.1 Parameterisations of the negative binomial.

There are at least three ways to formulate and interpret an experiment for the independent Bernoulli trials with a fixed number of outcomes (successes, failures or both) with the goal to estimate its probability distribution. The NB model which formalises this experiment is formulated using two out of three following parameters dependent on the formulation:

- number of trials (i.e. total number successes or failures)
- number of successes
- number of failures

The possible formulations for such experiment are

Option 1 observing k failures before obtaining the rth success (most common)

Option 2 obtaining r successes until kth failures have occurred

Option 3 number of trials (successes or failures), n, required before the rth success occurs.

Option 1 is the most common formulation in the literature, see from the table 2, while the remaining two options are encountered only in very few cases.

The UncertML, based on the english version of Wikipedia, uses the second interpretation of this distribution[1]. In the current version of ProbOnto we have reformulated it to align with the vast majority of sources. Otherwise, mistake are likely as the target tools e.g. Matlab, R and winBUGS, to mention here only few, use the most common form, see figure 1.

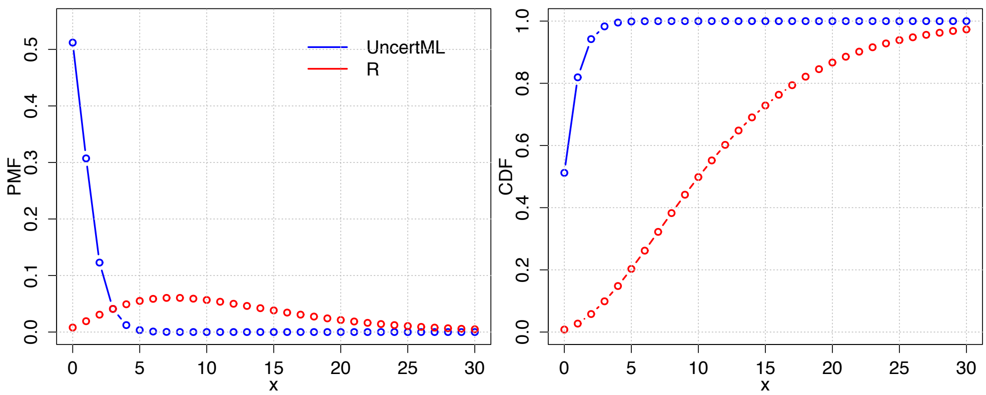

Figure 1: Plots of the negative binomial distribution functions, PMF and CDF, for the different formulations, here *Option 1* and *Option 2* compared. In (blue) the functions for the UncertML formulation, in red as implemented in R for parameters $r = n = 3$ and $p = .2$. See table 1 for the according PMFs and parameter definitions.

	UncertML	R
PMF	$f(x; r, p) = \binom{x+r-1}{x} p^x (1-p)^r$	$f(x; n, p) = \binom{x+n-1}{x} p^n (1-p)^x$
Support	$x \in 0, 1, 2, ...$ number of **successes**	$x \in 0, 1, 2, ...$ number of **failures**
Parameters	$p \in [0, 1]$ – probability of success	$p \in (0, 1]$ – probability of success
	r – number of **failures**	n – number of **successes**

Table 1: UncertML versus R: comparison of the model formulation, parameter used and the support variable. The mathematical formulation of the two models is almost identical, but because of the nuanced differences in support variable interpretation and meaning of the parameters, the resulting models provide very different results, see figure 1.

[1] This was not always the case as the change from the standard formulation to the current one happened around 2010. The majority of users who commented on this change are unhappy about it, see the English and French Wikipedia Talk/Discussion pages. Because there is no agreement to reverse these changes, which would require significant effort, the page remains as is for the time being.

Differences explained

The above mentioned formulation options, see also table 1, are better understood if one compares different language version for the articles on negative binomial distribution in Wikipedia, [4], [5] and [6]. All articles were accessed on 4th August 2015.

English Wikipedia version (Option 1) – supported in UncertML

Interpretation: distribution of the number of successes, k, until r failures have occurred.

$$P_{NB}(k; r, p) = \binom{k+r-1}{k} p^k (1-p)^r, \quad \mathrm{E}(X = k) = \frac{rp}{(1-p)}$$

- Support
  - $k \in \{0, 1, 2, 3, \dots\}$ – number of **successes**
- Parameters
  - $r > 0$ – number of **failures** until the experiment is stopped
  - $p \in (0, 1)$ – success probability in each experiment

French Wikipedia version (Option 2)

Interpretation: distribution of the number of failures, k, before obtaining n successes

$$P_{NB}(k; n, p) = \binom{k+n-1}{k} p^n (1-p)^k, \quad \mathrm{E}(X = k) = \frac{n(1-p)}{p}$$

- Support
  - $k \in \{0, 1, 2, 3, \dots\}$ – number of **failures**
- Parameters
  - $n > 0$ – number of **successes** until the experiment is stopped (fr: *le nombre de succès attendus*)
  - $p \in (0, 1)$ – success probability in each experiment (fr: *la probabilitè d'un succès*)

German Wikipedia version (Option 2)

The german Wiki page describes two alternative representations and interpolations of this distribution. We present here there one which is presented in the overview box on the right-hand side, denoted as the alternative representation.
Interpretation: distribution of the number of failures, k, before obtaining r successes. (ger.: *NB Distribution beschreibt die Anzahl, k, der Misserfolge bis zum Eintreten des r-ten Erfolgs.*)

$$P_{NB}(k; r, p) = \binom{k+r-1}{k} p^r (1-p)^k, \quad \mathrm{E}(X = k) = \frac{r(1-p)}{p}$$

- Support
  - $k \in \{0, 1, 2, 3, \dots\}$ – number of **failures** (ger: *Anzahl Misserfolge*)
- Parameters
  - $r > 0$ – number of **successes** until the experiment is stopped (ger: *Anzahl Erfolge bis zum Abbruch*)
  - $p \in (0, 1)$ – success probability in each experiment, (ger: *Einzel-Erfolgs-Wahrscheinlichkeit*)

$$P_{NB}(k; r, p) = \binom{k+r-1}{k} p^r (1-p)^k.$$

.1.1 Comparison of formulation support

The following table gives an overview of supported NB formulation options across 25 reference sources and specialised software tools.

Source	PMF	Parameter	Support	Support variable
	Option 1 *probability of observing a fixed number of failures before certain number of success*			
Hilbe [12]	$\binom{y+r-1}{y}p^r(1-p)^{\mathbf{y}}$	$0<p<1$	$0\leq y<\infty$	number of failures
Forbes et al. [11]	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	$0<p<1$	$0\leq x<\infty$	number of failures
Leemis [13]	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	$0<p<1$	$0\leq x<\infty$	–
Devroye [8]	$\binom{x+n-1}{x}p^n(1-p)^{\mathbf{x}}$	$p\in(0,1)$	$x\geq 0$	number of failures
Dobson [10]	$\binom{y+r-1}{y}p^r(1-p)^{\mathbf{y}}$	–	–	–
Bonate [2]	$\frac{\Gamma(y+k)}{\Gamma(k)y!}p^k(1-p)^{\mathbf{y}}$	–	$y\geq 0$	–
Cook [7]	$\binom{r+x-1}{x}p^r(1-p)^{\mathbf{x}}$	$0<p<1$	$x\geq 0$	number of failures
Matlab, Stats Toolbox	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	$0<p<1$	$0\leq x<\infty$	number of failures
R *stats* package [17]	$\frac{\Gamma(x+n)}{\Gamma(n)x!}p^n(1-p)^{\mathbf{x}}$	$0<p\leq 1$	$x\in 0,1,...$	number of failures
R *VGAM* package [20, 21]	$\binom{y+k-1}{y}p^k(1-p)^{\mathbf{y}}$	$0<p<1$	$y\in 0,1,2,...$	–
winBUGS [14]	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	–	$0\leq x<\infty$	–
JAGS [16]	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	$0<p\geq 1$	$x\geq 0$	–
UUPDE [15]	$\binom{m+n-1}{n-1}p^n(1-p)^{\mathbf{m}}$	$0<p\leq 1$	$m\in 0,1,2,...$	–
VoseSoftware.com	$\binom{s+x-1}{x}p^s(1-p)^{\mathbf{x}}$	$0<p\leq 1$	$x\in 0,1,...$	number of failures
boost.org	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	–	–	number of failures
Mathwave.com	$\binom{x+n-1}{x}p^n(1-p)^{\mathbf{x}}$	$0<p<1$	$x\in 0,1,...$	–
Wolfram	$\binom{x+r-1}{x}p^r(1-p)^{\mathbf{x}}$	–	–	number of failures
Wikipedia (French)	$\binom{k+r-1}{k}p^r(1-p)^{\mathbf{k}}$	$p\in(0,1)$	$k\in 0,1,2,3,...$	number of failures
Wikipedia (German)	$\binom{k+r-1}{k}p^r(1-p)^{\mathbf{k}}$	$p\in(0,1)$	$k\in 0,1,2,3,...$	number of failures
	Option 2 *probability of observing a fixed number of successes before certain number of failures*			
Agresti [1]	$\binom{y+k-1}{y}(1-p)^k p^{\mathbf{y}}$	–	$y\in 0,1,2,...$	number of successes
UncertML [19]	$\binom{x+r-1}{x}(1-p)^r p^{\mathbf{x}}$	$p\in[0,1]$	$x\in 0,1,2,3,...$	number of successes
Wikipedia (English)	$\binom{k+r-1}{k}(1-p)^r p^{\mathbf{k}}$	$p\in(0,1)$	$k\in 0,1,2,3,...$	number of successes
Consul & Famoye [3]	$\binom{x+n-1}{x}(1-p)^{\mathbf{n}}p^x$	$0<p<1$	$x=0,1,2,3,...$	–
	Option 3 *probability of number of trials required to achieve certain number of successes*			
Distributome [9]	$\binom{x-1}{k-1}p^k(1-p)^{\mathbf{x}-k}$	–	$x\in k,k+1,...$	number of trials
Song & Chen [18]	$\binom{x-1}{k-1}p^k(1-p)^{\mathbf{x}-k}$	$0\leq p\leq 1$	$x\in k,k+1,...$	–
massmatics.de	$\binom{n-1}{r-1}p^r(1-p)^{\mathbf{n}-r}$	$0\leq p\leq 1$	$n\in N, n\geq r$	number of trials

Table 2: Overview of different formulation types of the negative binomial described or supported in reference sources and specialised software tools. Note that *Option 1* and *Option 2* are parameterised with the number of success or failures, respectively, and the probability of success while *Option 3* formulation is parameterised with the total number of success and failures and the probability of success.

Note, the first term of the PMF, in Type 1 & 2, can be formulated in various ways and the following equation shows their equivalent forms

$$\binom{x+z-1}{x}=\binom{x+z-1}{z-1}=\frac{(x+z-1)!}{x!\,(z-1)!}=\frac{\Gamma(x+z)}{x!\,\Gamma(z)}.$$

Both, the UncertML and R version are stored in ProbOnto, as NB1 and NB4, respectively.

Bibliography

[1] Alan Agresti. *Categorical data analysis*, volume 792 of *Wiley series in probability and statistics*. Wiley, Hoboken, NJ, 3rd ed edition, 2013.

[2] Peter L Bonate. *Pharmacokinetic-pharmacodynamic modeling and simulation*. Springer, New York, 2011.

[3] P. C. Consul and Felix Famoye. *Lagrangian probability distributions*. Birkhäuser, Boston, 2006.

[4] Wikipedia contributors. Negative binomial distribution — Wikipedia, the free encyclopedia. Avaiable at `https://en.wikipedia.org/wiki/Negative_binomial_distribution`, Retrieved 4th August 2015.

[5] Wikipedia contributors. Negative binomial distribution — Wikipedia, the free encyclopedia. Avaiable at `https://fr.wikipedia.org/wiki/Loi_binomiale_n%C3%A9gative`, Retrieved 4th August 2015.

[6] Wikipedia contributors. Negative binomialverteilung — Wikipedia, the free encyclopedia. Avaiable at `https://de.wikipedia.org/wiki/Negative_Binomialverteilung`, Retrieved 4th August 2015.

[7] John D. Cook. Notes on the Negative Binomial Distribution. Avaiable at `http://www.johndcook.com/negative_binomial.pdf`, October 2009.

[8] Luc Devroye. *Non-uniform random variate generation*. Springer-Verlag, New York, 1986.

[9] Ivo D Dinov, Kyle Siegrist, Dennis K Pearl, Alexandr Kalinin, and Nicolas Christou. Probability distributome: a web computational infrastructure for exploring the properties, interrelations, and applications of probability distributions. *Computational Statistics*, pages 1–19, 2015.

[10] Annette J. Dobson. *An introduction to generalized linear models*. Chapman & Hall/CRC texts in statistical science series. Chapman & Hall/CRC, Boca Raton, 2nd ed edition, 2002.

[11] Catherine Forbes, Merran Evans, Nicholas Hastings, and Brian Peacock. *Statistical distributions*. John Wiley & Sons, 2011.

[12] Joseph M Hilbe. *Negative binomial regression*. Cambridge University Press, 2011.

[13] Lawrence M Leemis and Jacquelyn T Mcqueston. Univariate distribution relationships. *The American Statistician*, 62(1):45–53, 2008.

[14] David J Lunn, Nicky Best, Andrew Thomas, Jon Wakefield, and David Spiegelhalter. Bayesian analysis of population pk/pd models: general concepts and software. *J Pharmacokinet Pharmacodyn*, 29(3):271–307, Jun 2002.

[15] Oleg Marichev and Michael Trott. The Ultimate Univariate Probability Distribution Explorer. Avaiable at `http://blog.wolfram.com/2013/02/01/the-ultimate-univariate-probability-distribution-explorer/`, 2013.

[16] Martyn Plummer. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, March 2003.

[17] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. ISBN 3-900051-07-0.

[18] Wheyming Tina Song and Yi-Chun Chen. Eighty univariate distributions and their relationships displayed in a matrix format. *Automatic Control, IEEE Transactions on*, 56(8):1979–1984, 2011.

[19] UncertML Team. Uncertainty Markup Language: UncertML Version 3.0. Avaiable at `http://www.uncertml.org`, 2014.

[20] Thomas W. Yee. The VGAM Package. *R News*, 8(2):28–39, October 2008.

[21] Thomas W. Yee. Package VGAM. Avaiable at `https://cran.r-project.org/web/packages/VGAM/VGAM.pdf`, May 2015.