SOME LIMITING DISTRIBUTIONS RELATED TO THE SUM OF A RANDOM NUMBER OF RANDOM VARIABLES[1]

EDWARD PAULSON

1. **Introduction.** Let X_j $(j=1, 2, \cdots)$ be a sequence of independent random variables with the same distribution, and consider the random variable $Y=X_1+X_2+ \cdots +X_N$. The limiting distribution of Y (as $\lambda\to\infty$) when N is an integer-valued random variable whose distribution depends on a parameter λ has been studied in detail by Robbins [1].[2] In this paper we shall attempt to extend some of Robbins' results by considering a more general type of statistic Y_1, which reduces to Y in a special case. We shall also consider the limiting distribution of a second statistic Y_2, which reduces to the arithmetic mean Y/N in a special case.

Where possible, we shall follow Robbins' notation. The probability that $N=k$ $(k=0, 1, 2, \cdots)$ will be denoted by ω_k, where the $\omega_k=\omega_k(\lambda)$ are functions of λ so that, for all λ, $\omega_k\geqq 0$ and $\sum_{k=0}^{\infty}\omega_k=1$. Furthermore, we set

$$\alpha = E(N) = \sum_{k=0}^{\infty} \omega_k \cdot k,$$

$$y^2 = \text{Var}(N) = \sum_{k=0}^{\infty} \omega_k \cdot (k-\alpha)^2 \qquad (y^2 < \infty \text{ for all } \lambda),$$

$$\theta(t) = E\left[e^{it((N-\alpha)/y)}\right] = \sum_{k=0}^{\infty} \omega_k \cdot e^{it((k-\alpha)/y)}.$$

Clearly α, y^2, and $\theta(t)$ are functions of λ. We restrict ourselves in this paper to cases for which

$$\lim_{\lambda\to\infty} \theta(t) = g(t)$$

where $g(t)$ is a characteristic function, so that $(N-\alpha)/y$ has a limiting distribution as $\lambda\to\infty$.

2. **The statistic Y_1.** The statistic Y_1 to be considered is to have a distribution function $F(u)=P\{Y_1\leqq u\}$ of the form

$$F(u) = \sum_{k=0}^{\infty} \omega_k \cdot F_k(u)$$

where $F_k(u)$ denotes the conditional distribution of Y_1 when $N=k$;

Received by the editors April 15, 1949 and, in revised form, July 23, 1949.

[1] Work done under the sponsorship of the Office of Naval Research.

[2] Numbers in brackets refer to the references cited at the end of the paper.

this conditional distribution $F_k(u)$ shall be subject to the essential restriction that there exist constants a and c (with $c>0$) so that

$$\text{(1)} \qquad \lim_{k\to\infty}\int_{-\infty}^{\infty} e^{it((u-ak)/ck^{1/2})}dF_k(u) = q(t)$$

where $q(t)$ is another characteristic function. It is clear that Y_1 will reduce to Y when $F_k(u)$ is the kth convolution with itself of a random variable X such that $E(X)=a$, $E(X-a)^2=c^2$, $0<c^2<\infty$.

Now let $\sigma^2=\alpha c^2+y^2a^2$, and consider the random variable

$$Z_1 = \frac{Y_1 - a\alpha}{\sigma}.$$

THEOREM 1. *Suppose $a\neq 0$. If as $\lambda\to\infty$ limit $\alpha=\infty$, limit $\alpha/y^2=m$ $(0<m<\infty)$, then Z_1 has a limiting distribution whose characteristic function $\Phi(t)$ is given by*

$$\Phi(t) = g\left(t\left(\frac{a^2}{a^2+c^2m}\right)^{1/2}\right)\cdot q\left(t\left(\frac{c^2m}{a^2+c^2m}\right)^{1/2}\right).$$

The proof consists in showing that $\text{limit}_{\lambda\to\infty} E(e^{itZ_1})=\Phi(t)$ for every real t. In the proof we shall use $o(1)$ to denote a quantity which $\to 0$ as $\lambda\to\infty$, and $o^*(1)$ to denote a quantity which $\to 0$ as $k\to\infty$.

We have

$$E(e^{itZ_1}) = \sum_{k=0}^{\infty}\omega_k\cdot\int_{-\infty}^{\infty} e^{it((u-a\alpha)/\sigma)}dF_k(u).$$

Since the hypotheses of Theorem 1 imply $y\to\infty$ as $\lambda\to\infty$, we find that $P\{|(N-\alpha)/y|>y^{1/2}\}=o(1)$ from Tchebycheff's inequality. Hence

$$\begin{aligned} E(e^{itZ_1}) &= o(1) + \sum_{|k-\alpha|\leq y^{3/2}}\omega_k\cdot\int_{-\infty}^{\infty} e^{it((u-a\alpha)/\sigma)}dF_k(u)\\ &= o(1) + \sum_{|k-\alpha|\leq y^{3/2}}\omega_k e^{iat((k-\alpha)/\sigma)}\int_{-\infty}^{\infty} e^{it'((u-ak)/ck^{1/2})}dF_k(u)\end{aligned}$$

$$\text{where}\quad t' = t\left(\frac{ck^{1/2}}{\sigma}\right).$$

Since $t(ck^{1/2}/\sigma)=t(c^2m/(c^2m+a^2))^{1/2}+o(1)$ for $|k-\alpha|\leq y^{3/2}$, we see that t' lies in a finite interval for any fixed value of t. Since (1) implies uniform convergence in any finite interval, we have uniformly for all k such that $|k-\alpha|\leq y^{3/2}$

$$\begin{aligned}\int_{-\infty}^{\infty} e^{it'((u-ak)/ck^{1/2})}dF_k(u) &= q(t') + o^*(1)\\ &= q(t') + o(1)\\ &= q\left(t\left(\frac{c^2m}{c^2m + a^2}\right)^{1/2}\right) + o(1),\end{aligned}$$

making use of the fact that $\alpha - y^{3/2} \to \infty$ as $\lambda \to \infty$, and that a characteristic function is everywhere continuous. Combining these results we obtain

$$E(e^{itZ_1}) = o(1) + q\left(t\left(\frac{c^2m}{c^2m + a^2}\right)^{1/2}\right) \sum_{|k-\alpha| \leqq y^{3/2}} \omega_k \cdot e^{iat((k-\alpha)/\sigma)}.$$

Now

$$\begin{aligned}\sum_{|k-\alpha| \leqq y^{3/2}} \omega_k \cdot e^{iat((k-\alpha)/\sigma)} &= o(1) + \sum_{k=0}^{\infty} \omega_k \cdot e^{iat(y/\sigma)((k-\alpha)/y)}\\ &= o(1) + \left[o(1) + g\left(t\frac{ay}{\sigma}\right)\right]\\ &= o(1) + g\left(t\left(\frac{a^2}{c^2m + a^2}\right)^{1/2}\right).\end{aligned}$$

We now have

$$E(e^{itZ_1}) = o(1) + g\left(t\left(\frac{a^2}{c^2m + a^2}\right)^{1/2}\right) \cdot q\left(t\left(\frac{c^2m}{c^2m + a^2}\right)^{1/2}\right),$$

which proves Theorem 1.

By proceeding in a similar manner, it is easy to verify that if either of the two sets of conditions (1) $a \neq 0$, limit $\alpha = \infty$, limit $\alpha/y^2 = \infty$, or (2) $a = 0$, limit $\alpha = \infty$, limit $\alpha/y^2 = m'$ $(0 < m' \leqq \infty)$ hold, then Z_1 has a limiting distribution with characteristic function $q(t)$.

3. **The statistic** Y_2. The statistic Y_2 again has a distribution function $F(u) = P\{Y_2 \leqq u\}$ of the form

$$F(u) = \sum_{k=0}^{\infty} \omega_k \cdot F_k(u)$$

but the conditional distribution $F_k(u)$ of Y_2 when $N = k$ is now subject to the restriction that there exist a constant a such that

$$\lim_{k\to\infty} \int_{-\infty}^{\infty} e^{itk^{1/2}(u-a)}dF_k(u) = m(t)$$

where $m(t)$ is a characteristic function.

Now let $Z_2=\alpha^{1/2}(Y_2-a)$.

THEOREM 2. *If as* $\lambda\rightarrow\infty$ *limit* $\alpha=\infty$, *limit* $\alpha/y^2=m'$ $(0<m'\leqq\infty)$, *then* Z_2 *has a limiting distribution corresponding to the characteristic function* $m(t)$.

The proof again consists in showing that $\text{limit}_{\lambda\to\infty} E(e^{itZ_2})=m(t)$.

$$
\begin{aligned}
E(e^{itZ_2}) &= \sum_{k=0}^{\infty}\omega_k\cdot\int_{-\infty}^{\infty}e^{it\alpha^{1/2}(u-a)}dF_k(u)\\
&= o(1)+\sum_{|k-\alpha|\leqq y^{3/2}}\omega_k\cdot\int_{-\infty}^{\infty}e^{it(\alpha/k)^{1/2}[k^{1/2}(u-a)]}dF_k(u)\\
&= o(1)+\sum_{|k-\alpha|\leqq y^{3/2}}\omega_k\left[m\left(t\left(\frac{\alpha}{k}\right)^{1/2}\right)+o(1)\right]\\
&= o(1)+m(t)\sum_{|k-\alpha|\leqq y^{3/2}}\omega_k\\
&= o(1)+m(t),
\end{aligned}
$$

which proves Theorem 2.

4. Some examples. We conclude this paper by giving an example involving Theorem 1, and an example involving Theorem 2. First let N have a Poisson distribution with mean λ, so that $\omega_k=e^{-\lambda}\lambda^k/k!$, $\alpha=\lambda$, $y^2=\lambda$. Let $\bar{x}=(1/N)\sum_{i=1}^N x_i$ and put $Y_1=\sum_{i=1}^N (x_i-\bar{x})^2$ where $\{x_i\}$ $(i=1, 2, \cdots)$ is a sequence of independent observations on a random variable whose fourth central moment μ_4 is finite. If we make use of the known fact that the sample variance (based on k observations) is asymptotically normal with mean μ_2 and variance $(\mu_4-\mu_2^2)/k$, it follows from Theorem 1 that as $\lambda\rightarrow\infty$ the limiting distribution of $(Y_1-\lambda\mu_2)/(\lambda\mu_4)^{1/2}$ is normal with zero mean and unit variance.

For the second example, let N have a binomial distribution with parameters λ, p, $q=1-p$, so that $\omega_k=C_{\lambda,k}p^kq^{\lambda-k}$ $(k=0, 1, 2, \cdots, \lambda)$. Now let $Y_2=r/N$ where r is the number of successes in N independent trials with constant probability P of a success. It follows from Theorem 2 that $(\lambda p)^{1/2}[Y_2-P]$ has in the limit a normal distribution with zero mean and variance $P(1-P)$. The statistic in the second example arose in a problem considered by Birnbaum and Sirken where Y_2 represented the percentage of people voting yes when only N out of a random sample of λ were available for interviewing. The limiting distribution of Y_2 for this special case was obtained by Birnbaum and Sirken by a different method.

References

1. H. Robbins, *The asymptotic distribution of the sum of a random number of random variables*, Bull. Amer. Math. Soc. vol. 54 (1948) pp. 1151–1162.

2. H. Cramér, *Mathematical methods of statistics*, Princeton University Press, 1946, p. 365.

3. Z. W. Birnbaum and M. Sirken, *Bias due to non-availability in sampling surveys*, Journal of the American Statistical Association vol. 45 (1950) pp. 98–111.

University of Michigan

BOUNDS FOR THE COEFFICIENTS OF UNIVALENT FUNCTIONS

ARYEH DVORETZKY

1. The purpose of this note is to give some estimates for the moduli of the coefficients a_n of

$$w = f(z) = z + a_2z^2 + \cdots + a_nz^n + \cdots, \tag{1}$$

assumed regular and univalent in $|z| < 1$, in terms of the domain onto which $|z| < 1$ is mapped through (1). A typical result, cf. (27), is that if this domain does not cover arbitrarily large circles, then[1] $a_n = O(\log n)$.

Let W be the domain in the w-plane onto which $|z| < 1$ is mapped through (1) and denote by $A(R)$ the radius of the largest circle with center on $|w| = R$ the whole interior of which is contained in W, that is,

$$A(R) = \max_{|w|=R} \min_{w' \notin W} |w' - w| \qquad (0 \leqq R < \infty).$$

Our aim is to derive upper bounds for $|a_n|$ in terms of $A(R)$.

Always $A(R) \leqq R+1$ while $A(R) = 0$ for sufficiently large R if and only if (1) is bounded in $|z| < 1$. The function $A(R)$ measures, in some sense, the extension of W. This "extension" has, however, little to do with area, in fact whatever the positive function $a(R)$ there exists a function (1) mapping $|z| < 1$ on a slit domain and for which $A(R) = o(a(R))$.

Presented to the Society, August 30, 1949; received by the editors March 31, 1949 and, in revised form, July 6, 1949.

[1] Throughout the paper, O and o refer to $n \to \infty$ or $R \to \infty$ according to which of the variables n or R appears in the formula.