J. Appl. Prob. **6**, 682–686 (1969)
Printed in Israel

THE TOTAL PROGENY IN A BRANCHING PROCESS AND A RELATED RANDOM WALK

MEYER DWASS, *Northwestern University*

1. Introduction

This paper is a continuation of [1]. The techniques of [1] are used to get specific information about the distribution of the total progeny in a branching process. This distribution is also related to one which arises in a random walk problem.

2. Total progeny in a branching process

Consider a branching process $Y_0, Y_1, Y_2, \cdots$ with regeneration generating function

$$P(t) = p_0 + p_1 t + \cdots, \qquad p_0 > 0.$$

That is,

$$E(t^{Y_{n+1}} \mid Y_n = k) = [P(t)]^k, \qquad k = 0, 1, 2, \cdots.$$

Let A be the event that the process eventually becomes extinct. It is well-known that

$$P(A) = \begin{cases} 1 & \text{if } P'(1) \leqq 1, \\ q & \text{if } P'(1) > 1, \end{cases}$$

where q is the unique root in $(0,1)$ of $P(t) = t$. By the *total progeny* we mean the random variable Z whose value at a sample point w is

$$Z(w) = \begin{cases} Y_0(w) + Y_1(w) + \cdots & \text{if } w \text{ is in } A, \\ \infty & \text{if } w \text{ is not in } A. \end{cases}$$

Theorem

(1) $$P(Z = n; A \mid Y_0 = k) \equiv P_k(Z = n; A) = (k/n) p_{n-k}^{(n)}, \quad n = k, k+1, \cdots, \quad k = 1, 2, \cdots,$$

where

$$[P(t)]^n = p_0^{(n)} + p_1^{(n)} t + p_2^{(n)} t^2 + \cdots.$$

Proof. Let $X_1, X_2, \cdots,$ be a sequence of independent and identically distributed, random variables with

Received 10 February 1969.

$$P(X_i = n) = p_n, \; n = 0,1,2,\cdots.$$

Let $X_1^{(s)}, X_2^{(s)}, \cdots$ also be a sequence of independent and identically distributed random variables with

$$P(X_i^{(s)} = n) = p_n s^n / P(s), \quad 0 \leqq s \leqq 1, \quad n = 0,1,2,\cdots.$$

Define

$$T_n = X_1 + \cdots + X_n, \; T_n^{(s)} = X_1^{(s)} + \cdots + X_n^{(s)}, \quad n = 1,2,\cdots.$$

First suppose that $E(X_i) \leqq 1$. Then

$$E(X_i^{(s)}) = sP'(s)/P(s) \leqq 1 \quad \text{for } s \text{ in } [0,1].$$

As pointed out in [1],

$$P(T_n^{(s)} < n, \; n = 1,2,\cdots) = 1 - sP'(s)/P(s).$$

Let k be a positive integer. Then by the strong law of large numbers and a "last entry" argument,

$$\begin{aligned} P(X_1^{(s)} = k) &= \sum_{n=k}^{\infty} P(X_1^{(s)} = k, \; T_n^{(s)} = n, \; T_{n+i}^{(s)} < n+i, \quad i = 1,2,\cdots) \\ &= P(X_1^{(s)} = k) \sum_{n=k}^{\infty} P(T_{n-1} = n-k) \frac{s^{n-k}}{[P(s)]^{n-1}} [1 - sP'(s)/P(s)]. \end{aligned}$$

(This uses the fact that $P(T_n^{(s)} = k) = P(T_n = k)s^k/[P(s)]^n$. We define $T_0 = 0$.) Let $h(u)$ be the inverse of $s/P(s) = u$, s in $[0,1]$. Then, as shown in [1],

$$[1 - sP'(s)/P(s)]^{-1} = uh'(u)/h(u).$$

It follows that

$$(2) \qquad [h(u)]^{k-2}h'(u) = \sum_{n=k}^{\infty} P(T_{n-1} = n-k)u^{n-2}, \qquad 0 \leqq u < 1.$$

Integrating (2), we obtain for $k \geqq 2$,

$$[h(u)]^{k-1} = \sum_{n=k}^{\infty} [(k-1)/(n-1)]P(T_{n-1} = n-k)u^{n-1},$$

which is equivalent to (1), as it is well-known that when $E(X_i) \leqq 1$,

$$h(u) = E(u^Z), \qquad 0 \leqq u \leqq 1.$$

(See [3] and [2], page 298.) This completes the proof when $E(X_i) \leqq 1$, since $[h(u)]^k$ is the generating function of Z under the condition that $Y_0 = k$.

We now consider the case when $E(X_i) > 1$. An elementary computation shows that

$$(3) \qquad P_k(Z = n; A) = \sum_{m=1}^{\infty} \sum_{i_1 + \cdots + i_m = n} p_{i_1}^{(k)} p_{i_2}^{(i_1)} \cdots p_0^{(i_m)} .$$

We now consider the branching process which evolves according to the regenerating distribution whose generating function is

$$P(qt)/P(q) = P(qt)/q .$$

Let $Z^{(q)}$ be the total progeny for this process, which we shall refer to as the "q-process." Since

$$\frac{d}{dt}\left(\frac{P(qt)}{q}\right)\bigg|_{t=1} = P'(q) < 1 ,$$

it follows that the q-process becomes extinct with probability 1. Evaluating (3) for the q-process shows that

$$\begin{aligned} P_k(Z^{(q)} = n; A) &= \sum_{m=1}^{\infty} \sum_{i_1 + \cdots + i_m = n} p_{i_1}^{(k)} p_{i_2}^{i_1} \cdots p_0^{(i_m)} \frac{q^{i_1 + \cdots + i_m}}{[P(q)]^{k + i_1 + \cdots + i_m}} \\ &= P_k(Z = n; A)/q^k . \end{aligned}$$

By the first part of the proof,

$$\begin{aligned} P_k(Z^{(q)} = n) &= (k/n)P(T_n^{(q)} = n-k) = (k/n)P(T_n = n-k)q^{n-k}/[P(q)]^n \\ &= (k/n)P(T_n = n-k)/q^k. \end{aligned}$$

Hence,

$$P_k(Z = n; A)/q^k = P_k(Z^{(q)} = n; A) = P_k(Z^{(q)} = n) = (k/n)P(T_n = n-k)/q^k,$$

which completes the proof.

Remark. By integrating (2) for $k = 1$, one determines that

$$h(u) = u \exp\left\{ \sum_{n=1}^{\infty} \frac{P(T_n = n)}{n} (u^n - 1) \right\}.$$

(See [1] for the details.)

3. A related problem in random walk

The random variables $X_1 - 1, X_2 - 1, \cdots$, assume the values $-1, 0, 1, \cdots$. The sums

$$T_n - n = (X_1 - 1) + \cdots + (X_n - 1)$$

execute a "left-continuous" random walk since the only movements in the negative direction are single unit steps. Suppose that the walk starts at height 0. (That is $T_0 = 0$.) Let B be the event that the walk ever reaches -1 and let W be the number of steps required for this to happen. That is,

$$B' = (T_1 \geqq 1,\ T_2 \geqq 2, \cdots),$$

and

$$(W = n) \cap B = (T_1 \geqq 1, \cdots, T_{n-1} \geqq n-1,\ T_n = n-1),\ n = 1,2,\cdots.$$

We will show that

$$P(W = n;\, B) = P_1(Z = n;\, A), \quad n = 1,2,\cdots, \tag{4}$$

where Z and A are as in Section 2.

Proof of (4). First suppose that $E(X_i) \leqq 1$. Then $P(B) = 1$. For s in $[0,1]$, let $X_i^{(s)}, T_n^{(s)}$ be as defined in Section 2. Using a "first entry" argument we have the following.

$$\begin{aligned} 1 &= \sum_{n=1}^{\infty} P(T_1^{(s)} \geqq 1, \cdots, T_{n-1}^{(s)} \geqq n-1,\ T_n^{(s)} = n-1) \\ &= \sum_{n=1}^{\infty} P(T_1 \geqq 1, \cdots, T_{n-1} \geqq n-1, T_n = n-1)s^{n-1}/[P(s)]^n \\ &= \sum_{n=1}^{\infty} P(W = n)s^{n-1}/[P(s)]^n. \end{aligned}$$

Hence, using the same function $h(u)$ as in Section 2,

$$\sum_{n=1}^{\infty} P(W = n)u^n = E(u^W) = h(u).$$

This completes the proof when $E(X_i) \leqq 1$.

For $E(X_i) > 1$, let q be as in Section 2. Then

$$\begin{aligned} &P(T_1 \geqq 1, \cdots, T_{n-1} \geqq n-1,\ T_n = n-1) \\ &\qquad = P(T_1^{(q)} \geqq 1, \cdots, T_{n-1}^{(q)} \geqq n-1,\ T_n^{(q)} = n-1)q \\ &\qquad = P_1(Z^{(q)} = n)q = P_1(Z = n). \end{aligned}$$

This completes the proof.

Corollary. If $E(X_i) > 1$, then

$$P(T_1 \geqq 1,\ T_2 \geqq 1, \cdots) = q.$$

Remarks. By identifying the results of Sections 2 and 3 we have the following relationship.

$$P(T_m > m-k,\ m = 1, \cdots, n-1 \mid T_n = n-k) = k/n. \tag{5}$$

Replacing $X_1, \cdots, X_n$ by their reversal, $X_n, \cdots, X_1$, (5) is equivalent to

$$P(T_1 < 1, \cdots, T_n < n \mid T_n = n-k) = k/n. \tag{6}$$

The relation (6) is a kind of "ballot theorem" which is known and not hard to prove directly. This relation can be used as a basis for an alternate derivation of the results of this paper.

References

[1] Dwass, M. (1968) A theorem about infinitely divisible distributions. *Z. Wahrscheinlichkeitsth.* **9,** 287–289.

[2] Feller, W. (1968) *An Introduction to Probability Theory and Its Applications.* Vol. 1, 3rd edition. Wiley, New York.

[3] Good, I. J. (1949) The number of individuals in a cascade process. *Proc. Camb. Phil. Soc.* **45,** 360–363.