Statistics and Its Interface Volume 11 (2018) 699–707

Clinical trial design using a stopped negative binomial distribution

Michelle DeVeaux, Michael J. Kane*, and Daniel Zelterman

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. This report provides a closed-form expression for the mass function, moment generating function, and provides connections to other, standard distributions.

AMS 2000 subject classifications: Primary 62E15; secondary 62P10.
Keywords and phrases: Discrete distributions, Stopped negative binomial distribution, Early-stage clinical trials, Curtailed clinical trials.

1. INTRODUCTION AND MOTIVATION

Consider a prototypical early phase, single-arm clinical trial in which 17 patients are enrolled and treated. The trial is modeled as a sequence of independent Bernoulli(p) samples. Suppose the Bernoulli probability of a patient responding to treatment is $p = 0.2$ under the null hypothesis that the treatment is not any more effective than the current standard of care. If seven or more patients out of these 17 respond to the treatment then we reject this hypothesis and claim the treatment has successfully showed superiority at a significance level of 0.1. If fewer than seven respond then the null hypothesis is not rejected and the treatment is said to have failed to show superiority. The trial ends when either seven responders or 11 non-responders are observed.

If all 17 patients are enrolled at once, as in the classic design, then the sample size is 17; however, in most clinical trials the patients are enrolled sequentially over time. In the present example, observing seven successful patients ends the trial and so the number of enrollees required could be as small as seven. Similarly, 11 observed treatment failures would also end the trial. This sampling mechanism, in which the experiment ends as soon as either predefined endpoint is reached, is called *curtailed sampling*. Under curtailed sampling, the range of the sample size for this trial is seven through 17.

arXiv: 1508.01264
*Corresponding author.

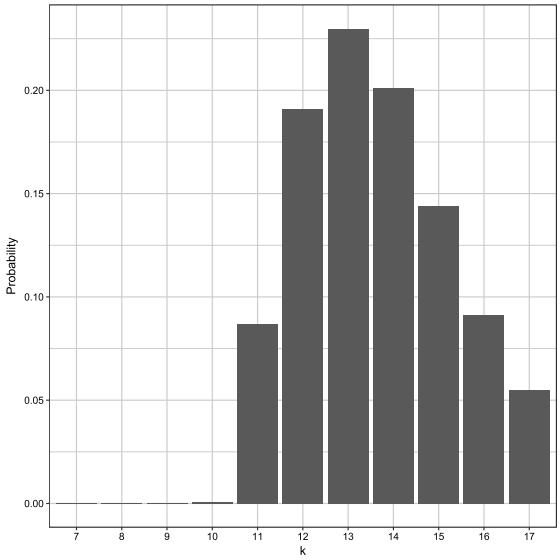


Figure 1. The distribution of the sample size in a trial that stops after either seven patients respond to treatment or 11 do not when $p = 0.2$.

The distribution of the number of trial enrollments is shown in Figure 1. There is relatively little probability mass for values of sample sizes from 7 through 10 since p is small and it is unlikely the treatment will succeed quickly. Figure 2 shows the expected value and variance for the number of trial enrollments varying the success rate p between zero and one. When p is small then the treatment is likely to fail shortly after the 11^{th} enrollment. When p is large then the treatment is more likely to succeed and the number of enrollees approaches seven from above.

When $p = 0$ or 1 then the processes are deterministic and variance is zero. Values of p between zero and one change the mixing proportions of the two endpoints. When p is close to zero, most of the variance is contributed from the failure endpoint. The saddle around $p = 0.25$ results from a trade-off between the success and failure endpoints.

In the rest of this work, we derive the distribution of the number of enrollees needed to observe either s successes or t failures. We refer to this distribution as the Stopped Negative Binomial (SNB). In Section 2, we derive the distribution function and explores its properties. Section 3 derives the

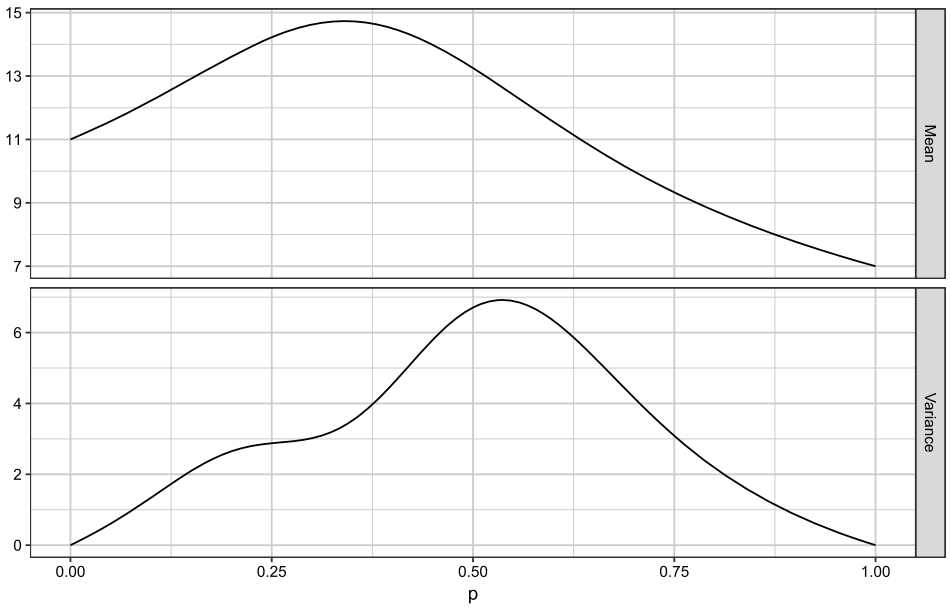


Figure 2. The mean and variance of the sample size for the probability of success p when the trial stops after $s = 7$ patients respond or $t = 11$ fail to respond.

moment generating function. Section 4 describes the likelihood function. Section 5 describes the posterior and predictive probability distributions. Section 6 relates the distribution to other standard distributions and connects the SNB tail probability to the binomial tail probability. Section 7 shows how to design a trial using the SNB in a prototypical setting.

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 the ith patient responding to treatment.

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 responders *or* t non-responders. 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 sequence of Bernoulli random variables can be modeled as the process $\mathbf{X} = \{X_k : k = 0, 1, ...\}$ with $X_0 = 0$ and

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

where $\mathbb{1}_{\{f\}}$ is the *indicator function*, taking the value of one if f is true and zero otherwise, and at each step a patient's outcome is measured. If it is a response, then the response count increases; otherwise, it stays the same. The process continues until either $X_k = s$ or $X_k = k - t$ corresponding to the success and failure boundaries

Figure 3 provides a graphical illustration of X_k against k for one possible realization where $s = 7$ and $t = 11$. The vertical axis denotes the number of successful outcomes. The vertical axis counts the number of responders observed. The horizontal axis counts the number of patients enrolled. It is strictly increasing in sequence size and the failure boundary is tilted so that no more than t (in this case 11) failure can occur. The horizontal and tilted boundaries represent the two endpoints for the trial. In this case, the seventh response was reached on the 15^{th} enrollment. Since the success boundary is reached, we would reject the null hypothesis in this example.

We next derive the probability mass function of Y. The distribution of Y has support on integer values in the range

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

In Proposition 1 below, we show the probability mass function of Y is

$$\begin{aligned} \mathbb{P}[Y = k] = {} & S(k, p, s) \, \mathbb{1}_{\{s\leq k\leq s+t-1\}} \\ & + S(k, 1-p, t) \, \mathbb{1}_{\{t\leq k\leq s+t-1\}} \end{aligned} \tag{1}$$

where

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

is a translated version of the negative binomial probability mass.

The negative binomial cumulative distribution function can be expressed in terms of the *regularized incomplete beta*

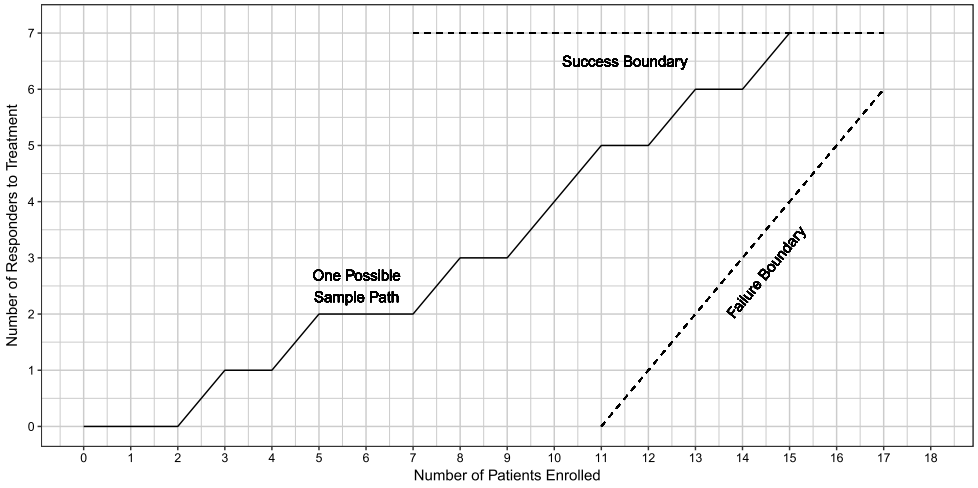


Figure 3. A hypothetical realization of the trial.

function [5]

$$\mathcal{I}_{1-p}(l-s+1,s) = \sum_{k=s}^{l} S(k,p,s)$$

for all integer values l satisfying $s \le l \le s+t-1$. It has property [8]

$$\mathcal{I}_{1-p}(l-s+1,s) = 1 - \mathcal{I}_p(s, l-s+1). \tag{3}$$

Proposition 1. *The distribution of the stopping time*

$$Y = \operatorname*{argmin}_k [X_k \ge s \cup X_k \le k-t]$$

is given at (1).

Proof. The endpoint $X_k = s$ can only occur if $X_{k-1} = s-1$ followed by a success. That is,

$$\begin{aligned} \mathbb{P}[X_k = s] &= p\,\mathbb{P}[X_{k-1} = s-1] \\ &= p\binom{k-1}{s-1} p^{s-1}(1-p)^{k-s} \\ &= \binom{k-1}{s-1} p^s (1-p)^{k-s}. \end{aligned} \tag{4}$$

This expression is given in (2). Similarly, the probability a given realization reaches the endpoint $X_k = k-t$ satisfies

$$\begin{aligned} \mathbb{P}[X_k = k-t] &= (1-p)\,\mathbb{P}[X_{k-1} = k-t] \\ &= (1-p)\binom{k-1}{t-1}(1-p)^{t-1}p^{k-t} \\ &= \binom{k-1}{t-1}(1-p)^t p^{k-t}. \end{aligned} \tag{5}$$

The result (1) follows by summing (4) and (5).

To show (1) sums to one, define

$$R = \sum_{k=s}^{s+t-1} S(k,p,s) + \sum_{k=t}^{s+t-1} S(k,1-p,t).$$

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 (CDF) of two negative binomial distributions:

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

using (3). This completes the proof that (1) is the distribution of the stopping time and it is a valid probability mass function. □

We next consider an interim analysis of a clinical trial after s' patients have responded to treatment and t' failed to respond for $s' < s$ and $t' < t$.

Corollary 1. *The number of subsequent enrollments needed to reach either s or t endpoints behaves as SNB(p, $s-s'$, $t-t'$).*

Having observed s' responders and t' non-responders, there are $s-s'$ additional responders needed to reach the success endpoint and $t-t'$ additional non-responders needed to reach the failure endpoint.

3. THE MOMENT GENERATING FUNCTION

Proposition 2. *Let Y be distributed SNB with parameters p, s, and t. Then the moment generating function (MGF)*

of Y is

(6)
$$\mathbb{E}\ e^{xY} = \left(\frac{pe^x}{1-qe^x}\right)^s \mathcal{I}_{1-qe^x}(s,t) + \left(\frac{qe^x}{1-pe^x}\right)^t \mathcal{I}_{1-pe^x}(t,s)$$

for $q = 1-p$ and is defined for $x < \min\{\log(1/p), \log(1/q)\}$.

The moment generating function for the SNB is calculated in a manner similar to that of two negative binomial distributions. Appendix 1 provides a proof for the derivation.

Proposition 3. *The MGF of the SNB converges to that of the negative binomial when either s or t gets large. That is*

$$lim_{t\to\infty}\ \mathbb{E}e^{xY} = \left(\frac{pe^x}{1-qe^x}\right)^s$$

as. The analogous result holds when $s \to \infty$.

Proof. The second incomplete beta function in (6) can be written in terms of a cumulative binomial distribution

$$\mathcal{I}_{1-pe^x}(t,s) = \mathbb{P}\left[B \le s-1\right]$$

where B is distributed as Binomial$(t-k, pe^x)$. From Chebychev's inequality it follows that

$$\mathbb{P}\left[B \le s-1\right] \le \frac{(t-k)pe^x(1-pe^x)}{(s-(t-k)pe^x)^2} \tag{7}$$

As t gets large $\mathcal{I}_{1-pe^x}(t,s)$ tends to zero and $\mathcal{I}_{1-qe^x}(s,t)$ approaches one. The proof follows by realizing

$$0 < \frac{qe^x}{1-pe^x} < 1$$

over the support of x. □

4. THE LIKELIHOOD FUNCTION

In this section we derive the likelihood function for the SNB for a single trial. In early-stage clinical trial only a single trial is performed, usually because of resource constraints, and the object of interest is p parameter, which determines if a therapy will be delivered to the market. A multi-sample version is less relevant for this use case, but is represented as a product of mixtures of Beta distributions. Deriving it's theoretical characteristics not straight-forward.

Let Y be distributed SNB(p,s,t). Then, the likelihood of Y is proportional to a mixture of Beta distributions.

$$\mathcal{L}(p|s,t,Y) = B_1\,\mathbb{1}_{\{s\le Y_1\}} + B_2\,\mathbb{1}_{\{t\le Y_1\}}$$

where $B_1 = \text{Beta}(s, Y-s)$ and $B_2 = \text{Beta}(Y-t, t)$ and

$$\text{Beta}\,(\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}.$$

Proposition 4. *The mode of B_1 occurs at a value of p greater than that of B_2 in the likelihood function.*

Proof. The mode of the Beta(α, β) distribution is $(\alpha - 1)/(\alpha + \beta - 2)$. Plugging in the shape parameters of B_1 and B_2 into the expression of the mode, the proposition is equivalent to showing

$$\frac{s-1}{Y-2} > \frac{Y-t-1}{Y-2}$$

which is true when $s > Y - t$. The maximum value of the right hand of the inequality occurs when $Y_1 = s+t-1$ and the inequality is equivalent to $s > s-1$, which is true. □

Proposition 5. *The difference between the modes of B_1 and B_2 is bounded by*

$$MODE(B_1) - MODE(B_2)\ \ge\ \frac{1}{s+t-3}.$$

Proof. Proposition 4 shows that the mode of B_1 is greater than that of B_2. The difference between the two can be expressed as

$$\frac{s-1}{Y-2} - \frac{Y-t-1}{Y-2} = \frac{s+t-Y}{Y-2}.$$

This function is strictly increasing in Y over its support and obtains its minimum at $s+t-1$. The result follows. □

As an example, the likelihood function for $Y = 7, 11, 13$ and 17 is shown in Figure 4. When $Y = 7$ the trial must have ended in success and the likelihood function concentrates near 1. The success and failure endpoints can be reached for any value of $Y \ge 11$. When $Y = 11$ we see a bimodal likelihood function which one mode, at $p = 0.6$ provided by the success endpoint and the other, at $p = 0$, from the failure endpoint where no responses are observed. Similarly, when $Y_1 = 13$ we see contributions from both the success and failure endpoints but the two modes are converging. At $Y = 17$ the endpoints contributed likelihoods with similar modes and the result is unimodal.

After an endpoint has been reached, the resulting conditional likelihood is either B_1 or B_2, depending if the trial was a success or failure. However, when the endpoint is not known, such as the planning phase of a trial, unintuitive situations may arise. Since the likelihood is bimodal, there are even settings where we may reject the null despite a poorer alternative. Suppose in the hypothetical trial $p_0 = 0.25$ and $p_1 = 0.5$, the outcome is unknown, and the trial completes after 11 enrollees, as shown in Figure 4 labeled $Y = 11$. Since the null is at a "trough" in the likelihood we fail to reject even though it is closer to the most likely value of $p = 0$.

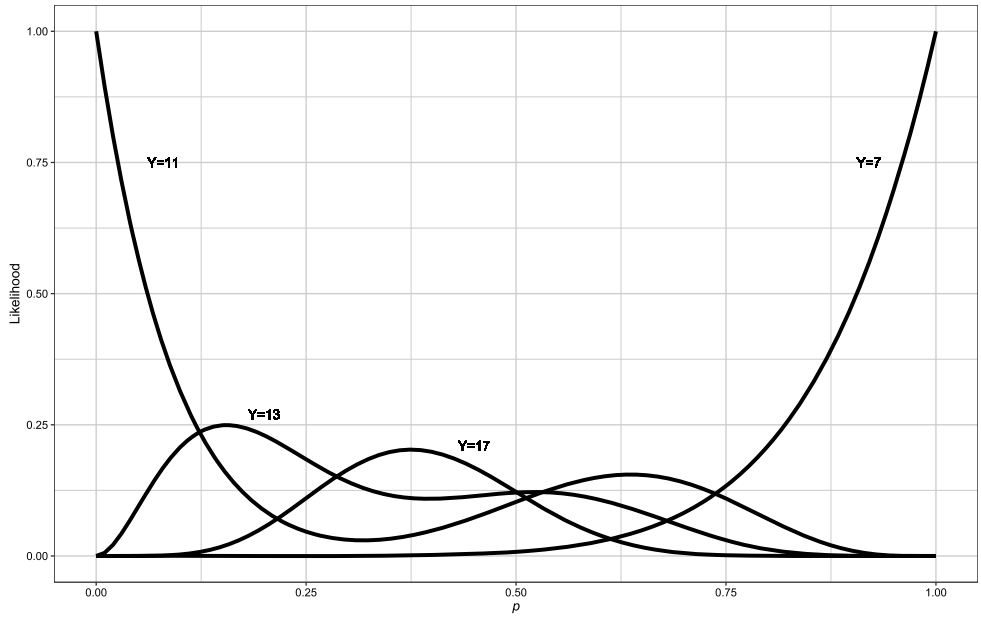


Figure 4. Shapes of the likelihood function for given values of Y with $s = 7$ and $t = 11$.

5. THE POSTERIOR AND PREDICTIVE PROBABILITY DISTRIBUTION

Consider the Bayesian setting where Y is an SNB(P, s, t) distribution and the rate parameter, P is distributed as Beta(α, β). The posterior distribution $P|Y$ is proportional to the likelihood, given by the function

(8)
$$f_{P|Y}(p,k,s,t,\alpha,\beta) \propto \frac{\binom{k-1}{s-1}}{B(\alpha,\beta)} p^{\alpha+s-1}(1-p)^{k+\beta-s-1} + \frac{\binom{k-1}{t-1}}{B(\alpha,\beta)} p^{k+\alpha-t-1}(1-p)^{\beta+t-1}$$

where $0 \le p \le 1$ and $\min(s,t) \le k \le s+k-1$.

The predictive distribution of the SNB can be found as by integrating p over the interval zero to one and applying the definition of the beta function.

(9)
$$f_Y(k,s,t,\alpha,\beta) = \int_0^1 f_P(p|\alpha,\beta) f_{Y|P}(p,k,s,t)dp = \binom{k-1}{s-1}\frac{B(\alpha+s,k-s+\beta)}{B(\alpha,\beta)} + \binom{k-1}{t-1}\frac{B(\alpha+k-t,t+\beta)}{B(\alpha,\beta)}$$

If both α and β are non-negative integers then the predictive distribution is a mixture of negative-hypergeometric distributions.

$$f_Y(k,s,t,\alpha,\beta) = \frac{\binom{k-1}{s-1}\binom{\alpha+\beta}{\alpha}}{\binom{\alpha+\beta+k-1}{\alpha+s}} \frac{\alpha}{\alpha+\beta}\frac{\beta}{k-s+\beta} + \frac{\binom{k-1}{t-1}\binom{\alpha+\beta}{\beta}}{\binom{\alpha+\beta+k-1}{t+\beta}} \frac{\beta}{\alpha+\beta}\frac{\alpha}{k-t+\alpha}$$

The ratio of combinations in the first term can be interpreted as the probability of $s-1$ responders from $k-1$ patients in $\alpha+s$ draws from a population size of $\alpha+\beta+k-1$. This value is multiplied by $\alpha/(\alpha+\beta)$, the expected response rate of the prior. The final term in the product weights the prior based on the number of non-responders $(k-s)$. Terms in the second summand are interpreted similarly for non-responders.

The ratio of (8) divided by (9) gives the posterior distribution of P. It is a mixture of beta distributions. The mixing parameter depend on the endpoints (s and t), the number of enrollees needed to reach an endpoint (k), and the prior parameters (α and β).

6. CONNECTIONS AND APPROXIMATIONS TO OTHER DISTRIBUTIONS

Examples of different shapes of the SNB are shown in in Figure 5 varying parameters p, s, and t. The SNB distribution is a generalization of the negative binomial. As a result, the SNB can approximate other distributions in the same manner as the negative binomial. When t is large then

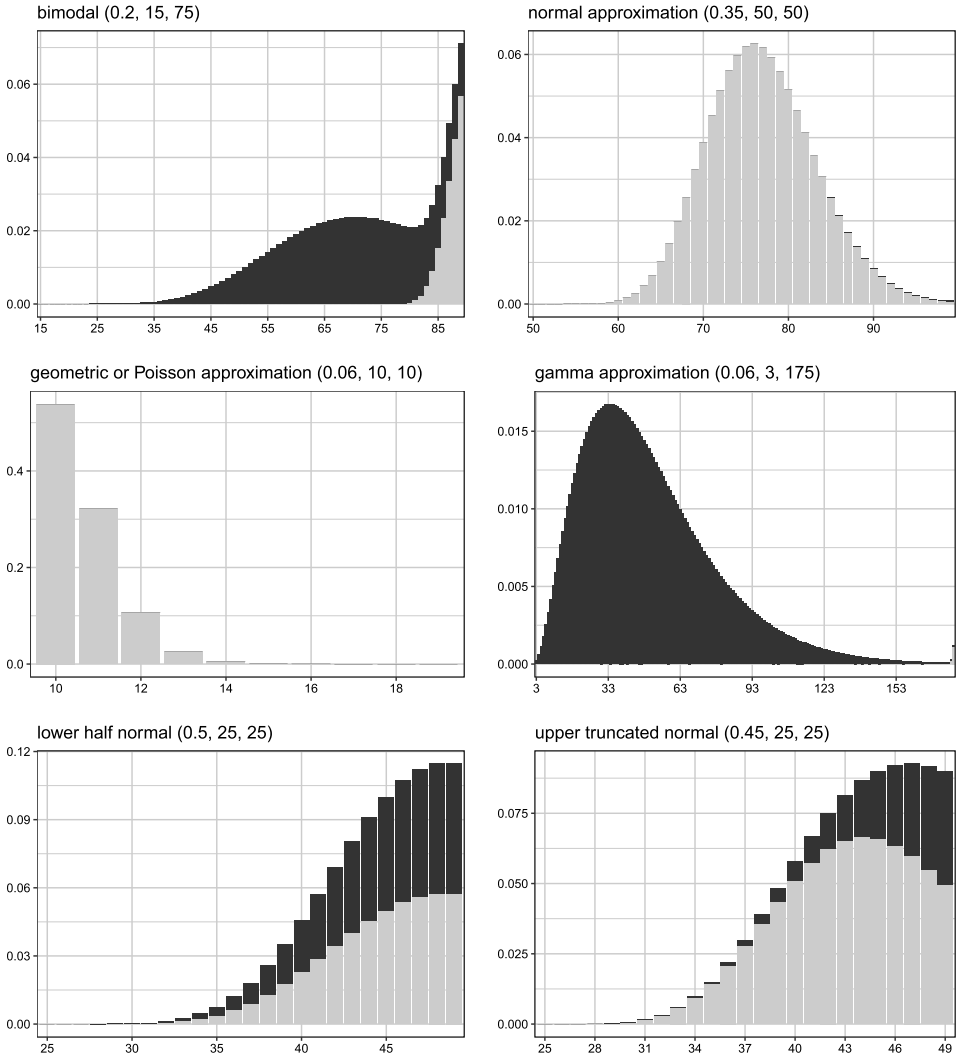


Figure 5. Different shapes of the SNB mass function with parameters $(p,\ s,\ t)$ given. Black areas indicate the mass contributed by reaching s responders before t non-responders. Grey indicates mass contributed by reaching t non-responders first.

$Y - s$ has a negative binomial distribution with

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

for $j = 0, 1, \ldots$. A similar statement can be made when s is large and t is moderate. As a result, with proper parameter choice, the SNB can mimic other probability distributions in a manner similar to those described in [2] and [7]. As a generalization of the negative binomial distribution the SNB inherits the ability to approximate other distributions. For example, when $s = 1$ and $t \to \infty$, the SNB$(p,\ s,\ t)$ converges to an negative binomial distribution with index parameter s and rate parameter p. When $s = 1$, this is the geometric distribution. The connection between the negative binomial and the gamma distribution are well-studied in the literature (see [2, 6, 3] for examples) as well the connection to the Poisson [1].

The SNB generalizes both the minimum (riff-shuffle) and maximum negative binomial distributions up to a translation of the support. For the special case of $s = t$, the SNB distribution is the maximum negative binomial [4, 9, 10] - the smallest number of outcomes necessary to observe at least s responders *and* s non-responders. This is equivalent to a translated version of the riff-shuffle or minimum negative binomial distribution [4, 8].

There is also an equivalence between the probability of reaching an endpoint in the SNB model and the tail probability of the binomial distribution. Specifically, the probability the number of responders is at least s in the binomial model is the same as the probability the treatment succeeds (reaches s) in the SNB model.

Proposition 6. *Let Y be distributed as SNB$(p,\ s,\ t)$ and let X_Y correspond to the number of responders at the end of*

the trial. Let B be distributed binomial with index parameter $n = s + t - 1$ and response probability p. Then

$$\mathbb{P}[B \geq s] = \mathbb{P}[X_Y = s]. \tag{10}$$

Proof. The binomial tail probability is

$$\mathbb{P}[B \geq s] = 1 - \mathcal{I}_{1-p}(s, t)$$

The corresponding success probability is

$$\mathbb{P}[X_Y = s] = \sum_{k=s}^{s+t-1} \binom{k-1}{s-1} p^s (1-p)^{k-s}. \tag{11}$$

Let $i = k - s$. Since

$$\binom{i+s-1}{s-1} = \binom{i+s-1}{i},$$

the summation in (11) can be rewritten as

$$\begin{aligned} \mathbb{P}[X_Y = s] &= \sum_{i=0}^{t-1} \binom{i+s-1}{i} p^s (1-p)^i \\ &= 1 - \mathcal{I}_{1-p}(t, s) \end{aligned}$$

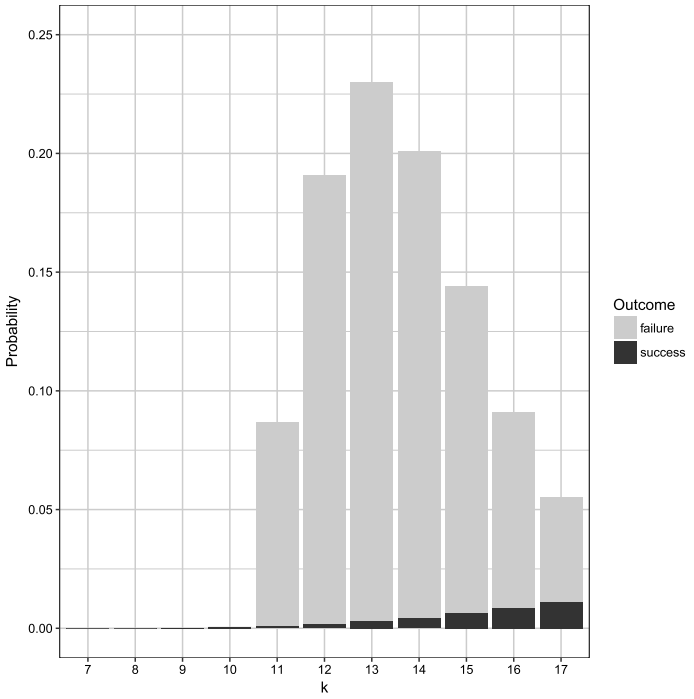
completing the proof. □

To illustrate this result, let us return to our initial example where $s = 7$, $t = 11$, and $p = 0.2$. The probability masses in Figure 6 represented in black are equal in panels (a) and (b) as are the masses in grey. The probability s responders are reached in the SNB process is the same as the binomial probability of at least seven responders. Likewise, the probability t non-responders reached in the SNB process is the same as the binomial probability of zero through six responders.

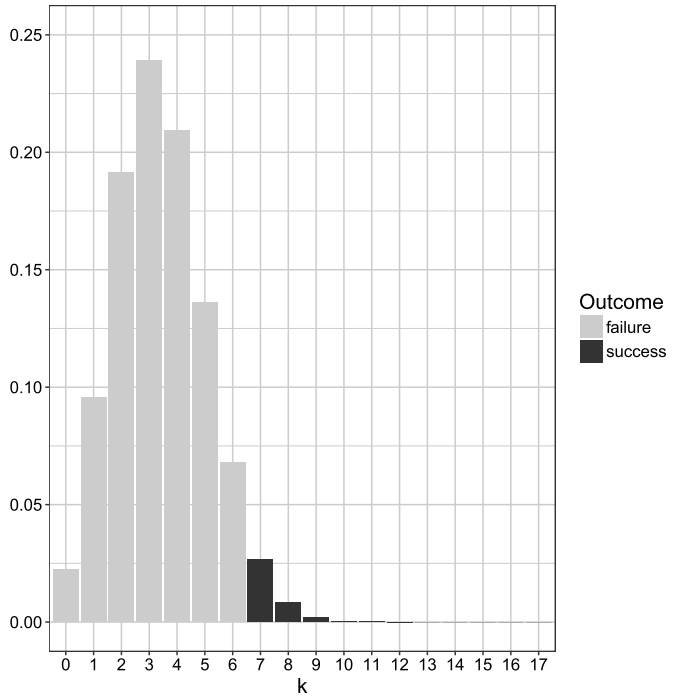
7. TRIAL DESIGN WITH THE SNB

Consider the problem of designing a trial using curtailed sampling and the SNB. Assume that the maximum number of patients n is given along with the null and alternative hypotheses. In this case, the parameter s tells how many responses are needed to reach the success endpoint, can vary between 1 and $n - 1$. For fixed s and n the value of t is determined by the relation $t = n - s + 1$. The ROC curves for all trial designs in the prototype, where $n = 17$, $p = 0.2$ under the null, and $p = 0.4$ under the alternative, are shown in Figure 7. In designing these trials, small significance and large power values are attained by either increasing the value of n or increasing the difference between the alternative and null response rates. Figure 8 shows the expected number of enrollees for each of these trials as a function of the number of responses required to reach the success endpoint.

The expected number of enrollees for each trial is shown in Figure 8. The curve reaches its maximum of 15 patients when $s = 5$. The prototype design ($p = 0.2$, $s = 7$, and $t = 11$) has an expected sample size of 14 enrollees.



(a) SNB distribution



(b) binomial distribution

Figure 6. SNB(0.2, 7, 11) with mass contributed from 7 responders (black) or 11 non-responders (grey) along with Binomial(0.2, 17) with at least 2 responders (black) or fewer (grey).

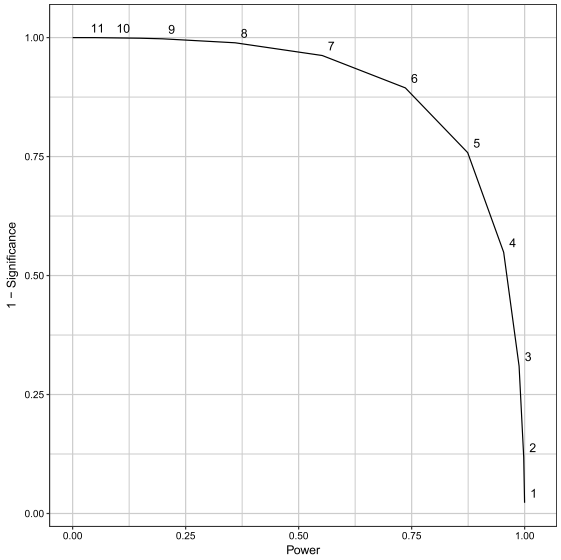


Figure 7. ROC curve of the all trial designs where the maximum number of patients is 17, $p = 0.2$ under the null and $p = 0.4$ under the alternative. The numerical values indicate the number of responses s required reach the success endpoint.

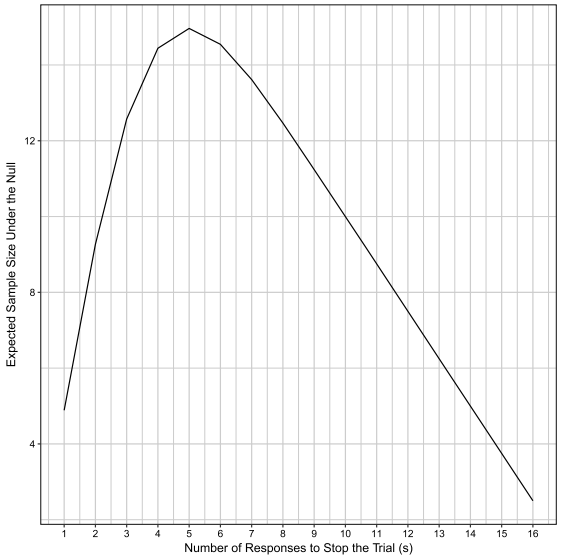


Figure 8. The expected sample size all trial designs where the maximum number of patients is 17 and $p = 0.2$.

APPENDIX: PROOF OF PROPOSITION 2

The MGF of the SNB is, by definition:

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

and can be rewritten as:

$$\mathbb{E}\ e^{xY} = \sum_{k=s}^{s+t-1} \binom{k-1}{s-1} (pe^x)^s (qe^x)^{k-s} + \sum_{k=t}^{s+t-1} \binom{k-1}{t-1} (qe^x)^t (pe^x)^{k-t}. \tag{12}$$

The first summation in (12) satisfies

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

Since the p parameter in $\mathcal{I}_p$ has domain zero to one, we have $0 \le pe^x < 1$. This implies $x < -\log(p)$. A similar expression can be derived for the second summation in (12) and results in the constraint $x < -\log(1-p)$.

ACKNOWLEDGEMENTS

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 LaJolla Pharmaceutical for his suggestions.

Received 10 August 2017

REFERENCES

[1] Anscombe, F. J. Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika 37*, 3/4 (1950), 358–382. MR0039193

[2] Best, D., and Gipps, P. An improved gamma approximation to the negative binomial. *Technometrics 16*, 4 (1974), 621–624. MR0362635

[3] Guenther, W. C. A simple approximation to the negative binomial (and regular binomial). *Technometrics 14*, 2 (1972), 385–389. MR0383608

[4] Johnson, N. L., Kemp, A. W., and Kotz, S. *Univariate Discrete Distributions*, vol. 444. John Wiley & Sons, 2005.

[5] Olver, F. W. *NIST handbook of mathematical functions*. Cambridge University Press, 2010, ch. 8, pp. 189–190.

[6] Ord, J. Approximations to distribution functions which are hypergeometric series. *Biometrika 55*, 1 (1968), 243–248.

[7] Peizer, D. B., and Pratt, J. W. 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.

[8] Uppuluri, V., and Blot, W. Probability distribution arising in a riff-shuffle. Tech. rep., Oak Ridge National Lab., Tenn., 1967.

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

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

Michelle DeVeaux
Regeneron Pharmaceuticals Inc.
Tarrytown, NY
USA
E-mail address: michelle.deveaux@regeneron.com

Michael J. Kane
Department of Biostatistics
School of Epidemiology and Public Health
Yale University
New Haven, CT
USA
E-mail address: michael.kane@yale.edu

Daniel Zelterman
Department of Biostatistics
School of Epidemiology and Public Health
Yale University
New Haven, CT
USA
E-mail address: daniel.zelterman@yale.edu