For small n these equations can be solved by iteration, which was done in constructing Table 1. Initial conditions an $U_k(0) = 1$, $U_i(0) = 0$ for $i \neq k$. It might be noted that the $U_i(j + 1)$ are subtotals of the $U_i(j)$ so that the iteration proceeds very rapidly on an adding machine. The probability that $d \leq k/n$ is $[U_0(n) + U_1(n) + U_2(n) \cdots + U_k(n)]n!n!/(2n)!$.

REFERENCES

[1] W. Feller, "On the Kolmogorov-Smirnov theorems," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 177–189.

[2] F. Massey, "A note on the estimation of a distribution function by confidence limits," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 116–119.

[3] F. Massey, "A note on the power of a non-parametric test," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 440–443.

[4] N. Smirnov, "On the estimation of the discrepancy between empirical curves of distribution for two independent samples," *Bulletin Mathématique de l'Université de Moscou*, Vol. 2 (1939), fasc. 2.

[5] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Annals of Math. Stat.*, Vol. 19 (1948), pp. 279–281.

A NOTE ON THE SURPRISE INDEX

By R. M. Redheffer

Harvard University

Let $p_m (m = 0, 1, 2, \cdots)$ be a set of probabilities of events E_m, and suppose that the event E_i, with probability p_i, actually occurred. Is the fact that E_i occurred to be regarded as surprising? In a recent article [1] this question is answered by introducing the surprise index S_i,

$$S_i = (\Sigma p_m^2)/p_i, \tag{1}$$

which gives a comparison between the probability expected and that actually observed.[1] The event is to be regarded as surprising when S_i is large.

The author remarks on the difficulty of computing (1) for the Poisson and binomial distribution. The problem consists in evaluating the numerator, which we shall express here in terms of tabulated functions. The Poisson case leads to Bessel functions, the binomial case to Legendre or hypergeometric functions, and the asymptotic behavior involves square roots only.

1. *The Poisson case.* For the Poisson case we have $p_m = \lambda^m e^{-\lambda}/m!$ so that the generating function is

$$e^{-\lambda} e^{\lambda x} = \Sigma p_m x^m. \tag{2}$$

Let $x = e^{i\theta}$, then $e^{-i\theta}$; multiply; integrate from 0 to 2π; and simplify slightly to obtain

$$\Sigma p_m^2 = (e^{-2\lambda}/\pi) \int_0^\pi e^{2\lambda \cos\theta}\, d\theta. \tag{3}$$

[1] Cf. also [6].

The integral on the right is recognized[2] as the zero-order Bessel function [2] so that we have

(4) $$\Sigma p_m^2 = e^{-2\lambda} J_0(-2i\lambda) = e^{-2\lambda} I_0(-2\lambda)$$

as the final answer. The relevant tables are listed on pages 271, 272, and 275 of [5].

2. *The binomial case.* When $p_m = C_m^n p^m q^{n-m}$ with $q = 1 - p$, the value of Σp_m^2 for $p = q = \frac{1}{2}$ is given in the literature [3]; it is the product of the first n odd integers, divided by the product of the first n even integers. For general p,

(5) $$(q + px)^n = \Sigma p_m x^m$$

is the equation corresponding to (2). Following through the derivation of (3), we get

(6) $$\Sigma p_m^2 = \frac{1}{2\pi} \int_0^{2\pi} (p^2 + 2pq \cos \theta + q^2)^n \, d\theta$$

which is recognized as the n^{th} order Legendre function [4],

(7) $$\Sigma p_m^2 = |p - q|^n P_n\left(\left|\frac{p^2 + q^2}{p - q}\right|\right) \qquad (p \neq q).$$

For tables see [5], pages 232–235, 242.

The result (6) is also expressible as a hypergeometric function, and this without intervention of (7). The change of variable $u = p^2 + 2pq \cos \theta + q^2$ leads to

(8) $$\Sigma p_m^2 = (1/\pi) \int_a^1 u^n (u - a)^{-1/2} (1 - u)^{-1/2} \, du$$

with $a = (p - q)^2$, and letting $u = a + (1 - a)x$ gives an integral which turns out to be [4]

$$\Sigma p_m^2 = (p - q)^{2n} F[-n, \tfrac{1}{2}; 1; -4pq/(p - q)^2].$$

It was brought to the author's attention, by Weaver himself via Mosteller, that (7) is given in Pólya-Szegö, Vol. II, p. 92. There, however, the point of view is to evaluate the integral rather than the sum, and hence the above derivation is the more natural here.

3. *Approximation.* For large values of λ, (4) gives [2]

(9) $$\Sigma p_m^2 \sim \frac{1}{2\sqrt{\pi\lambda}}.$$

To obtain the asymptotic behavior in the binomial case, note that if the limits of integration in (8) were $0 - 1$, and if the factor $(u - a)^{-1/2}$ were absent, we should have the Beta function $\mathrm{B}(n + 1, \frac{1}{2})$. Because u^n emphasizes the region

---

[2] This connection between (3) and the Bessel function was pointed out to the author by by M. V. Cerrillo of M. I. T.

near $u = 1$, this resemblance may be exploited to give (after elementary but tedious calculations)

$$\pi \Sigma p_m^2 = B(n + 1, \tfrac{1}{2}) + e \tag{10}$$

with

$$0 < e < 2e^{-n\delta} + (\tfrac{2}{3})[\delta/(1 - a)]^{3/2}$$

whenever $n > a/(1 - a)$. Here δ is any number $< pq$. Picking $\delta = n^{-\theta}$, $\theta < 1$, shows that the error goes to zero almost as fast as $n^{-3/2}$. A similar result may be obtained by the methods of Uspensky.

From (10) we have easily

$$\Sigma p_m^2 \sim 1/(2\sqrt{\pi npq}) \qquad (n \to \infty), \tag{11}$$

which is correct even for $p = q$.

It was pointed out by the referee that (9) and (11) are special cases of the relation

$$\Sigma p_m^2 \sim (\tfrac{1}{2}) \sqrt{\text{variance}}$$

which generally holds whenever the shape of the distribution curve approaches a limit.

REFERENCES

[1] W. WEAVER, "Probability, rarity, interest, and surprise," *The Scientific Monthly*, Vol. 67 (1948), p. 390.
[2] JAHNKE AND EMDE, *Tables of Functions*, 3rd rev. ed., Dover Publications, 1943, p. 149, p. 117.
[3] HALL AND KNIGHT, *Higher Algebra*, Macmillan Co., 1936, p. 148.
[4] WHITAKER AND WATSON, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, 1940, p. 312, p. 293.
[5] FLETCHER, MILLER AND ROSENHEAD, *An Index to Mathematical Tables*, Science Computing Service, Ltd., London, 1946.
[6] T. C. FRY, *Probability and Its Engineering Uses*, D. Van Nostrand Co., 1928, pp. 199–200.

APPROXIMATION TO THE POINT BINOMIAL

BY BURTON H. CAMP
Wesleyan University

The following approximation to the sum of the first $(t + 1)$ terms of the point binomial appears to be useful. Let this sum be denoted by S_{t+1}, and let the point binomial be the expansion of $(p + q)^N$; *i.e.*, let

$$S_{t+1} = p^N + Np^{N-1}q + \cdots + \binom{N}{t} p^{N-t}q^t. \tag{1}$$