"Mircea cel Batran" Naval Academy Scientific Bulletin, Volume XVIII – 2015 – Issue 2
Published by "Mircea cel Batran" Naval Academy Press, Constanta, Romania // The journal is indexed in:
PROQUEST SciTech Journals, PROQUEST Engineering Journals, PROQUEST Illustrata: Technology, PROQUEST Technology Journals, PROQUEST Military Collection PROQUEST Advanced Technologies & Aerospace

BAYESIAN ESTIMATIONFOR A NEGATIVE-BINOMIAL MODEL WITH A WEIGHTED PROBABILITY DISTRIBUTION

Ioana DANILA[1]
[1]University of Bucharest, Doctoral School of Mathematics, Bucharest, Romania

Abstract: *In the frame of the Negative-Binomial Model, we consider a Bayesian approach using Beta and Pearson type VI as priors. Taking them into consideration in terms of posterior densities, we shall reach the closed-form integration.*
From there on, an expansion of polynomial type for the gamma function is introduced and further, different parameters of interest to re-parameterize the model are taken into account.
Often scientists cannot select sampling units in observational studies with equal probability. Well defined sampling frames often do not exist for human, wildlife, insect, plant, or fish populations. Recorded observations on individuals in these populations are biased and will not have the original distribution unless every observation is given an equal chance of being recorded.
Therefore, when data are recorded according to a certain stochastic model, the recorded observations will not have the original distribution unless every observation is given an equal chance of being recorded.
Biased data arise in all domains of science. Often, sampling units cannot be selected with equal probability for statistical studies. The importance of using weighted distributions arises in such kind of situations. For this, we'll introduce a weighted distribution – with a new positive parameter. An apriori Gamma distribution for the new parameter c exists, and given that, the new formed distribution is also of Gamma type. For this, the line of Line is considered.
For each Poisson process with a rate parameter, it will be given a probability function and a marginal distribution which has a simple mixture probability function.
Beta and gamma priors are introduced for these parameters and afterwards, these parameters are estimated by using the Bayesian approach. Comparisons with the maximum likelihood and moment method estimates are performed.
The Bayesian estimates for the parameters of interest are analyzed via mean squared error and variance through computer simulation, the first model being used in order to model accident statistics. Since then, it has been applied to model phenomena as diverse as the purchasing of consumer package goods, salesperson productivity, in the Biological sciences and so on.
All in all, the Bayesian method for the Negative-binomial model provides alternatives to the maximum likelihood approach. Unlike the maximum likelihood estimates and the moment method estimates, the Bayes estimation produces values in the feasible regions of parameters. By using the proposed sampling procedure, the Bayesian approach for the Negative-binomial Model can be implemented with success in real life applications.

1. Introduction to weighted distributions

Traditional environmetric theory and practice have been occupied with randomization and replication. But in environmental and ecological work, observations also fall in the nonexperimental, nonreplicated, and nonramdom categories. The problems with the model specification and data interpretation then acquire special importance. The theory of weighted distributions provides an approach for all these problems. Weighted distribution takes into account the method of ascertainment, by adjusting the probabilities at a specification of the probabilities of those events as observed and recorded.

The concept of weighted distributions can be traced to the study of the effect of methods of acertainment upon estimation of frequencies by Fisher. In a series of, Patil has pursued weighted distributions for purposes of encountered data analysis, equilibrum population analysis subject to harvesting and predation, meta-analysis incorporating publication bias and heterogeneity, etc.

To introduce the concept of a weighted distribution, suppose X is a non-negative random variable (r.v.) with its natural probability function (p.d.f.) $f(x;\theta)$, where the natural parameter $\theta \in \Omega;\ \Omega$ is the parameter space.

Suppose a realization x of X under $f(x;\theta)$ enters the investigator's record from a the series of events with probability proportional to $w(x,\beta)$, so that:

$$\frac{\Pr\,(recording|X=y)}{\Pr\,(recording|X=x)}=\frac{w(y,\beta)}{w(x,\beta)}$$

The recording (weight) function $w(x,\beta)$ is a non-negative function with the parameter β representing the recording (sighting) mechanism. The recorded x is not an observation of X, but on the r.v. X^w, say, having a pd.f.:

$$f^w(x;\theta,\beta)=\frac{w(x,\beta)f(x;\theta)}{\omega}$$

where ω is the normalizing factor obtained to make the total probability equal to unity by choosing $\omega=E[w(X,\beta)]$. The r.v. X^w is called the weighted version of X, and its distribution in relation to that of X is called the weighted distribution with weight function w. Note that the weight distribution $w(x,\beta)$ need not lie between zero and one, and actually may exceed unity, as, for example, when $w(x,\beta)=x$, in which case $X^*=X^w$ is called the size-biased version of X. The distribution of X^* is called the size-biased distribution with p.d.f. $f^*(x;\theta)=\frac{xf(x;\theta)}{\mu}$ where $\mu=E[X]$.

The p.d.f. f^* is called the length-biased or size-biased version of f, and the corresponding observational mechanism is called length-or size-biased sampling. Weighted distributions were used as a tool in the selection of appropriate models for observed data drawn without a proper frame. In many situations the model given above is appropriate, and the statistical problems that arise are the determination of a suitable weight function, $w(x,\beta)$, and drawing inferences on θ. Appropriate statistical modeling helps accomplish unbiased inference in spite of the biased data and, at times, even provides a more informative and economic setup.

The following examples may help illustrate a few situations generating weighted distributions and their applications.

Example 1 [from G.P.Patil, *Weighted distributions,* Volume 4]

Analysis of family data, $w(x,\beta)=w(x)=x$.

Various demographic studies involve family size and sex ratio as important factors that can have some bearing on the main study.

This example shows how a weighted distribution arises as a result of the size-biased sampling.

Consider the data in Table 0 relating to brothers and sisters in families of 104 boys admitted to a postgraduate course. Assume that in families of given size n, the probability of a family with x boys coming into the record is proportional to x. Also, suppose that the number of boys follows a binomial distribution with probability parameter π.

Then $f(x;\,\pi)=\binom{n}{x}\pi^x(1-\pi)^{n-x}, w(x)=x, \omega=n\pi, f^w(x;\pi)=\binom{n-1}{x-1}\pi^{x-1}(1-\pi)^{n-x}, E\left(\frac{X^W}{n}\right)=\pi+\frac{1-\pi}{n}>\pi,$ and $E\left[\frac{X^W-1}{n-1}\right]=\pi.$

If k boys representing families of size $n_1,n_2,\dots n_k$ report $x_1,x_2,\dots x_k$ boys, an unbiased estimate of π is $\tilde{\pi}=\frac{\sum x_i-k}{\sum n_i-k}=\frac{414-104}{726-104}=\frac{1}{2}$.

Table 0: Family data from Example 1.

Family size:	1	2	3	4	5	6	7	8	9	10	11	12	13	15	Total
No. of families:	1	6	6	13	12	7	14	11	12	8	6	5	2	1	104
Brothers:	1	8	12	34	34	29	59	50	54	46	32	31	16	8	414
Sisters:	0	4	6	18	26	13	39	38	54	34	34	29	10	7	312

Example 2 [from G.P.Patil, *Weighted distributions,* Volume 4]

Analysis of intervention data, $w(x,\beta)=w(x)=x$. The expected value of the duration to completion of a random event sampled randomly at the end of its duration turns out to be approximately equal to the expected duration of its random intervention. This can be explained using the concept of size-biased/ length-biased sampling with weight function $w(x,\beta)=w(x)=x$, where x represents the duration of the random event in the life cycle assessment (LCA). Applications in medicine and environmental health include: (a) cell cycle analysis and pulse labeling; (b) efficacy of early screening for disease and scheduling of examinations; (c) cardiac transplantation; (d) estimation of antigen frequencies; and (e) ascertainment studies in genetics.

Briefly, the weighted distributions, can be summarized as follows:

The weighting scheme - hypotheses:

The original observation x_0 is modeled by a distribution with p.d.f. $f_0(x_0,f_1)$with θ_1 parameter vector.

Observation $x$is recorded according to a probability re-weighted by a weight function w(x; *q*2) >0, with *q*2 parameter vector.

The weighted model:
Observation $x$is modeled using the distribution with p.d.f. $f(x) = Aw(x;\theta_2)f_0(x;\theta_1)$ - weighted distribution; A normalizing constant.
Importance:
- provides a new understanding of standard distributions;
- provides methods of extending distributions for added flexibility in fitting data.

2. Introduction into the Negative-binomial model – reparameterized, and weight distributed.
The negative-binomial model is generated in the following manner. Consider a population in which the count data, X_i, for each individual member has a Poisson process with the rate parameter λ. Thus given, λ, $X_i$has the probability function as follows:

$$p(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0,1,2 \ldots$$

Suppose that λvaries across the population and has a weighted probability distribution gamma with shape parameter $\gamma > 0$and scale parameter $\alpha > 0$. The $c > 0$ constant is introduced:

$$f(\lambda|\gamma,\alpha,c) = \frac{\alpha^{\gamma+c}\lambda^{\gamma+c-1}e^{-\alpha\lambda}}{\Gamma(\gamma+c)}, \qquad \lambda,\alpha,c > 0,$$

where $\Gamma(\gamma+c)$is the gamma function. Since $\lambda$is not observable, the marginal distribution for Xhas the following simple mixture probability function:

$$Pr_{NB}(x|\gamma,\alpha,c) = \int_0^\infty f(\lambda|\gamma,\alpha,c)p(x|\lambda,c)d\lambda\, dc.$$

Therefore, the negative binomial model can be rewritten as,

$$Pr_{NB}(x|\gamma,\alpha,c) = \frac{\Gamma(\gamma+c+x)}{x!\,\Gamma(\gamma+c)}\alpha^{\gamma+c}(1+\alpha)^{\gamma+c+x}, \qquad x = 0,1,2 \ldots$$

Letting $\mu = \frac{\gamma+c}{\alpha}$, the Negative-binomial canbe re-parameterized as:

$$Pr_{NB}(x|\mu,\gamma,c) = \frac{\Gamma(\gamma+c+x)}{x!\,\Gamma(\gamma+c)}\left(1+\frac{\mu}{\gamma+c}\right)^{-\gamma-c}\left(\frac{\mu}{\gamma+c+\mu}\right)^x, \qquad x = 0,1,2 \ldots$$

Here, $\mu = E(X)$and $var(X) = \mu + \frac{\mu^2}{\gamma+c}$.

Letting $\beta = {}^1/_{\gamma + c}$, the Negative-binomial probability function $Pr_{NB}(x|\mu,\gamma,c)$
can be re-parameterized as follows:

$$Pr_{NB}(x|\mu,\gamma,c) = \frac{\Gamma(\beta^{-1}+x)}{x!\,\Gamma(\beta^{-1})}(1+\beta\mu)^{-1/\beta}\left(\frac{\beta\mu}{1+\beta\mu}\right)^x, \qquad x = 0,1,2 \ldots$$

Bradlow et al. (2002) introduced a beta-prime prior on α and a Pearson type VI prior on γ, cin the Negative-binomial:

$$Pr_{NB}(x|\gamma,\alpha,c) = \frac{\Gamma(\gamma+c+x)}{x!\,\Gamma(\gamma+c)}\alpha^{\gamma+c}(1+\alpha)^{\gamma+c+x}, \qquad x = 0,1,2 \ldots$$

Letting $p = {}^\alpha/_{\alpha + 1}$, the equation above can be re-parameterized as:
$Pr_{NB}(x|\gamma,p,c) = \frac{\Gamma(\gamma+c+x)}{x!\Gamma(\gamma+c)}(p)^{\gamma+c}(1-p)^x$, $x = 0,1,2 \ldots$where $0 < p < 1$ and $\gamma + c > 0$.
In this note, the Beta prior and the Gamma prior are introduced for p, γ and c, respectively.
The Bayes estimations for the parameters will be studied. Section 2 describes the models in a Bayesian framework with input data. Section 3 presents and discusses the results of the simulation.

3. FORMULAS FOR ESTIMATIONS AND PREDICTIONS

Let a sample of size Nbe selected from the model given in

$$Pr_{NB}(x|\gamma,p,c) = \frac{\Gamma(\gamma+x)}{x!\,\Gamma(\gamma+c)}(p)^{\gamma+c}(1-p)^x$$

Then the data will consist of the number of occurrences, k, for each of the Nunits.

The data can be summarized as $n_k, k = 0,1,2,\ldots,r$, where $n_k$is the number of units with k occurrences, r is the largest possible occurrence for each sample and

$$N = \sum_{k=0}^{r} Pr_{NB}(k|\gamma,p,c)^{n_k} \ (\bullet)$$

Hence, the likelihood function is given as:

$$L(\gamma,p,c) = \prod_{k=0}^{r} Pr_{NB}(k|\gamma,p,c)^{n_k}.$$

Assume that γ, c and p are independent a priori and have a Gamma distribution with shape parameter δ_2, scale parameter $\delta_1$and a beta distribution with shape parameters $\alpha_1$and α_2, respectively. Then the joint prior distribution of γ, c and p is expressed as:

$g(\gamma,p,c) = p^{\alpha_1-1}(1-p)^{\alpha_2-1}\gamma^{\delta_2-1}e^{-\delta_1\gamma}c^{\delta_3-1}e^{-\delta_4 c},\ 0 < p < 1,\ \gamma,\ c > 0,\ \delta_1,\delta_2,\delta_3,\delta_4 > 0.$ (••)
Combining the likelihood function ($\bullet$)with the joint prior $g(\gamma,p,c)$(••), the joint posterior distribution of $\gamma,\ c$and p, given data $n_k,\ k = 0,1,2,\ldots,r,$ can be presented as follows:

$$Pr(\gamma,p,c|r,n_k) \propto \prod_{k=0}^{r}\left[\frac{\Gamma(\gamma+c+k)}{k!\,\Gamma(\gamma+c)}(p)^{\gamma+c}(1-p)^{k}\right]^{n_k} g(\gamma,p,c)\Big/\Phi(r,n_k),$$

$$0 < p < 1, \qquad \gamma,c > 0 \text{ where} \Phi(r,n_k) = \iiint L(\gamma,p,c|r,n_k)g(\gamma,p,c)dpd\gamma dc.$$

$Pr(\gamma,p,c|r,n_k)$can also be rewritten as:

$$Pr(\gamma,p,c|r,n_k) \propto \prod_{k=1}^{r}\left[(\gamma+c+k-1)^{N-\sum_{i=0}^{k-1}n_i}(1-p)^{kn_k}\right] \cdot$$
$$p^{N(\gamma+c)+\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\gamma^{\delta_2-1}e^{-\delta_1\gamma}c^{\delta_3-1}e^{-\delta_4 c}, \quad 0 < p < 1, \gamma, c > 0.$$

The equation from above can be represented as follows:

$Pr(\gamma,p,c,|r,n_k) \propto f_1(\gamma,p,c,n_k)p^{\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\gamma^{\delta_2-1}e^{-\delta_1\gamma}c^{\delta_3-1}e^{-\delta_4 c},$
$0 < p < 1, \qquad \gamma, c > 0.$
Let

$$K(r,n_k) = \int_0^1\int_0^\infty\int_0^\infty f_1(\gamma,p,c,n_k)\,p^{\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\gamma^{\delta_2-1}e^{-\delta_1\gamma}c^{\delta_3-1}e^{-\delta_4 c}\,dp\,d\gamma\,dc$$

The marginal posterior distribution of γ and $c$could be obtained as:
$P\,r(\gamma|r,n_k) = \int_0^1\int_0^\infty Pr(\gamma,p,c|r,n_k)dp\,dc$;$Pr(c|r,n_k) == \int_0^1\int_0^\infty Pr(\gamma,p,c|r,n_k)dp\,d\gamma.$
The marginal posterior distribution of p could be obtained as:

$$Pr(p|r,n_k) = \int_0^\infty\int_0^\infty Pr(\gamma,c,p|r,n_k)\,d\gamma\,dc.$$

Since the marginal posterior distributions for $\gamma, c$and for $p$arenot amenable to closed-form integration, the moments of the marginal posteriorcan be computed in the following ways:

$E(\gamma^l) = \int_0^1\int_0^\infty\int_0^\infty \gamma^l Pr(\gamma,p,c|r,n_k)\,dp\,d\gamma\,dc$
$\propto \int_0^1\int_0^\infty\int_0^\infty \gamma^l f_1(\gamma,p,c,n_k)p^{\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\,\gamma^{\delta_2-1}e^{-\delta_1\gamma}c^{\delta_3-1}\,e^{-\delta_4 c}dp\,d\gamma\,dc$ (*)
$E(c^l) = \int_0^1\int_0^\infty\int_0^\infty c^l\,Pr(\gamma,p,c|r,n_k)\,dp\,d\gamma\,dc$
$\propto \int_0^1\int_0^\infty\int_0^\infty c^l f_1(\gamma,p,c,n_k)p^{\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\gamma^{\delta_2-1}\,e^{-\delta_1\gamma}\,c^{\delta_3-1}\,e^{-\delta_4 c}dp\,d\gamma\,dc.$ (*')
and
$E(p^l) = \int_0^1\int_0^\infty\int_0^\infty p^l\,Pr(\gamma,p,c|r,n_k)\,dp\,d\gamma\,dc$
$\propto \int_0^1\int_0^\infty\int_0^\infty p^l f_1(\gamma,p,c,n_k)p^{\alpha_1-1}(1-p)^{\alpha_2-1}\delta_1^{\delta_2}\gamma^{\delta_2-1}\,e^{-\delta_1\gamma}\,c^{\delta_3-1}\,e^{-\delta_4 c}dp\,d\gamma\,dc.$ (**)
When $l = 1$, (*) and (**) are the Bayesian estimates of $\gamma,\ c$and p, respectively (under quadratic loss).

Let $Y_i$be the count variable for individual $i$in a non-overlapping time period, which is equal to the length of the time period for the count variable, X_i. For the Bayesian prediction in the negative binomial model, mean, $E(Y_i|x)$and variance, $Var(Y_i|x)$of the predictive distribution are emphasized in this note.

Given $\gamma,\ c$ and p, the mean and variance of Y_i, conditional on x_i, are

$$E(Y_i|\gamma, x_i, c) = (\gamma + c + x_i)(1-p)$$

and

$$Var(Y_i|\gamma, x_i, c) = (\gamma + c + x_i)\big((1-p) + (1-p)^2\big)$$

Hence,

$$E(Y_i|x_i) = \int_0^1 \int_0^\infty \int_0^\infty (\gamma + c + x_i)(1-p)\, Pr(\gamma, p, c|r, n_k)\, dp\, d\gamma\, dc$$

$$\propto \int_0^1 \int_0^\infty \int_0^\infty (\gamma + x_i + c)(1-p) f_1(\gamma, p, c, n_k) p^{\alpha_1 - 1}(1-p)^{\alpha_2 - 1} \delta_1^{\delta_2} \gamma^{\delta_2 - 1} e^{-\delta_1 \gamma}\, c^{\delta_3 - 1}\, e^{-\delta_4 c}\, dp\, d\gamma\, dc$$

(***)

and

$$Var(Y_i|x_i) = \int_0^1 \int_0^\infty \int_0^\infty (\gamma + x_i + c)\big((1-p) + (1-p)^2\big)\, Pr(\gamma, p, c|r, n_k)\, dp\, d\gamma\, dc$$

$\alpha\ \int_0^1 \int_0^\infty \int_0^\infty (\gamma + x_i + c)\big((1-p) +$
1−p2f1γ,p,c,nkpα1−11−pα2−1δ1δ2 γδ2−1e−δ1γ cδ3−1 e−δ4c dp dγ dc (****)

It should be mentioned that the proportionality for (*), (**),(***) and (****) is the reciprocal of $K(r, n_k)$.

4. Simulation studies

As the closed forms for the posterior l th moments of $\gamma,\ c$ and $p$and the conditional mean and variance of $Y_i$are not available given x, an importance sampling method can be used to estimate those parameters. The importance sampling process for Bayesian estimations is described in the following steps.

1. Randomly generate observations, p_i, of size $n$from a Beta distribution with parameters $\alpha_1$and α_2, and randomly generate observations, $\gamma_j,\ c_s$of size m from a Gamma distribution with scale parameter $\delta_1$and shape parameter δ_2.

2. Calculate

$$\hat{K}(r, n_k) = \sum_j \sum_s \sum_i f_1(\gamma_j, c_s, p_i, n_k)$$

as the estimate of $K(r, n_k)$;

$$\rho_l(r, n_k) = \sum_j \sum_s \sum_i \gamma_j^l f_1(\gamma_j, c_s, p_i, n_k);$$

$$\theta_l(r, n_k) = \sum_j \sum_s \sum_i p_i^l f_1(\gamma_j, c_s, p_i, n_k);$$

$$\omega_l(r, n_k) = \sum_j \sum_s \sum_i c_s^l f_1(\gamma_j, c_s, p_i, n_k).$$

3. $E(\gamma_j)^l$ and $E(c_s)^l$ and are estimated by $\rho_l(r, n_k)/\hat{K}(r, n_k)$ and $\omega_l(r, n_k)/\hat{K}(r, n_k)$;
and $E(p^l)$is estimated by$\theta_l(r, n_k)/\hat{K}(r, n_k)$.

Similarly, $E(Y_i|x)$can be estimated by $\sum\sum\sum((\gamma_j, c_s) + x)(1-p) f_1(\gamma_j, c_s, p_i, n_k)/\hat{K}(r, n_k)$and $Var(Y_i|x)$ can be estimated by $\sum\sum\sum\big((\gamma_j, c_s) + x\big)\big((1-p) + (1-p)^2\big)\, f_1(\gamma_j, c_s, p_i, n_k)/\hat{K}(r, n_k)$.

The main purpose of this section is to estimate the parameters, $\gamma,\ c$and $p,\ \omega$ for each of the nine models by using a computing simulation process. These nine models represent quite distinct negative-binomial models which have γ, cselected from 1.00, 2.00 and 4.82 and p, ωselected from 0.25, 0.50 and 0.75.

Assume that the prior for $p$is non-informative prior which has $\alpha_1 = 1$ and $\alpha_2 = 1$ and the

prior of γ and $c$is the Gamma distribution which has $\delta_1$selected from 0.5, 1.0 or 2.0 and δ_2 selected from 0.5, 1.0, 2.0 or 4.5. Given a Negative-binomial model mentioned in this section, 1000 samples of size 100 are generated. For each random sample of size 100, we have $r, n_k, k = 0,1,2, \ldots, r$ and $100 = \sum_{k=0}^{r} n_k$.
A Bayes' estimate,$\hat{\gamma}$, $\hat{c}$, for γ, $c$and a Bayes' estimate, $\hat{p}$, for $p$are calculated via Steps 1-3 of the simulation procedure using a sample of size 100 from a non-informative Beta distribution of p and a sample of size 1000 from one of the Gamma prior distributions. Then the mean squared error (MSE), the variance (VAR) and the mean absolute deviation (MAD) for the Bayes' estimator of γ, c and the mean squared error (MSE), the variance (VAR) and the mean absolute deviation (MAD) for the Bayes' estimator of $p$are calculated from the 1000 Bayes' estimates of γ, c and the 1000 Bayes' estimates of p, respectively.
In general, the MSEs, Mads and VARs for the Bayes' estimates of $p$are consistently small and are less sensitive to the priors than the MSEs, MADs and VARs for the Bayes' estimates of γ and c. When misinformed prior are given for γ and c such as a Gamma prior of $\delta_1 = 1.5$ and $\delta_2 = 4.5$or a Gamma prior of $\delta_1 = 1$ and $\delta_2 = 1$, the resulting Bayes' estimator of γ and $c$has largest MSEs, VARs and MADs among all the Bayes' estimators of γ and c.

Acknowledgements:
Ioana Danila and her work was supported by the strategic grant POSDRU/187/1.5/S/155559 Competitive
multidisciplinary doctoral research at European level (CdocMD) *cofinanced by thl7e European Social Found within the Sectorial Operational Program Human Resources Development 2007 – 2013*

BIBLIOGRAPHY:

[1] F.J. Anscombe (1950), *Sampling Theory of The Negative Binomial And Logarithmic Series Distributions*, Biometrika 36;
[2] C.I. Bliss, R.A. Fisher (1953*), Fitting the negative binomial distribution to biological data*, Biometrics 9;
[3] R. Ihaka, R. Gentleman (1996), *R: A language for data analysis and graphics*, Journal of Computational and Graphical Statistics 5;
[4] B.H. Margolin, N. Kaplan, E. Zeiger (1981), *Statistical Analysis of the Ames Salmonella/microsome Test. Proc. Nat. Acad. Sci.* USA 76;
[5] W.W. Piegorsch (1990), *Maximum Likelihood Estimation for the Negative Binomial Dispersion Parameter*, Biometrics 46;
[6] Y. L. Lio (2009), *A note on Bayesian Estimation for the Negative-Binomial Model,* Pliska Stud. Math. Bulgar. 19;
[7] G.P.Patil, *Weighted distributions,* Volume 4, pp. 2369-2377 in Encyclopedia of Environmetrics; Edited by Abdel H. El – Shaarawi and Walter W. Piegorsch, John Wiley & Sons, Ltd. Chichester, 2002.
[8] L.L.McDonald, *The need for teaching weighted distribution theory: Illustrated with application in environmental statistics*, Western EcoSystems Technology, Inc., United States of America, 2010.