Journal of Hyperstructures 2 *(2)* (*2013*)*, 185-200.*
ISSN: 2251-8436 print/2322-1666 online

BAYES AND EMPIRICAL BAYES ESTIMATION OF PARAMETER k IN NEGATIVE BINOMIAL DISTRIBUTION

MASOUD GANJI *, NASRIN EGHBALI AND MASOUD AZIMIAN

Abstract. In this paper, the problem of estimating the number of successes, k, in a negative binomial distribution for both known and unknown probability p of success are examined by a Bayesian point of view. Also, we introduce two estimations for the parameter of negative bionomial distribution.

Key Words: Bayes estimation, negative binomial distribution, Empirical Bayes estimation, hyper parameter.
2010 Mathematics Subject Classification: Primary: 13A15; Secondary: 13F30, 13G05.

1. Introduction

Let $X_1, X_2, ..., X_n$ be a random Variables of size n from a negative binomial distribution $NB(K, P)$, where K and P are independent. The probability density function (briefly, pdf) of X_i is

$$f(x_i|k,p) = \binom{x_i + k - 1}{x_i} p^k (1-p)^{x_i}, \tag{1.1}$$

$$x_i = 0, 1, 2, ... \ , \ \ k \in \{1, 2, 3, ...\}, \ \ p \in (0, 1).$$

The negative binomial distribution is an univariate discrete probability model with one variable, that has many applications in different fields such as ecology, biology and genetics. Also for modeling count data with over dispersion (that is sample variance exceeds the mean

Received: 7 April 2013, Accepted: 11 October 2013. Communicated by H. R. Nilisani
∗Address correspondence to Masoud Ganji; E-mail:mganji@uma.ac.ir

sample), a popular and convenient model is the negative binomial distribution (see [1], [2], [3]). The usual problem in the negative binomial situation is to estimate the probability of success p. However, in other instances, the number of success k may be the unknown parameter. There are different ways to estimate this parameter, for example the method of moments estimation (MME), maximum likelihood estimation (MLE) and the Bayes method. Many researchers in each turn, have obtained estimations for parameters of negative binomial distribution. For example, Piegorsch [9] used the maximum likelihood estimation for parameter k in this distribution; Clark and Perry [3] used the moment and maximum extended quasi-likelihood estimators for estimation k and Adamids [1], used the EM algorithm for the parameters of the negative binomial distribution. These methods are classical methods. Ganji [6], in a more general case, investigated Bayes estimation of parameters of generalized negative binomial distribution and it's applications; he has fixed the parameter k and then estimated the parameter p (probability of success). In this research, we are going to estimate the parameter k. We are interested in Bayes and Empirical Bayes estimations of parameter k in two cases of known and unknown parameter p. We consider a left-truncated prior distribution for k and a beta prior distribution for p and suggest choosing values for the hyper parameters of the prior distribution. In Section 2, we describe the probability models that are needed in this work. Bayes estimations for k are proposed in section 3. Empirical Bayes estimations for k are investigeted in section 4. We provide an example in section 5. In section 6 simulation study are presented.

2. The Models

Let $X_1, X_2, ..., X_n \sim NB(k, p)$ and the pdf of X_i is given in (1.1). The likelihood function of $\{x_1, x_2, ..., x_n\}$ given $k \geq 1$ and p, is given by:

$$L(x_1, x_2, ..., x_n|k, p) = \prod_{i=1}^{n} \binom{x_i + k - 1}{x_i} p^{nk} (1-p)^{\sum_{i=1}^{n} x_i}. \tag{2.1}$$

Since K is the number of success in the negative binomial experiment, so $k \geq 1$. We consider a left-truncated distribution for K for which the probability density function, is given by:

$$g(k) = \frac{e^{-\lambda}\lambda^k}{k!(1-e^{-\lambda})}, \qquad k \in \{1, 2, 3, ...\}, \tag{2.2}$$

where the hyper parameter λ is a preselected integer that is chosen to reflect our beliefs about the expected value of k, because the expected value of k equals approximately λ. In order to construct the Bayes estimation for k, the posterior pdf of K given the observations $\{x_1, x_2, \ldots, x_n\}$ is derived in this section under the assumption of known and unknown p, separately.

2.1. **Case 1: p is unknown.** Assume that p is an unknown parameter. A suitable prior distribution for p is the beta distribution with hyper parameters alpha and beta, that is $P \sim beta(\alpha, \beta)$. The prior pdf of P is given by:

$$w(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}; \qquad p \in (0,1) \tag{2.3}$$

where the hyper parameters α and β are preselected positive numbers that should be chosen according to our beliefs about the expected value of the true p, because the expected value of p equals $\frac{\alpha}{\alpha+\beta}$. If we believe that the actual value of the p is small, less than 0.5 then, let $\frac{\alpha}{\alpha+\beta} < 1/2$, therefore, it is recommended to set $\beta > \alpha$. If the actual value of the p is moderate, around 0.5, then $\frac{\alpha}{\alpha+\beta} = 1/2$, therefore, it is recommended to set $\beta = \alpha$, and if the the actual value of the p is large, greater than 0.5, then $\frac{\alpha}{\alpha+\beta} > 1/2$, therefore it is recommended to set $\beta < \alpha$.

It is mentioned that in Bayesian Statistics we use the Bayes rule to obtain posterior distribution. Here we introduce this rule in the following theorem.

Theorem 2.1. *(Bayes rule) Let $A_1, A_2, ...$ be a partition of the sample space, and let B be any set. Then for each $i = 1, 2, ...,$*

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{\sum_{j=1}^{\infty} p(B|A_j)p(A_j)}.$$

Now for a parameter θ, if we denote the prior distribution by $\pi(\theta)$ and the sampling distribution by $f(x|\theta)$, then the posterior distribution of θ for the given sample x is:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)},$$

where $m(x)$ is the marginal distribution of X, that is

$$m(x) = \int f(x|\theta)\pi(\theta)d\theta.$$

In this case, the parameter p is unknown, so we have two parameters k and p in the probability density function of X_i. Now to obtain the Bayes point estimation for this parameters, we have to first find the joint posterior pdf of K and P:

Lemma 2.2. *The joint posterior pdf of K and P given observations $\{x_1, x_2, \cdots, x_n\}$, is given by:*

(2.4)
$$h(k,p|x_1,...,x_n) = \tfrac{1}{C_1} p^{nk+\alpha-1}(1-p)^{T_n+\beta-1} \tfrac{\lambda^k}{k!} \prod_{i=1}^{n} \left(\begin{array}{c} x_i + k - 1 \\ x_i \end{array} \right).$$

where $T_n = \sum_{i=1}^{n} x_i$ and C_1 is normalizing constant:

$$C_1 = \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \frac{\Gamma(T_n+\beta)\Gamma(nk+\alpha)}{\Gamma(T_n+\beta+\alpha+nk)} \prod_{i=1}^{n} \left(\begin{array}{c} x_i + k - 1 \\ x_i \end{array} \right), \tag{2.5}$$

Proof. Using theorem 2.1 we have:

$$\begin{aligned} h(k,p|x_1,x_2,...,x_n) &= \frac{L(x_1,x_2,...,x_n|k,p)g(k)w(p)}{\int_0^1 \sum\limits_{k=1}^{\infty} L(x_1,x_2,...,x_n|k,p)g(k)w(p)dp} \\ &= \frac{1}{C_1} p^{\alpha+kn-1}(1-p)^{T_k+\beta-1} \frac{\lambda^k}{k!} \prod_{i=1}^{n} \left(\begin{array}{c} k + x_i - 1 \\ x_i \end{array} \right). \end{aligned}$$

□

Lemma 2.3. *The marginal posterior pdf of K given observations*

$$\{x_1, x_2, \cdots, x_n\},$$

is given by:

$$h_K(k|x_1,...,x_n) = \tfrac{1}{C_1} \tfrac{\lambda^k}{k!} \tfrac{\Gamma(nk+\alpha)\Gamma(T_n+\beta)}{\Gamma(nk+\alpha+T_n+\beta)} \prod_{i=1}^{n} \left(\begin{array}{c} x_i + k - 1 \\ x_i \end{array} \right), \tag{2.6}$$

where $k \geq 1$.

Proof. By taking the integral of $h(k,p|x_1,x_2,...,x_n)$ over p, the marginal density function of K is obtained:

$$h_K(k|x_1,x_2,...,x_n) = \int_0^1 h(k,p|x_1,x_2,...,x_n)dp$$
$$= \frac{1}{C_1}\frac{\lambda^k}{k!}\prod_{i=1}^{n}\left(\begin{array}{c} k+x_i-1 \\ x_i \end{array}\right)\frac{\Gamma(kn+\alpha)\Gamma(T_n+\beta)}{\Gamma(\alpha+\beta+T_n+kn)}.$$

□

Lemma 2.4. *The marginal posterior pdf of P given observations*

$$\{x_1,x_2,\cdots,x_n\},$$

is given by:

$$h_P(p|x_1,...,x_n) = \tfrac{C_2(p)}{C_1}p^{\alpha-1}(1-p)^{T_n+\beta-1}, \tag{2.7}$$

where $k \geq 1$, $p \in (0,1)$ *and* $C_2(p) = \sum\limits_{k=1}^{\infty}\frac{(p^n\lambda)^k}{k!}\prod\limits_{i=1}^{n}\left(\begin{array}{c} x_i+k-1 \\ x_i \end{array}\right)$.

Proof. The proof is similar to the proof of the Lemma 2.3, but here we have to replace the integral by sum. □

Since the likelihood function $L(x_1,x_2,\cdots,x_n|k,p)$ is less than one and, also the prior $g(k)$ and $w(p)$ are proper, so the denominator of $h(k,p|x_1,x_2,\ldots,x_n)$ which equals C_1, is less than 1. So the normalizing constant C_1 is convergent. Similarly, the convergence of $C_2(p)$ can be obtained. As a result C_1 and $C_2(p)$ can be approximated by the finite sums:

$$C_1 = \sum_{k=1}^{z} S(k), \tag{2.8}$$

and

$$C_2(p) = \sum_{k=1}^{q} Y(k), \tag{2.9}$$

where $S(k) = \frac{\lambda^k}{k!}\frac{\Gamma(nk+\alpha)\Gamma(T_n+\beta)}{\Gamma(nk+\alpha+T_n+\beta)}\prod\limits_{i=1}^{n}\binom{x_i+k-1}{x_i}$, $Y(k) = \frac{(p^n\lambda)^k}{k!}\prod\limits_{i=1}^{n}\binom{x_i+k-1}{x_i}$, z is the first integer that satisfies the inequality $|S(z+1)-S(z)| \leqslant \epsilon$ and q is the first integer that satisfies in the inequality $|Y(q+1)-Y(q)| \leqslant \epsilon$; where ϵ is a very small positive number, for instance 10^{-6}.

2.2. **Case 2: p is known:** Suppose that p is known. In this case the posterior pdf of K is given in the following Lemma:

Lemma 2.5. *The posterior pdf of K given observations $\{x_1, x_2, \cdots, x_n\}$ and p, is given by:*

$$h_K(k|x_1,...,x_n,p) = \frac{1}{C_2(p)}\frac{(p^n\lambda)^k}{k!}\prod_{i=1}^{n}\binom{x_i+k-1}{x_i}, \tag{2.10}$$

where $k \geq 1$ and $C_2(p)$ is the normalizing constant defined in Lemma 2.4.

Proof. Using Theorem 2.1, the posterior density function of K is obtained:

$$\begin{aligned} h_K(k|x_1,...,x_n,p) &= \frac{L(x_1,...,x_n|k,p)g(k)}{\sum\limits_{k=1}^{\infty} L(x_1,...,x_n|k,p)g(k)} \\ &= \frac{1}{C_2(p)}\frac{(p^n\lambda)^k}{k!}\prod_{i=1}^{n}\binom{x_i+k-1}{x_i}. \end{aligned}$$

□

2.3. **Case 2: p is known:** Suppose that p is known. In this case the posterior pdf of K is given in the following Lemma:

Lemma 2.6. *The posterior pdf of K given observations $\{x_1, x_2, \cdots, x_n\}$ and p, is given by:*

$$h_K(k|x_1,...,x_n,p) = \frac{1}{C_2(p)}\frac{(p^n\lambda)^k}{k!}\prod_{i=1}^{n}\binom{x_i+k-1}{x_i}, \tag{2.11}$$

where $k \geq 1$ and $C_2(p)$ is the normalizing constant defined in Lemma 2.4.

Proof. Using Theorem 2.1, the posterior density function of K is obtained:

$$\begin{aligned} h_K(k|x_1,...,x_n,p) &= \frac{L(x_1,...,x_n|k,p)g(k)}{\sum\limits_{k=1}^{\infty} L(x_1,...,x_n|k,p)g(k)} \\ &= \frac{1}{C_2(p)}\frac{(p^n\lambda)^k}{k!}\prod_{i=1}^{n}\binom{x_i+k-1}{x_i}. \end{aligned}$$

$\square$

3. Bayes Point Estimation of k

In this section, Bayes estimation for k are proposed under the squared error loss function, that is the mean of the posterior density function. Both cases, known and unknown p are considered, separately.

3.1. Case 1: p is Unknown.

Assume that p is an unknown parameter.

Lemma 3.1. *The Bayes point estimation ($\hat{k}_1$) of k is:*

$$\hat{k}_1 = \frac{\Gamma(T_n+\beta)}{C_1} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} \frac{\Gamma(nk+\alpha)}{\Gamma(T_n+\beta+\alpha+nk)} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i} . \tag{3.1}$$

Proof. As it is mentioned, the Bayes point estimation of k under the squared error loss function is the the mean of the marginal posterior pdf of k in (2.6). So we have:

$$\begin{aligned}\hat{k}_1 &= E_{h_K}(K)\\ &= \sum_{k=1}^{\infty} k h_K(k|x_1, ..., x_n)\\ &= \frac{\Gamma(T_n+\beta)}{C_1} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} \frac{\Gamma(nk+\alpha)}{\Gamma(T_n+\beta+\alpha+nk)} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i}.\end{aligned}$$

$\square$

Bayes point estimation of p is given in the following results:

Lemma 3.2. *The Bayes point estimation ($\hat{p}_1$) of p is:*

$$\hat{p}_1 = \frac{\Gamma(T_n+\beta)}{C_1} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \frac{\Gamma(nk+\alpha+1)}{\Gamma(T_n+\beta+\alpha+nk+1)} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i}. \tag{3.2}$$

Proof. Bayes point estimation of p under the squared error loss function is the mean of the marginal posterior pdf of p in (2.7). So we have:

$$\begin{aligned}\hat{p}_1 &= E_{h_P}(P)\\ &= \int_0^1 p h_p(p|x_1, ..., x_n) dp\\ &= \frac{\Gamma(T_n+\beta)}{C_1} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \frac{\Gamma(nk+\alpha+1)}{\Gamma(T_n+\beta+\alpha+nk+1)} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i}.\end{aligned}$$

$\square$

3.2. **Case 2: p is known:** Assume that p is a known parameter.

Lemma 3.3. *The Bayes point estimation ($\hat{k}_2$) of k is:*

$$\hat{k}_2 = \frac{1}{C_2(p)} \sum_{k=1}^{\infty} k \frac{(p^n\lambda)^k}{k!} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i}. \tag{3.3}$$

Proof. Bayes point estimation of k under the squared error loss function is the mean of the marginal posterior pdf of k in (2.10). So we have:

$$\begin{aligned}\hat{k}_2 &= E_K(k)\\ &= \sum_{k=1}^{\infty} kK(k|x_1, ..., x_n, p)\\ &= \frac{1}{C_2(p)} \sum_{k=1}^{\infty} k \frac{(p^n\lambda)^k}{k!} \prod_{i=1}^{n} \binom{x_i+k-1}{x_i}.\end{aligned}$$

□

It should be mentioned that the proposed Bayes estimations $\hat{k}_1$, $\hat{k}_2$ and $\hat{p}_1$ are finite because the priors of k and p are proper with the finite mean, and the likelihood is bounded by one. Wald [7] showed that, Bayesian posterior mean estimations that arise from proper priors, are always admissible. So $\hat{k}_1$, $\hat{k}_2$ and $\hat{p}_1$ are admissible estimators.

4. Empirical Bayes Estimation of k

In this section, a procedure for constructing empirical Bayes estimations for k is provided in both cases: known and unknown p, separately. MME technique is used to estimate the hyper parameters λ, α and β.

4.1. **Case 1: p is unknown.** By evaluating

$$\int_0^1 \sum_{k=1}^{\infty} f(x|k,p)g(k)w(p)dp,$$

where $f(x|k,p)$, $g(k)$ and $w(p)$ are defined in (1.1), (2.2) and (2.3) respectively, the marginal pdf of X, is:

$$\begin{aligned}f_X(x) &= \int_0^1 \sum_{k=1}^{\infty} f(x|k,p)w(p)g(k)dp\\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \sum_{k=1}^{\infty} \binom{x+k-1}{x} \times \frac{e^{-\lambda}\lambda^k}{k!(1-e^{-\lambda})} \frac{\Gamma(\alpha+k)\Gamma(\beta+x)}{\Gamma(x+\beta+\alpha+k)}.\end{aligned}$$

It is difficult to calculate the moments by using the pdf of X, so in the following Lemma we obtain it:

Lemma 4.1. *First, second and third moment of X are:*

$$
\begin{aligned}
E(X) &= \frac{\lambda}{(1-e^{-\lambda})}\frac{\beta}{(\alpha-1)},\\
E(X^2) &= \frac{\lambda}{(1-e^{-\lambda})}\frac{\beta}{(\alpha-1)}[1+(2+\lambda)\frac{(\beta+1)}{(\alpha-2)}],\\
E(X^3) &= \frac{\lambda}{(1-e^{-\lambda})}\frac{\beta}{(\alpha-1)}[1+(6+3\lambda)\frac{(\beta+1)}{(\alpha-2)}\\
&\quad +(\lambda^2+6\lambda+6)\frac{(\beta+2)(\beta+1)}{(\alpha-3)(\alpha-2)}].
\end{aligned} \tag{4.1}
$$

Proof. Using the well-known identity $E_X(X) = E_{K,P}[E(X|k,p)]$, first, second, and third moment of X are obtained. □

In order to construct the empirical Bayes estimation for k, it is necessary to estimate the hyper parameters λ, α and β. The MME values for λ, α and β can be obtained by solving the following system of nonlinear equations:

$$
\begin{aligned}
E(X) &= \frac{\sum_{i=1}^{n} X_i}{n},\\
E(X^2) &= \frac{\sum_{i=1}^{n} X_i^2}{n},\\
E(X^3) &= \frac{\sum_{i=1}^{n} X_i^3}{n}.
\end{aligned} \tag{4.2}
$$

The empirical Bayes estimation $\hat{k}_1(emp)$ is same as Bayes estimation $\hat{k}_1$ in (3.1), but here, hyper parameters λ, α and β are replaced by MME values.

4.2. **Case 2: p is known:** By evaluating $\sum_{k=1}^{\infty} f(x|k,p)g(k)$, where $f(x|k,p)$ and $g(k)$ are defined in (1.1) and (2.2) respectively, the marginal pdf of X is:

$$
f_X(x) = \sum_{k=1}^{\infty} \binom{x+k-1}{x} p^k (1-p)^x \frac{e^{-\lambda}\lambda^k}{k!(1-e^{-\lambda})}. \tag{4.3}
$$

In this case in order to obtain the empirical bayes estimation of k, we have to estimate the hyper parameter λ, so the MME of λ is obtained in the following lemma.

Lemma 4.2. *The MME of λ is:*

$$\hat{\lambda} = \frac{pT_n + (1-p)n\Omega(\frac{-pT_n e^{-\frac{pT_n}{(1-p)n}}}{(1-p)n})}{(1-p)n}, \tag{4.4}$$

where $T_n = \sum_{i=1}^{n} X_i$ and $\Omega(A)$ is the product log function A, which is defined as the value of x that satisfies the equation $xe^x = A$.

Proof. Using the well-known identity $E_X(X) = E_{K,P}[E(X|k,p)]$, first moment of X is obtained and the MME of λ is obtained by solving the following equation:

$$\frac{(1-p)}{p}\frac{\lambda}{1-e^{-\lambda}} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

□

The empirical Bayes point estimation $\hat{k_2}(emp)$ is $\hat{k_2}$, defined in (3.3) with $\hat{\lambda}$ instead of λ.
It should be mentioned that, since k is an integer, the nearest integer for the proposed estimations $\hat{k_1}$, $\hat{k_1}(emp)$, $\hat{k_2}$, and $\hat{k_2}(emp)$ is taken as the estimator of k.

5. **Illustrative Example**

Example 5.1. Let $\{23, 29, 47, 25, 54\}$ be a random sample of size $n = 5$ that was generated using Mathematica7 from $NB(20, 0.4)$. Suppose that $\alpha = \beta = 1$ and λ is assumed to be 23. The marginal posterior pdf of k given the sample $\{23, 29, 47, 25, 54\}$, is given by:

$$h_K(k|23, 29, 47, 25, 54)$$

$$= \frac{1}{C_1}\frac{\lambda^k}{k!}\frac{\Gamma(5k+1)\Gamma(179)}{\Gamma(5k+180)}\binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47}\binom{k+24}{25}\binom{k+53}{54},$$

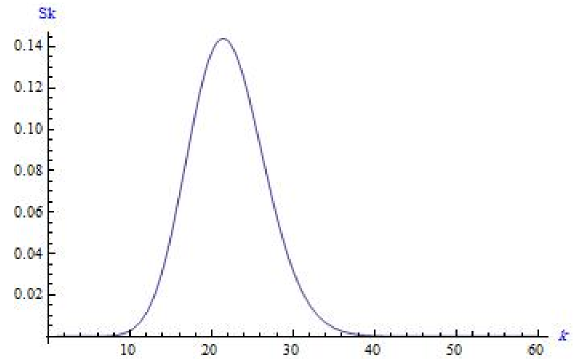

FIGURE 1. Plot of S(k) in 2.8 for the Example 5.1

where

$$C_1 = \sum_{k=1}^{\infty} S(k) \backsimeq \sum_{k=1}^{z} S(k) = \sum_{k=1}^{z} \frac{23^k}{k!} \frac{\Gamma(5k+1)\Gamma(179)}{\Gamma(5k+180)} \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47} \binom{k+24}{25}\binom{k+53}{54} = \sum_{k=1}^{40} \frac{\lambda^k}{k!} \frac{\Gamma(5k+1)\Gamma(179)}{\Gamma(5k+180)} \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47}\binom{k+24}{25}\binom{k+53}{54}$$

$$= 1.66.$$

Hence, the marginal posterior pdf of K is:

$$h_K(k|23, 29, 47, 25, 54) = \frac{1}{1.66} \frac{23^k}{n!} \frac{\Gamma(5k+1)\Gamma(179)}{\Gamma(5k+180)} \times \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47} \binom{k+24}{25}\binom{k+53}{54},$$

so when p is unknown, the Bayes estimation of k is given by the following sum:

$$\hat{k}_1 \backsimeq \frac{1}{1.66} \sum_{k=1}^{40} k \frac{23^k}{k!} \frac{\Gamma(5k+1)\Gamma(179)}{\Gamma(5k+180)} \times \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47} \binom{k+24}{25}\binom{k+53}{54} = 22.0264.$$

In order to choose z for notice after equation 2.8 and 2.9 had expressed, we design the plot of S(k) in figure 1 and we see after point 40 plot is flat, so we choose 40 that satisfy $|S(41) - S(40)| \leqslant 10^{-6}$. Also for calculating $\hat{k}_1$ for notice Figure 2 that shows the values of $h_K(k)$ when p is unknown we see that after point 40 this values are almost 0 so in calculating $\hat{k}_1$ we choose the point 40. The posterior distribution for k is given in the table 1.

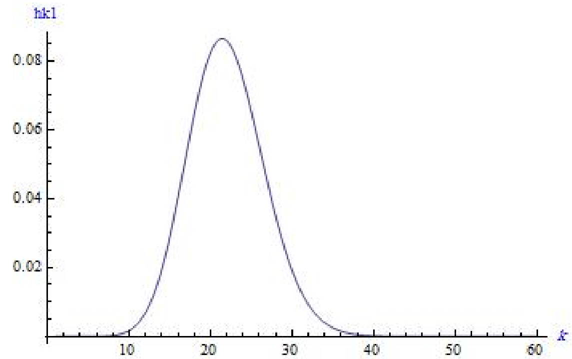

FIGURE 2. plot of posterior pdf $h_K(k)$ when p is unknown for Example 5.1

k	1	$\cdots$	6	$\cdots$	10	11	12	13
$h_K(k)$	$7.72e-11$	$\cdots$	0.000021	$\cdots$	0.0017	0.0035	0.0068	0.012
k	14	15	16	17	18	19	20	21
$h_K(k)$	0.02	0.03	0.04	0.053	0.065	0.075	0.082	0.086
k	22	23	24	25	26	27	28	29
$h_K(k)$	0.0859	0.082	0.075	0.065	0.053	0.04	0.03	0.02
k	30	$\cdots$	35	40	45	50	60	>60
$h_K(k)$	0.0068	$\cdots$	0.0035	0.0017	0.000021	$7.72e-11$	$7.02e-11$	0.0000

TABLE 1. Values of marginal posterior of example 1 when p is unknown

When p is known, the marginal posterior pdf of K given sample $\{23, 29, 47, 25, 54\}$, is given by:

$$h_K(k|23, 29, 47, 25, 54) = \frac{1}{C_2(p)} \frac{(p^5\lambda)^k}{k!} \binom{k+22}{23} \binom{k+28}{29} \binom{k+46}{47} \binom{k+24}{25} \binom{k+53}{54},$$

where

$$C_2(p) = \sum_{k=1}^{\infty} Y(k) \backsimeq \sum_{k=1}^{q} Y(k) = \sum_{k=1}^{q} \frac{(23\times 0.4^5)^k}{k!} \binom{k+22}{23} \binom{k+28}{29}$$
$$\binom{k+46}{47} \binom{k+24}{25} \binom{k+53}{54} = \sum_{k=1}^{40} \frac{(23\times 0.4^5)^k}{k!} \binom{k+22}{23} \binom{k+28}{29} \binom{k+46}{47} \binom{k+24}{25} \binom{k+53}{54}$$

$$= 3.4292 \times 10^{40}.$$

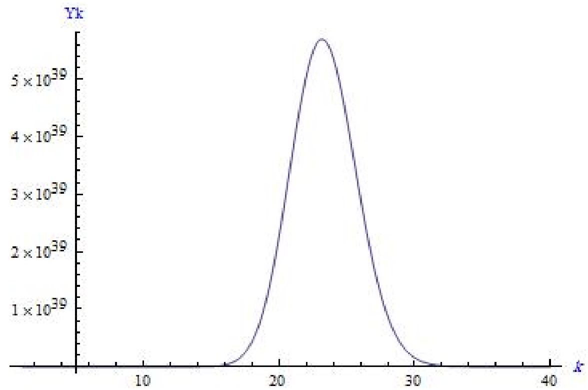

FIGURE 3. Plot of Y(k) in 2.9 for the Example 5.1

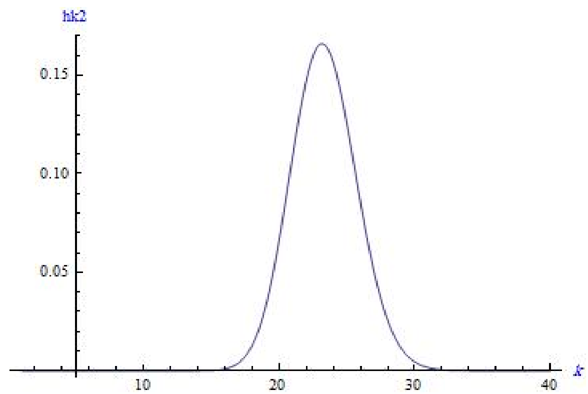

FIGURE 4. plot of posterior pdf $h_K(k)$ when p is known for Example 5.1

So when p is known, the Bayes estimator of k ($\hat{k}_2$) is given by the following sum:

$$\hat{k}_2 = \frac{1}{3.4292\times 10^{40}} \sum_{k=1}^{40} k \frac{(23\times 0.4^5)^k}{k!} \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47}\binom{k+24}{25}\binom{k+53}{54}$$

$$= 23.32.$$

Note that choosing q is similar to choosing z in case of unknown p. The values of posterior distribution for k, when p is known are given in the table 2.

k	1	$\cdots$	6	$\cdots$	10	11	12
$h_K(k)$	$6.86e-42$	$\cdots$	$3.5e-19$	$\cdots$	$1.18e-10$	$6.38e-9$	$1.43e-7$
k	13	14	15	16	17	18	19
$h_K(k)$	$2.16e-6$	0.000023	0.00017	0.00094	0.0039	0.01279	0.03254
k	20	21	22	23	24	25	26
$h_K(k)$	0.06614	0.1089	0.1474	0.1656	0.1563	0.1250	0.08549
k	27	28	29	30	31	32	33
$h_K(k)$	0.05038	0.02577	0.01152	0.0045	0.00157	0.00048	0.000135
k	34	35	$\cdots$	40	50	60	>60
$h_K(k)$	0.000033	$7.47e-6$	$1.0092e-9$	$\cdots$	$4.89e-20$	$5.058e-33$	0.0000

TABLE 2. Values of marginal posterior of example 1 when p is known

When p is unknown, the empirical Bayes estimation $\hat{k}_1(emp)$ can't be calculated, because solving the system of nonlinear equation in (4.3) gives bad estimation for the hyperparameters α, β and λ, so we omitted this estimator.
When p is known, as we said, the empirical Bayes estimation $\hat{k}_2(emp)$ is obtained by replacing $\hat{\lambda}$ in (4.5) with λ in $\hat{k}_2$; that is, $\hat{\lambda} = 23.73$. So:

$$\hat{k}_2(emp) = \frac{1}{7.12718\times 10^{40}} \sum_{k=1}^{40} k \frac{(23.73\times 0.4^5)^k}{k!} \binom{k+22}{23}\binom{k+28}{29}\binom{k+46}{47}\binom{k+24}{25}\binom{k+53}{54}$$

$$= 23.505.$$

6. Simulation Study

In this section, the performance of the proposed estimators of k is investigated through a simulation study. The simulation study is carried out for different values of the combinations (n, k, p). In this study, in the case of known p, we suggested the $\hat{\lambda} = \frac{\bar{X}p}{1-p}$ to calculating the estimator $\hat{k}_2$, that we choose this value for hyperparameter λ, since if $100(1-p)$ percent of the number of trials is failure, so $100p$ leaving of trials are the number of success(k). Hence by a simple proportion we have $\hat{\lambda} = \frac{\bar{X}p}{1-p}$ and by using this value for λ we obtained good estimator similar of that in empirical estimator. It has to mentioned that, in this proportion we replace the k by λ, because in prior distribution its mean is equal to λ approximately. In these cases we have generated 500 random

k	n	θ		
		0.2	0.5	0.8
5	10	5(0.6)	5(1.01)	5(2.62)
	20	5(0.32)	5(.43)	5(1.37)
	30	5(.197)	5(0.33)	5(0.85)
20	10	20(2.32)	20(3.97)	20(10.013)
	20	20(1.34)	20(1.8)	20(4.92)
	30	20(0.75)	20(1.25)	20(3.28)
50	10	48(4.61)	46(18.43)	47(16.32)
	20	49(2.30)	48(8.8)	47(9.8)
	30	49(1.56)	48(6.67)	48(5.73)

TABLE 3. expected and MSE valuses of the stimator $\hat{k}_2(emp)$ for k when p is known

k	n	θ		
		0.2	0.5	0.8
5	10	5(0.63)	5(1.00)	5(2.23)
	20	5(0.23)	5(.45)	5(1.14)
	30	5(0.2)	5(0.27)	5(0.85)
20	10	20(2.48)	20(3.89)	20(9.68)
	20	20(1.25)	20(1.66)	20(5.27)
	30	20(0.85)	20(1.37)	20(3.26)
50	10	48(4.43)	48(7.2)	47(16.47)
	20	49(2.6)	48(3.87)	47(9.1)
	30	49(1.57)	49(2.66)	48(6.2)

TABLE 4. expected and MSE valuses of the estimator $\hat{k}_2$ for k when p is known

sample of size $n(= 10, 20\, and\, 30)$ from a negative binomial distribution with parameters $p(= 0.2, 0.5\, and\, 0.8)$ and $k(= 5, 20\, and\, 50)$ by using the mathematica 7. In order to see how the estimators of k perform with respect to sample size, the average value of estimators $\hat{k}_2(emp)$ and $\hat{k}_2$ along with MSE in parentheses are reported in Table 3 and Table 4, respectively. From tables 3 and 4, we see that when θ increases, the MSE also increases. For each k, when the size of sample is increased, then the MSE decreases.

Acknowledgments

The authors thank the referee for his (her) helpful comments.

References

[1] K. Adamids, *An EM Algoritm for estimation negative binomial parameters*, Australian New zealand Journal of Statistics, **41**(2) (1999), 213–221.
[2] N. Breslow, *Extra-Poisson variation in log-linear models*, Applied Statistics, **3**3 (1984), 38–44.
[3] S. J. Clark and J. N. Perry, *Estimation of the negative binomialparameter k by maximum quasi-likelihood*, Biometrics,**45** (1989), 309–316.
[4] G. Casella and R. L. Berger, *Statistical Inference*, Duxbury Thomson Learning Publishing, (2001).
[5] J. Engel, *Models for respons data showing extra-poisson variation*, Statistical Neerlandica, **38** (1984), 159–167.
[6] M. Ganji, *Bayes estimation of parameters of generalized negative binomial distribution and it's applications*, Iranian Journal of Science and Technology (in persion), **3,4** (2005), 47–53.
[7] J. F. Lawless, *Negative Binomial and mixed poisson regression*, Canadian Jornal of Statistics, **15** (1987), 209–225.
[8] K. G. Manton, M. A. Woodbury, and E. Stallard, *A variance components approch to categorical data models with heterogeneos cell populations: Analysis of spatial gradiants in lung canser mortality rates in North Carolina counties.*, Biometrics, **37** (1981), 259–269.
[9] W. W. Piegorsch, *Maximum likelihood estimation for the negative binomial dispersion parameter k*, Biometrics, **46** (1990), 863–867.
[10] A. Wald, *Statistical Decision Functions.* Wiley , New York, (1950).

Masoud Ganji
Department of Statistics, Facualty of Mathematical Sciences, University of Mohaghegh Ardabili, 56199-11367, Ardabil, Iran.
Email: mganji@uma.ac.ir

Nasrin Eghbali
Department of Mathematics and Applications, Facualty of Mathematical Sciences, University of Mohaghegh Ardabili, 56199-11367, Ardabil, Iran.
Email: nasrineghbali@gmail.com;eghbali@uma.ac.ir

Masoud Azimian
Department of Statistics, Facualty of Mathematical Sciences, University of Mohaghegh Ardabili, 56199-11367, Ardabil, Iran.
Email: masoudazimian65@gmail.com