Biostatistics Department Technical Report

BST2015-002

The Derivations of the Edgeworth Expansion: A Tutorial

Charles R. Katholi, Ph.D.

March 26, 2015

The Derivation of the Edgeworth Expansion: A Tutorial

Charles R. Katholi

March 26, 2015

1 Background Material:

A general introduction to asymptotic approximations to distributions is given in Wallace [5] and the notation of his article will be used at least initially. The Central Limit Theorem has its limitations, especially when the sample size ,n, is small. The Edgeworth expansion is aimed at improving the approximation for such samples. The expansion is derived from a formal identity which relates the characteristic functions of two distributions. For simplicity we shall assume that both distributions are continuous and share the same domain on the real line. Let $F(x)$ be the CDF of the distribution to be approximated and let $\Psi(x)$ be the CDF of the distribution to be used in making the approximation. In general, $\Psi(x)$ need not be the normal distribution, but later it will be what we use to obtain the Edgeworth expansion. Let $\psi(t)$ be the characteristic function of $\Psi(x)$ and let $\gamma_1, \gamma_2, \cdots$ be its cumulants. Then we know that $\psi(t)$ can be written as

$$\psi(t) = \exp(\sum_{r=1}^{\infty} \gamma_r \frac{(it)^r}{r!}) \tag{1.1}$$

where i is the complex number defined as $i = \sqrt{-1}$. It follows then that it is the case that

$$\psi(t) \exp(\sum_{r=1}^{\infty} -\gamma_r \frac{(it)^r}{r!}) = 1 \tag{1.2}$$

If we denote the characteristic function for the distribution $F(x)$ by $f(t)$ and its associated cumulants by $\kappa_1, \kappa_2, \cdots$ then

$$f(t) = \exp(\sum_{r=1}^{\infty} \kappa_r \frac{(it)^r}{r!}) \tag{1.3}$$

Combining equation (1.3) and equation(1.2) the characteristic functions satisfy the formal identity

$$f(t) = \exp\left(\sum_{r=1}^{\infty} [\kappa_r - \gamma_r] \frac{(it)^r}{r!}\right) \psi(t) \tag{1.4}$$

Next quoting from Wallace [5]

> If now, Ψ and all its derivatives vanish at the extreme range of x and exist for all x in that range, then by integration by parts, $(it)^r \psi(t)$ is the characteristic function of $(-1)^r \Psi^{(r)}(x)$. Introducing the differential operator D to represent differentiation with respect to x, the formal identity corresponds term-wise in any formal expansion to the formal identity
>
> $$F(x) = \exp\left(\sum_{r=1}^{\infty} [\kappa_r - \gamma_r] \frac{(-D)^r}{r!}\right) \Psi(x) \tag{1.5}$$
>
> One can formally and apparently construct a distribution with prescribed cumulants by choosing Ψ and formally expanding.
> The most important developing function $\Psi(x)$ is a normal distribution and with that choice, the formal expansion has been given earlier by Chebyshev,[2] Edgeworth [3] and Charlier [1].

2 Developing the Expansion

Let $X_1, X_2, \cdots, X_n$ be an i.i.d. sample from a distribution with $E(X_i) = \mu = \kappa_1$ and $Var(X_i) = \sigma^2 = \kappa_2$ and higher cumulants denoted by κ_r and with $-\infty < X < \infty$. Our objective will be to develop an approximation for the distribution of the statistic

$$Y_n = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right) \tag{2.1}$$

where $Y_n \xrightarrow{L} N(0,1)$ as $n \to \infty$ which is better than $N(0,1)$ when n is not too large. In our development we shall use the standard normal distribution with CDF $\Phi(x)$ as the approximating distribution $\Psi(x)$ of equation (1.5). For the standard normal it is well known that the cumulants are $\gamma_1 = 0, \gamma_2 = 1$ and $\gamma_r = 0, r \geq 3$. Next we must find the characteristic function for Y_n. Let $g(t)$ be the characteristic function for the distribution of the X_i, Then it is well known by properties of the characteristic function that for any X_i the characteristic function of

$$\frac{1}{\sqrt{n}}\left(\frac{X_i - \mu}{\sigma}\right)$$

is

$$\begin{aligned} h(t) &= \exp\left(-t\frac{\mu}{\sqrt{n}\sigma}\right) g\left(\frac{1}{\sqrt{n}\sigma}t\right) \\ &= \exp\left(-[\frac{\mu}{\sqrt{n}\sigma}](it) + [\frac{\kappa_1}{\sqrt{n}\sigma}](it) + \frac{\kappa_2}{n\sigma^2}\frac{(it)^2}{2!} + \sum_{r=3}^{\infty}\frac{\kappa_r}{(\sqrt{n}\sigma)^r}\frac{(it)^r}{r!}\right) \end{aligned} \tag{2.2}$$

Since $\kappa_1 = \mu$ and $\kappa_2 = \sigma^2$ we see that the $h(t)$ reduces to

$$h(t) = \exp\left(\frac{1}{2n}(it)^2 + \sum_{r=3}^{\infty}\frac{\kappa_r}{(\sqrt{n}\sigma)^r}\frac{(it)^r}{r!}\right)$$

Finally, the characteristic function of Y_n is just equal to $h(t)^n$ so we have

$$w(t) = h(t)^n = \exp\left(\frac{1}{2}(it)^2 + \sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(it)^r}{r!}\right)$$

Next we note that for the standard normal distribution, $\gamma_1 = 0, \gamma_2 = 1$ and $\gamma_r = 0\,,\, r \geq 3$, so that the characteristic function of the standard normal is $\exp[(1/2)(it)^2]$. Plugging into equation (1.5) leads to the expression

$$F_{Y_n}(x) = \exp\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right)\Phi(x) \tag{2.3}$$

where as has been previously stated D is the differential operator, $D^r = d^r/dx^r$. The next step is to expand the exponential function in its MacLaurin expansion; that is

$$\exp\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right) = 1+\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right) + \frac{1}{2!}\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right)^2 + \frac{1}{3!}\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right)^3 + \frac{1}{4!}\left(\sum_{r=3}^{\infty}\frac{1}{n^{r/2-1}}\frac{\kappa_r}{\sigma^r}\frac{(-D)^r}{r!}\right)^4 + \cdots \tag{2.4}$$

At this point each term in the formal relationship is expanded and the terms gathered in powers of $1/\sqrt{n}$. This is a daunting task even with today's modern computer tools for doing the algebra involved. This can be accomplished using a program like Maple or Mathematica. Written in terms of the cumulants and powers of D the first few terms of this expansion are,

$$1-\frac{1}{\sqrt{n}}\left[\frac{\kappa_3}{6\sigma^3}D^3\right]+\frac{1}{n}\left[\frac{1}{72}\left(\frac{\kappa_3}{\sigma^3}\right)^2 D^6+\frac{1}{24}\left(\frac{\kappa_4}{\sigma^4}\right)D^4\right] - \frac{1}{n^{3/2}}\left[\frac{1}{144}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_4}{\sigma^4}\right)D^7+\frac{1}{120}\left(\frac{\kappa_5}{\sigma^5}\right)D^5+\frac{1}{1296}\left(\frac{\kappa_3}{\sigma^3}\right)^3 D^9\right] + \frac{1}{n^2}\left[\frac{1}{1152}\left(\frac{\kappa_4}{\sigma^4}\right)^2 D^8+\frac{1}{720}\left(\frac{\kappa_6}{\sigma^6}\right)D^6+\frac{1}{1728}\left(\frac{\kappa_3}{\sigma^3}\right)^2\left(\frac{\kappa_4}{\sigma^4}\right)D^{10} + \frac{1}{31104}\left(\frac{\kappa_3}{\sigma^3}\right)^4 D^{12}+\frac{1}{720}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_5}{\sigma^5}\right)D^8\right]+\bigcirc\left(\frac{1}{n^{5/2}}\right) \tag{2.5}$$

When this lengthy expression is applied as an operator on $\Phi(x)$ the result is the Edgeworth expansion. This can be written in a number of ways, where the effects of the differentiation can be expressed most simply in terms of the Hermite polynomials $He_n(x)$ defined by the relationship

$$He_n(x) = (-1)^n\frac{\phi^{(n)}(x)}{\phi(x)} \quad \text{or} \quad \phi^{(n)}(x) = (-1)^n\phi(x)He_n(x)$$

where $\phi(x) = d\Phi(x)/dx$. The polynomial functions, $He_n(x)$ can be found recursively as

$$He_{n+1}(x) = xHe_n(x) - nHe_{n-1}(x) \text{ for } n \geq 1 \tag{2.6}$$

given that $He_0 = 1$ and $He_1 = x$. Application of equation(2.5) to the function $\Phi(x)$ leads to the Expansion for the distribution of the random variable Y_n,

$$\begin{aligned} F_{Y_n}(x) = \Phi(x) &- \frac{1}{\sqrt{n}}\left[\frac{\kappa_3}{6\sigma^3}He_2(x)\right]\phi(x) + \frac{1}{n}\left[\frac{1}{72}\left(\frac{\kappa_3}{\sigma^3}\right)^2(-He_5(x))\right. \\ &+ \left.\frac{1}{24}\left(\frac{\kappa_4}{\sigma^4}\right)(-He_3(x))\right]\phi(x) - \frac{1}{n^{3/2}}\left[\frac{1}{144}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_4}{\sigma^4}\right)He_6(x)\right. \\ &+ \left.\frac{1}{120}\left(\frac{\kappa_5}{\sigma^5}\right)He_4(x) + \frac{1}{1296}\left(\frac{\kappa_3}{\sigma^3}\right)^3 He_8(x)\right]\phi(x) \\ &+ \frac{1}{n^2}\left[\frac{1}{1152}\left(\frac{\kappa_4}{\sigma^4}\right)^2(-He_7(x)) + \frac{1}{720}\left(\frac{\kappa_6}{\sigma^6}\right)(-He_5(x))\right. \\ &+ \frac{1}{1728}\left(\frac{\kappa_3}{\sigma^3}\right)^2\left(\frac{\kappa_4}{\sigma^4}\right)(-He_9(x)) + \frac{1}{31104}\left(\frac{\kappa_3}{\sigma^3}\right)^4(-He_{11}(x)) \\ &+ \left.\frac{1}{720}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_5}{\sigma^5}\right)(-He_7(x))\right]\phi(x) \\ &+ \bigcirc\left(\frac{1}{n^{5/2}}\right) \end{aligned} \tag{2.7}$$

Or after adjusting for the minus signs on the odd ordered Hermite polynomials, the expansion becomes

$$
\begin{aligned}
F_{Y_n}(x) = \Phi(x) &- \frac{1}{\sqrt{n}}\left[\frac{\kappa_3}{6\sigma^3}He_2(x)\right]\phi(x) - \frac{1}{n}\left[\frac{1}{72}\left(\frac{\kappa_3}{\sigma^3}\right)^2 He_5(x)\right. \\
&+ \left.\frac{1}{24}\left(\frac{\kappa_4}{\sigma^4}\right) He_3(x)\right]\phi(x) - \frac{1}{n^{3/2}}\left[\frac{1}{144}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_4}{\sigma^4}\right) He_6(x)\right. \\
&+ \left.\frac{1}{120}\left(\frac{\kappa_5}{\sigma^5}\right) He_4(x) + \frac{1}{1296}\left(\frac{\kappa_3}{\sigma^3}\right)^3 He_8(x)\right]\phi(x) \\
&- \frac{1}{n^2}\left[\frac{1}{1152}\left(\frac{\kappa_4}{\sigma^4}\right)^2 He_7(x) + \frac{1}{720}\left(\frac{\kappa_6}{\sigma^6}\right) He_5(x)\right. \\
&+ \frac{1}{1728}\left(\frac{\kappa_3}{\sigma^3}\right)^2\left(\frac{\kappa_4}{\sigma^4}\right) He_9(x) + \frac{1}{31104}\left(\frac{\kappa_3}{\sigma^3}\right)^4 He_{11}(x) \\
&+ \left.\frac{1}{720}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_5}{\sigma^5}\right) He_7(x)\right]\phi(x) \qquad (2.8) \\
&+ \bigcirc\left(\frac{1}{n^{5/2}}\right)
\end{aligned}
$$

Some sources express the expansion in terms of the central moments of the the distribution of X and these first few of the relationships between the cumulants and the central moments are

$$
\begin{aligned}
\kappa_1 &= \mu \\
\kappa_2 &= \mu_2 = \sigma^2 \\
\kappa_3 &= \mu_3 \\
\kappa_4 &= \mu_4 - 3\mu_2^2 \\
\kappa_5 &= \mu_5 - 10\mu_2\mu_3
\end{aligned}
\qquad (2.9)
$$

Given that $\kappa_3 = \mu_3$ and that $He_2(x) = x^2 - 1$ the term of order $1/\sqrt{n}$ in equation (2.7) can be written as

$$
-\frac{1}{\sqrt{n}}\left[\frac{\mu_3}{6\sigma^3}(x^2-1)\right]\phi(x) = \frac{1}{\sqrt{n}}\left[\frac{\mu_3}{6\sigma^3}(1-x^2)\right]\phi(x) \qquad (2.10)
$$

Similarly, given that $He_3 = x^3 - 3x$, $He_4 = x^4 - 6x^2 + 3$ and $He_5 = x^5 - 10x^3 + 15x$ the second term in the expansion (term of order $1/n$) becomes

$$
-\frac{1}{n}\left[\frac{1}{72}\left(\frac{\mu_3}{\sigma^3}\right)^2\left(x^5 - 10x^3 + 15x\right) + \frac{1}{24}\left(\frac{\mu_4 - 3\mu_2^2}{\sigma^4}\right)\left(x^3 - 3x\right)\right]\phi(x) \qquad (2.11)
$$

The terms given in equations (2.10) and (2.11) correspond to those given by Lehmann [4] on pages 81 and 83.

3 Remarks

In this report we have developed the Edgeworth expansion for the case of a random variable Y_n which has the property that

$$Y_n = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \xrightarrow{L} N(0,1) \text{ as } n \to \infty$$

where $\bar{X}_n$ is the sample mean of an i.i.d. sample from a distribution with cumulants $\kappa_1, \kappa_2, \cdots$. If the moment generating function, $M(t)$ of the distribution of a random variable is known then the cumulant generating function is

$$K(t) = \ln(M(t)) = \kappa_1 t + \frac{\kappa_2}{2!}t^2 + \frac{\kappa_3}{3!}t^3 + \cdots$$

The cumulants are then found by differentiating repeatedly by t and evaluating the respective derivatives at $t = 0$. From a computational point of view, the recursion given in equation (2.6) is very handy, since for any particular value of x, the numerical values of the polynomials $He_r(x)$ can be found without actually finding the polynomial form.

Appendix

In this appendix we give the expressions for the terms of order $n^{-5/2}$ and n^{-3}. These are long and involve high order cumulants and high order Hermite polynomials.

$$\begin{aligned}
-\frac{\phi(x)}{n^{5/2}} \Bigg[& \frac{1}{933120}\left(\frac{\kappa_3}{\sigma^3}\right)^5 He_{14}(x) + \frac{1}{31104}\left(\frac{\kappa_3}{\sigma^3}\right)^3 \left(\frac{\kappa_4}{\sigma^4}\right) He_{12}(x) \\
& + \frac{1}{8640}\left(\frac{\kappa_3}{\sigma^3}\right)^2 \left(\frac{\kappa_5}{\sigma^5}\right) He_{10}(x) + \frac{1}{6912}\left(\frac{\kappa_3}{\sigma^3}\right) \left(\frac{\kappa_4}{\sigma^4}\right)^2 He_{10}(x) \\
& + \frac{1}{4320}\left(\frac{\kappa_3}{\sigma^3}\right) \left(\frac{\kappa_6}{\sigma^6}\right) He_8(x) + \frac{1}{2880}\left(\frac{\kappa_4}{\sigma^4}\right) \left(\frac{\kappa_5}{\sigma^5}\right) He_8(x) \\
& - \frac{1}{5040}\left(\frac{\kappa_7}{\sigma^7}\right) He_6(x) \Bigg]
\end{aligned}$$

and

$$
\begin{aligned}
&-\frac{\phi(x)}{n^3}\left[\frac{1}{33592320}\left(\frac{\kappa_3}{\sigma^3}\right)^6 He_{17}(x) + \frac{1}{746496}\left(\frac{\kappa_3}{\sigma^3}\right)^4\left(\frac{\kappa_4}{\sigma^4}\right) He_{15}(x)\right. \\
&\quad + \frac{1}{155520}\left(\frac{\kappa_3}{\sigma^3}\right)^3\left(\frac{\kappa_5}{\sigma^5}\right) He_{13}(x) + \frac{1}{82944}\left(\frac{\kappa_3}{\sigma^3}\right)^2\left(\frac{\kappa_4}{\sigma^4}\right)^2 He_{13} \\
&\quad + \frac{1}{51840}\left(\frac{\kappa_3}{\sigma^3}\right)^2\left(\frac{\kappa_6}{\sigma^6}\right) He_{11}(x) + \frac{1}{17280}\left(\frac{\kappa_3}{\sigma^3}\right)^4\left(\frac{\kappa_6}{\sigma^6}\right) He_{11}(x) \\
&\quad + \frac{1}{17280}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_4}{\sigma^4}\right)\left(\frac{\kappa_6}{\sigma^6}\right) He_{11}(x) + \frac{1}{82944}\left(\frac{\kappa^4}{\sigma^4}\right)^3 He_{11}(x) \\
&\quad - \frac{1}{30240}\left(\frac{\kappa_3}{\sigma^3}\right)\left(\frac{\kappa_7}{\sigma^7}\right) He_9(x) + \frac{1}{17280}\left(\frac{\kappa_4}{\sigma^4}\right)\left(\frac{\kappa_6}{\sigma^6}\right) He_9(x) \\
&\quad \left. + \frac{1}{28800}\left(\frac{\kappa_5}{\sigma^5}\right)^2 He_9(x) + \frac{1}{40320}\left(\frac{\kappa_8}{\sigma^8}\right) He_7(x)\right]
\end{aligned}
$$

References

[1] C. V. L. Charlier, “Über das Fehlergesets”, *Ark. Mat. Astr. och Fys.*, Vol **2**, No. 8,9 pp, 1905.

[2] P. L. Chebyshev, “Sur deux théorèmes relatifs aux probabilités”, *Acta Math.*, Vol. **14**, pp. 305-315, 1890.

[3] F. Y. Edgeworth, “The law of error”. *Cambridge Philos. Trans,*, Vol. **20**, pp. 36-66 and 114-141, 1905.

[4] E. L. Lehmann, *Elements of Larg-Sample Theory*, Springer-Verlag New York, Inc, New York, 1999.

[5] D. L. Wallace, “Asymptotic Approximations to Distributions”, *The Annals of Mathematical Statistics*, **29**, pp 635-654, 1958.