Z. Wahrscheinlichkeitstheorie 2, 118—121 (1963)

The Distribution of Majority Times in a Ballot

By

LAJOS TAKÁCS

1. Introduction

Suppose that in a ballot candidate A scores a votes and candidate B scores b votes and that all the possible $\binom{a+b}{a}$ voting records are equally probable. Denote by α_r and β_r the number of votes registered for A and B respectively among the first r votes recorded. Let c and μ be nonnegative integers. Denote by P_j the probability that the inequality $\alpha_r > \mu\beta_r$ holds for exactly j values among $r = 1, 2, \ldots, a+b$, and by P_j^* the probability that the inequality $\alpha_r > \mu\beta_r - c$ holds for exactly j values among $r = 1, 2, \ldots, a+b$.

In 1887 É. BARBIER [3] found that

$$P_{a+b} = \frac{a - \mu b}{a + b} \tag{1}$$

if $a \geqq \mu b$ and this was proved by A. AEPPLI [1] in 1924. However, it is of some interest to find the complete distribution $\{P_j\}$. In this paper we shall prove the following theorems:

Theorem 1. *If* $a > \mu b + 1$, *then*

$$P_j = \frac{(a - b\mu - 1)}{(a+b)(a+b-1)} \sum_{\frac{a+b-j}{\mu+1} \leqq s \leqq b} \frac{\binom{a}{s\mu+1}\binom{b}{s}}{\binom{a+b-2}{s(\mu+1)}} \tag{2}$$

for $j = 0, 1, \ldots, a+b-1$, *and if* $a = \mu b + 1$, *then*

$$P_j = \frac{1}{(a+b)} \tag{3}$$

for $j = 1, 2, \ldots, a+b$.

Theorem 2. *If* $a > \mu b - c$, *then*

$$P^*_{a+b} = 1 - \frac{a + c - b\mu}{a+b} \sum_{\frac{c+1}{\mu+1} \leqq s \leqq b} \frac{\binom{a}{s\mu - c}\binom{b}{s}}{\binom{a+b-1}{s(\mu+1)-c}}\,. \tag{4}$$

The proofs are based on two combinatorial theorems which have special interest in fluctuation theory, in order statistics and in the theory of queues.

2. Two combinatorial theorems

Let $\nu_1, \nu_2, \ldots, \nu_n$ be interchangeable random variables that assume nonnegative integer values. Define $N_r = \nu_1 + \cdots + \nu_r$ for $r = 1, \ldots, n$. Denote by Δ_r the number of subscripts $i = 1, \ldots, r$ for which the inequality $N_i < i$ holds.

In [6] we proved that

$$P\{\Delta_n = n \mid N_n = k\} = 1 - \frac{k}{n} \tag{5}$$

for $k = 0, 1, \ldots, n$, provided that the left hand side is defined. Further we have

$$P\{\Delta_n = j \mid N_n = n-1\} = \frac{1}{n} \tag{6}$$

for $j = 1, 2, \ldots, n$, provided that the left hand side is defined. The latter follows immediately from Theorem 1 of [7]* or can be deduced from Theorem 3 of E. Sparre Andersen [2] or from Theorem 2.1 of F. Spitzer [5].

The following two theorems can easily be proved by using (5) and (6).

Theorem 3. *If $k < n-1$, then we have*

$$P\{\Delta_n = j \mid N_n = k\} = \begin{cases} 1 - \dfrac{k}{n} \quad \text{if} \quad j = n, \\ \displaystyle\sum_{i=n-j+1}^{k+1} \frac{(n-k-1)}{i(n-i)} P\{N_i = i-1 \mid N_n = k\} \\ \qquad \text{if} \quad j = n-k, \ldots, n-1, \\ 0 \quad \text{if} \quad j = 0, 1, \ldots, n-k-1. \end{cases} \tag{7}$$

Proof. Without loss of generality we may suppose that $N_n = k$ is fixed. If $\Delta_n = j < n-1$ and $N_n < n-1$, then there exists an r such that $N_r = r-1$. Denote by $r = i$ $(i = 1, \ldots, k+1)$ the greatest r with this property. Then $N_i = i-1$ and $N_r - N_i < r - i$ for $r = i+1, \ldots, n$. Thus

$$P\{\Delta_n = j\} = \sum_{i=1}^{k+1} P\{N_i = i-1\} P\{\Delta_i = i+j-n \mid N_i = i-1\} \times \\ \times P\{\Delta_n - \Delta_i = n-i \mid N_i = i-1\}. \tag{8}$$

By (6) $P\{\Delta_i = i+j-n \mid N_i = i-1\} = 1/i$ for $n-j < i \leqq n$, and 0 otherwise. By (5)

$$P\{\Delta_n - \Delta_i = n-i \mid N_i = i-1\} = P\{\Delta_n - \Delta_i = n-i \mid N_n - N_i = k-i+1\} \\ = (n-k-1)/(n-i) \quad \text{for} \quad i = 1, \ldots, k+1.$$

This proves (7) for $j = n-k, \ldots, n-1$. If $j = n$, then (7) reduces to (5). The case $j < n-k$ is obvious.

Theorem 4. *For a fixed $c \geqq 0$ denote by Δ_n^* the number of subscripts $r = 1, 2, \ldots, n$ for which $N_r < r + c$ holds. We have*

$$P\{\Delta_n^* = n \mid N_n = k\} = 1 - \sum_{i=1}^{k-c} \binom{n+c-k}{n-i} P\{N_i = i+c \mid N_n = k\} \tag{9}$$

for $c < k < n+c$. $P\{\Delta_n^* = n \mid N_n = k\} = 1$ if $k \leqq c$, and 0 if $k \geqq n+c$.

Proof. Let $N_n = k$ be fixed and $c < k < n+c$. Then

$$1 - P\{\Delta_n^* = n\} = \sum_{i=1}^{k-c} P\{N_i = i+c\} P\{\Delta_n - \Delta_i = n-i \mid N_i = i+c\}. \tag{10}$$

---

* I should like to note here that when I proved Theorem 1 of [7], it escaped my attention that it can be deduced from Theorem 2.1 of F. Spitzer [5].

The left hand side is the probability that at least one of the inequalities $N_r < r + c$, $r = 1, \ldots, n$, is violated. If $r = i$, $i = 1, \ldots, n$, is the greatest r for which $N_r < r + c$ is violated, then necessarily $N_i = i + c$ and $N_r - N_i < r - i$ for $r = i + 1, \ldots, n$. Thus we get the right hand side of (10). By (5)

$$\begin{aligned} P\{\Delta_n - \Delta_i = n - i \mid N_i = i + c\} &= P\{\Delta_n - \Delta_i = n - i \mid N_n - N_i = k - c - i\} \\ &= (n + c - k)/(n - i) \quad \text{if} \quad 0 \leqq i \leqq k - c, \end{aligned}$$

and 0 otherwise. This proves (9). The cases $k \leqq c$ and $k \geqq n + c$ are obvious. If $c = 0$, then (9) reduces to $\left(1 - \frac{k}{n}\right)$.

3. Proof of the ballot theorems

Define ν_r, $r = 1, \ldots, a + b$, as follows: $\nu_r = 0$ if the r-th vote is cast for A and $\nu_r = (\mu + 1)$ if the r-th vote is cast for B. Now $\nu_1, \nu_2, \ldots, \nu_{a+b}$ are interchangeable random variables that assume nonnegative integer values and

$$\nu_1 + \cdots + \nu_{a+b} = b(\mu + 1).$$

Since $\nu_1 + \cdots + \nu_r = (\mu + 1)\beta_r$ and $r = \alpha_r + \beta_r$, the inequality $\alpha_r > \mu\beta_r$ holds if and only if $\nu_1 + \cdots + \nu_r < r$, and $\alpha_r > \mu\beta_r - c$ holds if and only if $\nu_1 + \cdots + \nu_r < r + c$. Thus $P_j = P\{\Delta_n = j \mid N_n = k\}$ and $P_j^* = P\{\Delta_n^* = j \mid N_n = k\}$ where $n = a + b$, $k = (\mu + 1) b$ and obviously

$$P\{N_i = j\} = \frac{\binom{a}{i-s}\binom{b}{s}}{\binom{a+b}{i}} \tag{11}$$

if $j = s(\mu + 1)$, and 0 otherwise. Formulas (1), (2), (3), and (4) can be obtained from (5), (7), (6), and (9) respectively. If, in particular, $c = 0$, then

$$P_{a+b}^* = P_{a+b} = (a - \mu b)/(a + b),$$

and if $\mu = 1$ and $c > 0$, then

$$P_{a+b}^* = 1 - \frac{\binom{a+b}{a+c}}{\binom{a+b}{a}}$$

which can be proved directly.

Finally we note that K. L. CHUNG and W. FELLER [*4*] found the distribution of the number of subscripts for which either

$$\alpha_r > \beta_r \quad \text{or} \quad \alpha_r = \beta_r \quad \text{but} \quad \alpha_{r-1} > \beta_{r-1}, \quad r = 1, 2, \ldots, a + b.$$

4. Further generalizations

Theorems 3 and 4 can be further generalized for stochastic processes. Suppose that $\{\chi(t), 0 \leqq t < \infty\}$ is a separable stochastic process with nonnegative, stationary, independent increments, $\chi(t)$ is increasing only in jumps and $\chi(0) = 0$. Denote by $\varrho(t)$ the measure of the set $\{u : \chi(u) < u, 0 \leqq u \leqq t\}$.

If in (6), (7), and (9) we write

$$\nu_r = \left[\frac{n}{t}\left(\chi_a\left(\frac{rt}{n}\right) - \chi_a\left(\frac{rt - t}{n}\right)\right)\right], \quad r = 1, 2, \ldots, n,$$

where $\chi_a(t)$ is the total amount of jumps of magnitude $\geqq a > 0$ occurring in the interval $(0, t)$ in the process $\{\chi(t), 0 \leqq t < \infty\}$, and if we let $n \to \infty$ and $a \to 0$, then we get the following results: By (6)

$$P\{\varrho(t) \leqq x \mid \chi(t) = t\} = \frac{x}{t} \tag{12}$$

if $0 \leqq x \leqq t$. By (7)

$$P\{\varrho(t) \leqq x \mid \chi(t) = y\} = \iint\limits_{\substack{t \leqq u+v \\ u \leqq y,\, v \leqq x}} \frac{1}{u}\left(\frac{t-y}{t-u}\right) P\{u < \chi(u) \leqq u + du \mid \chi(t) = y\}\, dv \tag{13}$$

if $0 \leqq y \leqq t$ and $t - y \leqq x \leqq t$, and

$$P\{\varrho(t) = t \mid \chi(t) = y\} = 1 - \frac{y}{t} \tag{14}$$

if $0 \leqq y \leqq t$. By (9)

$$\begin{aligned} &P\{\chi(u) < u + x \quad \text{for} \quad 0 \leqq u \leqq t \mid \chi(t) = y\} \\ &\qquad = 1 - \int_0^{y-x} \left(\frac{t + x - y}{t - u}\right) P\{u + x < \chi(u) \leqq u + x + du \mid \chi(t) = y\} \end{aligned} \tag{15}$$

if $x < y < t + x$.

References

[1] Aeppli, A.: Zur Theorie verketteter Wahrscheinlichkeiten. Zürich: Thèse 1924.

[2] Andersen, E. Sparre: On the fluctuations of sums of random variables. Math. Scandinavica **1**, 263—285 (1953).

[3] Barbier, É.: Généralisation du problème résolu par M. J. Bertrand. C. R. Acad. Sci. (Paris) **105**, 407 (1887).

[4] Chung, K. L., and W. Feller: Fluctuations in coin tossing. Proc. nat. Acad. Sci. USA **35**, 605—608 (1949).

[5] Spitzer, F.: A combinatorial lemma and its application to probability theory. Trans. Amer. math. Soc. **82**, 323—339 (1956).

[6] Takács, L.: The probability law of the busy period for two types of queuing processes. Operations Res. **9**, 402—407 (1961).

[7] — Ballot problems. Z. Wahrscheinlichkeitstheorie **1**, 154—158 (1962).

Department of Mathematical Statistics
Columbia University, New York 27, N.Y.

(Received January 14, 1963)