STIRLING BEHAVIOR IS ASYMPTOTICALLY NORMAL[1]

By L. H. Harper

The Rockefeller University

0. Summary. The Stirling numbers $\{\sigma_n^j\}$ of the second kind are asymptotically normal. This result is similar to results achieved by Feller [1] and Gončarov [2] for other combinatorial distributions. Here the technique of proof is different; one of the most general forms of the central limit theorem is used.

Interesting qualitative information about the Stirling numbers is also obtained from this result. Asymptotic estimates on the value of $\max_j \{\sigma_n^j\}$ are given.

1. Introduction. Mathematicians have been aware for quite a while that probability theory has combinatorial applications. The classical De Moivre-Laplace theorem, for example, can be interpreted as a theorem about binomial coefficients, the binomial coefficients being the solution to the difference equation

$$A_{nj} = A_{n-1,j} + A_{n-1,j-1}$$

with the boundary conditions

$$\begin{aligned} A_{0j} &= 1, \qquad j = 0, \\ &= 0, \qquad j \neq 0. \end{aligned}$$

Feller ([1], p. 241) uses more general versions of the central limit theorem to show that the distributions given by B_{nj}, the number of permutations of n elements with j inversions, and C_{nj}, the number of permutations of n elements with j cycles, are asymptoticlly normal. There he defines random variables on the set of all permutations to count either inversions or cycles. He shows that these random variables are independent and satisfy the "Lindeberg condition" ([1], p. 239), and thus have asymptotically normal distributions.

It can be verified that B_{nj} and C_{nj} are the solutions of the difference equations

$$B_{nj} = \textstyle\sum_{k=m}^{j} B_{n-1,k}\,, \qquad M = \max(0, j - n + 1)$$

and

$$C_{nj} = (n-1)C_{n-1,j} + C_{n-1,j-1}$$

respectively, with boundary conditions

$$\begin{aligned} B_{oj} &= 1, \qquad j = 0, \\ &= 0, \qquad j \neq 0, \end{aligned}$$

Received 11 October 1966.

[1] This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

and

$$C_{0j} = 1, \qquad j = 0,$$
$$= 0, \qquad j \neq 0.$$

In view of the similarity of these three results, it seems appropriate to call any "generalized Pascal's triangle" on the lattice points of the positive quadrant, defined by a difference equation with the common boundary condition above, a *combinatorial distribution.*

V. Gončarov [2] has shown asymptotic normality for B_{nj}, C_{nj} and other combinatorial distributions by a different and less elegant method. By brute force he tortuously manipulates the characteristic functions of the distributions until they approach $\exp(-x^2/c)$, c a positive constant. His attack is certainly the most general conceptually, but hardly the most efficient, and not even feasible in cases where there is lack of knowledge about the characteristic functions involved.

This paper presents a program for showing asymptotic normality of combinatorial distributions, much like Feller's in that it gives the problem a probabilistic interpretation, and uses the central limit theorem in an essential way.

2. The Stirling numbers of the second kind. The Stirling numbers of the second kind are combinatorially distributed by the following difference equation:

$$\sigma_n^{\,j} = j\sigma_{n-1}^{j} + \sigma_{n-1}^{j-1}.$$

Several preliminary lemmas are needed.

LEMMA 1. *If* $P_n(x) = \sum_{j=0}^{n} \sigma_n^{\,j} x^j$, *then the roots of* P_n *are real, distinct, and non-positive for all* $n = 1, 2, \cdots$.

PROOF. By induction $P_0(x) = 1$, so the statement is vacuously true for $n = 0$; for other values of n,

$$\begin{aligned} P_n(x) &= \sum_{j=0}^{n} \sigma_n^{\,j} x^j = \sum_{j=0}^{n} j\sigma_{n-1}^{j} x^j + \sum_{j=0}^{n} \sigma_{n-1}^{j-1} x^j \\ &= x\left[\sum_{j=0}^{n-1} j\sigma_{n-1}^{j} x^{j+1} + \sum_{j=0}^{n-1} \sigma_{n-1}^{j} x^{j}\right] \\ &= x[dP_{n-1}(x)/dx + P_{n-1}(x)]. \end{aligned}$$

Therefore $P_1(x) = x$, and $P_2(x) = x(1 + x) = x + x^2$, so the statement still holds for $n = 1, 2$. Now suppose $n > 2$. By hypothesis, P_{n-1} has $n - 1$ distinct real non-positive roots.

If we define

$$H_n(x) = P_n(x)e^x,$$

then H_n has exactly the same finite zeroes that P_n does, and the identity for P_n above becomes

$$H_n(x) = x \cdot dH_{n-1}(x)/dx.$$

H_{n-1} also has a zero at $-\infty$, and by Rolle's theorem between any two zeroes of

H_{n-1}, dH_{n-1}/dx will have a zero. This places $n - 1$ distinct zeroes of H_n on the negative axis; $x = 0$ is obviously another one making n altogether. Since P_n is of degree n by induction, we have found all roots and the lemma is proved.

A curious property of the Stirling numbers appears as a consequence of Lemma 1: A combinatorial distribution is called *unimodal* if all local maxima are consecutive. It is easily seen that the product of the generating functions for a unimodal distribution and a distribution concentrated on two consecutive integers is the generating function of a unimodal distribution. By repeated application of this the Stirling numbers are unimodal and in fact have either one or two maximum points. By the same analysis the second difference of the Stirling numbers changes sign just twice.

3. Bell numbers. The sum $B_n = \sum_{j=0}^{n} \sigma_n^{\,j}$ is called the Bell number of order n. We now show

LEMMA 2.

$$B_{n+2}/B_n - (B_{n+1}/B_n)^2 \to \infty \quad as \quad n \to \infty.$$

PROOF. Actually[2]

$$B_{n+2}/B_n - (B_{n+1}/B_n)^2 = n/R(R+1) + o(1)$$

where R is the unique, real solution of $Re^R = n$. This is straight forward to establish from a formula due to Moser and Wyman [4]:

$$B_n \sim (R+1)^{-\frac{1}{2}} \exp\,[n(R + R^{-1} - 1) - 1]$$
$$(1 - R^2(2R^2 + 7R + 10)/24n(R+1)^3).$$

4. Main theorem. We are now in a position to prove the main theorem.

THEOREM. *The Stirling numbers of the second kind are asymptotically normal in the sense that*

$$B_n^{-1} \sum_{j=1}^{x_n} \sigma_n^{\,j} \to (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-t^2/2}\,dt \quad as \quad n \to \infty$$

where

$$x_n = \{B_{n+2}/B_n - (B_{n+1}/B_n)^2 - 1\}^{\frac{1}{2}} x + \{B_{n+1}/B_n - 1\}.$$

PROOF. The result is an application of the "bounded variance normal convergence criterion" in [3], p. 295. It is stated: "Let the independent summands $\{X_{nk}\}_{k=1}^{m_n}$, centered at expectations, be such that $\sum \operatorname{Var}(X_{nk}) = 1$ for all n. Let F_{nk} be the distribution function of X_{nk}. Then $S_n = \sum_k X_{nk}$ converges normally with mean zero, unit variance and $(\max_k \operatorname{Var}(X_{nk})) \to 0$, if and only if: for all $\epsilon > 0$,

$$\lim_{n\to\infty} g_n(\epsilon) = \lim_{n\to\infty} \sum_k \int_{|x| \geqq \epsilon} t^2\, dF_{nk} = 0.$$

In order to use this theorem, we need a "hat" from which to pull a double

[2] I am indebted to J. Haigh for bringing this improvement to my attention.

sequence of independent (for fixed n) random variables. The hat is Lemma 1. That the roots of the polynomial P_n are real and non-positive is equivalent to the fact that P_n can be factored into linear terms with real non-negative coefficients. If we normalize each of these terms suitably, we see that the distribution whose density function is $\{\sigma_n^j/B_n\}_{j=0}^n$ is the distribution of a sum of independent random variables taking on only the values zero and one. If $-x_{nj}$ is a root of P_n, then define the random variable X'_{nk} by

$$\Pr[X'_{nk} = y] = x_{nk}/(1 + x_{nk}) \qquad \text{if} \quad y = 0,$$
$$= (1 + x_{nk})^{-1} \qquad \text{if} \quad y = 1.$$

Letting $S_n' = \sum_k X'_{nk}$ we have

$$E(S_n') = \sum_{j=0}^n j\sigma_n^j/B_n = B_{n+1}/B_n - 1$$

and

$$\mathrm{Var}(S_n') = \sum_{j=0}^n j^2\sigma_n^j/B_n - (B_{n+1}/B_n - 1)^2$$
$$= B_{n+2}/B_n - (B_{n+1}/B_n)^2 - 1.$$

Thus, by Lemma 2, $\mathrm{Var}(S_n') \to \infty$ as $n \to \infty$.

Now we normalize and let

$$S_n = (\mathrm{Var}(S_n'))^{-\frac{1}{2}}(S_n' - E(S_n'))$$
$$= \sum_k (\mathrm{Var}(S_n'))^{-\frac{1}{2}}(X'_{nk} - E(X'_{nk})) = \sum_k X_{nk}.$$

Since $0 \leqq X'_{nk} \leqq 1$ and $-1 \leqq X'_{nk} - E(X'_{nk}) \leqq 1$, and since given $\epsilon > 0$ there exists N such that $|X_{nk}| < \epsilon$ for all $n \geqq N$, we conclude:

$$g_n(\epsilon) = \sum_k \int_{|x|>\epsilon} x^2\, dF_{nk} = 0, \qquad \text{for all} \quad n \geqq N.$$

Thus, the hypotheses of the normal convergence criterion are fulfilled. This finally proves the main theorem.

5. Corollaries. Since the Stirling numbers are unimodal, the maximum must lie near $E(S_n')$. With a bit more effort we can gain more qualitative information from the theorem.

COROLLARY 1.

$$B_n^{-2}(B_nB_{n+2} - B_{n+1}^2 - B_n^2)^{\frac{1}{2}}\sigma_n^{[x_n]} \to (2\pi)^{-\frac{1}{2}}e^{-x^2/2}$$

where $[x_n]$ *is the greatest integer* $\leqq x_n$.

COROLLARY 2. *Given* $\epsilon > 0$, *there exists* N *such that* $n \geqq N$ *implies*

$$|J_n - (B_{n+1}/B_n - 1)| < \epsilon(B_{n+2}/B_n - (B_{n+1}/B_n)^2 - 1)^{\frac{1}{2}}$$

where J_n *is defined as that integer* j *for which* $\sigma_n^j = \max_j \sigma_n^j$. *Thus* $J_n \sim B_{n+1}/B_n - 1 \sim n/e \ln n$. Also

$$\max_j \sigma_n^j \sim B_n^2/(2\pi)^{\frac{1}{2}}(B_nB_{n+2} - B_{n+1}^2 - B_n^2).$$

These corollaries follow from the fact that the convergence of the Stirling numbers to the Gaussian function is actually uniform. Given that the integrals converge as in Theorem 1, they can escape from uniform convergence only by wild local oscillations, but this is ruled out by the fact that the second differences only change sign twice.

REFERENCES

[1] FELLER, W. (1957). *An Introduction to Probability Theory and Its Applications*, **1** (Second Edition). Wiley, New York.

[2] GONČAROV, V. (1944). Du domaine d'analyse combinatoire. *Bull. Acad. Sci. USSR Ser. Math.* **8** 3–48. (In Russian with French summary.)

[3] LOÈVE, M. (1960). *Probability Theory* (Second Edition). Van Nostrand, Princeton.

[4] MOSER, L. and WYMAN, M. (1955). An asymptotic formula for the Bell numbers. *Trans. Roy. Soc. Canad.* **49** 49–53.

[5] ROTA, G. C. (1964). The number of partitions of a set. *Amer. Math. Monthly* **71** 498–503.