-1))$$

for even n, and

$$-\frac{(n-1)!x}{\left(\frac{n-1}{2}\right)!2^{(n-1)/2}} = -(1 \cdot 3 \cdot 5 \cdots (n-1))x$$

for all odd n ([5]).

By definition, we know that the coefficient for the x^n term of H_n will always be one. Therefore, if we completely factor this polynomial, we should get

$$H_n(x) = \prod_{i=1}^{n}(x - x_i).$$

Combining this with what we know about the last term of the polynomial, we obtain

$$\prod_{X_i: root of H_n} (x_i) = (x^0 \text{ term of } H_n(x))$$

for even n,

$$\prod_{X_i: non-zero root of H_n} (x_i) = (\text{coefficient of } x \text{ term of } H_n(x)).$$

for odd n. This holds true since all of the zeros are real (by Theorem 3.3).

Chapter 4: Edgeworth Expansion

The Edgeworth expansion was derived in 1905, and was named in honor or Francis Ysidro Edgeworth. It relates the probability function of a random variable to that of a standard normal distribution. The expansion itself expresses the probability distribution in terms of its cumulants, and the Hermite polynomials. Its associated series is a rearrangement of the Gram-Charlier A Series (see [1]).

Suppose $\{X_n\}_{n\geq 1}$ is a sequence of independent random variables satisfying $\mathbf{E}[X] = \mu_i$ and $Var(X) = \sigma_i^2 < \infty$ for $i \geq 1$. Then we have

$$\begin{aligned} Var(X_i) &= \mathbf{E}[X_i^2] - \mu_i^2 \\ \alpha_2 = \alpha_{2,i} &= \mathbf{E}[X_i^2] = Var(X_i) + \mu_i^2 < \infty \end{aligned}$$

for all i.

Before we can define the expansion, we first need to introduce some notation. Let $\gamma_{v,i}$ be the cumulant of order v of our random variable X_i. Then for $n \geq 1$, let

$$B_n = \sum_{j=1}^{n} \sigma_j^2$$

$$\lambda_{v,n} = \frac{n^{(v-2)/2}}{B_n^{v/2}} \sum_{j=1}^{n} \gamma_{v,j} \tag{4.1}$$

$$\lambda_{m+2,n}^* = \frac{\lambda_{m+2,n}}{(m+2)!}. \tag{4.2}$$

We now combine these and define

$$q_{v,n}(x) \stackrel{def}{=} \sum H_{v+2s}(x) \prod_{m=1}^{v} \frac{1}{k_m!} (\lambda_{m+2,n}^*)^{k_m} \tag{4.3}$$

where the summation in (4.3) goes over all nonnegative integer solutions, $(k_1, k_2, ..., k_r)$, satisfying

$$v = \sum_{i=1}^{v} ik_i, \tag{4.4}$$

and

$$s = \sum_{i=1}^{v} k_i. \tag{4.5}$$

The following table below illustrates the solutions (as well as the corresponding value of $v + 2s$) for $v = 1, 2, 3, 4$.

Table 4.1: Computation of k_i for q_v

r \i	1	2	3	4	s	**v+2s**
1	1	0	0	0	1	**3**
2	2	0	0	0	2	**6**
	0	1	0	0	1	**4**
3	3	0	0	0	3	**9**
	1	1	0	0	2	**7**
	0	0	1	0	1	**5**
4	4	0	0	0	4	**12**
	2	1	0	0	3	**10**
	1	0	1	0	2	**8**
	0	2	0	0	2	**8**
	0	0	0	1	1	**6**

Now, suppose that X has probability function p, where p is either a density or a probability mass function, and let $p^{(n)}$ be the probability function for the partial sum $S_n = X_1 + X_2 + ... + X_n$. By the central limit theorem,

$$S_n = n^{1/2} \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma}$$

is asymptotically normally distributed with zero mean and unit variance. Further,we have under mild conditions

$$p_n(x) = \phi_{\sigma,n}(x) + \sum_{v=1}^{k-2} \frac{q_{v,n}(x)}{n^{v/2}} + o\left(\frac{1}{n^{(k-2)/2}}\right) \tag{4.6}$$

(see Theorem 5.1).

Now consider the particular case in which the random variables $X_1, X_2, ..., X_n$ have identical distributions with a positive variance σ^2, zero mean, and $\mathbf{E}[X] = \mu$. Since all random variables under consideration have identical distributions, Equations (4.1), (4.2), and (4.3) no longer depend on n. Therefore, with

$$\lambda_v = \frac{\gamma_v}{\sigma^v} \tag{4.7}$$

$$\lambda_v^* = \frac{\lambda_v}{v!} \tag{4.8}$$

$$q_v(x) = \sum H_{v+2s}(x) \prod_{m=1}^{v} \frac{1}{k_m!} (\lambda_{m+2}^*)^{k_m}. \tag{4.9}$$

We obtain

$$p_n(x) = \phi_{\sigma,n}(x) + \sum_{v=1}^{k-2} \frac{q_v(x)}{n^{v/2}} + o\left(\frac{1}{n^{(k-2)/2}}\right) \tag{4.10}$$

for $k \geq 3$.

Chapter 5: New Expansion

As in the previous chapter, suppose X is a random variable satisfying $\mathbf{E}[X] = \mu$ and $Var(X) = \sigma^2 < \infty$ and that $\{X_n\}_{n \geq 1}$ is a sequence of independent random variables. Then we have

$$\begin{aligned} Var(X_i) &= \mathbf{E}[X_i^2] - \mu^2 \\ \alpha_2 = \alpha_{2,i} &= \mathbf{E}[X_i^2] = Var(X_i) + \mu^2 < \infty \end{aligned}$$

for all i. For simplistic purposes, from here forward, we will assume $\mu = 0$. The results still hold for $\mu \neq 0$ by a shift in variable. Also assume that every moment of X exists and is finite.

Let

$$\eta_h = \eta_h^{(n)} \tag{5.1}$$

be the probability function for

$$\tilde{S_n} = X_1 + ... + X_h + N_{h+1} + ... + N_n \tag{5.2}$$

where the N_i are independent identically distributed $N(\mu, \sigma^2)$ random variables also independent of the X_i. Note that for small h and small support sets, it is quite simple to compute η_h via conditioning.

We will now state a theorem of Petrov (see [2]) regarding Edgeworth expansions for not necessarily independent identically distributed random variables.

Theorem 5.1 *Let X_n be a sequence of independent random variables with zero mean, and set*

$$w_n(t) = \mathbf{E}[e^{itX_n}]$$

$$B_n = \sum_{j=1}^{n} \mathbf{E}[X_j^2].$$

In addition, suppose that

$$\liminf \frac{B_n}{n} > 0 \tag{5.3}$$

$$\limsup \frac{1}{n} \sum_{j=1}^{n} \mathbf{E}[|X_j|^k] < \infty \tag{5.4}$$

and

$$\frac{1}{n} \sum_{j=1}^{n} \int_{|x|>n^{\tau}} |x|^k dT_j(x) \to 0 \tag{5.5}$$

for some integer $k \geq 3$ and some positive $\tau < 1/2$. Suppose also that

$$\int_{|t|>\epsilon} \prod_{j=1}^{n} |w_j(t)| dt = o\left(\frac{1}{n^{(k-1)/2}}\right) \tag{5.6}$$

for every fixed $\epsilon > 0$. Then the probability function p_n of the random variable

$$\frac{1}{\sqrt{B_n}} \sum_{j=1}^{n} X_j \tag{5.7}$$

satisfies

$$p_n(x) = \phi_{\sigma,n}(x) + \sum_{v=1}^{k-2} \frac{q_{v,n}(x)}{n^{v/2}} + o\left(\frac{1}{n^{(k-2)/2}}\right) \tag{5.8}$$

uniformly in x.

With the goal of applying the above theorem of Petrov (Theorem 5.1), we will prove some technical lemmas.

Not that $\liminf \frac{B_N}{n} > 0$ and (5.3) is satisfied.

Lemma 2 *Define*

$$T_j(x) = P(W_j \leq x)$$

$$W_i = \begin{cases} X_i & i \leq h \\ N_i & i > h \end{cases}.$$

Then we have that for $m > 0$

$$\limsup \frac{1}{n} \sum_{j=1}^{n} \mathbf{E}[|T_j|^m] < \infty.$$

Proof. In the case of both X_i and N_i, the mth moments will all be finite for all i. We will call these moments $\mu_m^{X_i}$ and $\mu_m^{N_i}$. Then we have

$$\frac{1}{n}\sum_{j=1}^{n}\mathbf{E}[|T_j|^m] = \frac{1}{n}(h\mu_m^{X_1} + (n-h)\mu_m^{N_1}) < \max(\mu_m^{X_1}, \mu_m^{N_1}) < \infty$$

for all m, i. Thus,

$$\limsup \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}[|T_j|^m] < \limsup \max(\mu_m^{X_1}, \mu_m^{N_1}) = \max(\mu_m^{X_1}, \mu_m^{N_1}) < \infty$$

for all m, i. □

Lemma 3 *For $k = 4$ and $\tau = 1/4$,*

$$\frac{1}{n}\sum_{j=1}^{n}\int_{|x|>n^\tau}|x|^k dW_j(x) \to 0. \tag{5.9}$$

Proof. Let

$$\alpha(x) = P(X_i \leq x), \quad \beta(x) = P(N_i \leq x)$$

and

$$a_n = \int_{|x|>n^{\frac{1}{4}}}|x|^4 da(x), \quad b_n = \int_{|x|>n^{\frac{1}{4}}}|x|^4 db(x).$$

Then

$$\frac{1}{n}\sum_{j=h+1}^{n}\int_{|x|>n^{\frac{1}{4}}}|x|^4 dW_j(x) = \frac{(n-h)b_n}{n}. \tag{5.10}$$

Clearly

$$\frac{ha_n}{n} \to 0.$$

Note that since the random variables considered in the computation of b_n are $N(0, \sigma^2)$, and $x^4 e^{-\frac{x^2}{\sigma^2}} \leq e^{-x}$ for sufficiently large x,

$$\begin{aligned} b_n &\leq C\int_{|x|>n^{\frac{1}{4}}} e^{-x}dx \\ &= 2e^{-n^{\frac{1}{4}}}. \end{aligned}$$

for sufficiently large n and some constant C. This tends to zero as n tends to infinity, and the result follows. □

Lemma 4 *Let* $w_n(t) = \mathbf{E}[e^{itW_n}]$. *Then for any* $\epsilon > 0$,

$$\int_{|t|>\epsilon} \prod_{j=1}^{n} |w_j(t)| dt = o\left(\frac{1}{n^{(k-1)/2}}\right) \tag{5.11}$$

Proof. Let

$$a(t) = \mathbf{E}[e^{itX_1}]$$

and

$$b(t) = \mathbf{E}[e^{itN_{h+1}}].$$

Define

$$w_n(t) = a(t)^h b(t)^{n-h}.$$

We know $|a(t)|^h \leq 1$ (see for instance [8]). Similar to the last proof, we can use the fact that

$$b(t) = e^{-\frac{1}{2}t^2} < e^{-t}.$$

Then

$$\begin{aligned} \int_{|t|>\epsilon} \prod_{j=1}^{n} |b(t)| dt &= \int_{\epsilon}^{z} (e^{-t})^{n-h} \\ &= \frac{e^{-\epsilon(n-h)} - e^{-z(n-h)}}{n-h}. \end{aligned}$$

This gives

$$\int_{|t|>\epsilon} \prod_{j=1}^{n} |w_j(t)| dt = \lim_{z \to \infty} \frac{e^{-\epsilon(n-h)} - e^{-z(n-h)}}{n-h} = \frac{1}{(n-h)e^{\epsilon(n-h)}} \leq \frac{1}{n^{\frac{k-1}{2}}}$$

for $\epsilon > 0$. □

Before we begin, we will need to regroup the variables for both the Edgeworth and new expansions. In the Edgeworth expansion, given a specific order of n, we are looking to group each term of that order into its own individual variable. So let

$$\beta_{a,b} = \left(\frac{1}{\sqrt{n}}\right)^a c_{ab}, \tag{5.12}$$

where c_{ab} is everything in the bth order of the λ's in the ath order of $n^{-1/2}$ coefficient in the expansion. From this we can give a general form of f_n via

$$f_n(M) = \sum_{a=0}^{s}\sum_{b=0}^{a} \beta_{a,b} + o\left(\frac{1}{n^{s/2}}\right) \tag{5.13}$$

for $s \geq 1$

The new expansion will be grouped a little differently. Based on the definition of η_h, for a given term of a specific order of $n^{-1/2}$, every γ_k^* in (4.8) will be multiplied by a factor of h/n due to the definitions in (4.1) through (4.3) and the fact that every cumulant for the N_i in (5.2) is zero.

Now let $\eta_h^*(M) = \frac{n}{h}(\eta_h(M) - \phi_{\sigma,n}(x))/\phi_{\sigma,n}(x)$. Then we can denote individual terms of η_k^* as

$$\nu_{a,b}^h(M) = \left(\frac{1}{\sqrt{n}}\right)^a \frac{c_{a,b}h^{b-1}}{n^{b-1}}. \tag{5.14}$$

The general form of η^* can now be defined by

$$\eta_h^*(M) = \sum_{a=0}^{s}\sum_{b=0}^{a} \nu_{a,b}^h(M) + o\left(\frac{1}{n^{s/2}}\right). \tag{5.15}$$

We then apply the delta difference operator to each individual term with respect to h to obtain

$$\Delta^{j-1}\nu_{a,b}^1(M) = \sum_{i=1}^{j}(-1)^{j-i-1}\binom{j-1}{i-1}\nu_{a,b}^i(M),$$

which then gives

$$\Delta^{j-1}\eta_1^*(M) = \sum_{i=1}^{j}(-1)^{j-i-1}\binom{j-1}{i-1}\eta_i^*(M).$$

From all of this, we can introduce the following proposition.

Proposition 6 *For any a,b, we have*

$$\sum_{j=1}^{\infty} \binom{n-1}{j-1} \Delta^{j-1} \nu^1_{a,b}(M) = \beta_{a,b}. \tag{5.16}$$

where $\beta_{a,b}$ *is as in (5.12).*

Proof. This follows via the following straightforward computations.

$$\begin{aligned}
\sum_{j=1}^{\infty} \binom{n-1}{j-1} \Delta^{j-1} \nu^1_{a,b}(M) &= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} \sum_{j=1}^{\infty} \sum_{i=0}^{j} \binom{n-1}{j-1} (-1)^{j-i-1} \binom{j-1}{i-1} \frac{i^{b-1}}{n^{b-1}} \\
&= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} \sum_{j=1}^{b} \binom{n-1}{j-1} \frac{(j-1)!S(b,j)}{n^{b-1}} \\
&= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} \sum_{j=1}^{b} \frac{(n-1)!S(b,j)}{((n-1)-(j-1))!n^{b-1}} \\
&= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} \sum_{j=1}^{b} \frac{(n)_j S(b,j)}{n^b} \\
&= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} \frac{1}{n^b} \sum_{j=1}^{b} (n)_j S(b,j) \\
&= \left(\frac{1}{\sqrt{n}}\right)^a c_{a,b} = \beta_{a,b}.
\end{aligned}$$

□

Referring back to (5.13) we have

$$\begin{aligned}
f_n(M) &= \sum_{a=1}^{\infty} \sum_{b=1}^{a} \sum_{j=1}^{n} \binom{n-1}{j-1} \Delta^{j-1} \nu^1_{a,b}(M) \\
&= \sum_{j=1}^{n} \binom{n-1}{j-1} \Delta^{j-1} \eta^*_1(M).
\end{aligned}$$

From this, we have the estimate

$$V_s(M) = \sum_{j=1}^{s} \binom{n-1}{j-1} \Delta^{j-1} \eta^*_1(M). \tag{5.17}$$

Now we only need to account for the error term of V. We see from above that the error term of the Edgeworth expansion is $o\left(\frac{1}{n^{(k-2/2)}}\right)$. Based on the definition above, the error term of V_s will depend on s in a similar way. However, here we are also affected by the n appearing in (5.17). This can be accounted for by simply carrying out the expansion until the necessary error term is reached. This will require investigation of further terms in the expansion. But the new terms will be of smaller order then the error term, so they will be absorbed back into it. This gives us the following theorem.

Theorem 5.2 *For any $s \in \mathbf{Z}$ such that $s \leq n-1$, we have the following equality:*

$$\begin{aligned} f_n(M) &= \sum_{j=1}^{s} \binom{n-1}{j-1} \Delta^{j-1} \eta_1^*(M) + o\left(\frac{1}{n^{s/2}}\right) \\ &= V_s(M) + o\left(\frac{1}{n^{s/2}}\right). \end{aligned}$$

Notice the error term here is somewhat different than for the Edgeworth expansion. Let us examine what makes up the error term of V_s. If expanded out to $s = 1$, we simply have η_1^*. If we look back at the expansion of η_1^*, we see that every term with exactly k γ_i's in any coefficient will be accompanied by a factor of

$$\frac{h}{n}^{k-1} = \frac{1}{n}^{k-1}.$$

However, for $k = 1$ the terms match the corresponding one from f_n. So the error term of V_1 equals that of E_1 minus all terms containing a single γ_i. This can be generalized to say that the error term of V_s will contain all remaining terms in the error of E_k minus such terms containing s or fewer γ_i. Therefore, the above error term of V_s will in general be different the error term of the Edgeworth expansion. As we shall see later this can prove advantageous.

Chapter 6: Discussion

6.1 Lattice Case

In this section, we examine the lattice distribution as defined in Definition 2 with respect to the expansions introduced earlier. Let X have support set $\{0, 1, 2, ..., t\}$ with

$$P(X = i) = p_i$$

for $i = 0, 1, 2, ..., t$. Assuming that X is aperiodic, we have $\gcd\{i : p_i > 0\} = 1$. Now consider the partial sum $S_n = \sum_{i=1}^{n} X_i$. Note that $\mathbf{E}[S_n] = n\mu$ and $Var(S_n) = n\sigma^2$. Note that every moment of X exists and is finite. Set

$$x = x(M) = \frac{M - n\mu}{\sigma\sqrt{n}}.$$

Then by the Edgeworth expansion, we have

$$p_n = \phi_{\sigma,n}(x)(f_n(M) + \phi_{\sigma,n}(x)) \tag{6.1}$$

where

$$f_n(M) = \sum_{v=1}^{k-2} \frac{q_v(x)}{n^{v/2}} + o\left(\frac{1}{n^{\frac{k-2}{2}}}\right) = E_k(M) + o\left(\frac{1}{n^{\frac{k-2}{2}}}\right) \tag{6.2}$$

and

$$\phi_{\sigma,n}(x) = \frac{1}{\sigma\sqrt{2n\pi}} e^{-x^2/2}. \tag{6.3}$$

From this we see that the first few lower order estimates of E_k are as follows.

For $k = 2$, (i.e. the Central Limit Theorem)

$$f_n(M) = E_2(M) + o(1) = o(1).$$

For $k = 3$

$$
\begin{aligned}
f_n(M) &= E_3(M) + o\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

For $k = 4$

$$
\begin{aligned}
f_n(M) &= E_4(M) + o\left(\frac{1}{n}\right) \\
&= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{H_6(x)(\lambda_3^*)^2 + H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right).
\end{aligned}
$$

For $k = 5$

$$
\begin{aligned}
f_n(M) &= E_5(M) + o\left(\frac{1}{n^{3/2}}\right) \\
&= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{H_6(x)(\lambda_3^*)^2 + H_4(x)\lambda_4^*}{n} \\
&\quad + \frac{H_9(x)\lambda_3^3 + H_7(x)\lambda_4^*\lambda_3^* + H_5(x)\lambda_5^*}{n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right).
\end{aligned}
$$

This represents the Edgeworth expansion for our lattice case. Notice that k just represents the number of terms in which this Edgeworth expansion is carried out. Then the error term represents all the remaining terms not included in this partial expansion, i.e.

$$
f_n(M) = \sum_{v=1}^{\infty} \frac{q_v(x)}{n^{v/2}}.
$$

Now let us assume that $\eta_h = \eta_h^{(n)}$ is the density function for $X_1 + X_2 + ... + X_h + N_{h+1} + ... + N_n$ where $n > h$ and N_i are independent identically distributed $N(0, \sigma^2)$ random variables also independent of X_i.

Using the same logic as above, it can be easily proven that the requirements of Theorem 5.1 are met. Therefore, we can use this to obtain

$$\eta_h(M) = \frac{(h/n)H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{(h^2/n^2)H_6(x)(\frac{\lambda_3^*}{2})^2 + (h/n)H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right).$$

This pattern will continue for however long we choose to expand out the series. Now using the above definition, we obtain the new expansion for η_h^* as

$$\eta_h^*(M) = \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{(h/n)H_6(x)(\frac{\lambda_3^*}{2})^2 + H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right).$$

Notice here the factors of h/n have all been reduced by a power of one. These will now be used to obtain the new expansion. First, let us compare the first term of this partial sum to the Edgeworth expansion. From above we see that.

$$\begin{aligned} V_1(M) &= \eta_1^*(M) + o\left(\frac{1}{\sqrt{n}}\right) \\ &= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right) \\ &= f_n(M) + o\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Let us carry out this expansion to match the $(1/n)$ order term for the Edgeworth expansion given by the formula in Theorem 5.2.

$$\begin{aligned} V_2(M) &= \eta_1^*(M) + (n-1)(\eta_2^*(M) - \eta_1^*(M)) + o\left(\frac{1}{n}\right) \\ &= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{H_6(x)(\frac{\lambda_3^*}{2})^2 + H_4(x)\lambda_4^*}{n} + o(1) \end{aligned}$$

Notice that multiplying the expansion by n has raised the power of the error term, therefore, we need to account for this by expanding out a few more terms. In general, we will need to consider $2(k-2)$ additional terms to attain the proper order of the error term. One can see that by doing this, the extra terms do not match their corresponding terms of f_n.

$$
\begin{aligned}
&= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{H_6(x)(\frac{\lambda_3^*}{2})^2 + H_4(x)\lambda_4^*}{n} + \frac{\frac{7n-6}{n^3}H_9(x)(\frac{\lambda_3}{3!})^3 + H_7(x)\lambda_4^*\lambda_3^* + H_5(x)\lambda_5^*}{n^{3/2}} \\
&\quad + \frac{\frac{16n-15}{n^4}H_{12}(x)(\frac{\lambda_3^*}{4!})^4 + H_8(x)((\frac{\lambda_4^*}{2})^2 + \lambda_5^*\lambda_3^*) + H_6(x)\lambda_6^*}{n^2} + o\left(\frac{1}{n}\right) \\
&= E_(4)(M) + o\left(\frac{1}{n}\right).
\end{aligned}
$$

The error term is now of the order $(1/n)$. Notice that the final two terms of the expansion are of an order smaller than the error term. This means these terms will have less effect on the result than the error term. Therefore, the fact that the extra terms didn't match their corresponding terms of f_n was of no consequence, as they can now be absorbed back into the error term. This leaves us with

$$
\begin{aligned}
V_2(M) &= \frac{H_3(x)\lambda_3^*}{\sqrt{n}} + \frac{H_6(x)(\frac{\lambda_3^*}{2})^2 + H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right) \\
&= f_n(M) + o\left(\frac{1}{n}\right).
\end{aligned}
$$

Therefore, the two expansions up to the error term are the same. This can be expressed with the following relation (since $s = k - 2$)

$$
E_{k+2}(M) = V_s(M) + o\left(\frac{1}{n^{(k+2)/2}}\right).
$$

However, the error terms themselves are very different. The claim is that the error term in the new expansion is smaller than the error term in the Edgeworth expansion, even though they are of the same order in some cases. If we compare the Edgeworth expansion to our new expansion by order, we see from the following graphs that the new expansion can give much more accurate results.

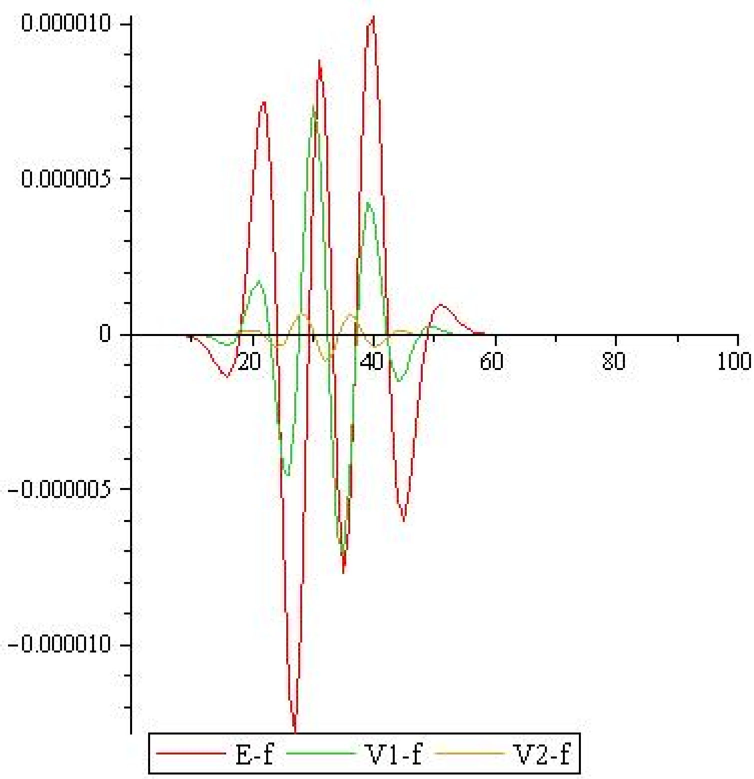


Figure 6.1: Comparing the differences between the target function and various estimates in the lattice case (E: First order Edgeworth; V1: First order new expansion where $h = 1$; V2: First order new expansion where $h = 2$).

6.2 Symmetric Random Variable Case

We will now test the effects of a restriction to the symmetric case for our new expansion compared with the Edgeworth. Suppose X is a random variable satisfying $\mathbf{E}[X] = \mu$ and $Var(X) = \sigma^2 < \infty$ and that $\{X_n\}_{n\geq 1}$ is a sequence of independent and identically distributed random variables each with the same symmetric distribution as X. We assume again that all moments exists and are finite. Note that for every odd j, $\gamma_j^* = 0$. Constructing both the Edgeworth and the new expansions will require the same process. However, many of the terms in both expansions will be zero. Recall that

$$p_n(x) = \phi_{\sigma,n}(x) + \sum_{v=1}^{k-2} \frac{q_v(x)}{n^{v/2}} + o\left(\frac{1}{n^{(k-2)/2}}\right). \tag{6.4}$$

Now if we look at the error terms of the $E_k(M)$ for the first few terms we see for $k = 2, 3$,

$$f_n(M) = E_2(M) + o\left(\frac{1}{\sqrt{n}}\right).$$

For $k = 4, 5$,

$$\begin{aligned} f_n(M) &= E_4(M) + o\left(\frac{1}{n^{3/2}}\right) \\ &= \frac{H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n^{3/2}}\right). \end{aligned}$$

For $k = 6, 7$,

$$\begin{aligned} f_n(M) &= E_6(M) + o\left(\frac{1}{n^{5/2}}\right) \\ &= \frac{H_4(x)\lambda_4^*}{n} + \frac{H_8(x)(\lambda_4^*)^2 + H_6(x)\lambda_6^*}{n^2} \\ &\quad + \frac{H_{12}(x)(\lambda_4^*)^3 + H_{10}(x)\lambda_4^*\lambda_6^* + H_8(x)\lambda_8^*}{n^3} + o\left(\frac{1}{n^{5/2}}\right). \end{aligned}$$

For $k = 8, 9$,

$$\begin{aligned} f_n(M) &= E_8(M) + o\left(\frac{1}{n^{7/2}}\right) \\ &= \frac{H_4(x)\lambda_4^*}{n} + \frac{H_8(x)(\lambda_4^*)^2 + H_6(x)\lambda_6^*}{n^2} \\ &\quad + \frac{H_{12}(x)(\lambda_4^*)^3 + H_{10}(x)\lambda_4^*\lambda_6^* + H_8(x)\lambda_8^*}{n^3} \\ &\quad + \frac{H_{16}(x)(\lambda_4^*)^4 + H_{14}(x)(\lambda_4^*)^2\lambda_6^* + H_{12}(x)(\lambda_4^*\lambda_8^* + (\lambda_6^*)^2) + H_{10}(x)\lambda_{10}^*}{n^3} + \\ &\quad o\left(\frac{1}{n^{7/2}}\right). \end{aligned}$$

The error terms for the Edgeworth expansion are much better in the symmetric case than in previous case. Notice that only the even values of k need be considered, which bring us to the following proposition:

Proposition 7 *In the symmetric case for the Edgeworth expansion, we have*

$$E_k = E_{k+1}$$

for all even k.

Proof. This is easy to see once we look at the summation over v in (4.4). If our value of v is odd, then at least one odd number will be needed in the summation. Therefore, every term in q_v will contain at least one λ_i for odd i. Since we defined $\lambda_j = 0$ for all odd j, then every term in q_v will be 0. Hence,

$$\sum_{v=1}^{k} \frac{q_v(x)}{n^{v/2}} = \sum_{v=1}^{k+1} \frac{q_v(x)}{n^{v/2}}$$

for even k.

This just leaves the error term. Clearly if $q_v = 0$ for all odd v, then the first term in the error for E_k will also be zero. This causes the error term to be pushed out to the next non-zero term. Therefore,

$$f_n(M) = E_k(M) + o\left(\frac{1}{n^{k-2/2}}\right) = E_{k+1}(M) + o\left(\frac{1}{n^{k-2/2}}\right).$$

□

Now, as before, let us assume that $\eta_h = \eta_h^{(n)}$ is the probability function for $X_1 + X_2 + ... + X_h + n_{h+1} + ... + N_n$ where N_i are independent identically distributed $N(0, \sigma^2)$ random variables also independent of the X_i. In this case, as with above, we assume that X_i has a symmetric distribution for all i. Then, as with the previous example, it can be easily proven that the requirements from Theorem 5.1 are met. Now we can expand η_h to $k = 4$ to obtain

$$\eta_h(M) = \frac{(h/n)H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right).$$

Using this, we have

$$\eta_h^*(M) = \frac{n}{h}\eta_h(M) = \frac{H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n}\right).$$

Finally, we can consider the new expansion for the symmetric case.

For $s = 1$,

$$\begin{aligned} V_1(M) &= \eta_1^*(M) + o\left(\frac{1}{n^{3/2}}\right) \\ &= \frac{H_4(x)\lambda_4^*}{n} + o\left(\frac{1}{n^{3/2}}\right) \\ &= f_n(M) + o\left(\frac{1}{n^{3/2}}\right). \end{aligned}$$

For $s = 2$,

$$\begin{aligned} V_2(M) &= \eta_1^*(M) + (n-1)(\eta_2^*(M) - \eta_1^*(M)) + o\left(\frac{1}{n^{5/2}}\right) \\ &= \frac{H_4(x)\lambda_4^*}{n} + \frac{H_8(x)(\frac{\lambda_4^*}{2})^2 + H_6(x)\lambda_6^*}{n^2} + o\left(\frac{1}{n^{5/2}}\right) \\ &= f_n(M) + o\left(\frac{1}{n^{5/2}}\right). \end{aligned}$$

For $s = 3$,

$$\begin{aligned}
V_3(M) &= \eta_1^*(M) + (n-1)(\eta_2^*(M) - \eta_1^*(M)) + \\
&\quad (n-2)(n-1)(\eta_3^*(M) - 2\eta_2^*(M) + \eta_1^*(M)) + o\left(\frac{1}{n^{7/2}}\right) \\
&= \frac{H_4(x)\lambda_4^*}{n} + \frac{H_8(x)(\frac{\lambda_4^*}{2})^2 + H_6(x)\lambda_6^*}{n^2} \\
&\quad + \frac{H_{12}(x)(\frac{\lambda_4 3!}{)}^3 + H_{10}(x)\lambda_4^*\lambda_6^* + H_8(x)\lambda_8^*}{n^3} + o\left(\frac{1}{n^{7/2}}\right) \\
&= f_n(M) + o\left(\frac{1}{n^{7/2}}\right).
\end{aligned}$$

Now if we compare what we have from the E_k with our new value for V_{k-2}, we see that all the terms in the new expansion are the same as those of the Edgeworth. The difference is that the new expansion gives us more terms much sooner. This gives us a much more accurate approximation. The following table displays the order of the error terms for the first few k.

Table 6.1: Order of Error Terms in the Symmetric Case

	3	4	5	6	7
$E_k(M)$	$\frac{1}{n^{1/2}}$	$\frac{1}{n^{3/2}}$	$\frac{1}{n^{3/2}}$	$\frac{1}{n^{5/2}}$	$\frac{1}{n^{5/2}}$
$V^{k-2}(M)$	$\frac{1}{n^{3/2}}$	$\frac{1}{n^{5/2}}$	$\frac{1}{n^{7/2}}$	$\frac{1}{n^{9/2}}$	$\frac{1}{n^{11/2}}$

Therefore, we can see that the new expansion is much more accurate than the Edgeworth. This can easily be seen by the following graph

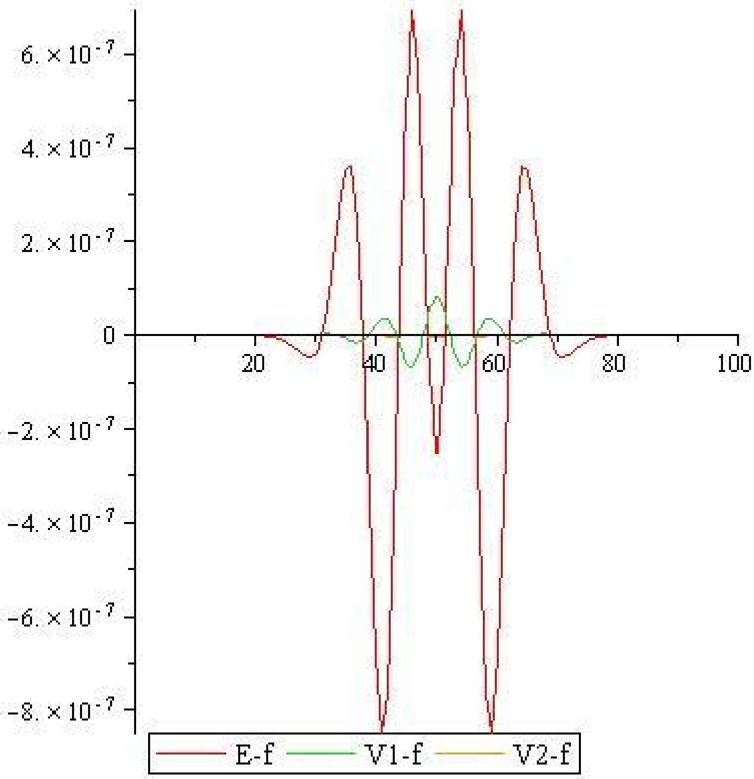


Figure 6.2: Comparing the differences between the target function and various estimates in the symmetric random variable case (E: First order Edgeworth; V1: First order new expansion where $h = 1$; V2: First order new expansion where $h = 2$).

This graph shows the difference between both expansions and the distribution function f_n. Notice that for $k = 2$ the new expansion is significantly better. Even expanding out to $k = 3$ gives us an approximation that is so accurate it can't even be read on the same scale. From this there is a subtle though clear benefit to using this new expansion as apposed to the Edgeworth.

6.3 Crossing Points

For both the Edgeworth expansion, and our new expansion, we are simply estimating our function f_n. Neither of these estimates are exact, and we have addressed the issue of the error terms in each. If our estimates are not equal to the function f_n, then we can ask the following questions: Where are the two estimates greater than or less then our function f_n? Where are they equal? We know that the function f_n is asymptotically similar in distribution to a normal distribution. If we were to graph one of the estimates along with normal, the majority of the estimate would be either above or below the function f_n, crossing the function only a few times. Now we can ask how many times will the estimate cross the target? And can we figure out exactly where they will cross? The answers to those two questions and more lie in the estimates themselves. If we examine the first terms in the errors of both expansions, we can see how many crossing points there will be, and which x represent those crossing points.

In the new expansion, for $s = 2$, we can examine the first error term by taking

$$f_n(M) - V_2(M) = \frac{H_9(x)\lambda_3^3}{n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right).$$

In this case the dominate term of the error only involves a single Hermite polynomial. We don't need to factor in for any γ_i. Therefore, the crossing points of our expansion occur for all x such that

$$H_9(x) = x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x = 0.$$

Here we get a polynomial of degree 9. We know that there must be 9 crossing points, since all the zeros of a Hermite polynomial are simple and real. We can use Maple to solve for our zeros as

$$x \approx 0, \pm 1.023255664, \pm 2.076847979, \pm 3.205429003, \pm 4.512745863.$$

In the Edgeworth expansion, for $k = 4$, we can examine the error term in the same way. From this we get

$$f_n(M) - E_4(M) = \frac{H_9(x)(\frac{\lambda_3}{3!}^3 + H_7(x)\lambda_4^*\lambda_3^* + H_5(x)\lambda_5^*}{n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right).$$

Similar to above, the crossing points of the estimate with the target will occur for all x such that

$$H_9(x)(\frac{\lambda_3}{3!})^3 + H_7(x)\lambda_4^*\lambda_3^* + H_5(x)\lambda_5^* = 0. \tag{6.5}$$

Again we get a polynomial of degree 9. We know that we can have as many as 9 crossing points in this case. However, do not know for sure if all of the zeros will be simple and real. Based on our definitions of the Hermite polynomials, and taking into account all the γ_n's, we can see that finding these crossing points can be somewhat difficult. Also notice that we now have a polynomial that depends on our probability p as well as x. All these factors may affect our result. In fact, we can use Maple to solve for the solutions to equation (6.5) for $p = .25$ as

$$x \approx 0, \pm 1.474652131, \pm 2.758663088, \pm 4.189641889, \pm .9489046833i.$$

Here there are only 7 crossing points since there are only 7 real zeros of our polynomial. So not only was this polynomial more difficult to compute, but we can't even rely on its order to describe the number of crossing points. This is an example of how the new expansion has benefits not available in the case of the Edgeworth.

Bibliography

[1] Hall, Peter *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York, 1992.

[2] Petrov, Valentin V. *Sums of Independent Random Variables.* Springer-Verlag, Heidelberg, 1975.

[3] Szego, Gabor *Orthogonal Polynomials.* American Mathematical Society, Providence, RI, 1939.

[4] Smith, E. R. *Zeros of the Hermitian Polynomial.* The American Mathematical Monthly, Vol. **43**, No. 6 (Jun. - Jul., 1936).

[5] Greenwood, Robert E., Miller, J. J. *Zeros of the Hermite Polynomials and Weights for Gauss' Mechanical Quadrature Formula* American Mathematical Society, August 1948.

[6] Riordan, J. *Combinatorial Identities.* Wiley, New York, 1979.

[7] Attar, Rafeet E. *Special Functions and Orthogonal Polynomials* Lulu Press, Morrisville, NC, 2006.

[8] Severini, Thomas Alan *Elements of Distribution Theory* Cambridge University Press, New York, NY, 2005.

Vita

James Chernesky Jr. was born July 25, 1985 to James and Lisbeth Chernesky of Oxford, CT. He graduated from Notre Dame Catholic High School, Fairfield, CT in May 2003. He then went on to complete a Bachelor of Science in May 2007 from the University of Connecticut, with a major in Mathematics and a minor in Physics. After graduating, he worked for one year as a Project Engineer for Pratt & Whitney of East Hartford, CT. There he worked in the design and production of the F-35 Joint Strike Fighter, now employed by the U.S. military. In August 2008, he enrolled in Wake Forest University as a graduate student. In May 2010, he will be awarded the degree of Master of Arts in Mathematics by Wake Forest.