JOURNAL OF COMBINATORIAL THEORY 7, 258–261 (1969)

NOTE

More Birthday Surprises

MEYER DWASS

Department of Mathematics, Northwestern University, Evanston, Illinois 60201

Communicated by Gian-Carlo Rota

Received March 21, 1969

1. The phrase "birthday surprise" refers to the elementary fact that the expected number of persons that one must stop to discover two who have the same birthday is only about 24. In general, given n equally likely alternatives, we choose among them repeatedly until any one of the alternatives has appeared k times. Using the notation $E(n, k)$ for the expected number of choices required, Klamkin and Newman [1] have shown the following interesting formula:

$$E(n, k) = \int_0^\infty [S_k(t/n)]^n \, e^{-t} \, dt, \tag{1}$$

where $S_k(t) = \sum_{j<k} t^j/j!$ Using (1), Klamkin and Newman then deduced the asymptotic relationship,

$$\lim_{n\to\infty} E(n, k)/n^{1-1/k} = \int_0^\infty \exp - (u^k/k!) \, du = (k!)^{1/k}\Gamma(1 + 1/k) \tag{2}$$

Our purpose in this note is to obtain asymptotic expressions for the higher moments and consequently to discover the asymptotic distribution of the waiting time. Specifically, we show the following. Let W_n be the number of choices required when there are n alternatives in order for any one of the alternatives to appear k times. We suppose that k is greater than S. Then,

$$\lim_{n\to\infty} E[(W_n/n^{1-1/k})^r] = \int_0^\infty u^r f(u) \, du, \qquad r = 0, 1, 2,..., \tag{3}$$

where

$$f(u) = \begin{cases} [(k-1)!]^{-1} \, u^{k-1} \exp - (u^k/k!) & \text{if } u \geqslant 0, \\ 0 & \text{if } u < 0. \end{cases} \tag{4}$$

Since f describes probability distribution which is completely determined by its moments, it follows that for $t \geqslant 0$,

$$\lim_{n\to\infty} P(W_n/n^{1-1/k} \leqslant t) = \int_0^t f(u)\,du = 1 - \exp - (t^k/k!)$$

(Notice that (3) agrees with (2), after integration by parts, for $r = 1$.)

2. PROOFS. Relation (1) is proved in [1] by an extension of a combinatorial method of Newman and Shepp [2]. Our first task to provide a somewhat different and direct probabilistic method which allows the evaluation of higher moments of W_n. We assume throughout that $k > 1$.

LEMMA 1.

$$\sum_{m=0}^{\infty} P(W_n + r > m)\, m^{(r)} = \int_0^{\infty} [S_k(t/n)]^n t^r e^{-t}\,dt, \qquad r = 0, 1, 2, \ldots, \tag{5}$$

where $m^{(r)} = m(m-1)\cdot \ldots \cdot (m-r+1)$. $(0^{(0)} = 0)$. *Notice that, for* $r = 0$, (5) *agrees with the Klamkin-Newman formula* (1).

PROOF OF LEMMA 1: Suppose that $X_1, X_2, \ldots, X_n$ are independent, Poisson-distributed random variables, each with parameter t/n, where $t > 0$. Then $X_1 + \ldots + X_n = T$ is Poisson distributed with parameter t. Also, the conditional distribution of $X_1, \ldots, X_n$, given that T is fixed, is multinomial and does not depend on t. Specifically,

$$\begin{aligned} &P(X_1 = i_1, \ldots, X_n = i_n \mid T = m = i_1 + \cdots + i_n) \\ &= \frac{m!}{i_1! \cdot \cdots \cdot i_n!} (1/n)^m, \qquad m = 0, 1, 2, \ldots. \end{aligned}$$

Hence, the conditional probability of the event that each of the X_i's is less than k, given that $T = m$, is the same as $P(W_n > m)$. It follows that

$$\Big[\sum_{j<k} e^{-t/n}(t/n)^j/j!\Big]^n = \sum_{m=0}^{\infty} P(W_n > m)\, e^{-t} t^m/m!$$

or, equivalently,

$$[S(t/n)]^n = \sum_{m=0}^{\infty} P(W_n > m)\, t^m/m!. \tag{6}$$

Now multiply both sides of (6) by $t^r e^{-t}$ and integrate out t from 0 to ∞ to obtain (5).

LEMMA 2. *Suppose that W is a non-negative, integer-valued random variables having all moments. Then we have the following expression for the factorial moments of W:*

$$E(W^{(r)}) = r \sum_{m=0}^{\infty} m^{(r-1)} P(W > m), \qquad r = 1, 2, \ldots \tag{7}$$

PROOF: It is well known that, for $|s| < 1$,

$$E(s^W) = \sum_{r=0}^{\infty} E(W^{(r)})(s-1)^r/r!.$$

Hence,

$$\frac{1 - E(s^W)}{1 - s} = \sum_{r=1}^{\infty} E(W^{(r)})(s-1)^{r-1}/r!. \tag{8}$$

But, also,

$$\frac{1 - E(s^W)}{1 - s} = \sum_{m=0}^{\infty} P(W > m)\, s^m. \tag{9}$$

The relation (7) now follows by successively differentiating (8) and (9) with respect to s and letting $s \nearrow 1$.

LEMMA 3.

$$E((W_n + r - 1)^{(r)}) = r \int_0^{\infty} [S_k(t/n)]^n t^{r-1} e^{-t}\, dt, \qquad r = 1, 2, \ldots \tag{10}$$

PROOF: Use (5) with $r - 1$, and let $W_n + r - 1$ play the role of W in (7).

Notice that (10) agrees with (1) for $r = 1$.

LEMMA 4.

$$\lim_{n\to\infty} E(W_n + r - 1)^{(r)})/n^{(1-1/k)r} = r \int_0^{\infty} u^{r-1} \exp - (u^k/k!)\, du. \tag{11}$$

PROOF: The calculation needed is practically the same as that done in [1] for the special case of $r = 1$. We refer the reader to that paper for the details.

PROOF OF (3): The proof is now completed by observing that the left side of (11) is the same as

$$\lim_{n\to\infty} E((W_n/n^{1-1/k})^r).$$

Moreover, the right side of (11) is the same as $\int_0^\infty u^r f(u)\, du$ after integrating by parts. This completes the proof.

REFERENCES

1. M. S. KLAMKIN AND D. J. NEWMAN, Extensions of the Birthday Surprise, *J. Combinatorial Theory* **3** (1967), 279–282.
2. D. J. NEWMAN AND L. SHEPP, The Double Dixie Cup Problem, *Amer. Math. Monthly* **67** (1960), 58–61.