Journal of the Math Citadel
Original Research Article

Generalizing the Negative Binomial Distribution via First-Kind Dependence

Rachel Traylor*, Ph.D.

Abstract

This paper generalizes the negative binomial random variable by generating it from a sequence of first-kind dependent Bernoulli trials under the identity permutation. The PMF, MGF, and various moments are provided, and it is proven that the distribution is indeed an extension of the standard negative binomial random variable. We examine the effect of complete dependence of the Bernoulli trials on the generalized negative binomial random variable. We also show that the generalized geometric random variable is a special case of the generalized negative binomial random variable, but the generalized negative binomial random variable cannot be generated from a sum of i.i.d. generalized geometric random variables.

Keywords

dependency — probability theory

Contents

Introduction

A binomial random variable Z_n is constructed from a sequence of n Bernoulli random variables $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, and counts the number of 1s, or ”successes” in the sequence. Mathematically, a binomial random variable is given by $Z_n = \sum_{i=1}^{n} \epsilon_n$. A traditional binomial random variable requires an i.i.d sequence of Bernoulli random variables. Korzeniowski [1] developed a generalized binomial distribution under the condition that the Bernoulli sequence is *first-kind dependent*.

A ”different perspective”, as it were, to the binomial random variable is the negative binomial random variable. With a binomial random variable, we fix the number of trials and count the number of

[*] The Math Citadel

”successes”. Suppose now we fix the number of successes as k and continue to run Bernoulli trials until the kth success. The random number of trials necessary to get k successes is a *negative binomial random variable*, and may be formulated mathematically as $V_k = min_{n\geq k}(n : Z_n = k)$. The sequence is halted when the kth success appears, which will always be on the last trial. Thus, the event $(V_k = n)$ is equivalent to $(Z_{n-1} = k-1 \wedge \epsilon_n = 1)$. A standard negative binomial distribution is constructed from an i.i.d. sequence of Bernoulli random variables, just like the binomial random variable. The PMF is given by

$$P(V_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \quad n \geq k$$

We may also characterize the negative binomial distribution in a different way. Let Y denote the number of additional trials beyond the minimum possible k required for k successes. Since the trials are Bernoulli trials, Y denotes the random number of failures that will occur before the kth success is observed. Thus, if we denote y as the number of failures, $n = k + y$, where k is fixed, and thus the random variable Y with support $\{0,1,2,\ldots\}$ is equivalent to the previous characterization of V_k. The PMF of Y is easily derived and given by

$$P(Y = y) = \binom{k+y-1}{y} p^k (1-p)^y$$

A first-kind dependent sequence of Bernoulli trials is identically distributed but dependent [1, 3, 4], and thus can generate generalized versions of random variables that are functions of sequences of identically distributed Bernoulli or categorical random variables [3, 2]. This paper generalizes the negative binomial distribution given above by allowing the Bernoulli sequence to be first-kind dependent.

1. Derivation of the PMF

Theorem 1. *Let $k \in \mathbb{N}$ be fixed, and $\epsilon = \{\epsilon_1, \epsilon_2, \ldots\}$ be a sequence of first-kind dependent Bernoulli trials under the identity permutation with $P(\epsilon_k = 1) = p$ and dependency coefficient δ. Define $q = 1-p$,$p^+ = p + \delta q$, [$p^- = p - \delta p$, $q^+ = q + \delta p$, and $q^- = q - \delta q$. Let Z_n denote a generalized binomial random variable of length n. Let V_k denote the random variable that counts the number of first-kind dependent Bernoulli trials until the kth success. That is,$V_k = \min_{n\geq k}(n : Z_n = k)$.Then the PMF of V_k is given by*

$$P(V_k = n) = p\binom{n-2}{k-2}(p^+)^{k-1}(q^-)^{n-k} + q\binom{n-2}{k-1}(p^-)^k(q^+)^{n-k-1} \tag{1}$$

Proof. The negative binomial random variable V_k is equivalent to ”stitching” a binomial random variable Z_n in $n-1$ Bernoulli trials together with a Bernoulli random variable whose outcome is 1. Thus,

$$P(V_n = k) + P(Z_{n-1} = k-1 \wedge \epsilon_n = 1)$$

Under first-kind dependence, $P(\epsilon_n = 1|\epsilon_1 = 1) = p^+$, and $P(\epsilon_n = 1|\epsilon_1 = 0) = p^-$. So we have two possibilities: either the kth success occurs after $\epsilon_1 = 1$, or the kth success occurs after $\epsilon_1 = 0$. Thus

$$\begin{aligned}P(V_n = k) &= P(\epsilon_n = 1 \wedge Z_{n-1} = k-1 \wedge \epsilon_1 = 1) + P(\epsilon_n = 1 \wedge Z_{n-1} = k-1 \wedge \epsilon_1 = 0)\\ &= P(\epsilon_n = 1|Z_{n-1} = k-1 \wedge \epsilon_1 = 1)P(Z_{n-1} = k-1 \wedge \epsilon_1 = 1)\\ &\quad + P(\epsilon_n = 1|Z_{n-1} = k-1 \wedge \epsilon_1 = 0)P(Z_{n-1} = k-1 \wedge \epsilon_1 = 0)\end{aligned}$$

Now, from the generalized binomial distribution in [1], $P(Z_{n-1} = k-1 \wedge \varepsilon_1 = 1) = p\binom{n-2}{k-2}(p^+)^{k-1}(q^-)^{n-k}$ and $P(Z_{n-1} = k-1 \wedge \varepsilon_1 = 0) = q\binom{n-2}{k-1}(p^-)^{k-1}(q^+)^{n-k-1}$. Then

$$P(V_n = k)p\binom{n-2}{k-2}(p^+)^{k-1}(q^-)^{n-k} + q\binom{n-2}{k-1}(p^-)^{k}(q^+)^{n-k-1}, \qquad n = k, k+1, k+2, \ldots$$

□

It will be more helpful to characterize the generalized negative binomial random variable our alternative way, by letting $V_k = Y + k$, where Y is the random variable that counts the number of additional trials beyond the minimum possible k necessary to achieve the kth success or, equivalently, the number of failures in a sequence of FK-dependent Bernoulli variables with k successes. The PMF of Y is given in the following corollary

Corollary 1. *Let y be the random variable described here, with support $\{0, 1, 2, \ldots\}$. Then Y is equivalent to V_k, and the PMF of Y is given by*

$$P(Y = y) = p\binom{y+k-2}{y}(p^+)^{k-1}(q^-)^{y} + q\binom{y+k-2}{y-1}(p^-)^{k}(q^+)^{y-1} \tag{2}$$

When $\delta = 0$, a FK-dependent Bernoulli sequence becomes a standard i.i.d. Bernoulli sequence. Thus, when $\delta = 0, Y$ reverts to a standard negative binomial distribution.

Corollary 2. *Let Y be a generalized negative binomial distribution constructed via FK-dependency under the identity permutation with dependency coefficient δ. When $\delta = 0$, Y is a standard negative binomial random variable.*

Proof. When $\delta = 0$, $p = p^+ = p^-$, and $q = q^+ = q^-$. Thus,

$$\begin{aligned} P(Y = y) &= p\binom{y+k-2}{y}(p^+)^{k-1}(q^-)^{y} + q\binom{y+k-2}{y-1}(p^-)^{k}(q^+)^{y-1} \\ &= \binom{y+k-2}{y}p^k q^y + \binom{y+k-2}{y-1}p^k q^y \\ &= p^k q^y\left(\binom{y+k-2}{y} + \binom{y+k-2}{y-1}\right) \\ &= p^k q^y\left(\frac{(y+k-2)!}{y!(k-2)!}\frac{(y+k-2)!}{(y-1)!(k-1)!}\right) \\ &= p^k q^y\left(\frac{(k-1)(y+k-2)!}{y!(k-1)!} + \frac{y(y+k-2)!}{y!(k-1)!}\right) \\ &= p^k q^u\binom{y+k-1}{y} \end{aligned}$$

which is indeed the PMF of a standard negative binomial random variable. □

2. The Moment Generating Function and various moments

Theorem 2. *The moment generating function of the generalized negative binomial distribution is given by*

$$M_Y(t) = \frac{p(p^+)^{k-1}}{(1-e^t q^-)^{k-1}} + \frac{q(p^-)^k e^t}{(1-e^t q^+)^k} \tag{3}$$

Remark: his is indeed a generalization of the standard negative binomial distribution, as it is now a special case of the generalized negative binomial distribution when $\delta = 0$. To see this, we will show that the MGF of the generalized negative binomial distribution becomes the MGF for the standard negative binomial distribution.

When $\delta = 0$, a FK-dependent sequence reverts back to a standard i.i.d. sequence. That is, $p^+ = p^- = p$, and $q^+ = q^- = q$. So in this case,

$$\begin{aligned} M_Y(t) &= \frac{p^k}{(1-e^tq)^{k-1}} + \frac{qp^ke^t}{(1-e^tq)^k} \\ &= \frac{(1-e^tq)p^k + qp^ke^t}{(1-e^tq)^k} \\ &= \frac{p^k}{(1-e^tq)^k} \end{aligned}$$

The proof of Theorem 2 is straightforward from the definition of a moment generating function.

Proof.

$$\begin{aligned} M_Y(t) &:= \mathrm{E}\left[e^{tY}\right] \\ &= \sum_{y=0}^{\infty} e^{ty} p \binom{y+k-2}{y} (p^+)^{k-1}(q^-)^y + \sum_{y=0}^{\infty} e^{ty} q \binom{y+k-2}{y-1} (p^-)^k (q^+)^{y-1} \end{aligned}$$

Now, $\binom{y+k-2}{y} = \binom{y+(k-1)-1}{y} = (-1)^y\binom{-(k-1)}{y}$, so the first sum becomes

$$\begin{aligned} \sum_{y=0}^{\infty} e^{ty} p \binom{y+k-2}{y} (p^+)^{k-1}(q^-)^y &= p(p^+)^{k-1} \sum_{y0}^{\infty} \binom{-(k-1)}{y} (-e^tq^-)^y \\ &= \frac{p(p^+)^{k-1}}{(1-e^tq^-)^{k-1}} \end{aligned} \tag{4}$$

For the second sum, note that $\binom{k-2}{-1} = 0$. Now, let $z = y-1$. Then the second sum is

$$\begin{aligned} \sum_{y=0}^{\infty} e^{ty} q \binom{y+k-2}{y-1} (p^-)^k (q^+)^{y-1} &= q(p^-)^k e^t \sum_{z=0}^{\infty} \binom{z+k-1}{z} (e^tq^+)^z \\ &= q(p^-)^k e^t \sum_{z=0}^{\infty} \binom{-k}{z} (-e^tq^+)^z \\ &= \frac{q(p^-)^k e^t}{(1-e^tq^+)^k} \end{aligned} \tag{5}$$

Combining (4) and (5) yields the result. □

We may now derive the various moments of the generalized negative binomial distribution using the moment generating function.

2.1 Mean of generalized negative binomial distribution

The mean of the generalized negative binomial distribution is given by

$$\mu_Y = \mathrm{E}[Y] = \frac{kpq + k\delta q^2 - (k-1)\delta pq(1-\delta)}{p^2(1-\delta) + \delta pq(1-\delta)} \tag{6}$$

Remark:Note the reduction of the GNB mean to that of the standard negative binomial distribution when the sequence is independent. For $\delta = 0$, $\mathrm{E}[Y] = \frac{kq}{p}$.

2.2 Variance of the generalized negative binomial distribution

After many attempts to distill the formula to a palatable expression, the variance of the generalized negative binomial distribution is given by

$$\mathrm{Var}[Y] = \frac{kq}{(p+q\delta)^2(1-\delta)^2} + \frac{\delta(p^3q + kpq^2 - kq^3) + \delta^2(k^2pq + kq - 3kpq^2 - 2p^3q) + \delta^3 pq(p^2 - k - 2kq)}{p^2(1-\delta^2)(p+q\delta)^2} \tag{7}$$

Remark: Once again, under independence, $\mathrm{Var}[Y] = \frac{kq}{p^2}$, the variance of the standard negative binomial. Other higher order moments can also be obtained from the moment generating function and copious amounts of tedious arithmetic.

3. Other Considerations

3.1 The effect of complete dependence

It's also worth exploring the other extremes, like complete dependence. As illustrated in [3], complete dependence under FK-dependence implies that every Bernoulli trial will be identical to the outcome of the first trial. Thus, if $\varepsilon_1 = 0$, the entire sequence will be all 0s, and vice versa if $\varepsilon_1 = 1$. What does that mean for the generalized negative binomial distribution? If $\varepsilon_1 = 0$, the sequence will never end; k successes will never happen. On the other hand, if $\varepsilon_1 = 1$, then you are guaranteed to reach k successes in k trials. This results in both an infinite mean and variance, seen by plugging in $\delta = 1$.

Exploring the PMF under $\delta = 1$, $P(Y = 0) = p$, because $P(\varepsilon_1 = 1) = p$. If $\varepsilon_1 = 1$, and $\delta = 1$, then there will be only 1s in the sequence, and no 0s, and thus $P(Y = 0) = P(\varepsilon_1 = 1)$. If $\varepsilon_1 = 0$, then there are only 0s in the FK-dependent Bernoulli sequence, and no 1s. Thus, Y can only be ∞, and $P(Y = \infty) = P(\varepsilon_1 = 0) = q$.

Remark: When we say $Y = \infty$, we mean that the sequence of trials has no halting point. That is, the counting process never ends.

Thus, under complete dependence of the first kind, the support of Y has two points $\{0, \infty\}$, with probabilities p and q respectively. This is another way to confirm that Y will have infinite mean and variance in this case.

3.2 The Negative Binomial Random Variable as a Sum of Geometric Random Variables

The standard negative binomial distribution with k fixed successes can be derived as a sum of independent standard geometric random variables. One shows this by showing the moment generating function of the standard negative binomial distribution is equal to the product of k i.i.d. standard geometric random variables. Moreover, it can also be shown that the standard geometric random variable is a special case of the standard negative binomial distribution when $k = 1$.

How much of this carries over to the generalized versions of both distributions? The generalized geometric distribution was introduced and detailed by Traylor in [2]. Here, we are concerned with the PMF of "Version 2" of the generalized geometric distribution, as the "shifted" generalized geometric distribution counts the number of failures prior to the first success, and is analogous to counting the number of failures in a sequence of trials before the kth success. We reproduce Proposition 2 from [2]

Proposition 2, [2]: *Suppose $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots)$ is a FK-dependent sequence of Bernoulli random variables. Let $Z = X - 1$ be the count of failures prior to the first success. Then Z has a shifted generalized geometric distribution with PMF*

$$f_Z(z) = \begin{cases} p, & z = 0 \\ q(q^+)^{z-1}p^-, & z \geq 1 \end{cases}$$

We can quickly derive its MGF:

Proposition 1. *The moment generating function of the shifted generalized geometric distribution is given by*

$$M_Z(t) = p + \frac{e^t q p^-}{1 - e^t q^+}$$

Proof.

$$\begin{aligned} M_Z(t) &= \mathrm{E}\left[e^{tZ}\right] \\ &= p + \sum_{z=1}^{\infty} e^{tz} q p^- (q^+)^{z-1} \\ &= p + e^t q p^- \sum_{\lambda=0}^{\infty} \left(e^{t\lambda} q^+\right)^{\lambda} \\ &= p + \frac{e^t q p^-}{1 - e^t q^+} \end{aligned}$$

□

3.2.1 Generalized geometric RV as a special case of generalized negative binomial RV

It is true that, for $k = 1$, the generalized negative binomial distribution under FK-dependence reduces to the generalized geometric distribution. This is given in the following theorem.

Theorem 3. *When $k = 1$, the generalized negative binomial random variable reduces to a generalized geometric random variable.*

Proof. Simply plug in $k = 1$ to the PMF of the generalized negative binomial distribution:

$$\begin{aligned} P(Y = y) &= p\binom{y-1}{y}(q^-)^y + q\binom{y-1}{y-1}p^-(q^+)^{y-1} \\ &= \begin{cases} p, & y = 0 \\ q(q^+)^{y-1}p^-, & y \geq 1 \end{cases} \end{aligned}$$

because $\binom{b}{0} = 1$ for any b, and we take $\binom{b-1}{b} = 0$ □

3.2.2 Sum of independent generalized geometric random variables does not yield a generalized negative binomial random variable

Unlike the standard case, a sum of i.i.d. generalized geometric random variables does not yield a generalized negative binomial random variable. First, we note what we mean by a set of i.i.d. generalized geometric random variables. Suppose we have a set of generalized geometric random variables $\{X_1, X_2, \ldots, X_k\}$, each with the same p, q, and δ, and all first-kind dependent. Thus, to say that each of these geometric random variables is mutually independent of the others is to say that nothing about the other geometric random variables has any probabilistic bearing on the variable in question. That is, the dependency structure is not changed or altered, and $P(X_i|X_j) = P(X_i)$. $i \neq j$, $i = 1, 2, \ldots k$. The Bernoulli random variables that make up each geometric random variable still remain FK-dependent among themselves. That is, if $\varepsilon_i = (\varepsilon_1^{(i)}, \varepsilon_2^{(i)}, \ldots, \varepsilon_{n_i}^{(i)})$ is the sequence of Bernoulli trials that comprises X_i, each ε_i is FK-dependent among its elements, but independent of the other sequences ε_j.

One can easily see that, if the generalized negative binomial distribution were able to be generated by the sum of i.i.d. generalized geometric random variables, then $M_Y(t) = \prod_{i=1}^{k} M_{X_i}(t)$. But

$$\begin{aligned}\prod_{i=1}^{k} M_{X_i}(t) &= \left(p + \frac{e^t q p^-}{1 - e^t q^+}\right) \\ &\neq \frac{p(p^+)^{k-1}}{(1 - e^t q^-)^{k-1}} + \frac{q(p^-)^k e^t}{(1 - e^t q^+)^k} \\ &= M_Y(t)\end{aligned}$$

Why is this? The answer is quite intuitive. The generalized negative binomial distribution under FK-dependence is one sequence under a FK dependency structure. That is, all Bernoulli trials after the first depend directly on the first. Summing generalized geometric random variables under FK dependence is equivalent to constructing a sequence of generalized geometric random variables, one after the other. Since these are themselves comprised of FK-dependent Bernoulli trials, each time a success is observed, the dependency structure ″starts over″ with the next geometric random variable.

For example, suppose the first geometric random variable has a success on the third trial. Then the fourth Bernoulli trial is starting an entirely new sequence of FK-dependent Bernoulli variables, and does not depend on the outcome of the first. This is not equivalent to the definition of a generalized negative binomial distribution.

This property held for the standard versions of the geometric and negative binomial random variables because every Bernoulli trial in each geometric sequence is i.i.d. Thus, there is no ″transition″ from one geometric random variable to another; it's as if it was all one big sequence of Bernoulli trials to begin with. We lose that when we introduce dependency structures.

4. Conclusion

This paper introduced the generalized negative binomial distribution built from a sequence of FK- dependent random variables. The PMF, MGF, and various moments were derived. It was also noted that the generalized geometric distribution is a special case of the generalized negative binomial distribution, but one cannot construct a generalized negative binomial random variable from the sum of i.i.d. generalized geometric random variables.

References

[1] Andrzej Korzeniowski. On correlated random graphs. *Journal of Probability and Statistical Science*, pages 43–58, 2013.

[2] Rachel Traylor. A generalized geometric distribution from vertically dependent categorical random variables. *Academic Advances of the CTO*, 2017.

[3] Rachel Traylor. A generalized multinomial distribution from dependent categorical random variables. *Academic Advances of the CTO*, 2017.

[4] Rachel Traylor and Jason Hathcock. Vertical dependency in sequences of categorical random variables. *Academic Advances of the CTO*, 2017.