THE MOMENT GENERATING FUNCTION OF THE GEOMETRIC DISTRIBUTION OF ORDER k

Michael J. J. Barry and Anthony J. Lo Bello
Department of Mathematics, Allegheny College, Meadville, PA 16335
(Submitted August 1991)

Let X be the random variable denoting the number of trials until the occurrence of the k^{th} consecutive success; the trials are independent with constant success probability p $(0 < p < 1)$. The probability density function f of X has been determined by Philippou and Muwafi [3]. (See also Philippou, Georghiou, and Philippou [4].) In this note we show that the moment generating function of X exists, and we determine a formula for it by means of the following recurrence. For two other recursive formulas of f, see [1] and [2].

Proposition: The probability density f of X satisfies the following relations. (Here, $q = 1 - p$.)

(a) $f(k) = p^k$.
(b) $f(n) = qp^k$ if $k+1 \le n \le 2k-1$.
(c) $f(n) = qf(n-1) + qpf(n-2) + qp^2 f(n-3) + \cdots + qp^{k-1} f(n-k)$ if $n \ge 2k$.

Hence, for $n \ge 2k$, the terms $f(n)$ satisfy a linear recursive relation of order k whose auxiliary equation is $x^k - qx^{k-1} - qpx^{k-2} - \cdots - qp^{k-1} = 0$.

Proof: Clearly the formula holds for $n = k$. Suppose now that $k+1 \le n \le 2k-1$. The first run of k consecutive successes ends on the n^{th} trial. These k successes are preceded by a failure, which in turn is preceded by any sequence of $n-k-1$ outcomes. Thus, $f(n) = qp^k$. Now let $n \ge 2k$, and consider a sequence of n Bernoulli trials where the first run of k consecutive successes ends on the n^{th} trial. The first failure must occur on or before the k^{th} trial and may occur on any of the first k trials. For $1 \le j \le k$, let E_j be the event that the first run of k consecutive successes occurs on the n^{th} trial and that the first failure occurs on the j^{th} trial. Clearly $f(n)$ equals the sum of the probabilities of the E_j. We claim that the probability of E_j is $qp^{j-1} f(n-j)$. Points in E_j consist of $j-1$ successes, followed by a failure, followed by any sequence of $n-j$ outcomes consistent with the first run of k consecutive successes ending on the n^{th} trial, and so by independence the probability of E_j is as claimed.

Having established these properties for f, we proceed to our main result.

Theorem: The moment generating function $M(t)$ of X exists on some open interval containing 0 and is given by

$$M(t) = \frac{p^k e^{kt}}{1 - qe^t - qpe^{2t} - \cdots - qp^{k-1} e^{kt}} = \frac{p^k e^{kt}(1 - pe^t)}{1 - e^t + (1-p)p^k e^{(k+1)t}}.$$

The proof of the theorem will be given after establishing the following lemma.

Lemma: The roots of the auxiliary equation are distinct and have absolute value less than 1.

Proof: We have seen that, for $n \ge 2k$, the terms $f(n)$ satisfy a linear recursive relation of order k whose auxiliary equation is

$$x^k - qx^{k-1} - qpx^{k-2} - \cdots - qp^{k-1} = 0.$$

We now investigate this equation. Let $e(x)$ denote the polynomial

$$x^k - qx^{k-1} - qpk^{k-2} - \cdots - qp^{k-1} \text{ in } x,$$

and let $f(x) = (x-p)e(x) = x^{k+1} - x^k + qp^k$. Now

$$f'(x) = (k+1)x^k - kx^{k-1} = (k+1)x^{k-1}\left(x - \frac{k}{k+1}\right)$$

has roots 0 and $\frac{k}{k+1}$. Since 0 is not a root of f, f has a repeated root if and only if $\frac{k}{k+1}$ is a root of f. But

$$f\left(\frac{k}{k+1}\right) = \left(\frac{k}{k+1}\right)^{k+1} - \left(\frac{k}{k+1}\right)^k + qp^k = \frac{-1}{k+1}\left(\frac{k}{k+1}\right)^k + qp^k.$$

Since

$$(1-x)x^k - \frac{1}{k+1}\left(\frac{k}{k+1}\right)^k \le 0$$

on [0, 1] with equality if and only if $x = \frac{k}{k+1}$, we see that $\frac{k}{k+1}$ is a root of f if and only if $p = \frac{k}{k+1}$. Thus, f has a repeated root (of order 2) if and only if $p = \frac{k}{k+1}$. Hence, the roots of e are distinct.

We turn now to the absolute values of the roots of $e(x)$. We will show that if z is a (complex) number with $|z| \ge 1$, then z is not a root of the equation $e(x) = 0$.

$$\begin{aligned} |z^k - qz^{k-1} - qpz^{k-2} - \cdots - qp^{k-1}| &\ge |z|^k - q|z|^{k-1} - qp|z|^{k-2} - \cdots - qp^{k-1} \\ &\ge |z|^k - q|z|^k - qp|z|^k - \cdots - qp^{k-1}|z|^k \\ &\ge |z|^k - q|z|^k \frac{1-p^k}{1-p} \\ &= |z|^k - |z|^k(1-p^k) = p^k|z|^k > 0. \end{aligned}$$

Thus, all roots of the equation $e(x) = 0$ have absolute value less than 1.

Proof of the Theorem: Let $z_1, z_2, \ldots, z_k$ be the distinct roots of the auxiliary equation; then, from the theory of difference equations, we know that there exist (complex) constants $c_1, c_2, \ldots, c_k$ such that

$$f(n) = c_1 z_1^n + c_2 z_2^n + \cdots + c_k z_k^n \text{ if } n \ge k.$$

Now the series $\sum_{n=k}^{\infty} c_i z_i^n e^{nt} = c_i \sum_{n=k}^{\infty} (z_i e^t)^n$ converges to $\frac{c_i (z_i e^t)^k}{1 - z_i e^t}$ if $|z_i e^t| < 1$, that is, if $t < -\ln|z_i|$. Let $m = \min\{-\ln|z_1|, -\ln|z_2|, \ldots, -\ln|z_k|\}$ Then the moment generating function

$$M(t) = \sum_{n=k}^{\infty} e^{nt} f(n)$$

exists on the interval $(-\infty, m)$. The proof of the theorem now follows by substituting e^t for s in the formula of the probability generating function $\gamma_k(s)$ of [4, Lemma 2.3]. Alternatively, recasting the proposition above, we have

$$(*) \qquad f(n+k) = qf(n+k-1) + qp(n+k-2) + \cdots + qp^{k-1}f(n), \; n \ge 1,$$

with $f(1)=f(2)=\cdots=f(k-1)=0$ and $f(k)=p^k$. Therefore,

$$\begin{aligned}
M(t) &= \sum_{n=k}^{\infty} e^{nt} f(n) = e^{kt} f(k) + \sum_{n=1}^{\infty} e^{(n+k)t} f(n+k) \\
&= e^{kt} p^k + q\sum_{n=1}^{\infty} e^{(n+k)t} f(n+k-1) + qp\sum_{n=1}^{\infty} e^{(n+k)t} f(n+k-2) + \cdots + qp^{k-1}\sum_{n=1}^{\infty} e^{(n+k)t} f(n), \text{ by } (*), \\
&= e^{kt} p^k + qe^t\sum_{n=1}^{\infty} e^{(n+k-1)t} f(n+k-1) + qpe^{2t}\sum_{n=1}^{\infty} e^{(n+k-2)t} f(n+k-2) + \cdots + qp^{k-1}e^{kt}\sum_{n=1}^{\infty} e^{nt} f(n) \\
&= e^{kt} p^k + qe^t M(t) + qpe^{2t} M(t) + \cdots + qp^{k-1} e^{kt} M(t),
\end{aligned}$$

from which the proof follows.

Final Comment: From the moment generating function, one can calculate all the moments that are of interest. For example, when $p = \frac{1}{2}$, the mean of X is given by $\mu = 2(2^k - 1)$, and the variance of X by $\sigma^2 = 4(2^k-1)^2 - (4k-6)(2^k-1) - 4k$; the following table displays the skewness factor α_3 and the kurtosis factor α_4 for k = 1, ..., 10. Note that as k increases, α_3 and α_4 approach the skewness factor 2 and the kurtosis factor 9, respectively, of the Exponential Distribution.

k	α_3	α_4
1	2.211320344	9.5
2	2.035097747	9.144628099
3	2.010489423	9.042749454
4	2.003133201	9.012677353
5	2.000918388	9.003699063
6	2.000262261	9.00105334
7	2.000072886	9.00029223
8	2.000019756	9.00007913
9	2.000005243	9.000020986
10	2.000001368	9.000005473

ACKNOWLEDGMENT

The authors with to thank the referee for drawing their attention to two references, and for shortening their computation of the moment generating function.

REFERENCES

1. S. Aki, H. Kuboki, & K. Hirano. "On Discrete Distributions of Order k." *Annals of the Institute of Statistical Mathematics* **36A.3** (1984):431-40.
2. A. N. Philippou & F. S. Makri. "Longest Success Runs and Fibonacci-Type Polynomials." *Fibonacci Quarterly* **23.4** (1985):338-46.
3. A. N. Philippou & A. A. Muwafi. "Waiting for the k^{th} Consecutive Success and the Fibonacci Sequence of Order k." *Fibonacci Quarterly* **20.1** (1982):100-05.
4. A. N. Philippou, C. Georghiou, & G. N. Philippou. "A Generalized Geometric Distribution and Some of Its Properties." *Statistics and Probability Letters* **1.4** (1983):171-75.

AMS Classification number: 60E99

❖❖❖