A Stopped Negative Binomial Distribution

Michelle DeVeaux[a], Michael J. Kane[a,*], Daniel Zelterman[a]

[a] *Department of Biostatistics*
School of Epidemiology and Public Health
Yale University, New Haven, CT

Abstract

We introduce a discrete distribution suggested by curtailed sampling rules common in early-stage clinical trials. We derive the distribution of the smallest number of independent and identically distributed Bernoulli trials needed to observe either s successes or t failures. The closed-form expression for the distribution as well as its characteristics are derived and properties of the distribution are explored.

Keywords: discrete distribution, curtailed sampling

1. Introdution and Motivation

Consider a prototypical Phase II single-arm clinical trial in which 12 patients are enrolled and treated. If two or more respond then we reject the null hypothesis that the treatment is not effective and the trial proceeds to the next stage. If fewer that two respond then the null is not rejected and the trial is terminated.

If all 12 patients are enrolled at once, as in the classic design, then the sample size is 12. However, as in most clinical trials, the patients are enrolled sequentially, often with one patient's outcome realized before the next one enters

☆This research was supported by grants R01CA131301, R01CA157749, R01CA148996, R01CA168733, and PC50CA196530 awarded by the National Cancer Institute along with support from the Yale Comprehensive Cancer Center and the Yale Center for Outcomes Research. We would also like to thank Rick Landin at Ignyta Inc. for his suggestions.

*Corresponding author

Email addresses: `michelle.deveaux@yale.edu` (Michelle DeVeaux), `michael.kane@yale.edu` (Michael J. Kane), `daniel.zelterman@yale.edu` (Daniel Zelterman)

Table 1: Stopped Negative Binomial Distribution Characteristics

Notation	SNB(p, s, t)
Parameters	p the success probability ($q = 1 - p$) s the number of successes before stopping t the number of failures before stopping
Support	$\min(s,t) \leq k \leq s + t - 1$
PMF	$\binom{k-1}{s-1} p^s (1-p)^{k-s} + \binom{k-1}{t-1} (1-p)^s p^{k-t}$
CDF	$2 - \mathcal{I}_{1-p}(k+1, s) - \mathcal{I}_p(k+1, t)$ where $\mathcal{I}$ is the regularized incomplete beta function.
Mean	$\frac{s}{p} \mathcal{I}_p(s,t) + \frac{q^{t-1} p^{s-1}}{B(s,t)} + \frac{t}{q} \mathcal{I}_q(t,s) + \frac{q^{s-1} p^{t-1}}{B(s,t)}$ where B is the beta function
MGF	$\left(\frac{pe^x}{1-qe^x}\right)^s \mathcal{I}_{1-qe^x}(s,t) + \left(\frac{qe^x}{1-pe^x}\right)^t \mathcal{I}_{1-pe^x}(t,s)$

the trial. In the present example, observing two successful patients allows us reach one endpoint so the sample required could be as small as two. Similarly 11 observed treatment failures also ends the stage. This sampling mechanism, in which the experiment ends as soon as any of the endpoints is reached, is call *curtailed sampling.* Under curtailed sampling the range of the sample size is between two and 12.

Assume each of patient outcome can be modeled as an independent, identically distributed Bernoulli(p) random variable. The trial is realized as a sequence of these random variables that stops when either a specified number of success or failures is reached. In the previous example suppose two successes were reached after enrolling 10 patients (one in the third step and one at the 10^{th}). The sample path is illustrated in Fig. 1. The vertical axis denotes the number of successful outcomes. The horizontal axis counts the number of patients that have been enrolled. The horizontal and vertical boundaries represent endpoints for the trial.

In this work we derive the distribution of the number of enrollees needed for

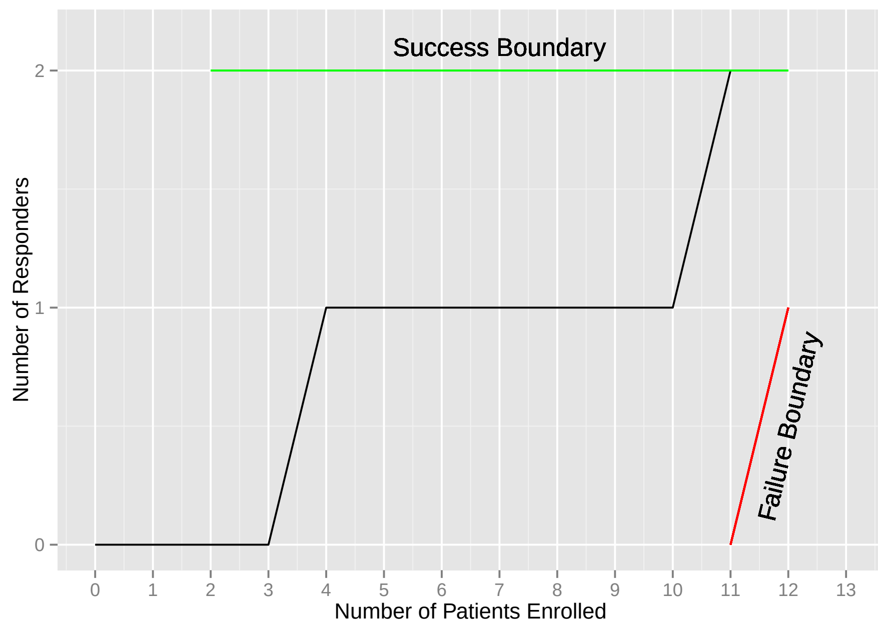


Figure 1: A hypothetical realization of a trial.

either s successes or t failures. We refer to this distribution as the Stopped Negative Binomial (SNB) and some of its characteristics are summarized in Tab. 1. The rest of this paper derives these results and explores properties of the distribution. The next section introduces our notation and basic results including the density of the distribution along with a description of it's relation to other distributions. Section 2 derives the distribution based on a defined Bernoulli process and gives some basic properties. Section 3 provides a connection to the Binomial tail probability. Section 4 derives the posterior distribution using a Beta prior. Section 5 provides a brief discussion on the use of the distribution in clinical trials along with future avenues for generalization.

2. Probability Mass Function

Let $b_1, b_2, \ldots$ denote a sequence of independent, identically distributed, Bernoulli random variables with $\mathbb{P}[b_i = 1] = p$ and $\mathbb{P}[b_i = 0] = 1 - p$, for probability parameter $0 \leq p \leq 1$. In the clinical trial setting $b_i = 1$ corresponds to a successful outcome. Let s and t be positive integers. Define the SNB random variable Y as the smallest integer value such that $\{b_1, \ldots, b_Y\}$ contains *either* s successes *or* t failures. That is, the SNB distribution of Y is the smallest integer such that either $\sum_i^Y b_i = s$ or $\sum_i^Y 1 - b_i = t$.

The distribution of Y has support on integer values in the range

$$\min(s,t) \leq \ Y \ \leq s + t - 1$$

and it is distributed as

$$\mathbb{E}\ I_{\{Y=k\}} = S(k,p,s)\ I_{\{s \leq k\}} + S(k, 1-p, t)\ I_{\{t \leq k\}} \tag{1}$$

where $I_{\{f\}}$ is one if f is true and zero otherwise and

$$S(k,p,s) = \binom{k-1}{s-1} p^s (1-p)^{k-s} \tag{2}$$

is the negative binomial probability

To prove (1), consider the process $\mathbf{X} = \{X(k) : k = 0, 1, ...\}$ with $X(0) = 0$ and

$$X_{k+1} = X_k + b_{k+1}\ I_{\{k-t < X_k < s\}}.$$

At each step a patient's outcome is measured. If it is success, the process advances one diagonally in the positive horizontal and vertical direction. Otherwise, it advances in the positive horizontal direction only. The process continues until either $X_k = s$ or $X_k = k - t$.

Proposition 1. *The distribution of the stopping time* $\underset{k}{\operatorname{argmin}} [X_k \geq s \cup X_k \leq k - t]$ *is given at (1).*

Proof. The probability a given realization of $\mathbf{X}$ reaches s at the kth outcome is the probability that, at time $k-1$ there are $s-1$ successful outcomes and $k-s$ unsuccessful outcomes multiplied by the probability of a success at time k. This expression is given in (2). Similarly, probability a given realization reaches $k-t$ is the probability that, at outcome $k-1$ there are $k-t$ successful outcomes and $t-1$ unsuccessful outcomes multiplied by the probability of an unsuccessful outcome at time k.

Next, define

$$S'(k,p,t) = \binom{k-1}{k-t} p^{k-t}(1-p)^t \tag{3}$$

and notice that $S(k,p,s) = S'(k,1-p,s)$ by writing $\binom{k-1}{k-s} = \binom{k-1}{s-1}$.

To show that (2) and (3) sum to one over their support, let

$$\begin{aligned} R &= \sum_{k=s}^{s+t-1} S(k,p,s) + \sum_{k=t}^{s+t-1} S(k,1-p,t) &(4)\\ &= \sum_{k=s}^{s+t-1} \binom{k-1}{s-1} p^s(1-p)^{k-s} + \sum_{k=t}^{s+t-1} \binom{k-1}{k-t} p^{k-t}(1-p)^t &(5) \end{aligned}$$

where we substitute $i = k-s$ in the first summation and $j = k-t$ in the second.

Then R can be written as the cumulative distribution function of two negative binomial distributions:

$$R = \sum_{i=0}^{t-1} \binom{i+s-1}{i} p^s(1-p)^i \; + \sum_{j=0}^{s-1} \binom{j+t-1}{j} p^j(1-p)^t. \tag{6}$$

Let $\mathcal{I}_p(s,t)$ be the *regularized incomplete beta function* [1] and recall this function satisfies $\mathcal{I}_p(s,t) = 1 - \mathcal{I}_{1-p}(t,s)$ [2].

$$\begin{aligned} R &= \sum_{i=0}^{t-1} \binom{i+s-1}{i} p^s(1-p)^i + \sum_{j=0}^{s-1} \binom{j+t-1}{j} p^j(1-p)^t \\ &= 1 - \mathcal{I}_p(s,t) + 1 - \mathcal{I}_{1-p}(t,s) \\ &= 1. \end{aligned}$$

This completes the proof that (1) is a valid probability mass function. □

3. Shape and Basic Properties

The SNB is a generalization of the negative binomial distribution. If t is large then the $Y-s$ has a negative binomial distribution with

$$\mathbb{P}[Y = s+j] = \binom{s+j-1}{s-1} p^s(1-p)^j$$

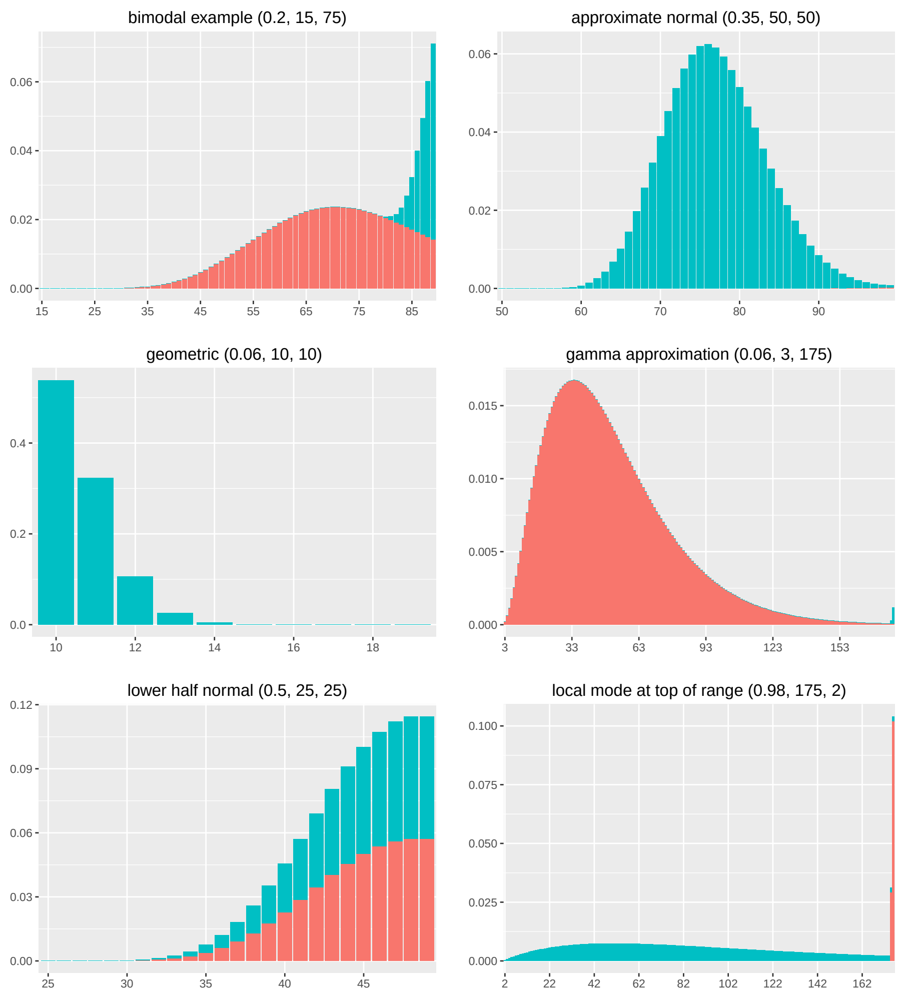


Figure 2: Different shapes of the SNB distribution with parameters (s, t, p), as given. Red indicates mass contributed by hitting s, teal indicates mass contributed by hitting t.

for $j = 0, 1, \ldots$. A similar statement can be made when s is large and t is small. As a result, with proper parameter choice, the SNB can mimic other probability distributions in a manner similar to those described in [3] and [4]. Examples are shown in in Fig. 2.

For the special case of $s = t$, the distribution of Y is the riff-shuffle, or minimum negative binomial distribution [5]. Similar derivations of the closely-related maximum negative binomial discrete distributions also appear in [6] and [7]. The maximum negative binomial is the smallest number of outcomes necessary to observe at least s successes *and* s failures. The SNB is the number of coin flips to observe *either* s heads or t tails.

4. Connection Between the SNB and the Binomial Tail Probability

Proposition 2. *Let Y be distributed as SNB(p, s, t) and let B be distributed Binomial with size $n = s + t - 1$ and success probability p. Then*

$$\mathbb{P}[B \geq s] = \mathbb{P}[Y \leq n \mid \#\textit{success} = s]. \tag{7}$$

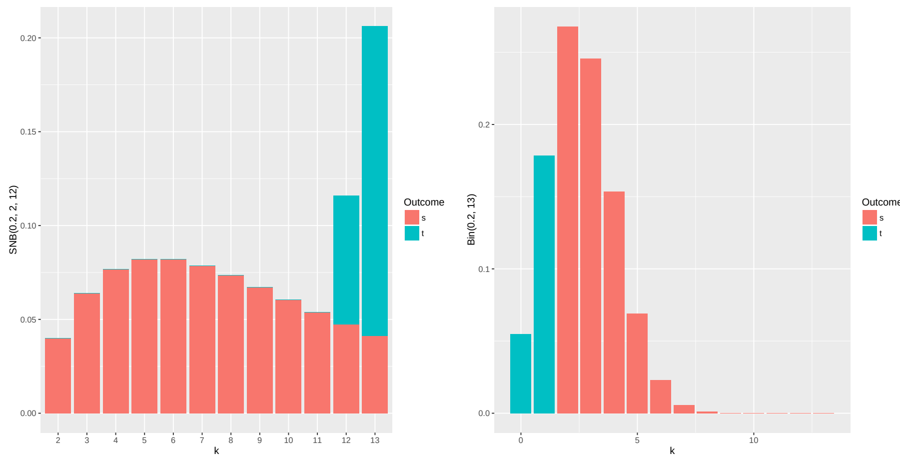
That is, the probability that the number of successes is at least s in the Binomial model is the same that the trial stops with s successes in the SNB model.

Proof. The Binomial tail probability is

$$\begin{aligned}\mathbb{P}[B \geq s] &= \sum_{k=s}^{s+t-1} \binom{n}{k} p^k (1-p)^{n-k} \\ &= 1 - \sum_{k=0}^{s-1} \binom{n}{k} p^k (1-p)^{n-k} \\ 1 - \mathcal{I}_{1-p}.\end{aligned}$$

The corresponding SNB probability is

$$\begin{aligned}\mathbb{P}[Y \leq n \mid \#\text{success} = s] &= \sum_{k=s}^{s+t-1} S(k, p, s) \\ &= \binom{k-1}{s-1} p^s (1-p)^{k-s}.\end{aligned}$$



(a) The SNB distribution.

(b) The Binomial distribution.

Figure 3: SNB(0.2, 2, 12) with mass contributed from s successes (red) or t failures (teal) along with Bin(0.2, 13) with at least 2 successes (red) or fewer (teal).

Let $i = k - s$. Use $\binom{i+s-1}{s-1} = \binom{i+s-1}{i}$ so the last summation can be rewritten as

$$\mathbb{P}[Y \leq n \mid \#\text{success} = s] = \sum_{i=0}^{t-1} \binom{i+s-1}{i} p^s (1-p)^i \tag{8}$$

$$= 1 - \mathcal{I}_{1-p}(t, s) \tag{9}$$

completing the proof. $\square$

Fig. 3 shows the case where $s = 2$, $t = 12$, and $p = 0.2$. The probability masses represented in red are equal as are the masses in teal. The probability that s successes are reached in the SNB process is the same as the binomial probability of at least two successes. Likewise, the probability that t failures are reached in the SNB process is the same as the binomial probability of zero or one successes.

5. The Moment Generating Function

Proposition 3. *Let Y be distributed SNB with parameters p, s, and t. Then the moment generating function (MGF) of Y is*

$$\mathbb{E}\ e^{xY} = \left(\frac{pe^x}{1-qe^x}\right)^s \mathcal{I}_{1-qe^x}(s,t) + \left(\frac{qe^x}{1-pe^x}\right)^t \mathcal{I}_{1-pe^x}(t,s) \tag{10}$$

for $q = 1-p$ when $x \leq \min\{\log(1/p), \log(1/q)\}$.

Proof. The MGF of the SNB is:

$$\mathbb{E}\ e^{xY} = \sum_{k=s}^{s+t-1} \binom{k-1}{k-s} p^s q^{k-s} e^{kx} + \sum_{k=t}^{s+t-1} \binom{k-1}{k-t} p^{k-t} q^t e^{kx}$$

and can be rewritten as:

$$\mathbb{E}\ e^{xY} = \sum_{k=s}^{s+t-1} \binom{k-1}{k-s} (pe)^{sx} (qe^x)^{k-s} + \sum_{k=t}^{s+t-1} \binom{k-1}{k-t} (qe^x)^t (pe^x)^{k-t}. \tag{11}$$

Taking the first summation in Equation 11:

$$\begin{aligned}\sum_{k=s}^{s+t-1} \binom{k-1}{k-s} (pe)^{sx} (qe^x)^{k-s} &= \left(\frac{pe^x}{1-qe^x}\right)^s \sum_{k=s}^{s+t-1} \binom{k-1}{k-s} (qe^x)^{k-s} (1-qe^x)^s \\ &= \left(\frac{pe^x}{1-qe^x}\right)^s \mathcal{I}_{qe^x}(s,t).\end{aligned}$$

Since the incomplete beta function has support on zero to one, we have $qe^x \leq 1$. This also shows $x \leq -\log(q)$.

A similar expression can be derived using the same calculation with the constraint that $x \leq -\log(p)$. The result follows from the afore mentioned property of the regularized incomplete beta function. $\square$

6. The Posterior Distribution

Let us consider a Bayesian analysis where there is a Beta prior distribution on p.

Proposition 4. *The posterior PMF of the Stopped Negative Binomial distribution with a Beta(α, β) prior is:*

$$f(k|s,t,\alpha,\beta) = \binom{k-1}{s-1} \frac{B\left(\alpha+s, k-s+\beta\right)}{B(\alpha,\beta)} \; I_{\{s \le k \le s+k-1\}} \; + \\ \binom{k-1}{k-t} \frac{B\left(\alpha+k-t, t+\beta\right)}{B(\alpha,\beta)} \; I_{\{t \le k \le s+k-1\}}. \tag{12}$$

Proof. For notational simplicity, assume that $s, t \le k \le s+t-1$. When this is not the case appropriate terms should be removed as dictated by the indicator functions.

$$\begin{aligned} f(k|s,t,\alpha,\beta) &= \frac{1}{B(\alpha,\beta)} \int_0^1 \binom{k-1}{s-1} p^{\alpha+s-1} \left(1-p\right)^{k-s+\beta-1} + \\ &\qquad \binom{k-1}{k-t} p^{k-t+\alpha-1} \left(1-p\right)^{t+b-1} \, dp \\ &= \frac{1}{B(\alpha,\beta)} \binom{k-1}{s-1} \int_0^1 p^{\alpha+s-1} \left(1-p\right)^{k-s+\beta-1} dp + \\ &\qquad \frac{1}{B(\alpha,\beta)} \binom{k-1}{k-t} \int_0^1 p^{k-t+\alpha-1} \left(1-p\right)^{t+b-1} dp \end{aligned}$$

The result follows by applying the definition of the Beta function to the integral terms. □

7. Discussion and Conclusion

We have presented a new discrete distribution by curtailed sampling rules common in early-stage clinical trials, which we refer to as the Stopped Negative Binomial distribution. The distribution models the stopping time of a sequential trial where the trial is stopped when a number of events are accumulated. A connection between the Binomial tail probability and the SNB distribution was shown; It's MGF was derived; and the posterior distribution was derived for the case when the event probability p has a Beta distribution.

Current work focuses on two different areas. First, the distribution is being applied to clinical trial design where the number of enrollees is smaller than traditional trials. This scenario is increasingly common in targeted cancer therapy where treatment may only be appropriate for small subpopulations. Second, generalizations and extensions of the distribution are being pursued for applications power analysis. The posterior construction allows us to quantify the uncertainty both in the response probability as well as the time until a trial is complete. This suggests new approaches to sample size calculations as well as trial monitoring.

References

[1] F. W. Olver, NIST handbook of mathematical functions, Cambridge University Press, 2010.

[2] M. Abramowitz, I. A. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables, Vol. 55, Courier Corporation, 1964.

[3] D. B. Peizer, J. W. Pratt, A normal approximation for binomial, f, beta, and other common, related tail probabilities, i, Journal of the American Statistical Association 63 (324) (1968) 1416–1456.

[4] D. Best, P. Gipps, An improved gamma approximation to the negative binomial, Technometrics 16 (4) (1974) 621–624.

[5] V. Uppuluri, W. Blot, A probability distribution arising in a riff-shuffle, Random Counts in Scientific Work, 1: Random Counts in Models and Structures (1970) 23–46.

[6] Z. Zhang, B. A. Burtness, D. Zelterman, The maximum negative binomial distribution, Journal of Statistical Planning and Inference 87 (1) (2000) 1–19.

[7] D. Zelterman, Discrete distributions: applications in the health sciences, John Wiley & Sons, 2005.