The Lord-Wingersky Algorithm

Shelby J. Haberman
Haberman Statistics

Abstract

The Lord-Wingersky algorithm is a common tool for exact computation of the distribution of a Poisson binomial random variable. It has also been generalized to sums of multinomial variables. The generalized algorithm and methods to reduce computation are provided.

Keywords: Poisson binomial distribution, multinomial distribution, recursive algorithms, normal approximation, large deviations, cumulants

Introduction

The Poisson binomial distribution is the distribution of the sum of independent Bernoulli random variables (Cramér, 1946, pp. 206–207). More generally, one may consider sums of independent multinomial variables. Let X_j, $j \geq 1$, be independent multinomial random variables such that, for $j \geq 1$, X_j has possible values between 0 and I_j for a positive integer I_j. For each positive integer j, let $X_j = x$ with probability $p_{xj} > 0$ for nonnegative integers $x < I_j$. Let $S_n = \sum_{j=1}^{n} X_j$ for each integer $n \geq 1$. If $I_j = 2$ for all positive integers j, then a simple recursive algorithm for computation of the probability mass function of S_n appears in Lord and Wingersky (1984). A generalization appears in Thissen et al. (1995). This procedure is relatively efficient; however, as evident in the algorithm description in section The Traditional Algorithm, computational labor is of order n^2, a far from negligible number for large n. In the section Reducing Computational Labor, a simple procedure for reducing computation is provided. The extent of reduction is considered in section Potential for Reduced Computation for large n. Throughout all sections, it is assumed that the variance $\sigma^2(S_n)$ of S_n approaches ∞ as n approaches ∞ and W, the supremum of I_j for $j \geq 1$, is finite.

Shelby J. Haberman https://orcid.org/0000-0002-5490-0405
Shelby Haberman is an independent statistical consultant whose website is https://www.habermanstatistics.com. He can also be reached at Barak 3/1, Jerusalem 9350276, Israel or at haberman.statistics@gmail.com.

The Traditional Algorithm

For each positive integer n, $0 \le S_n \le I_{Sn}$, where $I_{Sn} = \sum_{j=1}^{n} I_j$ for each positive integer n. The probability $P(S_1 = h)$ that $S_1 = X_1 = h$ is p_{h1} for $0 \le h \le I_1$. If n is a positive integer, $h \le I_{S(n+1)}$ is a nonnegative integer, $\eta(1, h, n) = \max(0, h - I_{Sn})$ and $\eta(2, h, n) = \min(I_n, h)$, then

$$P(S_{n+1} = h) = \sum_{a=\eta(1,h,n)}^{\eta(2,h,n)} P(S_n = h - a)p_{a(n+1)}. \tag{1}$$

Computational labor to find the probability mass function of S_n is of order n^2.

Reducing Computational Labor

For n large, substantial reduction in computational labor may be achieved by exploiting the fact that it is common for many values of $P(S_n = h)$, $0 \le h \le I_{Sn}$, to be very small. To avoid computations with very small numbers, let $c < 1$ be a very small positive number. Consider the revised algorithm with $r(h, 1) = r_0(h, 1) = P(X_1 = h)$ for nonnegative integers $h \le I_1$, $b(1) = 0$, and $t(1) = I_1$. Let $d(0) = 0$. Let n be a positive integer, and let $d(n) = nc/[2(n + 1)]$. For integers h such that $b(n) \le h \le t(n) + I_n$, let $\eta(3, h, n) = \max(0, h - t(n))$, $\eta(4, h, n) = \min(I_n, h - b(n))$, and

$$r_0(h, n + 1) = \sum_{a=\eta(3,h,n)}^{\eta(4,h,n)} r_0(h - a, n + 1)p_{a(n+1)} \tag{2}$$

If $r_0(b(n), n + 1) > d(n)$, let $b(n + 1) = b(n)$. Otherwise, let $b(n + 1)$ be the largest integer greater than $b(n)$ such that the sum of $r_0(h, n+1)$ for h from $b(n)$ to $b(n+1)-1$ is less than $d(n)$. If $r_0(t(n)+I_{n+1}, n+1) > d(n)$, let $t(n+1) = t(n)+I_{n+1}$. Otherwise, let $t(n + 1)$ be the smallest integer no greater than $t(n) + I_{n+1} - 1$ such that the sum of $r_0(h, n + 1)$ for h from $t(n) + 1$ to $t(n + 1)$ is less than $d(n)$. For integers h from 0 to $I_{S(n+1)}$, let $r(h, n + 1)$ be 0 if $h < b(n + 1)$ or $h > t(n + 1)$, and let $r(h, n + 1) = r_0(h, n + 1)$ otherwise. Thus computation of $r(h, n + 1)$ need only be considered for the $t(n)-b(n)+I_{n+1}+1$ integers h such that $b(n) \le h \le t(n)+I_{n+1}$. A simple induction shows that, for integers $n > 1$, $r(h, n) \le P(S_n = h)$ for $0 \le h \le n$, $r_+(n) = \sum_{h=0}^{n} r(h, n) > 1 - 2d(n - 1)$, and $b(n) < t(n)$. It follows that

$$0 \le 1 - r_+(n) = \sum_{h=0}^{I_{Sn}} [P(S_n = h) - r(h, n)] < 2d(n - 1). \tag{3}$$

For any subset H of the set $J(n)$ of nonnegative integers no greater than I_n, let $R_n(H)$ be the sum of $r_n(h)$ for h in H. Then the probability $P(S_n \in H)$ satisfies $0 \le P(S_n \in H) - R_n(H) < 2d(n - 1)$. If $\tilde{R}_n(H) = R_n(H)/r_+(n)$, then $\tilde{R}_n(J(n)) = 1$ and $P(S_n \in H) - 2d(n - 1) < \tilde{R}_n(H) < P(S_n \in H)/[1 - 2d(n - 1)]$.

Potential for Reduced Computation

For large values of n, it is common for $(I_{Sn}+1)^{-1}[t(n)-b(n)+I_n+1]$ to be small, so that actual computational labor to find the distribution of S_n is of order substantially less than n^2. This claim follows from a variety of arguments based on normal approximations for the distribution of S_n. For this purpose, it is helpful to note some basic definitions and standard results from complex analysis and probability theory (Ahlfors, 1979; Cramér, 1946; Haberman, 1996; Knapp, 2016).

Background Results from Complex Analysis

Let $\mathbb{R}$ denote the set of real numbers, and let $\mathbb{C}$ denote the set of complex numbers. For y in $\mathbb{C}$, let $\Re(y)$ denote the real part of y, and let $\Im(y)$ denote the imaginary part of y, so that $y=\Re(y)+\Im(y)i$ has absolute value equal to the nonnegative square root of $[\Re(y)]^2+[\Im(y)]^2$. If y is in $\mathbb{C}$, then $\exp(y)=\exp(\Re(y))[\cos(\Im(y))+\sin(\Im(y))i]$, so that $\exp(\Re(y))=|\exp(y)|$. If y_1 and y_2 are in $\mathbb{C}$, then $\exp(y_1+y_2)=\exp(y_1)\exp(y_2)$.

For any real $\delta>0$, let $\mathbb{B}(\delta,x)$ be the set of y in $\mathbb{C}$ such that $|y-x|\le\delta$, let $\mathbb{D}(\delta,x)$ be the set of y in $\mathbb{C}$ such that $|y-x|=\delta$, and let $\mathbb{O}(\delta,x)$ be the set of y in $\mathbb{C}$ such that $|y-x|<\delta$, so that $\mathbb{B}(\delta)$ is the union of $\mathbb{D}(\delta)$ and $\mathbb{O}(\delta)$. A complex function f on a nonempty open subset A of $\mathbb{C}$ is analytic if it is infinitely differentiable with kth derivative f_k for all positive integers and if, for every x in A and real $\delta>0$, if $\mathbb{D}(\delta,x)\subset A$, then the Taylor expansion

$$f(y)=f(x)+\sum_{k=1}^{\infty}f_k(x)(y-x)^k/k! \tag{4}$$

of f about x holds for all y in $\mathbb{O}(\delta,x)$, so that

$$\sum_{k=1}^{\infty}|f_k(x)|z^k/k! \tag{5}$$

for all positive real numbers $z<\delta$. If f is analytic on $\mathbb{C}$, then f is an entire function.

The Cauchy formulas for derivatives of analytic functions and for remainders of infinite series (Knapp, 2016, pp. 653-656) provides bounds on $f_k(x)$ for positive integers k and provides information concerning speed of convergence of the Taylor series about X of f. Define the functions $f_{KT}(\cdot,x)$ and $f_{KR}(\cdot,x)$ on A for all nonnegative integers K. For y in A,

$$f_{KT}(y,x)=\begin{cases} f(x), & K=0,\\ f(x)+\sum_{k=1}^{K}f_k(x)(y-x)^k/k!, & K>0,\end{cases} \tag{6}$$

and $f_{KR}(y,x)=f(y)-f_{KT}(y,x)$, so that $f_{KR}(x,x)=0$. For any analytic complex function g with domain that includes $\mathbb{B}(\delta,x)$, let $M(\delta,x,g)$ be the supremum of $|g(y)|$ for y in $\mathbb{D}(\delta,x)$. Then $|g(y)|\le M(\delta,x,g)$ for all y in $\mathbb{O}(\delta,x)$. For a positive real number $\delta_1 < \delta$, $M(\delta_1, x, g) \leq M(\delta, x, g)$. Let K be a nonnegative integer, and let L be a positive integer. If $L \leq K$, then 0 is the Lth derivative of $f_{KR}(\cdot, x)$ at x. If $L > K$, then the Lth derivative of $f_{KR}(\cdot, x)$ is f_L,

$$|f_L(x)| \leq L!M(\delta, x, f_{KR})/\delta^L \tag{7}$$

and, for y in $\mathbb{O}(\delta, x)$,

$$|f_{LR}(y, x)| \leq M(\delta, x, f_{KR})(|y - x|/\delta)^{L+1}/(1 - |y - x|/\delta). \tag{8}$$

On the other hand,

$$\begin{aligned} f_{KR}(y, x) &= \int_x^y (y - z)^K f_{K+1}(z) dz \\ &= \frac{f_{K+1}(x)(y - x)^{K+1}}{(K + 1)!} + \int_x^y (y - z)^K [f_{K+1}(z) - f_{K+1}(x)] dz. \end{aligned} \tag{9}$$

For $\delta \leq \delta_1$ and $\mathbb{B}(\delta_1, x) \subset A$,

$$M(\delta, x, f_{KR}(\cdot, x)) \leq M(\delta_1, x, f_{K+1})\delta^{K+1}/(K + 1)! \tag{10}$$

and

$$M(\delta, x, f_{(K+1)R}(\cdot, x)) \leq M(\delta_1, x, f_{K+1} - f_{K+1}(x))\delta^{K+1}/(K + 1)!. \tag{11}$$

Consider the case of $f(x) \neq 0$. Let $D(x, f, \mathrm{Lg})$ be ∞ if f is an entire function never equal to 0, and let $D(x, f, \mathrm{Lg})$ otherwise be the largest real number δ such that $\mathbb{O}(\delta, x) \subset A$ and $f(y) \neq 0$ for all y in $\mathbb{O}(\delta, x)$. Let $\mathrm{Dom}(x, f, \mathrm{Lg})$ be $\mathbb{C}$ if $D(x, f, \mathrm{Lg}) = \infty$. Otherwise, let $\mathrm{Dom}(x, f, \mathrm{Lg})$ be the set of y in A such that $|y-x| < D(x, f, \mathrm{Lg})$. Let $\mathrm{Lg}_1(\cdot, x, f)$ be the analytic function on $\mathrm{Dom}(x, f, \mathrm{Lg})$ such that $\mathrm{Lg}_1(y, x, f) = f_1(y)/f(y)$ for y in $\mathrm{Dom}(x, f, \mathrm{Lg})$. Define the complex analytic function $\mathrm{Lg}(\cdot, x, f)$ on $\mathrm{Dom}(x, f, \mathrm{Lg})$ for y in $\mathrm{Dom}(x, f, \mathrm{Lg})$ by

$$\mathrm{Lg}(y, x, f) = \int_x^y \mathrm{Lg}_1(z, x, f) dz = (y - x) \int_0^1 \mathrm{Lg}_1(x + u(y - x), x, f) du. \tag{12}$$

Then $f(y) = f(x) \exp(\mathrm{Lg}(y, x, f))$, so that $\mathrm{Lg}(y, x, f)$ is a possible logarithm of $f(y)/f(x)$. The kth derivative of $\mathrm{Lg}(\cdot, x, f)$ is $\mathrm{Lg}_k(\cdot, x, f)$ for all positive integers k, $\mathrm{Lg}(x, x, f) = 0$, and $\mathrm{Lg}(\cdot, x, f)$ has the Taylor expansion about x of

$$\mathrm{Lg}(y, x, f) = \sum_{k=1}^{\infty} \mathrm{Lg}_k(x, x, f)(y - x)^k/k!. \tag{13}$$

For example,

$$\mathrm{Lg}_2(y, x, f) = f_2(y)/f(z) - [\mathrm{Lg}_1(y, x, f)]^2 \tag{14}$$

for y in $\mathrm{Dom}(x, f, \mathrm{Lg})$. If f is an entire function that is never 0 and $|f(y)/f(x)| \leq \exp(u|y - x|)$ for some real $u > 0$, then $\mathrm{Lg}_k(x, x, f) = 0$ for $k > 1$, so that $\mathrm{Lg}(\cdot, x, f)$ is a linear function.

If g is an analytic function on an open set B in $\mathbb{C}$ that includes x and $g(x) \neq 0$, then let fg denote the analytic function on $A \cap B$ such that, for y in $A \cap B$, fg has value $f(y)g(y)$. For y in $A \cap B$, $f(y)g(y) = 0$ if either $f(y) = 0$ or $g(y) = 0$. Thus

$$D(x, fg, \mathrm{Lg}) = \min(D(x, f, \mathrm{Lg}), D(x, g, \mathrm{Lg})) \tag{15}$$

and

$$\mathrm{Dom}(x, fg, \mathrm{Lg}) = \mathrm{Dom}(x, f, \mathrm{Lg}) \cap \mathrm{Dom}(x, g, \mathrm{Lg}). \tag{16}$$

For y in $\mathrm{Dom}(x, fg, \mathrm{Lg})$,

$$\mathrm{Lg}_1(y, xfg) = [f_1(y)g(y) + f(y)g_1(y)]/[f(y)g(y)] = \mathrm{Lg}_1(y, x, f) + \mathrm{Lg}_1(y, x, g). \tag{17}$$

It follows that

$$\mathrm{Lg}(y, x, fg) = \mathrm{Lg}(y, x, f) + \mathrm{Lg}(y, x, g) \tag{18}$$

and

$$\mathrm{Lg}_k(y, x, fg) = \mathrm{Lg}_k(y, x, f) + \mathrm{Lg}_k(y, x, g) \tag{19}$$

for all integers $k > 1$. For some z in $\mathbb{C}$, if $g(y) = \exp(zy)$ for all y in $\mathbb{C}$, then $\mathrm{Lg}(y, x, g) = zy$ for y in $\mathbb{C}$, $\mathrm{Dom}(x, fg, \mathrm{Lg}) = \mathrm{Dom}(x, f, \mathrm{Lg})$, and

$$\mathrm{Lg}(y, x, fg) = \mathrm{Lg}(y, x, f) + zy \tag{20}$$

for y in $\mathrm{Dom}(x, f, \mathrm{Lg})$.

Moment Generating Functions

Let Y be a real random variable such that no real y exists such that $Y = y$ with probability 1. Let $\mathrm{Dom}(Y, \mathrm{Mgf})$ be the set of y in $\mathbb{C}$ such that $\exp(yY)$ has a finite expectation

$$\begin{aligned} \mathrm{Mgf}(y, Y) &= E(\exp(Yy)) \\ &= E(\exp(Y\Re(y))\cos(Y\Im(y)) + E(\exp(Y\Re(y))\sin(Y\Im(y))i. \end{aligned} \tag{21}$$

Assume that $\mathrm{Dom}(Y, \mathrm{Mgf}) = \mathbb{C}$, as is certainly the case if Y is bounded. Let the moment generating function $\mathrm{Mgf}(\cdot, Y)$ of Y be the function on $\mathbb{C}$ with value $\mathrm{Mgf}(y, Y)$ at y in $\mathbb{C}$. Then $\mathrm{Mgf}(\cdot, Y)$ is an entire function, $\mathrm{Mgf}(0, Y) = 1$, and, for each positive integer k, the kth derivative of $\mathrm{Mgf}(\cdot, Y)$ is $\mathrm{Mgf}_k(\cdot, Y)$. For y in $\mathbb{C}$, $\mathrm{Mgf}_k(y, Y) = E(Y^k \exp(Yy))$. Let $\mu_k(Y) = E(Y^k)$ denote the kth moment of Y, and let $\mu_{kc}(Y) = \mu_k(Y - E(Y))$ denote the kth central moment of Y. Thus $\mu_1(Y) = E(Y)$, $\mu_{1c}(Y) = E(Y) - E(Y) = 0$, and $\mu_{2c}(Y) = \sigma^2(Y)$, the variance of Y. Then $\mu_k(Y) = \mathrm{Mgf}_k(0, Y)$. For any real u, $\exp((Y - u)y) = \exp(-uy)\exp(Yy)$ has a finite expectation for all y in $\mathbb{C}$, so that $\mathrm{Dom}(Y - u, \mathrm{Mgf}) = \mathbb{C}$ and $\mathrm{Mgf}(y, Y - u) = \exp(-uy)M(y, Y)$. The case of $u = E(Y)$ then implies that $\mathrm{Mgf}_k(y, Y - E(Y)) = E([Y - E(Y)]^k \exp([Y - E(Y)]y)$ and $\mu_{kc}(Y) = \mathrm{Mgf}_k(0, Y - E(Y))$. For all y in $\mathbb{C}$,

$$\mathrm{Mgf}(y, Y) = 1 + \sum_{k=1}^{\infty} \mu_k y^k / k! \tag{22}$$

and

$$\mathrm{Mgf}(y, Y - E(Y)) = 1 + \sum_{k=2}^{\infty} \mu_{kc} y^k / k! \tag{23}$$

Because Y by assumption does not equal $E(Y)$ with probability 1, the variance $\sigma^2(Y)$ and the standard deviation $\sigma(Y)$ of Y are both positive. Corresponding to Y is the standardized random variable $\mathrm{Std}(Y) = [Y - E(Y)]/\sigma(Y)$ with expectation 0 and variance 1. If u is a nonzero real number, then $\mathrm{Dom}(uY, \mathrm{Mgf}) = \mathbb{C}$, and, for y in $\mathbb{C}$, $\mathrm{Mgf}(y, uY) = \mathrm{Mgf}(uy, Y)$. Thus $\mathrm{Dom}(\mathrm{Std}(Y), \mathrm{Mgf}) = \mathbb{C}$, and

$$\begin{aligned} \mathrm{Mgf}(y, \mathrm{Std}(Y)) &= \mathrm{Mgf}(y/\sigma(Y), Y) \exp(-E(Y)y/\sigma(Y)) \\ &= \mathrm{Mgf}(y/\sigma(Y), Y - E(Y)) \end{aligned} \tag{24}$$

for y in $\mathbb{C}$. For any integer k, let $\sigma^k(Y)$ denote $[\sigma(Y)]^k$. For each integer $k \geq 1$, $\mu_{kc}(\mathrm{Std}(Y)) = \mu_{kc}(Y)/\sigma^k(Y)$, and $\sigma^2(\mathrm{Std}(Y)) = \mu_{2c}(\mathrm{Std}(Y)) = 1$.

In the case of X_j, $j \geq 1$, $\mathrm{Dom}(X_j, \mathrm{Mgf}) = \mathbb{C}$,

$$\mathrm{Mgf}(y, X_j) = \sum_{x=0}^{I_j} \exp(xy) p_{xj} \tag{25}$$

for y in $\mathbb{C}$,

$$E(X_j) = \sum_{x=0}^{I_j} x p_{xj}, \tag{26}$$

$$\sigma^2(X_j) = \mu_{2c}(X_j) = \sum_{x=0}^{I_j} [x - E(X_j)]^2 p_{xj}, \tag{27}$$

and, for positive integers k,

$$\mu_k(X_j) = \sum_{x=1}^{I_j} x^k p_{xj}, \tag{28}$$

and

$$\mu_{kc}(X_j) = \sum_{x=0}^{I_j} [x - E(X_j)]^k p_{xj}. \tag{29}$$

If Z has a standard normal distribution with density $\phi(y) = \exp(-y^2/2)/(2\pi)^{1/2}$ for y in $\mathbb{R}$, then $\mathrm{Dom}(Z, \mathrm{Mgf}) = \mathbb{C}$, $\mathrm{Mgf}(y, Z) = \exp(y^2/2)$ for y in $\mathbb{C}$, $E(Z) = 0$, $\sigma^2(Z) = 1$, and $\mathrm{Std}(Z) = Z$. For an odd positive integer k, $\mu_k(Z) = \mu_{kc}(Z) = 0$. For an even positive integer k, $\mu_k(Z) = \mu_{kc}(Z) = k!/[(k/2)!2^k]$.

Let Y_j, $j \geq 1$, be independent real random variables such that, for all positive integers j, $\mathrm{Dom}(Y_j, \mathrm{Mgf}) = \mathbb{C}$ and no real y exists such that $Y_j = y$ with probability 1. For each integer $n \geq 1$, let $U_n = \sum_{j=1}^{n} Y_j$. Then $\mathrm{Dom}(U_n, \mathrm{Mgf}) = \mathbb{C}$, and

$$\mathrm{Mgf}(y, U_n) = \prod_{j=1}^{n} \mathrm{Mgf}(y, Y_j) \tag{30}$$

for all y in $\mathbb{C}$.

The characteristic function $\mathrm{Cf}(\cdot, Y)$ of Y is the function on $\mathbb{R}$ such that, for y in $\mathbb{R}$, $\mathrm{Cf}(y, Y) = \mathrm{Mgf}(yi, Y)$. Thus $\mathrm{Cf}(0, Y) = 1$ and $|\mathrm{Cf}(y, Y)| \le 1$ for all y in $\mathbb{R}$. For x in $\mathbb{R}$, $\cos(x) = \cos(-x)$ and $\sin(x) = -\sin(-x)$, Because

$$\mathrm{Cf}(y, Y) = E(\cos(yY)) + E(\sin(yY))i, \tag{31}$$

$$\mathrm{Cf}(-y, Y) = E(\cos(yY)) - E(\sin(yY))i, \tag{32}$$

and $|\mathrm{Cf}(y, Y)|^2 = \mathrm{Cf}(y, Y)\,\mathrm{Cf}(-y, Y)$. If Y and Y' are independent and identically distributed, then $Y - Y'$ and $Y' - Y = -(Y - Y')$ have the same distribution and

$$\mathrm{Cf}(y, Y - Y') = \mathrm{Cf}(-y, Y - Y') = \mathrm{Cf}(y, Y)\,\mathrm{Cf}(-y, Y) = E(\cos((Y - Y')y)). \tag{33}$$

To study convergence speed, for nonnegative integers K and y in $\mathbb{C}$, let

$$\mathrm{Mgf}_{KT}(y, Y - E(Y)) = \begin{cases} 1, & K \le 1, \\ 1 + \sum_{k=2}^{K} \mu_{kc}(Y) y^k / k!, & K > 1, \end{cases} \tag{34}$$

and

$$\mathrm{Mgf}_{KR}(y, Y - E(Y)) = \mathrm{Mgf}(y, Y) - \mathrm{Mgf}_{KT}(y, Y). \tag{35}$$

Let δ be a positive real number. Let $M(\delta, Y, \mathrm{Mgf}_{KR})$ be $M(\delta, 0, \mathrm{Mgf}_{KR}(\cdot, Y))$. If $K > 0$, let $M(\delta, Y, \mathrm{Mgf}_K)$ be $M(\delta, 0, \mathrm{Mgf}_K(\cdot, Y))$, and let $M(\delta, Y, \mathrm{Mgf}_K - \mu_{Kc})$ be $M(\delta, 0, \mathrm{Mgf}_K(\cdot, Y) - \mu_{Kc}(Y))$. If L is an integer greater than K, then

$$|\mu_{Lc}(Y)| \le L! M(\delta, Y, \mathrm{Mgf}_{KR})/\delta^L \tag{36}$$

and, for y in $\mathbb{O}(\delta, 0)$,

$$|\mathrm{Mgf}_{LR}(y, Y)| \le M(\delta, Y, \mathrm{Mgf}_{KR})(|y|/\delta)^{L+1}/(1 - |y|/\delta). \tag{37}$$

For nonnegative integers K and δ_1 a real number such that $\delta \le \delta_1$,

$$M(\delta, Y, \mathrm{Mgf}_{KR}) \le M(\delta_1, Y, \mathrm{Mgf}_{K+1})\delta^{K+1}/(K+1)! \tag{38}$$

and

$$M(\delta, Y, \mathrm{Mgf}_{(K+1)R}) \le M(\delta_1, Y, \mathrm{Mgf}_{K+1} - \mu_{(K+1)c})\delta^{K+1}/(K+1)!. \tag{39}$$

Cumulant Generating Functions

The cumulant generating function $\mathrm{Cgf}(\cdot, Y)$ of Y is helpful for independent random variables. Because $\mathrm{Mgf}(0, Y) = 1$,

$$D(Y, \mathrm{Cgf}) = D(0, \mathrm{Mgf}(\cdot, Y), \mathrm{Lg}) > 0, \tag{40}$$

$$\mathrm{Dom}(Y, \mathrm{Cgf}) = \mathrm{Dom}(0, \mathrm{Mgf}(\cdot, Y), \mathrm{Lg}) \tag{41}$$

is nonempty, and

$$\mathrm{Cgf}(\cdot, Y) = \mathrm{Lg}(\cdot, 0, \mathrm{Mgf}(\cdot, Y)) \tag{42}$$

is an analytic function on $\mathrm{Dom}(Y, \mathrm{Cgf})$ such that $\mathrm{Cgf}(0, Y) = 0$ and

$$\exp(\mathrm{Cgf}(y, Y)) = \mathrm{Mgf}(y, Y) \tag{43}$$

for y in $\mathrm{Dom}(Y, \mathrm{Cgf})$. For positive integers k, let $\mathrm{Cgf}_k(\cdot, Y)$ be the kth derivative of $\mathrm{Cgf}(\cdot, Y)$, so that

$$\mathrm{Cgf}_1(y, Y) = \frac{\mathrm{Mgf}_1(y, Y)}{\mathrm{Mgf}(y, Y)} \tag{44}$$

and

$$\mathrm{Cgf}_2(Y) = \frac{\mathrm{Mgf}_2(y, Y)}{\mathrm{Mgf}(y, Y)} - [\mathrm{Cgf}_1(y, Y)]^2 \tag{45}$$

for y in $\mathrm{Dom}(Y, \mathrm{Cgf})$. The kth cumulant $\kappa_k(Y)$ of Y is defined to be $\mathrm{Cgf}_k(0, Y)$, so that $\kappa_1 = E(Y)$ and $\kappa_2(Y) = \sigma^2(Y)$. For y in $\mathrm{Dom}(Y, \mathrm{Cgf})$,

$$\mathrm{Cgf}(y, Y) = \sum_{k=1}^{\infty} \kappa_k(Y) y^k / k!. \tag{46}$$

If u is real, then $\mathrm{Dom}(Y - u, \mathrm{Cgf}) = \mathrm{Dom}(Y, \mathrm{Cgf})$ and $\mathrm{Cgf}(y, Y - u) = \mathrm{Cgf}(y, Y) - uy$ for y in $\mathrm{Dom}(Y, \mathrm{Cgf})$. It follows that $\kappa_1(Y - u) = \kappa_1(Y) - u$ and $\kappa_k(Y - u) = \kappa_k(Y)$ for integers $k > 1$. If u is not zero, then y is in $\mathrm{Dom}(uY, \mathrm{Cgf})$ if, and only if, y/u is in $\mathrm{Dom}(Y, \mathrm{Cgf})$. If y is in $\mathrm{Dom}(uY, \mathrm{Cgf})$, then $\mathrm{Cgf}(y, uY) = \mathrm{Cgf}(uy, Y)$, $\mathrm{Cgf}_k(y, uY) = u^k \mathrm{Cgf}_k(uy, Y)$ and $\kappa_k(uY) = u^k \kappa_k(Y)$. Thus $\kappa_1(\mathrm{Std}(Y)) = 0$, $\kappa_2(\mathrm{Std}(Y)) = 1$, and for integers $k > 2$, $\kappa_k(\mathrm{Std}(Y)) = \kappa_k(Y)/\sigma^k(Y)$.

In the case of the standard normal variable Z, $\mathrm{Dom}(Z, \mathrm{Cgf}) = \mathbb{C}$, $\mathrm{Cgf}(y, Z) = y^2/2$ for y in $\mathbb{C}$, $\kappa_k(Z) = 0$ for all positive integers k not equal to 2, and $\kappa_2(Z) = 1$. It follows that, for y in $\mathrm{Dom}(Y, \mathrm{Cgf})$,

$$\mathrm{Cgf}(y, Y - E(Y)) = \sum_{k=2}^{\infty} \kappa_k(Y) y^k / k!, \tag{47}$$

while, for y in $\mathrm{Dom}(\mathrm{Std}(Y), \mathrm{Cgf})$,

$$\mathrm{Cgf}(y, \mathrm{Std}(Y)) = \mathrm{Cgf}(y, Z) + \sum_{k=3}^{\infty} \kappa_k(\mathrm{Std}(Y)) y^k / k!. \tag{48}$$

If Y is bounded, let $\sup(Y)$ be the supremum of Y, let $\inf(Y)$ be the infimum of Y, let the midpoint $\mathrm{Mid}(Y)$ of Y be the average of $\sup(Y)$ and $\inf(Y)$, and let the range $\mathrm{range}(Y)$ of Y be $\sup(Y) - \inf(Y)$. Then $\mathrm{Dom}(Y, \mathrm{Cgf})$ cannot be $\mathbb{C}$, for

$$|\mathrm{Mgf}(y, Y - \mathrm{Mid}(Y))| \leq \exp(\mathrm{range}(Y)|y|/2) \tag{49}$$

for all y in $\mathbb{C}$. If $\operatorname{Dom}(Y, \mathrm{Cgf}) = \mathbb{C}$, then $\mathrm{Cgf}(\cdot, Y)$ must satisfy

$$\mathrm{Cgf}(y, Y - \operatorname{Mid}(Y)) = [\kappa_1(Y) - \operatorname{Mid}(Y)]y \tag{50}$$

for all y in $\mathbb{C}$ and $\kappa_2(Y) = \sigma^2(Y)$ is 0. Because $\sigma^2(Y) > 0$, it follows that $\operatorname{Dom}(Y, \mathrm{Cgf})$ is not $\mathbb{C}$.

It must be the case that $D(Y, \mathrm{Cgf}) \geq \pi/\operatorname{range}(Y)$, for, if $|y| < \pi/\operatorname{range}(Y)$, then

$$\Re(\exp([Y - \operatorname{Mid}(Y)]y)) = \exp([Y - \operatorname{Mid}(Y)])\Re(y))\cos([Y - \operatorname{Mid}(Y)]\Im(y)). \tag{51}$$

Because

$$|[Y - \operatorname{Mid}(Y)]\Im(y)| \leq \operatorname{range}(Y)|y|/2 < \pi/2 \tag{52}$$

and

$$[Y - \operatorname{Mid}(Y)]\Re(y) \geq -\operatorname{range}(Y)|y|/2, \tag{53}$$

it follows that

$$\Re(\mathrm{Mgf}(y, Y - \operatorname{Mid}(Y))) \geq \exp(-\operatorname{range}(Y)|y|/2)\cos(\operatorname{range}(Y)|y|/2) > 0. \tag{54}$$

Because $|\mathrm{Mgf}(y, Y - \operatorname{Mid}(Y))| \geq |\Re(\mathrm{Mgf}(y, Y - \operatorname{Mid}(Y))|$,

$$D(Y - \operatorname{Mid}(Y), \mathrm{Cgf}) = D(Y, \mathrm{Cgf}) \geq \pi/\operatorname{range}(Y). \tag{55}$$

In addition,

$$\mathrm{Mgf}(y, Y - E(Y)) = \exp([\operatorname{Mid}(Y) - E(Y)]y)\,\mathrm{Mgf}(y, Y - \operatorname{Mid}(Y)). \tag{56}$$

Because $|E(Y) - \operatorname{Mid}(Y)| \leq \operatorname{range}(Y)/2$, $\mathrm{Mgf}(y, Y - E(Y)) \leq \exp(\operatorname{range}(Y)|y|)$ and

$$|\exp([\operatorname{Mid}(Y) - E(Y)]y)| \geq \exp(-\operatorname{range}(Y)|y|/2)\cos(\operatorname{range}(Y)|y|/2) \tag{57}$$

and

$$\begin{aligned} |\mathrm{Mgf}(y, Y - E(Y))| &= |\mathrm{Mgf}(y, Y - \operatorname{Mid}(Y)||\exp([\operatorname{Mid}(Y) - E(Y)]y)| \\ &\geq \exp(-\operatorname{range}(Y)|y|)[\cos(\operatorname{range}(Y)|y|/2)]^2. \end{aligned} \tag{58}$$

For example, for each positive integer j, $\sup(X_j) = I_n$, $\inf(X_j) = 0$, $\operatorname{Mid}(X_j) = I_n/2$, and $\operatorname{range}(X_j) = I_n$, so that $D(X_j, \mathrm{Cgf})$ must be at least π/I_n. If $I_j = 1$ and $p_{1j} = p_{0j} = 1/2$, then $D(X_j, \mathrm{Cgf}) = \pi$, so that the lower limit on $D(Y, \mathrm{Cgf})$ for bounded Y can be achieved.

For positive integers k,

$$\mathrm{Cgf}_k(y, U_n) = \sum_{j=1}^{n} \mathrm{Cgf}_k(y, Y_j) \tag{59}$$

and $\kappa_k(U_n) = \sum_{j=1}^n \kappa_k(Y_j)$.

To study convergence speed, for nonnegative integers K and y in $\mathrm{Dom}(Y, \mathrm{Cgf})$, let

$$\mathrm{Cgf}_{KT}(y, Y) = \begin{cases} 0, & K = 0, \\ \sum_{k=1}^K \kappa_k(Y) y^k/k!, & K > 0, \end{cases} \tag{60}$$

and

$$\mathrm{Cgf}_{KR}(y, Y) = \mathrm{Cgf}(y, Y) - \mathrm{Cgf}_{KT}(y, Y). \tag{61}$$

Let δ be a positive real number less than $D(Y, \mathrm{Cgf})$. Let $M(\delta, Y, \mathrm{Cgf}_{KR})$ be $M(\delta, 0, \mathrm{Cgf}_{KR}(\cdot, Y))$. If $K > 0$, let $M(\delta, Y, \mathrm{Cgf}_K)$ be $M(\delta, 0, \mathrm{Cgf}_K(\cdot, Y))$, and let $M(\delta, Y, \mathrm{Cgf}_K - \kappa_K)$ be $M(\delta, 0, \mathrm{Cgf}_K(\cdot, Y) - \kappa_K(Y))$. If L is an integer greater than K, then

$$|\kappa_L(Y)| \le L! M(\delta, Y, \mathrm{Cgf}_{KR})/\delta^L \tag{62}$$

and, for y in $\mathbb{O}(\delta, 0)$,

$$|\mathrm{Cgf}_{LR}(y, Y)| \le M(\delta, Y, \mathrm{Cgf}_{KR})(|y|/\delta)^{L+1}/(1 - |y|/\delta). \tag{63}$$

For nonnegative integers K and δ_1 a real number such that $\delta \le \delta_1 < D(Y, \mathrm{Cgf})$,

$$M(\delta, Y, \mathrm{Cgf}_{KR}) \le M(\delta_1, Y, \mathrm{Cgf}_{K+1})\delta^{K+1}/(K+1)! \tag{64}$$

and

$$M(\delta, Y, \mathrm{Cgf}_{(K+1)R}) \le M(\delta_1, Y, \mathrm{Cgf}_{K+1} - \kappa_{K+1})\delta^{K+1}/(K+1)!. \tag{65}$$

Additivity properties of cumulant generation functions imply that, for n a positive integer, δ a positive real number less than $D(U_n, \mathrm{Cgf})$, and K a nonnegative integer,

$$M(\delta, U_n, \mathrm{Cgf}_{KR}) \le \sum_{j=1}^n M(\delta, Y_j, \mathrm{Cgf}_{KR}), \tag{66}$$

$$M(\delta, U_n, \mathrm{Cgf}_K) \le \sum_{j=1}^n M(\delta, Y_j, \mathrm{Cgf}_K), \tag{67}$$

and

$$M(\delta, U_n, \mathrm{Cgf}_K - \kappa_K) \le \sum_{j=1}^n M(\delta, Y_j, \mathrm{Cgf}_K - \kappa_K). \tag{68}$$

For the standard normal random variable Z, results are trivial. Because $\mathrm{Cgf}_2(y, Z) = \kappa_2(Z) = 1$ for all y in $\mathbb{C}$, $M(\delta, Z, \mathrm{Cgf}_{2R}) = M(\delta, Z, \mathrm{Cgf}_2 - \kappa_2) = 0$ and $\kappa_k(Z) = 0$ for $k > 2$.

Let Y be bounded, and let $\delta < \pi/\mathrm{range}(Y)$. For y in $\mathbb{C}$ and $K \ge 0$, let

$$\exp_{KT}(y) = \sum_{k=0}^K y^k/k! \tag{69}$$

and

$$\exp_{KR}(y) = \exp(y) - \exp_{KT}(y) = \int_0^y (y-z)^K \exp(z)dz. \tag{70}$$

Then

$$|\exp_{KR}(y)| \le |y|^{K+1} \exp(\max(0, \Re(y))/(K+1)!. \tag{71}$$

In addition,

$$|\exp(y) - 1| \le \sum_{k=1}^{\infty} |y|^k/k! = \exp(|y|) - 1. \tag{72}$$

For y in $\mathbb{D}(\delta, 0)$, Equation 72 implies that

$$|\operatorname{Mgf}_2(y, Y - E(Y)) - \kappa_2(Y)| \le [\exp(\operatorname{range}(Y)\delta) - 1]\kappa_2(Y). \tag{73}$$

Because $\kappa_1(Y - E(Y)) = 0$, application of Equation 71 for $K = 1$ yields

$$|\operatorname{Mgf}_1(y, Y - E(Y))| \le |y| \exp(\operatorname{range}(Y)\delta)\kappa_2(Y). \tag{74}$$

In addition, for any integer $k \ge 3$, $\mu_{kc}(Y) \le \kappa_2(Y)[\operatorname{range}(Y)]^{k-2}$, so that

$$|\operatorname{Mgf}(y, Y - E(Y)) - 1| \le \kappa_2(Y)[\exp(\operatorname{range}(Y)\delta) - \operatorname{range}(Y)\delta - 1]/[\operatorname{range}(Y)]^2. \tag{75}$$

Recall that $\operatorname{Cgf}_2(y, Y) = \operatorname{Cgf}_2(y, Y - E(Y))$ for y in $\operatorname{Dom}(Y, \operatorname{Cgf})$. Application of Equations 44, 45, 58, 73, 74, and 75 demonstrates existence of a nonnegative continuously differentiable and nondecreasing real function β on $[0, \pi)$ with derivative β_1 such that $\beta(0) = 0$ and

$$M(\delta, Y, \operatorname{Cgf}_2 - \kappa_2) \le \beta(\operatorname{range}(Y)\delta)\kappa_2(Y) \tag{76}$$

for positive real $\delta < \pi/\operatorname{range}(Y)$. It follows that

$$M(\delta, Y, \operatorname{Cgf}_{2R}) \le \beta(\operatorname{range}(Y)\delta_1)\kappa_2(Y)\delta^2/2 \tag{77}$$

if δ_1 is a real number such that $\delta \le \delta_1 < \pi/\operatorname{range}(Y)$. In use of Equation 77, for x in $(0, \pi)$, let $M(x, \beta_1)$ be the supremum of $\beta_1(z)$ for z in $[0, x]$. Then for z in $[0, x]$, $\beta(z) \le M(x, \beta_1)z$.

If the random variables Y_j, $j \ge 1$, are all bounded and $V = \sup_{j \ge 1} \operatorname{range}(Y_j)$ is finite, then, for all positive integers j and n, $D(Y_j) \ge \pi/V$ and $D(U_n) \ge \pi/V$. For positive real $\delta < \pi/V$ and positive real δ_1 such that $\delta \le \delta_1 < \pi/V$,

$$M(\delta, U_n, \operatorname{Cgf}_2) - \kappa_2(U_n)| \le \beta(V\delta)\kappa_2(U_n), \tag{78}$$

$$M(\sigma(U_n)\delta, \operatorname{Std}(U_n), \operatorname{Cgf}_2 - \kappa_2) \le \beta(V\delta). \tag{79}$$

and

$$M(\sigma(U_n)\delta, \operatorname{Std}(U_n), \operatorname{Cgf}_{2R}) \le M(V\delta_1, \beta_1)V\delta^3/2. \tag{80}$$

Cumulants and Moments

In general, cumulants and moments are closely related; however, the relationship is somewhat complex for $k \geq 2$. For an integer v such that $2 \leq v \leq k$ and for positive integers $h \leq \mathrm{Fl}(k/v)$, let $\mathcal{K}(v,h,k)$ be the set of vectors $\mathbf{f}$ with $\lambda(\mathbf{f}) = h$ nonincreasing integer elements $f(a) \geq v$ such that k is the sum $\Sigma(\mathbf{f})$ of the $f(a)$, $1 \leq a \leq \lambda(\mathbf{f})$. Let $\mathcal{I}(\mathbf{f})$ be the set of integers such that each member of $\mathcal{I}(\mathbf{f})$ is equal to $f(a)$ for some positive integer no greater than $\lambda(\mathbf{f})$, and, for each positive integer $u \geq v$ in $\mathcal{I}(\mathbf{f})$, let $N(u,\mathbf{f})$ be the number of positive integers $a \leq \lambda(\mathbf{f})$ such that $f(a) = u$. For $\mathbf{f}$ in $\mathcal{K}(v,h,k)$, let

$$\theta_{v\kappa}(\mathbf{f},Y) = [\Sigma(\mathbf{f})]! \prod_{u\in\mathcal{I}(\mathbf{f})} \frac{[\kappa_u(Y)]^{N(u,\mathbf{f})}}{(u!)^{N(u,\mathbf{f})}[N(u,\mathbf{f})]!} \tag{81}$$

and

$$\theta_{v\mu}(\mathbf{f},Y) = [\Sigma(\mathbf{f})]! \prod_{u\in\mathcal{I}(\mathbf{f})} \frac{[\mu_{uc}(Y)]^{N(u,\mathbf{f})}}{(u!)^{N(u,\mathbf{f})}[N(u,\mathbf{f})]!}. \tag{82}$$

Then

$$\mu_{kc}(Y) = \sum_{h=1}^{\mathrm{Fl}(k/2)} \sum_{\mathbf{f}\in\mathcal{K}(2,h,k)} \theta_{2\kappa}(\mathbf{f},Y) \tag{83}$$

and

$$\kappa_k(Y) = -\sum_{h=1}^{\mathrm{Fl}(k/2)} (-1)^h(h-1)!] \sum_{\mathbf{f}\in\mathcal{K}(2,h,k)} \theta_{2\mu}(\mathbf{f},Y) \tag{84}$$

(Haberman, 1996, ch. 8). For example, $\mathcal{K}(2,1,2)$ includes only the vector $\mathbf{2}_1$ of length 1 with element 1, $\theta_{2\kappa}(\mathbf{2}_1) = \kappa_2(Y)$ and $\mu_{2c}(Y) = \kappa_2(Y)$. The set $\mathcal{K}(2,1,3)$ includes only the vector $\mathbf{3}_1$ with the single element 3, $\theta_{2v}(\mathbf{3}_1) = \kappa_3(Y)$, and $\mu_{3c}(Y) = \kappa_3(Y)$. The set $\mathcal{K}(2,1,4)$ includes only the vector $\mathbf{4}_1$ with the single element 4, and $\mathcal{K}(2,2,2)$ only contains the vector $\mathbf{2}_2$ with two elements equal to 2. The components of $\mu_{4c}(Y)$ are $\theta_{2\kappa}(\mathbf{2}_2,Y) = 3[\kappa_2(Y)]^2$ and $\theta_{2\kappa}(\mathbf{4}_1,Y) = \kappa_4(Y)$, so that $\mu_{4c}(Y) = \kappa_4(Y) + 3[\kappa_2(Y)]^2$. The components of $\kappa_4(Y)$ are $\theta_{2\mu}(\mathbf{2}_2,Y) = 3[\mu_{2c}(Y)]^2$ and $\theta_{2\mu}(\mathbf{4}_1,Y) = \mu_{4c}(Y)$. It follows that $\kappa_4(Y) = \mu_{4c}(Y) - 3[\mu_{2c}(Y)]^2$.

In addition, let $\Omega(y,Y) = \mathrm{Mgf}(y,\mathrm{Std}(Y))/\,\mathrm{Mgf}(y,Z)$ for y in $\mathbb{C}$. Then $\Omega(\cdot,Y)$ is an analytic function on $\mathbb{C}$ with kth derivative $\Omega_k(\cdot,Y)$ and with $\omega_k(Y) = \Omega_k(0,Y)$ for integers $k \geq 1$. Then $\Omega(0,Y) = 1$, $\omega_1(Y) = \omega_2(Y) = 0$, and

$$\Omega(y,Y) = 1 + \sum_{k=3}^{\infty} \omega_k(Y)y^k/k! \tag{85}$$

for y in $\mathbb{C}$. For integers $k \geq 3$,

$$\omega_k(Y) = \sum_{h=1}^{\mathrm{Fl}(k/3)} \sum_{\mathbf{f}\in\mathcal{K}(3,h,k)} \theta_{3\kappa}(\mathbf{f},\mathrm{Std}(Y)). \tag{86}$$

For example, $\omega_k(Y) = \kappa_k(\mathrm{Std}(Y))$ for $3 \le k \le 5$ and $\omega_6(Y) = \kappa_6(\mathrm{Std}(Y)) + 10[\kappa_3(\mathrm{Std}(Y))]^2$.

For an integer $K \ge 0$ and y in $\mathbb{C}$, let

$$\Omega_{KT}(y, Y) = \begin{cases} 1, & K \le 2, \\ 1 + \sum_{k=3}^{K} \omega_k(Y) y^k / k!, & K > 2, \end{cases} \tag{87}$$

and $\Omega_{KR}(y, Y) = \Omega(y, Y) - \Omega_{KT}(y, Y)$.

To study convergence speed, let $M(\delta, Y, \Omega_{KR})$ be $M(\delta, 0, \Omega_{KR}(\cdot, Y))$. If $K > 0$, let $M(\delta, Y, \Omega_K)$ be $M(\delta, 0, \Omega_K(\cdot, Y))$, and let

$$M(\delta, Y, \Omega_K - \omega_K) = M(\delta, 0, \Omega_K(\cdot, Y) - \omega_K(Y)). \tag{88}$$

Consider the case of $\delta < D(\mathrm{Std}(Y), \mathrm{Cgf})$. For y in $\mathbb{D}(\delta, 0)$, $\Omega(y, Y) = \exp(\mathrm{Cgf}_{2R}(y, \mathrm{Std}(Y)))$, so that

$$M(\delta, Y, \Omega_{0R}) = M(\delta, Y, \Omega_{2R}) \le \exp(M(\delta, \mathrm{Std}(Y), \mathrm{Cgf}_{2R})) - 1. \tag{89}$$

If $K > 2$ is an integer, then

$$|\omega_K(Y)| \le K! M(\delta, Y, \Omega_{0R}) / \delta^K \tag{90}$$

and, for y in $\mathbb{O}(\delta, 0)$,

$$|\Omega_{KR}(y, Y)| \le M(\delta, Y, \Omega_{0R}) (|y|/\delta)^{K+1} / (1 - |y|/\delta). \tag{91}$$

Normal Approximation

In the case of bounded Y_j, $j \ge 1$, such that $\mathrm{range}(Y_j) \le V$ for a real number $V > 0$, the normal approximation applies to $\mathrm{Std}(U_n)$ as long as $\sigma^2(U_n)$ approaches ∞ as n approaches ∞. For a real random variable Y, let $\mathrm{Cdf}(\cdot, Y)$ be its cumulative distribution function, so that $\mathrm{Cdf}(y, Y)$ is the probability $P(Y \le y)$ for y in $\mathbb{R}$. Let $\mathrm{Cdf}_-(\cdot, Y)$ be the alternative cumulative distribution function defined for y in $\mathbb{R}$ so that $\mathrm{Cdf}_-(y, Y)$ is the probability $P(Y < y)$ that $Y < y$. Then $\mathrm{Cdf}(\cdot, Y)$ and $\mathrm{Cdf}_-(\cdot, Y)$ are nondecreasing functions such that, for real y, $0 \le \mathrm{Cdf}_-(y, Y) \le \mathrm{Cdf}(y, Y) \le 1$ and $\mathrm{Cdf}_-(y, Y) = \mathrm{Cdf}(y, Y)$ if 0 is the probability $P(Y = y)$ that $Y = y$. For the standard normal random variable Z, $\mathrm{Cdf}(y, Z) = \mathrm{Cdf}_-(y, Z) = 1 - \mathrm{Cdf}(-y, Z)$.

If $\mathrm{Cdf}(\cdot, Y)$ is continuous and strictly increasing, $0 < \mathrm{Cdf}(y, Y) < 1$ for y in $\mathbb{R}$, $\inf(\mathrm{Cdf}(\cdot, Y)) = 0$, and $\sup(\mathrm{Cdf}(\cdot, Y)) = 1$, then the unique quantile function $Q(\cdot, Y)$ of Y is defined on the open interval $(0, 1)$ by the requirement that $Q(\mathrm{Cdf}(y, Y), Y) = y$ for y in $\mathbb{R}$. Alternatively, $\mathrm{Cdf}(Q(p, Y), Y) = p$ for all p in $(0, 1)$. In the case of the standard normal random variable Z, $Q(p, Z) = -Q(1 - p, Z)$ for all p in $(0, 1)$.

In general, a sequence G_n, $n \ge 1$, of real random variables convergences in distribution to a real random variable G if, for all bounded continuous real functions

f on $\mathbb{R}$, $E(f(G_n))$, $n \geq 1$, converges to $E(f(G))$. Equivalently, G_n, $n \geq 1$, converges in distribution to G if $\mathrm{Cdf}(y, G_n)$, $n \geq 1$, convergences to $\mathrm{Cdf}(y, G)$ for all y in $\mathbb{R}$ such that $\mathrm{Cdf}(y, G) = \mathrm{Cdf}_-(y, G)$. Thus $\mathrm{Std}(U_n)$, $n \geq 1$, converges in distribution to Z if $\mathrm{Cdf}(\cdot, \mathrm{Std}(U_n))$, $n \geq 1$, converges to $\mathrm{Cdf}(\cdot, Z)$, so that, for y in $\mathbb{R}$, $\mathrm{Cdf}(y, \mathrm{Std}(U_n))$, $n \geq 1$, converges to $\mathrm{Cdf}(y, Z)$. Equivalently, $\mathrm{Std}(U_n)$, $n \geq 1$, converges in distribution to Z if $\mathrm{Cdf}_-(\cdot, \mathrm{Std}(U_n))$, $n \geq 1$, converges to $\mathrm{Cdf}(\cdot, Z)$. Let $\mathbf{Y}_n$ be the n-dimensional vector with element Y_j, $1 \leq j \leq n$, and let

$$\Delta_n(\mathbf{Y}_n) = \frac{\sum_{j=1}^{n} E(|Y_j - E(Y_j)|^3}{\left[\sum_{j=1}^{n} \sigma^2(Y_j)\right]^{3/2}}. \tag{92}$$

If $\Delta_n(\mathbf{Y}_n)$, $n \geq 1$, approaches 0, then $\mathrm{Std}(U_n)$, $n \geq 1$, converges in distribution to Z (Cramér, 1946, pp. 213–218). If F is a bounded real function on $\mathbb{R}$, let $|F|$ be the supremum of $F(y)$ for y in $\mathbb{R}$. Then a real constant $\Gamma \geq (2\pi)^{-1/2}$ exists independent of the Y_j, $j \geq 1$, such that

$$|\mathrm{Cdf}(\cdot, \mathrm{Std}(U_n)) - \mathrm{Cdf}(\cdot, Z)| = (|\mathrm{Cdf}_-(\cdot, \mathrm{Std}(U_n)) - \mathrm{Cdf}_-(\cdot, Z)| \leq \Gamma \Delta_n(\mathbf{Y}_n) \tag{93}$$

(Berry, 1941; Esseen, 1945). In Equation 92, because $E(|Y_j - E(Y_j)|^3) \leq V\sigma^2(Y_j)$ for $j \geq 1$, $\Delta_n(\mathbf{Y}_n) \leq V/\sigma(U_n)$, so that convergence in distribution is assured as long as $\sigma^2(U_n)$, $n \geq 1$, converges to ∞. These results apply to $\mathrm{Std}(S_n)$, $n \geq 1$.

Large Deviations

Because $d(n)$ is normally much smaller than $V/\sigma(S_n)$, Equation 93 has only limited direct value in analysis of the modifed algorithm. Somewhat more information can be obtained by use of the theory of large deviations (Chernoff, 1952). The desired results involve convergence properties associated with moment generating functions. The general arguments apply to $\mathrm{Std}(U_n)$ when all Y_j, $j \geq 1$, are bounded and $\mathrm{range}(Y_j) \leq V < \infty$, $j \leq 1$.

For each integer $n \geq 1$, let δ_n be a positive real number less that $\sigma(U_n)\pi/V$. For some positive real $\delta' < \pi/V$, assume that $\delta_n^3/\sigma(U_n)$ converges to 0 and $\delta_n/\sigma(U_n) \leq \delta'$ for $n \geq 1$. Then

$$M(\delta_n, \mathrm{Std}(U_n), \mathrm{Cgf}_{2R}) \leq \beta_1(V\delta')V\delta_n^3/[2\sigma(U_n)]. \tag{94}$$

It follows that whenever $y_n > 0$ is in $\mathbb{O}(\delta_n, 0)$ for all positive integers n, $\Omega(y_n, U_n) = \mathrm{Mgf}(y_n, \mathrm{Std}(U_n)/\mathrm{Mgf}(y_n, Z)$, $n \geq 1$, converges to 1 as $\sigma(U_n)$, $n \geq 1$, approaches ∞. Let y be a positive real number. If z is real, then $z \geq y$ if, and only if $\exp(zy) \geq \exp(y^2)$. By the Markov inequality,

$$\exp(y^2)[1 - \mathrm{Cdf}_-(y, \mathrm{Std}(U_n)] < \mathrm{Mgf}(y, \mathrm{Std}(U_n)). \tag{95}$$

For any real $\epsilon > 0$, for n sufficiently large,

$$1 - \mathrm{Cdf}_-(y_n, \mathrm{Std}(U_n)) < (1 + \epsilon)/\mathrm{Mgf}(y_n, Z). \tag{96}$$

The same argument implies that, for n sufficiently large,

$$\mathrm{Cdf}(-y_n, \mathrm{Std}(U_n)) < (1+\epsilon)/\,\mathrm{Mgf}(y_n, Z). \tag{97}$$

In the case of $Z_n = \mathrm{Std}(S_n)$, $n \geq 1$, let $g(x) = [-2\log(x)]^{1/2}$ for positive real $x < 1$, so that $\mathrm{Mgf}(g(x), Z) = 1/x$. Let δ' be a positive real number less than $\pi/\sup_{j\geq 1} I_j$. For a real $z > 1$, let $\delta_n = \min(\delta'\sigma(S_n), g(zd(n)))$, $n \geq 1$. Because $d(n)$, $n \geq 0$, converges to $c/2$, δ_n, $n \geq 1$, converges to $g(zc/2)$. Let $z' < 1$ be a positive real number, and let y_n, $n \geq 1$, be positive real numbers such that $y_n = \min(g(z'd(n)), \delta_n)$ for $n \geq 1$. In Equations 96 and 97, let $\epsilon = (1 - z')/z'$. For positive integers n sufficiently large, $y_n = g(z'd(n))$ and $1 - \mathrm{Cdf}(g(z'd(n)), Z_n) < d(n)$ and $\mathrm{Cdf}(-g(z'd(n)), Z_n) < d(n)$. Because $r_0(h, n) \leq P(S_n = h)$ for $0 \leq h \leq I_{Sn}$ and z' is arbitrary, it follows that $[b(n) - E(S_n)]/[g(d(n))\sigma(S_n)]$, $n > 1$, has limit inferior of at least -1 and $[t(n) - E(S_n)]/[g(d(n))\sigma(S_n)]$, $n > 1$, has limit superior no greater than 1. Because $\sigma^2(X_j) \leq [\mathrm{range}(X_j)/2]^2 \leq V^2/4$ (Haberman, 1996, p. 276), $\sigma(S_n) \leq n^{1/2}V/2$. In addition, $g(d(n))$ converges to $g(c/2)$ as n approaches ∞. Because

$$\int_0^n x^{1/2}dx = 2n^{3/2}/3 < \sum_{m=1}^{n} m^{1/2} < \int_1^{n+1} x^{1/2}dx = 2[(n+1)^{3/2} - 1]/3, \tag{98}$$

the computational labor in determination of the distribution of S_n can be expected to be of order no greater than $n^{3/2}$ for large values of n.

References

Ahlfors, L. V. (1979). *An introduction to the theory of analytic functions of one complex variable* (3rd). McGraw-Hill.

Berry, A. (1941). The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, *49*, 122–136. https://doi.org/10.2307/1990053

Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, *23*, 493–507. https://doi.org/10.1214/aoms/1177729330

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton University Press.

Esseen, C.-G. (1945). Fourier analysis of distribution functions. a mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, *77*, 1–125. https://doi.org/10.1007/BF02392223

Haberman, S. J. (1996). *Advanced statistics, volume 1: Description of populations*. Springer-Verlag. https://doi.org/10.1007/978-1-4757-4417-0

Knapp, A. W. (2016). *Basic real analysis* (Digital 2nd). A. W. Knapp. https://doi.org/10.3792/euclid/9781429799997

Lord, F. M., & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings". *Applied Psychological Measurement*, *8*, 453–461. https://doi.org/10.1177/014662168400800409

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement*, *19*, 39–49. https://doi.org/10.1177/014662169501900105