# 1 Method Background

There exist several distinct but mathematically equivalent parameterizations of NB distributions. Perhaps the most common are: $X \sim NB(\phi, p)$, $X \sim NB(\phi, \mu)$, and $X \sim NB(\alpha, \mu)$. In these formulations, $p$ is the probability of success in a single trial, $\mu$ is the distribution mean, and $\phi$ and $\alpha$ are different representations of the dispersion parameter. Throughout this manuscript, we assume that $0 < p \leq 1$ and $\mu \geq 1$, and $\phi, \alpha > 0$. As the name suggests, the dispersion parameter describes the spread of the distribution and is related to the distribution variance through the equation, $Var(X) = \mu + \mu^2\phi^{-1}$. In addition, $\phi$ and $\alpha$ are related through the relationship $\alpha = \phi^{-1}$ and $p$ and $\mu$ are related through,

$$p = \frac{\phi}{\phi + \mu}. \tag{1}$$

The moment generating function (MGF) of a NB distribution parameterized in this way is,

$$M_X(t) = \Big(\frac{1-q}{1-qe^t}\Big)^{\phi},$$

where

$$q = 1 - p = \frac{\mu}{\phi + \mu}. \tag{2}$$

Due to the assumption of independence, it follows that the MGF of the convolution $Y = \sum_{i=1}^{n} X_i$ of NB r.v.s $(X_i)$ where $i = \{1, 2, 3, ..., n\}$ is,

$$M_Y(t) = \prod_{i=1}^{n} \Big(\frac{1-q_i}{1-q_ie^t}\Big)^{\phi_i}.$$

The cumulant generating function (CGF) of $Y$, expressed in terms of $\phi$ and $q$ is,

$$K_Y(t) = \log(M_Y(t)) = \sum_{i=1}^{n} \phi_i(\log(1-q_i) - \log(1-q_ie^t)) \tag{3}$$

Substituting (1) and (2) into (3) and simplifying yields the CGF used throughout `nbconv`,

$$K_Y(t) = \sum_{i=1}^{n} \phi_i(\log(\phi_i) - \log(\phi_i + \mu_i(1-e^t)). \tag{4}$$

The NB parameterization used in `nbconv` is the same parameterization used in R's `stats` package, as well as in [1].

## 1.1 Furman's exact method

The PMF derived by Furman is,

$$P(Y=y) = R\sum_{k=0}^{\infty}\delta_k \frac{\Gamma(\phi_s+y+k)}{\Gamma(\phi_s+k)y!}p_1^{\phi_s+k}(1-p_1)^y, \tag{5}$$

where,

$$R = \prod_{i=1}^{n}\Big(\frac{q_i p_1}{q_1 p_i}\Big)^{-\phi_i};$$

$$\delta_{k+1} = \frac{1}{k+1}\sum_{j=1}^{k+1} i\xi_j\delta_{k+1-j}, \text{ where } k=0,1,2,... \text{ and } \delta_0 = 1;$$

$$\xi_j = \sum_{i=1}^{n}\frac{\phi_i(1-q_1p_i/q_ip_1)^j}{j};$$

$$q_i = 1-p_i;$$

$$p_1 = \max(p_i);$$

$$q_1 = 1-p_1;$$

and

$$\phi_s = \sum_{i=1}^{n}\phi_i.$$

Evaluation of this PMF in `nbconv`, as well as evaluation of the parameters $\xi_j$ and $\delta_k$, are implemented on the log-scale to prevent numeric overflow. Importantly, Furman showed that $Y$ is a mixture NB distribution with parameters $Y \sim NB(\phi_s+K, p_1)$, where $K$ is an integer r.v. with PMF $P(K=k) = R\delta_k$ [1]. The shape of $K$, therefore, largely determines the shape of $Y$.

## 1.2 Saddlepoint approximation

The saddlepoint approximation is a convenient way to approximate the PMF of r.v.s when an exact expression cannot be easily derived or computed. The saddlepoint approximation requires knowledge of (4), as well as the first two derivatives thereof. The first two derivatives of (4) are,

$$K'_Y(t) = \sum_{i=1}^{n}\frac{\phi_i\mu_i e^t}{\phi_i+\mu_i(1-e^t)}, \tag{6}$$

$$K''_Y(t) = \sum_{i=1}^{n} \frac{\phi_i \mu_i e^t (\phi_i + \mu_i)}{(\phi_i + \mu_i(1 - e^t))^2}, \tag{7}$$

The saddlepoint approximation for the PMF of a discrete r.v. [**?**] is,

$$\hat{p}(x) = \frac{1}{\sqrt{2\pi K''_Y(\hat{t})}} \exp(K_Y(\hat{t}) - \hat{t}x), \tag{8}$$

where $\hat{t} = \hat{t}(x)$ represents the unique solution to

$$K'_Y(\hat{t}) = x. \tag{9}$$

In `nbconv`, the `stats::uniroot()` function is used to find the value of $\hat{t}$ that satisfies (9). In addition, $K_Y(t)$ only exists when $\phi_i + \mu_i(1 - e^t) > 0$. This constrains $t$ such that for given vectors of matched $\mu_i$ and $\phi_i$,

$$t < \min \log\Big(\frac{\phi_i}{\mu_i} + 1\Big).$$

This is used as the upper boundary when solving (9). As with the evaluation of (5), evaluation of (8) and (9) is done on the log-scale in `nbconv` to avoid numeric overflow.

## 1.3 Method of moments approximation

The method of moments approximation is the simplest method implemented in `nbconv`. It is based on the assumption that, under certain conditions (e.g. when the variance and/or skew of $K$ is small), $Y$ does not differ substantially from a NB distribution whose parameters can be derived from the moments of $Y$. Setting $t = 0$ in (6) and (7) yields the first two cumulants of the distribution, which are equal to the first two central moments. These are,

$$\kappa_1 = \bar{\mu} = \sum_{i=1}^{n} \mu_i, \tag{10}$$

and

$$\kappa_2 = \bar{\sigma}^2 = \sum_{i=1}^{n} \mu_i + \frac{\mu_i^2}{\phi_i}. \tag{11}$$

Under the assumption that the mean-variance relationship is the same for the convolution of NB r.v.s as it is for non-convoluted NB r.v.s, an expression for the estimation of the dispersion parameter can be derived by combining (10) and (11),

$$\bar{\phi} = \frac{\Big(\sum_{i=1}^{n} \mu_i\Big)^2}{\sum_{i=1}^{n} \frac{\mu_i^2}{\phi_i}}. \tag{12}$$

While the aforementioned assumption is not strictly true, there are certain instances where it might be a reasonably good assumption (see above). Under this assumption, calculated values of $\bar{\mu}$ and $\bar{\phi}$ can then be used to estimate the density, distribution, and quantile functions of $Y$ via the standard R functions `(d/p/q)nbinom()`.

### 1.4 Summary statistics

`nbconv` additionally contains a function that calculates the mean, variance, skewness, and excess kurtosis of $Y$, as well as the mean of the mixture r.v. $K$. These summary statistics can be useful when deciding which evaluation method to use. The mean and variance of $Y$ are defined in (10) and (11), respectively. To define the skewness and excess kurtosis of $Y$, the third and fourth cumulants of (4) must first be defined:

$$\kappa_3 = \sum_{i=1}^{n} \frac{(2\mu_i + \phi_i)(\mu_i + \phi_i)\mu_i}{\phi_i^2}$$

and

$$\kappa_4 = \sum_{i=1}^{n} \frac{(6\mu_i^2 + 6\mu_i\phi_i + \phi_i^2)(\mu_i + \phi_i)\mu_i}{\phi_i^3}.$$

The skewness ($\gamma_1$) and excess kurtosis ($\gamma_2$) can then be defined as,

$$\gamma_1 = \frac{\kappa_3}{\kappa_2^{3/2}} \tag{13}$$

and

$$\gamma_2 = \frac{\kappa_4}{\kappa_2^2}. \tag{14}$$

Finally, the mean of the mixture r.v. $K$ follows from the fact that $Y \sim NB(\phi_s + K, p_1)$ [1]:

$$\bar{K} = \left(\frac{\bar{\mu}p_1}{q_1}\right) - \phi_s. \tag{15}$$

## References


[1] Edward Furman. On the Convolution of the Negative Binomial Random Variables. In *Statistics & Probability Letters*, 77(2): 169–172. 2007.