PROBABILITY INEQUALITIES FOR THE SUM OF INDEPENDENT RANDOM VARIABLES

GEORGE BENNETT
University of New South Wales

This paper proves a number of inequalities which improve on existing upper limits to the probability distribution of the sum of independent random variables. The inequalities presented require knowledge only of the variance of the sum and the means and bounds of the component random variables. They are applicable when the number of component random variables is small and/or have different distributions. Figures show the improvement on existing inequalities.

THE distribution function for the sum of independent random variables, $x_1+x_2+\cdots+x_i+\cdots+x_n$, when some information about the distribution of the x_i is available "... may be regarded as the very starting point of a large number of those investigations by which the modern Theory of Probability was created" Cramér [5, p. 196].

Much work has been carried out on the asymptotic form of the distribution of such sums when the number of component random variables is large and/or when the component variables have identical distributions. The majority of this work, while being suitable for the determination of the asymptotic distribution of sums of random variables, does not provide estimates of the accuracy of such asymptotic distributions when applied to the summation of finite numbers of components. Godwin [7, pp. 935–8] reviews publications giving results which allow such numerical approximations to be obtained. However, a search of the literature reveals that there is little information on the distribution function of a sum when the number of component random variables is small and/or the variables have different distributions. Yet, for most practical problems, precisely this distribution function is required. This paper, therefore, deals with such cases and restricts its scope to bounded variables with known variances.

The inequalities presented are all one-sided, that is, of the form $P[a-E(a)\geqq t\sigma]<b$. Since reversing the sign of all the component variables does not alter the applicability of the ensuing analysis, it follows that $P[a-E(a)\leqq -t\sigma]<b$, and from this it is obvious that the two-sided inequality $P[|a-E(a)|\geqq t\sigma]<2b$ must hold. In general, this is a much weaker inequality than the one-sided inequality from which it was derived, since the latter is necessarily based on the combination of component distributions that maximizes $P[a-E(a)\geqq t\sigma]$ and this probability is then usually greater than $P[a-E(a)\leqq -t\sigma]$.

TERMINOLOGY

Let $x_1, x_2, \cdots, x_i, \cdots, x_n$ be totally independent random variables, with $\operatorname{Var}(x_i)=\sigma_i^2$ and bounded so that $|x_i-E(x_i)|\leqq M_i$. Let

$$M = \max(M_i), \quad S = \sum_{i=1}^{n} x_i \quad \text{and} \quad \sigma^2 = \operatorname{Var}(S) = \sum_{i=1}^{n} \sigma_i^2.$$

* Godwin [7] reports this incorrectly, due to misplaced brackets.

In what follows, $E(x_i)$ will be taken as zero, without any loss of generality.

EXISTING INEQUALITIES

The well-known Bienaymé-Tchebycheff inequality, that is, $P[|a-E(a)| \geqq t\sigma] \leqq 1/t^2$, holds for all probability distributions, and can, therefore, be applied to the distribution of the sum of random variables to give

$$P(S \geqq t\sigma) \leqq P(|S| \geqq t\sigma) \leqq \frac{1}{t^2}\cdot$$

Uspensky [13, pp. 198–9] improves this result to give the following one-sided inequality:

$$P(S \geqq t\sigma) \leqq \frac{1}{1+t^2}\cdot$$

The limits given by these two inequalities are generally weak when applied to the sum of random variables, but it will be shown that for small values of t they are lower than the limits given by alternative inequalities.

Bernstein [1; 2, pp. 159–65] proves that

$$P(S \geqq t\sigma) < \exp\left\{-\frac{t^2}{2+\frac{2}{3}\frac{M}{\sigma}t}\right\}$$

when the component random variables have restrictions on their absolute moments which are less than bounded. Bernstein's inequality leads to the obvious extension

$$P(|S| \geqq t\sigma) < 2\exp\left\{-\frac{t^2}{2+\frac{2}{3}\frac{M}{\sigma}t}\right\}.$$

Kolmogoroff [9, Satz II] provides limits for the probability that at least one of the partial sums of n random variables exceeds a given value. These limits give, *inter alia*,

$$P(|S| \geqq t\sigma) \leqq \left(\frac{t}{m}-\frac{M}{\sigma}\right)^{-2m}$$

where m is a positive integer.* Therefore,

$$P(S \geqq t\sigma) \leqq P(|S| \geqq t\sigma) \leqq \left(\frac{t}{m}-\frac{M}{\sigma}\right)^{-2m}.$$

Loève [10, Sec. 18.1, pp. 254–5] provides exponential limits for the distribution of sums of independent random variables. His proof contains several misprints and an error (p. 255), where he replaces X by X/σ, but does not alter the bound to c/σ. The corrected inequalities are:

If

$$t\frac{M}{\sigma} \leqq 1, \quad \text{then} \quad P(S \geqq t\sigma) < \exp\left(-\frac{t^2}{2}\left(1 - \frac{t}{2}\frac{M}{\sigma}\right)\right)$$

and if

$$t\frac{M}{\sigma} \geqq 1, \quad \text{then} \quad P(S \geqq t\sigma) < \exp\left(-\frac{1}{4}\frac{\sigma}{M}t\right).$$

Berry [3] proves that

$$P(S \geqq t\sigma) < 1 - \Phi(t) + 1.88\frac{M}{\sigma}$$

where $\Phi(t)$ is the distribution function for a unit normal random variable.

When these results are evaluated for values of t up to 15, it is found that, at least over this range, Kolmogoroff's inequality never provides lower limits than

(a) the lower of either Bernstein's one-sided or Uspensky's inequality, in the case of one-sided probability limits;

(b) the lower of either Bernstein's two-sided or Bienaymé-Tchebycheff's inequality, in the case of two-sided probability limits.

For all values of t and of M/σ, Loève's inequality produces higher limits to the probability than Bernstein's inequality.

Berry's inequality does not offer any improvement over the combination of Bernstein's one-sided and Uspensky's inequalities unless $M/\sigma < 0.27$; for lower values of M/σ Berry's inequality gives lower limits for small values of t only. Thus, for $M/\sigma = 0.25$ it gives lower limits for $t < 0.6$; for $M/\sigma = 0.1$ for $t < 1.8$ and for $M/\sigma = 0.01$ for $t < 2.9$ approximately. Such low values of M/σ require a large number of component random variables. For example, to obtain $M/\sigma = 0.1$ requires at least 100 component variables and $M/\sigma = 0.01$ requires at least 10,000 component variables.

Graphs of $P(S \geqq t\sigma)$ against t, for values of M/σ of 3, 2, 1, 0.5, 0.25, and 0, as given by the best combination of the above inequalities, (namely, Bernstein's one-sided and Uspensky's inequalities) are plotted on linear probability paper in Figure 1. Berry's inequality for values of M/σ of 0.25, 0.1, and 0.01 are also plotted. The distribution function of the normal distribution has been included for comparison purposes.

Bernstein's original work was published in Russian and appears to be unobtainable. It is reported—indirectly—by Craig [4], who eased the restrictions on the inequality, and by Uspensky [13, pp. 204–6] who indicates the proof in a series of exercises. The inequality is mentioned or quoted without proof by Savage [12, pp. 35–9], Godwin [7, p. 396], Mallows [11, p. 140], David [6, p. 172], and Kendall and Buckland [8, p. 25]. Apart from these brief references, Bernstein's inequality seems to have escaped notice in the English-speaking world. A version of the proof is given here because the original is inaccessible and the steps are similar to some in the improvements which will be presented later. Let

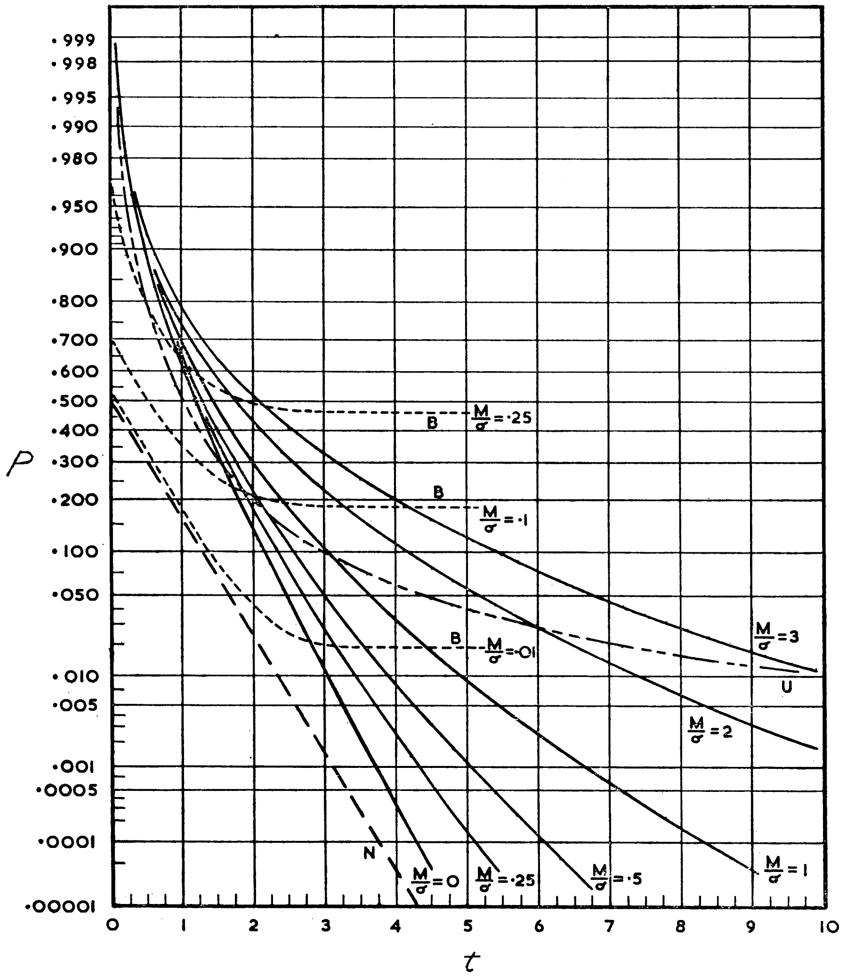

FIG. 1. Existing upper bounds on $P(S \geqq t\sigma)$.

Legend: ———— Bernstein's inequality
— – – — Uspensky's inequality
– – – – – – Berry's inequality
— — — — Normal Distribution

$$c > 0, \quad \text{then} \quad e^{cx_i} = 1 + cx_i + \sum_{r=2}^{\infty} \frac{c^r x_i^r}{r!}$$

and

$$E(e^{cx_i}) = 1 + cE(x_i) + \sum_{r=2}^{\infty} \frac{c^r E(x_i^r)}{r!}$$

$$= 1 + \frac{1}{2} c^2 \sigma_i^2 \sum_{r=2}^{\infty} \frac{c^{r-2} E(x_i^r)}{\frac{1}{2} r! \sigma_i^2} \qquad \text{since } E(x_i) = 0.$$

Let

$$F_i = \sum_{r=2}^{\infty} \frac{c^{r-2} E(x_i^r)}{\frac{1}{2} r! \sigma_i^2}, \tag{1}$$

then

$$E(e^{cx_i}) = 1 + \tfrac{1}{2}c^2 \sigma_i^2 F_i < \exp \tfrac{1}{2}c^2 \sigma_i^2 F_i$$

and since the x_i are independent,

$$\begin{aligned} E(e^{cS}) &= E(e^{c(x_1+x_2+\cdots+x_n)}) = E(e^{cx_1}) \cdot E(e^{cx_2}) \cdots E(e^{cx_n}) \\ &< \exp (\tfrac{1}{2}c^2 \sigma_1^2 F_1) \cdot \exp (\tfrac{1}{2}c^2 \sigma_2^2 F_2) \cdots \exp (\tfrac{1}{2}c^2 \sigma_n^2 F_n) \\ &< \exp (\tfrac{1}{2}c^2 \sigma^2 F) \qquad \text{where } F = \max (F_i). \end{aligned}$$

If $h(y)$ is a non-negative function of a random variable y with frequency function $f(y)$, and if $h(y) \geqq b$ when $y \geqq a$,

$$E[h(y)] = \int_{-\infty}^{\infty} h(y)f(y)dy \geqq \int_{y \geqq a} h(y)f(y)dy \geqq b \int_{y \geqq a} f(y)dy = bP(y \geqq a)$$

hence

$$P(y \geqq a) \leqq \frac{E[h(y)]}{b}.$$

If $h(y) = e^{cy}$, then for every positive c,

$$P(y \geqq a) \leqq \frac{E(e^{cy})}{e^{ca}}.$$

Let $y = S$ and $a = t\sigma$, then

$$P(S \geqq t\sigma) \leqq \frac{E(e^{cS})}{e^{ct\sigma}} < \exp (\tfrac{1}{2}c^2\sigma^2 F - ct\sigma). \tag{2}$$

The right-hand side of this equation is minimized by

$$c = \frac{t}{\sigma F} \tag{3}$$

therefore,

$$P(S \geqq t\sigma) < \exp \left(-\frac{1}{2} \frac{t^2}{F} \right) = \exp (-\tfrac{1}{2}ct\sigma). \tag{4}$$

Bernstein restricted himself to the summation of random variables whose absolute moments obeyed the following inequality,

$$\nu_r = E(|x_i|^r) \leqq \tfrac{1}{2}\sigma_i^2 W^{r-2} r! \qquad r \geqq 2, W \text{ being a constant.}$$

Since $E(x_i^r) \leqq E(|x_i|^r)$, substituting the above inequality into equation (1) gives

$$F_i \leqq \sum_{r=2}^{\infty} \frac{c^{r-2}\frac{1}{2}\sigma_i^2 W^{r-2} r!}{\frac{1}{2}\sigma_i^2 r!} = \sum_{r=2}^{\infty} (cW)^{r-2} = \sum_{s=0}^{\infty} (cW)^s = (1 - cW)^{-1}$$

$$\text{if } cW < 1 \qquad (5)$$

and since F_i does not depend on i,

$$F = \max(F_i) \leqq (1 - cW)^{-1}. \qquad (6)$$

From equations (3) and (6),

$$F = \frac{t}{c\sigma} \leqq (1 - cW)^{-1},$$

therefore,

$$c \geqq \frac{t}{\sigma + tW} \quad \text{and} \quad cW \geqq \frac{tW}{\sigma + tW};$$

the latter satisfies (5) if the equality sign is taken. Substituting this value of c in equation (4) leads to

$$P(S \geqq t\sigma) < \exp\left\{-\frac{t^2}{2 + 2\frac{W}{\sigma}t}\right\}. \qquad (7)$$

Since random variables bounded at M have $\nu_r \leqq M^{r-2}\sigma_i^2$, it follows that for such variables W can be taken as $M/3$ and $F \leqq (1 - cM/3)^{-1}$, where

$$cM/3 = cW < 1.$$

Substituting $W = M/3$ into equation (7) gives

$$P(S \geqq t\sigma) < \exp\left\{-\frac{t^2}{2 + \frac{2}{3}\frac{M}{\sigma}t}\right\}. \qquad (8)$$

Equations (6) and (8) are Bernstein's inequality.

FIRST IMPROVEMENT

Bernstein's inequality can be considerably improved at equation (2), without any further restriction upon the distributions of the component random variables. Since F is itself a function of c, equation (2) is actually

$$P(S \geqq t\sigma) < \exp\left(\frac{c^2\sigma^2}{2(1 - cW)} - ct\sigma\right). \qquad (2a)$$

The right-hand side of this equation is minimized by

$$cW = 1 - \left(1 + 2\frac{W}{\sigma}t\right)^{-1/2}$$

and substitution of this value of cW into equation (2a) gives

$$P(S \geqq t\sigma) < \exp\left\{-\frac{t^2}{1 + \frac{W}{\sigma}t + \sqrt{1 + 2\frac{W}{\sigma}t}}\right\}. \tag{7a}$$

If, as above, $W = M/3$,

$$P(S \geqq t\sigma) < \exp\left\{-\frac{t^2}{1 + \frac{1}{3}\frac{M}{\sigma}t + \sqrt{1 + \frac{2}{3}\frac{M}{\sigma}t}}\right\}. \tag{8a}$$

Since for identical values of M/σ and t, the denominator of (8a) is smaller than that of (8), equation (8a) will give lower values for $P(S \geqq t\sigma)$ than will Bernstein's inequality. Figure 2 shows graphs of the improved inequality and illustrates the degree of improvement.

SECOND IMPROVEMENT

The above inequality can be further improved when the component random variables are subject to $\nu_r \leqq M^{r-2}\sigma_i^2$, which is the case for random variables bounded at M. It follows from equation (1) that for such variables

$$\begin{aligned} F_i &= \sum_{r=2}^{\infty} \frac{c^{r-2}E(x_i^r)}{\frac{1}{2}r!\sigma_i^2} \leqq \sum_{r=2}^{\infty} \frac{c^{r-2}M^{r-2}\sigma_i^2}{\frac{1}{2}r!\sigma_i^2} = \sum_{r=2}^{\infty} \frac{(cM)^{r-2}}{\frac{1}{2}r!} = \sum_{s=0}^{\infty} \frac{(cM)^s}{\frac{1}{2}(s+2)!} \\ &= 2\frac{e^{cM} - 1 - cM}{(cM)^2} \end{aligned} \tag{9}$$

and since F_i is dependent only on M and c, which are independent of i,

$$F = \max(F_i) \leqq 2\frac{e^{cM} - 1 - cM}{(cM)^2}.$$

Equation (2) then is

$$P(S \geqq t\sigma) < \exp\left[\left(\frac{\sigma}{M}\right)^2(e^{cM} - 1 - cM) - ct\sigma\right]. \tag{2b}$$

The right-hand side of this equation is minimized when

$$cM = \ln\left(1 + \frac{M}{\sigma}t\right)$$

which when substituted into equation (2b) gives

$$P(S \geqq t\sigma) < e^{t(\sigma/M)}\left(1 + t\frac{M}{\sigma}\right)^{-[t(\sigma/M) + (\sigma/M)^2]} \tag{8b}$$

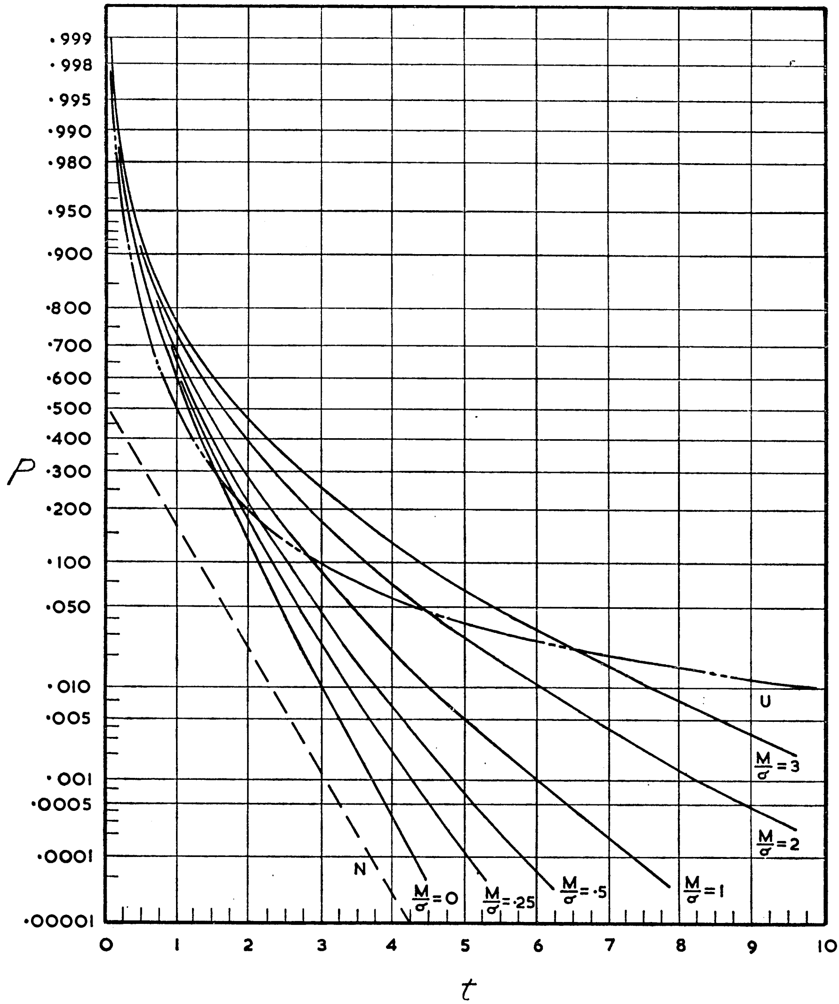

FIG. 2. Improved upper bounds on $P(S \geq t\sigma)$.
Legend: ——————— 1st Improvement
—— – – —— Uspensky's inequality
— — — — Normal Distribution

which provides lower probabilities than equations (8) or (8a). Figure 3 shows graphs of this inequality by unbroken lines (that is, $R=0$).

THIRD IMPROVEMENT

It was stated above that for random variables bounded at M, $E(x_i^r) \leq M^{r-2}\sigma_i^2$. While this equality can be attained, it cannot be attained simultaneously for all values of r. The equality sign will hold only if the distribution consists of at most three points (namely, $+M$, $-M$, and 0) with non-zero probabilities. In that case, $P(x_i=+M)$ must equal $P(x_i=-M)$ to give $E(x_i)=0$ and $E(x_i^r)$ must then be zero for odd values of r. Therefore, the equality sign in equation (9) cannot be attained under any conditions.

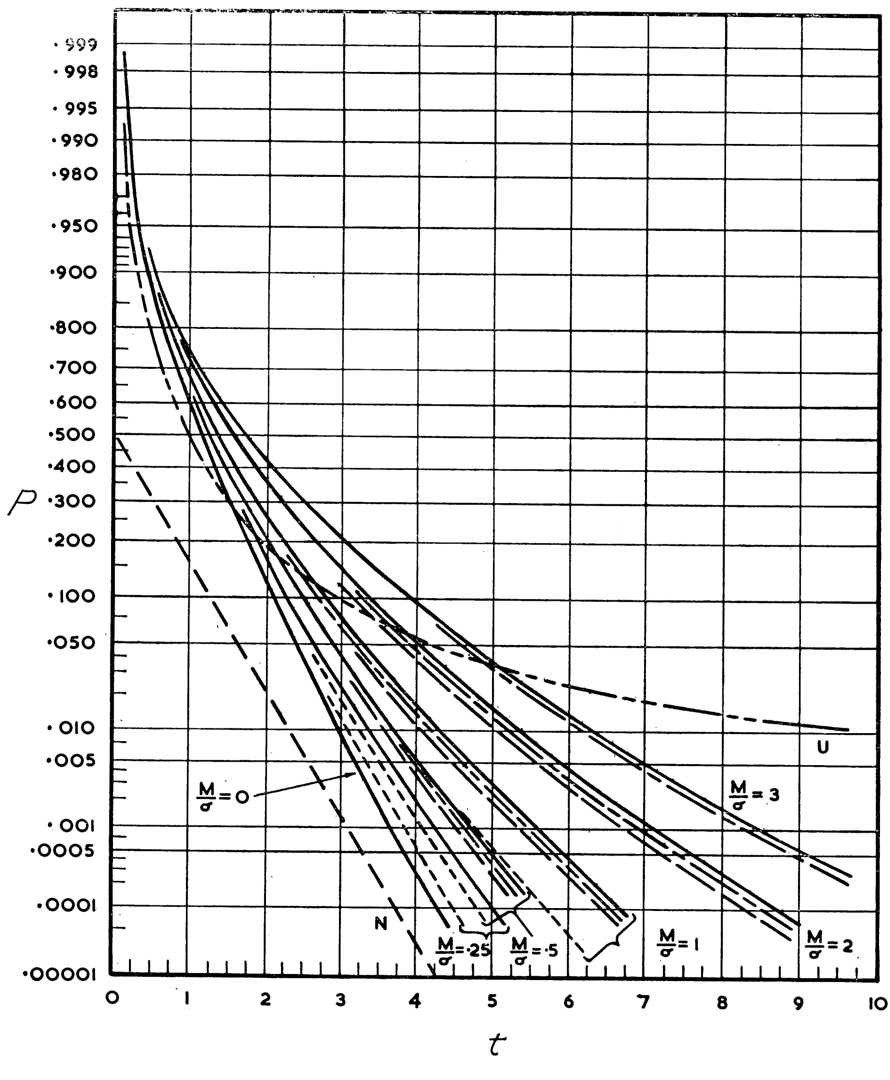

FIG. 3. Improved upper bounds on $P(S \geqq t\sigma)$.

Legend: ——————— 2nd Improvement $R=0$

—— – —— 3rd Improvement $R=\frac{1}{3}$

—— —— —— 3rd Improvement $R=\frac{1}{2}$

– – – – – – – 3rd Improvement $R=1$

—— – – – —— Uspensky's inequality

— — — — Normal Distribution

For a random variable x_i, with variance σ_i^2 and bounded at M_i, the value of $E(e^{cx_i})$ is maximized by the discrete distribution with

$$P(x_i = M_i) = \frac{\sigma_i^2}{M_i^2 + \sigma_i^2} \quad \text{and} \quad P\left\{x_i = -\frac{\sigma_i^2}{M_i}\right\} = \frac{M_i^2}{M_i^2 + \sigma_i^2} \tag{10}$$

when

$$E(e^{cx_i}) = \frac{\sigma_i^2 e^{cM_i} + M_i^2 e^{-c(\sigma_i^2/M_i)}}{M_i^2 + \sigma_i^2}. \tag{11}$$

This can be proved as follows (the suffix i has been dropped in this proof). Consider the function

$$\phi(x) = ux^2 + vx + w$$

where u, v and w are determined by the conditions

$$\phi(M) = e^{cM}, \qquad \phi\left(-\frac{\sigma^2}{M}\right) = \phi'\left(-\frac{\sigma^2}{M}\right) = \exp\left(-\frac{c\sigma^2}{M}\right).$$

Then, $e^{cx} \leqq \phi(x)$ when $x \leqq M$ and $c > 0$, so that

$$E(e^{cx}) \leqq E[\phi(x)] = E(ux^2 + vx + w) = u\sigma^2 + w. \tag{12}$$

This value of $E[\phi(x)]$ is the maximum value of $E(e^{cx})$ in equation (11), equality being attained in equation (12) with the distribution (10) since it has non-zero probability only where $\phi(x) = e^{cx}$. From equation (11)

$$\begin{aligned} E(e^{cx_i}) &= 1 + \frac{\sigma_i^2(e^{cM_i} - 1) + M_i^2(e^{-c(\sigma_i^2/M_i)} - 1)}{M_i^2 + \sigma_i^2} \\ &= 1 + \tfrac{1}{2}c^2\sigma_i^2\left[e^{cM_i} - 1 + \left(\frac{M_i}{\sigma_i}\right)^2 (e^{-cM_i(\sigma_i/M_i)^2} - 1)\right] \\ &\quad \cdot \frac{2}{(cM_i)^2} \cdot \frac{1}{1 + \left(\dfrac{\sigma_i}{M_i}\right)^2}. \end{aligned}$$

Thus,

$$F_i = \left[e^{cM_i} - 1 + \left(\frac{M_i}{\sigma_i}\right)^2 (e^{-cM_i(\sigma_i/M_i)^2} - 1)\right] \frac{2}{(cM_i)^2} \cdot \frac{1}{1 + \left(\dfrac{\sigma_i}{M_i}\right)^2}.$$

Since

$$\frac{dF_i}{d(\sigma_i^2)} < 0 \quad \text{and} \quad \frac{dF_i}{dM_i} > 0,$$

F_i is maximized by min (σ_i) and max $(M_i) = M$. Letting

$$\frac{\min(\sigma_i)}{M} = R,$$

$$F = \max(F_i) = \left[e^{cM} - 1 - \frac{1 - e^{-cMR^2}}{R^2}\right]\frac{2}{(cM)^2}\cdot\frac{1}{1 + R^2}.$$

Substituting this value of F into equation (2) gives

$$P(S \geqq t\sigma) < \exp\left[\left(\frac{\sigma}{M}\right)^2\left(e^{cM} - 1 - \frac{1 - e^{-cMR^2}}{R^2}\right)\frac{1}{1 + R^2} - ct\sigma\right]. \tag{2c}$$

The right-hand side is minimized when

$$e^{cM} - e^{-cMR^2} = t\frac{M}{\sigma}(1 + R^2),$$

which has no explicit solution. Re-arranging the last equation produces

$$t = \frac{\sigma}{M}\cdot\frac{e^{cM} - e^{-cMR^2}}{1 + R^2}$$

which in conjunction with equation (2c) provides limits to the probability in parametric form with cM as the parameter. It should be noted that these limits approach equation (8b) as R approaches zero, but elsewhere they provide a stronger inequality. Figure 3 shows the probability limits obtained for various values of M/σ and R. Values have been plotted for $R \leqq \sigma/M$ only, since $R = [\min(\sigma_i)]/M$ cannot be greater than σ/M. Graphs of $P(S \geqq t\sigma)$ for $M/\sigma = 0.25$, $R = \frac{1}{2}$ and $\frac{1}{3}$ have not been included because they cannot be distinguished on the graph.

When each component random variable has its own bound, M_i,—a case frequently arising in practical problems—the value of R is given by $\min(\sigma_i/M_i)$, which can never be less than $[\min(\sigma_i)]/M$, but may be greater. Therefore, this value of R may lead to a lower bound for $P(S \geqq t\sigma)$. This improvement is particularly useful when all the component random variables are bounded at a given number of standard deviations from their mean.

The third improvement is still applicable when the only information available about the dispersion of each component random variable is the maximum value of its variance and its bound. The value of $E(e^{cx_i})$, equation (11), is maximized by $\max(\sigma_i)$. If σ_i is below its maximum value, R will also be lower, leading to a possible increased value of F_i, but this increase is never sufficient to compensate for the decrease in σ_i^2 in equation (11). Therefore, the upper limit to $P(S \geqq t\sigma)$ obtained by using $\max(\sigma_i)$ will not be exceeded with any smaller value of σ_i. However, this does not apply to Berry's inequality, which was not designed for this extension. For example, if we have 400 component random variables, each with $\sigma_i \leqq 1$ and bounded at $M = 2$, then $\sigma \leqq 20$ and $M/\sigma \geqq 0.1$. If one random variable has the discrete distribution $P(x_1 = \frac{1}{2}) = 0.8$; $P(x_1 = -2) = 0.2$ and all the other variables have $P(x_i = 0) = 1$, then $P(S \geqq \frac{1}{2}) = P(S \geqq 0.025\sigma) = 0.8$. Berry's inequality for $M/\sigma = 0.1$ and $t = 0.025$ gives

$P<0.678$ whereas Uspensky's inequality gives $P<0.9994$.

The above analysis may be extended when more information about the distributions of the component random variables is available. Thus, for example, if all the components have symmetrical distributions

$$P(S \geqq t\sigma) < \exp\left(\frac{\sigma}{M}\right)^2\left(\cosh(cM) - 1 - cM\frac{M}{\sigma}t\right)$$

where

$$t = \frac{\sigma}{M}\sinh(cM)$$

and if all the components have uniform distributions between $+M_i$ and $-M_i$,

$$P(S \geqq t\sigma) < \exp\left(\frac{\sigma}{M}\right)^2\left(\frac{\sinh(cM)}{cM} - 1 - cM\frac{M}{\sigma}t\right)$$

where

$$t = \frac{cM\cosh(cM) - \sinh(cM)}{(cM)^2}.$$

In conclusion, it should be noted that none of the above inequalities is the best possible. The only cases in which the probability limits may be reached are the trivial ones when all the σ_i are zero; and when $M/\sigma \geqq 1$, $t \leqq M/\sigma$, one component has variance σ^2 and all other components have zero variance, Uspensky's inequality can be achieved.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to the referees for helpful suggestions. The proof that a particular distribution maximizes $E(e^{cx})$ was formulated by one of the referees and replaces a more cumbersome proof. The improvement obtained by substituting the value for F into equation (2) was discovered by the author after the original submission of the paper and was also pointed out by a second referee. The author also expresses his thanks to Mr. L. J. Birch for drawing three figures and to the staff of UTECOM (an English Electric "Deuce" computer at the University of New South Wales) for help in the calculations required for the graphs.

REFERENCES

*[1] Bernstein, S., "Sur une modification de l'inéqualité de Tchebichef," *Annals Science Institute Sav. Ukraine,* Sect. Math. I (1924). (Russian, French summary)

*[2] Bernstein, S., *Theory of Probability.* Moscow, 1927.

[3] Berry, A. C., "The accuracy of the Gaussian approximation to the sum of independent variates," *Transactions of the American Mathematical Society,* 49 (1941), 122–36.

[4] Craig, C. C., "On the Tchebychef inequality of Bernstein," *Annals of Mathematical Statistics,* 4 (1933), 94–102.

[5] Cramér, H., "Problems in probability theory," *Annals of Mathematical Statistics,* 18 (1947), 165–93.

* The author's knowledge of these two papers is limited to summaries from Craig [4] and Godwin [7].

[6] David, "Discussion on paper by Mallows (11)," *Journal of the Royal Statistical Society*, Series B, 18 (1956), 171–2.

[7] Godwin, H. J., "On generalizations of Tchebychef's inequality" *Journal of the American Statistical Association*, 50 (1955), 923–45.

[8] Kendall, M. G. and Buckland, W. R., *A Dictionary of Statistical Terms*. Edinburgh: Oliver and Boyd, 1957.

[9] Kolmogoroff, A., "Über die Summen durch den Zufall bestimmter unabhängiger Grössen," *Mathematische Annalen*, 99 (1928), 309–19.

[10] Loève, M., *Probability Theory*. New York: Van Nostrand, 1955.

[11] Mallows, C. L., "Generalizations of Tchebycheff's inequalities," *Journal of the Royal Statistical Society*, Series B, 18 (1956), 139–68; Discussion 168–76.

[12] Savage, I. R., "Probability inequalities of the Tchebycheff type," Washington: National Bureau of Standards Report No. 1744, 1952.

[13] Uspensky, J., *Introduction to Mathematical Probability*. New York: McGraw-Hill, 1937.