ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 1

CONJUGATE BAYESIAN ANALYSIS OF THE NEGATIVE BINOMIAL DISTRIBUTION

by

I. M. Morgan and J. C. Hickman

1. INTRODUCTION

The negative binomial distribution has frequently been suggested as a reasonable model for the number of insurance claims in a fixed time period. The suggestion has been motivated by several lines of reasoning: (a) With its two parameters, the negative binomial provides greater flexibility than the traditional Poisson. (b) The claims generating process may have contagion. (c) The claims generating process may be a gamma mixture of Poisson distributions because of imperfect classification of risks.

In any case, there is a large actuarial literature on the applications of the negative binomial distribution in insurance. See, for example, Ammeter (1948), Carlson (1962), Hewitt (1960), and Simon (1960).

The purpose of this paper is to perform conjugate Bayesian analysis on the negative binomial distribution with the goal of providing a standard by which approximate credibility formulas and approximate predictive distribution may be judged. The paper follows the outline of Daboni (1974).

2. THE LIKELIHOOD FUNCTION

The symbol N_i denotes the number of claims in period (year) i. We will assume that, given the parameters γ and p, the random variables N_i $(i=1,2,\ldots)$ are independent and identically distributed in accordance with the negative binomial distribution. That is

$$p_{N_i|\gamma,p}(n_i) = \Gamma(n_i+\gamma)p^{\gamma}(1-p)^{n_i}/\Gamma(\gamma)n_i! \ ; \ \gamma>0, \ 0<p<1$$

$$n_i = 0,1,2,\ldots \tag{2.1}$$

Because it is often of interest to make inferences about the mean, λ, let us reparameterize (2.1) to show λ in our likelihood function.

$$\text{Let } \lambda=\gamma(1-p)/p\ , \qquad 0<\lambda<\infty$$

$$\text{and } \gamma=\gamma \qquad 0<\gamma<\infty. \tag{2.2}$$

With this reparameterization, the probability mass function (2.1) becomes

$$p_{N_i|_{\gamma,\lambda}}(n_i)=\Gamma(n_i+\gamma)\gamma^{\gamma}\lambda^{n_i}(\gamma+\lambda)^{-(n_i+\gamma)}/n_i!\Gamma(\gamma)\ ;\ \gamma,\lambda>0,\quad n_i=0,1,2,\ldots \tag{2.3}$$

We have

$$\begin{aligned}
&E(N_i|\gamma,\lambda)=\lambda\\
&Var(N_i|\gamma,\lambda)=\lambda(1+\lambda/\gamma)\\
&\alpha_3(N_i|\gamma,\lambda)=(\gamma+2\lambda)/\sqrt{\gamma\lambda(\gamma+\lambda)}\\
&\alpha_4(N_i|\gamma,\lambda)=3+(1/\lambda)+(5/\gamma)+\lambda/\gamma(\gamma+\lambda)\\
&E_{N_i}|_{\gamma,\lambda}(e^{tN_i})=\{1-(\lambda/\gamma)(e^t-1)\}^{-\gamma}.
\end{aligned} \tag{2.4}$$

We observe that as $\gamma\to 0$, the negative binomial approaches the Poisson distribution.

3. THE PRIOR DISTRIBUTION

We will develop a proper prior distribution through conjugate analysis. In this analysis the parameters λ and γ are not independent a priori. A conjugate prior distribution for (2.3) will be constructed in a three step process.

The three steps will be denoted by (i), (ii) and (iii). If the parameter γ is known, the definition of the conjugate prior, and thereby the posterior distribution, can start with step (ii). This much simpler development is done by Morgan (1983). The prior and posterior distributions of p are members of the Beta family of distributions and the predictive distribution, that is, the unconditional distribution of N_i, is a member of the Beta-Binomial family of distributions.

(i) Let γ be distributed such that:

$$\xi_\gamma(\gamma) \propto p_{n_0}(\gamma) \;/\; p_{\phi_0}(\gamma\delta_0+m_0) \quad ; \quad \gamma, \delta_0 > 0 \tag{3.1}$$

where

$$p_n(\gamma) = \gamma(\gamma+1)\ldots(\gamma+n-1),$$

a polynomial of degree n in γ, n_0 and m_0 are non-negative integers and $\phi_0 = n_0+2, n_0+3, \ldots$. Notice that $\zeta_\gamma(\gamma)$ is a proper prior as a result of the restriction on ϕ_0.

The parameter ϕ_0 may be interpreted as the number of past (hypothetical) claims on which the prior is based. The parameter δ_0 may be interpreted as the number of past (hypothetical) observation periods on which the prior is based. The parameter n_0 can be set to any non-negative integer. Zero would be the simplest choice. The difference in degrees of the two polynomials, ϕ_0-n_0, determines the behavior of the prior in the right-hand tail. The parameter m_0 helps determine the behavior of the prior near zero. If one believes the negative binomial model is correct (the claims generating process has a mean less than the variance), then m_0 will be a positive integer.

The determination of the normalizing constant, denoted by C^{-1}, of the probability mass function (3.1) is a prototype for the evaluation of many integrals that occur in the sequel. Since $p_{n_0}(\gamma)$ is a polynomial in γ of degree n_0 and $p_{\phi_0}(\gamma\delta_0+m_0)$ is a polynomial in γ degree ϕ_0, then $p_{n_0}(\gamma)/p_{\phi_0}(\gamma\delta_0+m_0)$ is of the form of a quotient of two polynomials in γ, the numerator is of degree n_0 and the denominator is of degree ϕ_0. Since $\phi_0 \geq n_0+2$, we have a proper fraction of two polynomials, i.e., a rational function that can be expressed in terms of partial fractions. Notice that if δ_0 is such that we get some cancellations, the following technique is still applicable, the effect will be on the number and the value of the partial fraction constants. We will assume the general case. This is δ_0 will not cause any cancellations, and proceed as follows:

$$p_{n_0}(\gamma)/p_{\phi_0}(\gamma\delta_0+m_0)$$

can be written on the form

$$\sum_{i=1}^{\phi_0} a_i/(\gamma\delta_0+m_0-1+i),$$

where a_i's are the solution of

$$\prod_{s=1}^{n_0}(\gamma+s-1) = \sum_{i=1}^{\phi_0} a_i \prod_{i\neq r=1}^{\phi_0}(\gamma\delta_0+m_0-1+r).$$

This implies that

$$a_i = \prod_{s=1}^{n_0}(1+s\delta_0-\delta_0-m_0-i)/\delta_0^{n_0} \prod_{i\neq r=1}^{\phi_0}(r-i);\ i=1,2,\ldots,\phi_0 \tag{3.2}$$

and

$$\sum_{i=1}^{\phi_0} a_i = 0.$$

Using this partial fractions decomposition we can find the anti-derivative term by term to obtain the normalizing constant in (3.1) as follows:

$$C = \int_0^\infty [p_{n_0}(\gamma)\ /\ p_{\phi_0}(\gamma\delta_0+m_0)]d\gamma$$

$$= \int_0^\infty \left[\sum_{i=1}^{\phi_0} a_i\ /\ (\gamma\delta_0+m_0-1+i)\right] d\gamma$$

$$= \sum_{i=1}^{\phi_0} (a_i/\delta_0)\ \ln(\gamma\delta_0+m_0-1+i)\Big|_0^{\infty},$$

since $\sum_{i=1}^{\phi_0} a_i = 0$, therefore $\prod_{i=1}^{\phi_0} (\gamma\delta_0+m_0+i-1)^{a_i} \to 1$, as $\gamma \to \infty$, which implies that

$\sum_{i=1}^{\phi_0} (a_i/\delta_0)\ \ln(\gamma\delta_0+m_0+i-1) \to 0$, as $\gamma \to \infty$. This will enable us to

write $C = -\sum_{i=1}^{\phi_0} (a_i/\delta_0)\ \ln(m_0+i-1)$.

To evaluate $E_\gamma\{\gamma^k\ \Gamma(\gamma\delta_0+m_0-k)\ /\ \Gamma(\gamma\delta_0+m_0)\}$, which will be needed in the next development to find $E_\lambda(\gamma^k)$, we return to (3.1). The numerator of the function to be integrated is

$\gamma^k \prod_{s=1}^{n_0} (\gamma-1+s)$, a polynomial in γ of degree n_0+k. The denominator is

$\prod_{r=1}^{\phi_0+k} (\gamma\delta_0+m_0-k-1+r)$, a polynomial in γ of degree ϕ_0+k. The integrand

$\gamma^k \prod_{s=1}^{n_0} (\gamma-1+s) / \prod_{r=1}^{\phi_0+k} (\gamma\delta_0+m_0-k-1+r)$ is a quotient of two polynomials in γ, the denominator's degree exceeds the numerator's degree by at least two, and the denominator is greater than zero for $\gamma \geq 0$, as long as $k \leq m_0$. To evaluate the integral of this quotient we can use the fact that it is a rational function which can be expressed in terms of partial fractions and find the anti-derivative term by term.

The rational function

$$\gamma^k \prod_{s=1}^{n_0} (\gamma-1+s) \Big/ \prod_{r=1}^{\phi_0+k} (\gamma\delta_0+m-k-1+r)$$

can be written, using the partial fractions decomposition, as

$$\sum_{i=1}^{\phi_0+k} b_i(k) \,/\, (\gamma\delta_0+m_0-k-1+i)$$

where the $b_i(k)$'s satisfy

$$\gamma^k \prod_{s=1}^{n_0} (\gamma-1+s) \equiv \sum_{i=1}^{\phi_0+k} b_i(k) \prod_{i\neq r=1}^{\phi_0+k} (\gamma\delta_0+m_0-k-1+r).$$

This yields

$$b_i(k) = (k+1-m_0-i)^k \prod_{s=1}^{n_0} \{1+k+s\delta_0-\delta_0-m_0-i)/\delta_0^{n_0+k} \prod_{i\neq r=1}^{\phi_0+k} (r-i)$$

and

$$\sum_{i=1}^{\phi_0+k} b_i(k) = 0.$$

We can prove the following relationships linking the partial fraction decomposition of the last two rational functions:

$$b_{k+1-m_0}(k) = 0 \;;\; k=1,2,\ldots\ldots$$

$$b_{i+1}(1) = (1-m_0-i)a_i \ / \ (-i)\delta_0 \ ; \ i=1,2,\ldots,\phi_0$$

$$b_1(1) = \sum_{i=1}^{\phi_0} (1-m_0-i)a_i/i\delta_0 \tag{3.3}$$

$$b_{i+2}(2) = (1-m_0-i)^2 a_i \ / \ i(i+1)\delta_0^2 \ ; \ i=1,2,\ldots,\phi_0$$

$$b_1(2)+b_2(2) = \sum_{i=1}^{\phi_0} (1-m_0-i)^2 a_i \ / \ i(i+1)\delta_0.$$

These relationships will enable us to write

$$E_\gamma\{\gamma^k\Gamma(\gamma\delta_0+m_0-k) \ / \ \Gamma(\gamma\delta_0+m_0)\}$$

$$= C^{-1} \int_0^\infty \gamma^k\{\Gamma(\gamma_0+m_0-k)/\Gamma(\gamma\delta_0+m_0)\} \ \{p_{n_0}(\gamma) \ / \ p_{\phi_0}(\gamma\delta_0+m_0)\} \ d\gamma.$$

$$= \left\{\sum_{i=1}^{\phi_0+k} b_i(k) \ \ln(m_0-k-1+i)\right\} \Big/ \left\{\sum_{i=1}^{\phi_0} a_i \ \ln(m_0-1+i)\right\}. \tag{3.4}$$

(ii) Let $(\lambda|\gamma)$ be distributed such that $\{\gamma/(\gamma+\lambda)\}$ is Beta with parameters $\{\alpha_0(\gamma),\beta_0(\gamma)\}$. If $\alpha_0(\gamma) = \gamma\delta_0+m_0$, $\beta_0(\gamma) = \phi_0$ and $C_1 = \Gamma(\phi_0) \ / \ \gamma^{\gamma\delta_0+m_0} \ p_{\phi_0}(\gamma\delta_0+m_0)$, then

$$\xi_{\lambda|\gamma}(\gamma) = (C_1)^{-1} \ \{\lambda^{\phi_0-1} \ / \ (\gamma+\lambda)^{(\gamma\delta_0+\phi_0+m_0)}\} \tag{3.5}$$

$$\gamma,\lambda,\delta_0>0, \ \phi_0 = n_0+2, \ n_0+3, \ \ldots, \ m_0=1,2,\ldots$$

This implies that the conditional moments of $(\lambda|\gamma)$ can be written in the form

$$\xi_{\lambda|\gamma}(\lambda^k) = (C_1)^{-1} \int_0^\infty \lambda^{k+\phi_0-1} \,/\, (\gamma+\lambda)^{(\gamma\delta_0+\phi_0+m_0)} d\lambda$$

$$= \gamma^k \, \Gamma(\phi_0+k)\Gamma(\gamma\delta_0+m_0-k) \,/\, \Gamma(\phi_0)\Gamma(\gamma\delta_0+m_0). \tag{3.6}$$

Notice that for the small values of γ or δ_0, $E(\lambda^k|\gamma)$ exists only if $k \le m_0$.

Now using (3.3) and (3.5), we can determine the marginal moments of λ, $E_\lambda(\lambda^k)$, as follows:

$$E_\lambda(\lambda^k) = E_\gamma \, E_{\lambda|\gamma}(\lambda^k)$$

$$= E_\gamma\{\lambda^k\Gamma(\phi_0+k)\Gamma(\gamma\delta_0+m_0-k) \,/\, \Gamma(\phi_0)\Gamma(\gamma\delta_0+m_0)\}$$

$$= \{\Gamma(\phi_0+k)/\Gamma(\phi_0)\} \left\{ \sum_{i=1}^{\phi_0+k} b_i(k) \ln(m_0-k-1+i) \right\} \Big/ \left\{ \sum_{i=1}^{\phi_0} a_i \ln(m_0-1+i) \right\}. \tag{3.7}$$

That is:

(1) The mean, $k=1$, is

$$E_\lambda(\lambda) = (\phi_0) \left\{ \sum_{i=1}^{\phi_0+1} b_i(1) \ln(m_0-2+i) \right\} \Big/ \left\{ \sum_{i=1}^{\phi_0} a_i \ln(m_0-1+i) \right\}$$

and using (3.3) we have

$$E_\lambda(\lambda) = (\phi_0/\delta_0) \left[\sum_{i=1}^{\phi_0} a_i\{(i+m_0-1)/i\} \ln\{(i+m_0-1)/(m_0-1)\} \right] \Big/ \left\{ \sum_{i=1}^{\phi_0} a_i \ln(i+m_0-1) \right\}.$$

The special case $m_0 = 1$ yields

$$E_\lambda(\lambda) = \phi_0 \,/\, \delta_0. \tag{3.8}$$

(2) The second moment, k=2, is

$$E_\lambda(\lambda^2) = \phi_0(\phi_0+1) \left\{ \sum_{i=1}^{\phi_0+2} b_i(2)\ln(m_0-3+i) \right\} / \left\{ \sum_{i=1}^{\phi_0} a_i\ln(m_0-1+i) \right\}.$$

This means that if $m_0=1$,

$E_\lambda(\lambda^2)$, and the variance, will not exist.

(iii) Items (i) and (ii) imply that the joint prior distribution of γ and λ, the product of (3.1) and (3.5), is

$$\xi_{\gamma,\lambda}(\gamma,\lambda) = p_{n_0}(\gamma)\ \gamma^{\gamma\delta_0+m_0}\ \lambda^{\phi_0-1}(\gamma+\lambda)^{-(\gamma\delta_0+\phi_0+m_0)}/C\Gamma(\phi_0); \tag{3.9}$$

$\gamma,\lambda,\delta_0>0,\ n_0=0,1,2,\ldots,m_0=1,2,\ldots,\phi_0=n_0+2,n_0+3,\ldots.$

Notice that if we put $n_0=0$ and $m_0=1$ in this model we can get Daboni's (1974) model and results. If $m_0 \geq 2$ the mean formula will not be as simple as in Daboni's paper, but we can determine higher moments if the parameter m_0 is sufficiently large.

4. POSTERIOR DISTRIBUTIONS:

We collect t observations $\underline{n}_t = (n_1,n_2,\ldots,n_t)$ whose probability distributions is given by (2.3), which depend on the values of the two parameters (γ,λ). The posterior distribution of the parameters (γ,λ) is proportional to the product of the likelihood (2.3) and the prior (3.7). We have

$$\Xi_{\gamma,\lambda|\underline{n}_t}(\gamma,\lambda) \propto \left\{ \prod_{i=1}^{t} p_{n_i}(\gamma) \right\} \gamma^{\gamma\delta_t+m_0}\lambda^{\phi_t-1}(\gamma+\lambda)^{-(\gamma\delta_t+\phi_t+m_0)} \tag{4.1}$$

where

$$\delta_t = \delta_0 + t \quad \text{and} \quad \phi_t = \phi_0 + n = \phi_0 + \sum_1^t n_i .$$

Notice that $\xi(\lambda,\gamma|\underset{\sim}{n}_t)$ belongs to the same family of distributions as the prior. This was expected, because we selected our prior from the natural conjugate of the likelihood. Then we can assert that the posterior distributions and the posterior moments will be of the same form as the prior distributions and moments, we only need to use the modified updated parameters δ_t and ϕ_t. We follow the outline of Section 3.

(i) The parameter γ, posterior to the data $\underset{\sim}{n}_t$, will be distributed such that

$$\xi_{\gamma,|\underset{\sim}{n}_t}(\gamma) = \prod_{i=1}^{t} p_{n_i}(\gamma) \;/\; p_{\phi_t}(\gamma\delta_t+m_0)$$

$$n_0 = 0,1,\ldots,\ m_0=1,2,\ldots,\ \phi_0=n_0+2,n_0+3,\ldots$$

The normalizing constant, C'^{-1}, will take the form

$$C' = -\sum_{i=1}^{\phi_t} (a_i'/\delta_y)\ \ln(m_0+i-1)$$

where

$$a_i' = \prod_{j=0}^{t} \prod_{s=1}^{n_j} (1+s\delta_t-\delta_t-m_0-i)/\delta_t^{n'} \prod_{i\neq r=1}^{\phi_t} (r-i);\ n'=n+n_0,\ i=1,2,\ldots,\phi_t \tag{4.2}$$

and

$$\sum_{i=1}^{\phi_t} a_i' = 0.$$

Also we can show that if $k \le m_0$,

$$E_{\gamma|\underline{n}_t}\{\gamma^k\Gamma(\gamma\delta_t+m_0-k)/\Gamma(\gamma\delta_t+m_0)\}$$

$$= \left\{\sum_{i=1}^{\phi_t+k} b_i'(k)\ \ln(m_0-k-1+i)\right\} / \left\{\sum_{i=1}^{\phi_t} a_i'\ \ln(m_0-1+i)\right\}$$

where

$$b_i'(k) = (k+1-m_0-i)^k \prod_{j=0}^{t}\ \prod_{s=1}^{n_j} (1+k+s\delta_t-\delta_t-m_0-i)/\delta_t^{n'} \prod_{i\neq r=1}^{\phi_t+k} (r-i), \tag{4.3}$$

and

$$\sum_{i=1}^{\phi_t+k} b_i' = 0.$$

The relationships linking a_i' and b_i' will take the same form as those relationships linking a_i and b_i.

For net premium determination purposes, γ is a nuisance parameter. It is useful for model verification purposes.

(ii) The parameter λ posterior to the date $\underline{n}_t$ is distributed such that,

$$\xi_{\lambda|\gamma\underline{n}_t}(\lambda) = (C_1')^{-1}\ \{\lambda^{\phi_t-1}/(\gamma+\lambda)^{\gamma\delta_t+\phi_t+m_0}\}$$

$$\gamma,\lambda,\delta_0>0,\phi_0=n_0+2,n_0+3,\ldots,m_0=1,2,\ldots.$$

where

$$C_1' = \Gamma(\phi_t)\ /\ \{\gamma^{\gamma\delta_t+m_0}p_{\phi_t}(\gamma\delta_t+m_t)\}.$$

The posterior moments are of the form

$$E_{\lambda|\underline{n}_t}(\lambda^k) = \{\Gamma(\phi_t+k)/\Gamma(\phi_t)\}\left\{\sum_{i=1}^{\phi_t+k} b_i'(k)\ \ln(m_0-k-1+i)\right\} \Big/ \left\{\sum_{i=1}^{\phi_t} a_i'\ \ln(m_0-1+i)\right\};\ k \leq m_0 \tag{4.4}$$

In particular

$$\mathrm{Var}_{\lambda|\underline{n}_t}(\gamma) = (\phi_t/\delta_t^2)[1+\left\{(1-\phi_t)\sum_{i=1}^{\phi_t}(a_i'/i)\ \ln(i+1)/\sum_{i=1}^{\phi_t} a_i'\ \ln(i+1)\right\}$$

$$-\left\{\phi_t\sum_{i=1}^{\phi_t}(a_i'/i)\ \ln(i+1)/\sum_{i=1}^{\phi_t} a_i'\ \ln(i+1)\right\}^2]. \tag{4.5}$$

(iii) The joint posterior distribution of γ and λ is

$$\xi_{\gamma,\lambda|\underline{n}_t}(\gamma,\lambda) = \left\{\prod_{i=1}^{t} p_{n_i}(\gamma)\gamma^{\gamma\delta_t+m_0}\lambda^{\phi_t-1}(\gamma+\lambda)^{-(\gamma\delta_t+\phi_t+m_0)}\right\}/\{C'\Gamma(\phi_t)\} \tag{4.6}$$

$\gamma,\lambda,\delta_0>0,\ n_0=0,1,\ldots,m_0=1,2,\ldots,\phi_0=n_0+2,n_0+3,\ldots$

5. THE PREDICTIVE DISTRIBUTION

One of the important advantages of adopting a Bayesian approach in insurance is that we can obtain a predictive distribution of claims of the next year given past experience. The predictive distribution will enable us to set the next year net premium, the first moment, the security loading which usually depend on the second moment, and to make decisions about general risk management policies which may depend on probability statements about the claims process.

After adopting (2.3) as the probability distribution of the number of claims, and developing the form (4.5) for the distribution of the parameters γ and λ posterior to the data $\underset{\sim}{n}_t$, we can get the joint distribution for N_{t+1} and $\gamma,\lambda|\underset{\sim}{n}_t$. Then integration over the parameter space will produce the predictive distribution. We have

$$P_{N_{t+1}|\underset{\sim}{n}_t}(n_{t+1}) = \int_0^\infty \int_0^\infty \Pr\{N_{t+1} = n_{t+1}|\gamma,\lambda\}\ \xi_{\gamma,\lambda|n_t}(\gamma,\lambda)\ d\lambda\ d\gamma$$

$$= \int_0^\infty \int_0^\infty \Big[\prod_{i=0}^{t+1} P_{n_i}(\gamma)\gamma^{\gamma\delta_{t+1}+m_0}\lambda^{\phi_{t+1}-1}(\gamma+\lambda)^{-(\gamma\delta_{t+1}+\phi_{t+1}+m_0)}/ \tag{5.1}$$

$$n_{t+1}!\ C'\ \Gamma(\phi_t)\Big]d\lambda\ d\gamma.$$

Once again we integrate using partial fraction to obtain

$$P_{N_{t+1}|\underset{\sim}{n}_t}(n_{t+1}) = \{\Gamma(\phi_{t+1})/\Gamma(\phi_t)n_{t+1}!\}\{\delta_t/\delta_{t+1}\}$$

$$\left\{\sum_{i=1}^{\phi_{t+1}} a_i'' \ln(m_0-1+i)/\sum_{i=1}^{\phi_t} a_i' \ln(m_0-1+i)\right\}. \tag{5.2}$$

In (5.2) the constants a_i' are given by (4.2) and the constants a_i'' will satisfy the system

$$\prod_{i=0}^{t+1} P_{n_i}(\gamma) = \sum_{i=1}^{\phi_{t+1}} a_i'' \prod_{i\neq r=1}^{\phi_{t+1}} (\gamma\delta_{t+1}+m_0-1+r).$$

The solution is

$$a_i'' = \prod_{j=0}^{t+1} \prod_{s=1}^{n_j} (1-m_0-i-\delta_{t+1}+s\delta_{t+1})/\delta_{t+1}^{n''} \prod_{i\neq r=1}^{\phi_{t+1}} (r-i),\ \sum_{i=1}^{\phi_{t+1}} a_i'' = 0.$$

The predictive mean is the same as the posterior mean, given $\underline{n}_t$, and is displayed in (4.4). The predictive variance, of interest in determining security loadings, can be determined as follows:

$$\mathrm{Var}(N_{t+1}|\underline{n}_t) = \mathrm{Var}_{\gamma,\lambda|\underline{n}_t}\{E(N_{t+1}|\gamma,\lambda)\} + E_{\gamma,\lambda|\underline{n}_t}\{\mathrm{Var}(N_{t+1}|\gamma,\lambda)\}$$

$$= \mathrm{Var}_{\gamma,\lambda|\underline{n}_t}(\lambda) + E_{\gamma,\lambda|\underline{n}_t}\{(\lambda + (\lambda^2/\gamma)\}$$

$$= \mathrm{Var}_{\lambda|\underline{n}_t}(\lambda) + E_{\lambda|\underline{n}_t}(\lambda) + E_{\gamma,\lambda|\underline{n}_t}(\lambda^2/\gamma),$$

The first two components, $\mathrm{Var}(\lambda|\underline{n}_t)$ and $E(\lambda|\underline{n}_t)$, are given by (4.5) and (4.4). The last $E_{\gamma,\lambda|\underline{n}_t}(\gamma^2/\gamma)$ can be determined as follows:

$$E_{\gamma,\lambda|\underline{n}_t}(\lambda^2/\gamma) = E_{\gamma|\underline{n}_t}\{\gamma^{-1}E_{\lambda|\gamma,\underline{n}_t}(\lambda^2)\},$$

substitute from (3.6) and

$$E_{\gamma,\lambda|\underline{n}_t}(\lambda^2/\gamma) = \{\Gamma(\phi_t+2)/\Gamma(\phi_2)\}\ E_{\gamma|\underline{n}_t}\{\gamma\Gamma(\gamma\delta_t+m_0-2)/\Gamma(\gamma\delta_t+m_0)\}$$

which can be shown to be

$$\phi_t(\phi_t+1)\sum_{i=1}^{\phi_t+2}\{b_i'(2)/(3-m_0-i)\}\ \ln(m_0+i-3).$$

6. APPROXIMATING THE PREDICTIVE DISTRIBUTION

One of the major reasons for going through the somewhat tedious mathematics of Section 5 is to have available the exact predictive distribution so that various approximations can be evaluated. There has recently been a great deal of interest in approximating predictive distributions. Dunsmore (1976), Lejeune and Falkenberry (1982) and Tiernay and Kadane (1984) discuss this issue. The method of Tiernay and Kadane is conceptually rather simple and has an error term of order $O(n^{-2})$, where n is the number of observations. However, the method requires that the posterior mode be calculated and the Hessian matrix of the log likelihood plus the log prior divided by n, evaluated at the posterior mode, be available. We have not completed an evaluation of their approximation in this case.

A predictive distribution is an expected value of the likelihood for N_{t+1} with respect to the posterior distribution after $\underline{n}_t$ has been observed. Therefore, one is led to ask if the predictive distribution can be approximated by a member of the negative binomial family. A series of manipulations with (5.2) lead us to suggest, for large t the following approximation

$$P_{N_{t+1}|\underline{n}_t}(n_{t+1}) = \{\Gamma(\phi_{t+1})/n_{t+1}!\Gamma(\phi_t)\}\ \{\delta_t/\delta_{t+1}\}^{\phi_t}\{1/\delta_{t+1}\}^{n_{t+1}};\ n_{t+1} = 0,1,2,\ldots \tag{6.1}$$

This approximation can be seen as substituting ϕ_t/δ_t for λ and ϕ_t for γ in (2.3).

Then we can approximate the predictive (or credible) mean by

$$E(N_{t+1}|\underline{n}_t) \doteq \phi_t/\delta_t = (\phi_0+n)/(\delta_0+t), \tag{6.2}$$

which is exact if $m_0=1$, was we showed previously, (3.8). The prediction (or credible) variance can be approximated by

$$Var(N_{t+1}|\underline{n}_t) \doteq \phi_t\delta_{t+1}/\delta_t^2 = (\phi_t/\delta_t^2) + (\phi_t/\delta_t). \tag{6.3}$$

The credible mean, in this case, takes the credibility linear form,

$$E(N_{t+1}|\underline{n}_t) \doteq z\bar{N} + (1-z)\ E(N_1) \tag{6.4}$$

where the credibility factor, $z=t/(t+k_0)$, with $k=\delta_0$, and $\bar{N}=n/t$.

Two comparisons of the exact predictive distribution (5.2) and the approximate predictive distribution (6.1) are found in Table 1.

TABLE 1

The exact and the approximate predictive distributions:

(1) If $n=2$, $\phi=3$, $\delta=25$ and $m_0=1$,

n	Pr(N=n) Exact	Pr(N=n) Approximate	Difference	n	Pr(N=n) Exact	Pr(N=n) Approximate	Difference
0	.892897	.888996	.003901	9	.000013	.000000	.000013
1	.095565	.102577	-.007012	10	.000009	.000000	.000009
2	.005126	.007891	-.002765	11	.000007	.000000	.000007
3	.000351	.000506	-.000155	12	.000007	.000000	.000007
4	.000055	.000029	.000025	13	.000005	.000000	.000005
5	.000027	.000002	-.000026	14	.000004	.000000	.000004
6	.000018	.000000	.000018	15	.000004	.000000	.000004
7	.000016	.000000	.000016				
8	.000013	.000000	.000013				

(2) If $n=2$, $\phi=3$, $\delta=164$ and $m_0=1$,

n	Pr(N=n) Exact	Pr(N=n) Approximate	Difference
0	.977654	.981928	-.004274
1	.017414	.017853	-.000439
2	.000165	.000216	-.000051
3	.000098	.000002	.000096
4	.000000	.000000	.000000
5	.000000	.000000	.000000

REFERENCES

Ammeter, H. (1948), "A Generalization of the Collective Theory of Risk in Regard to Fluctuating Basic Probabilities," Skandivarisk Aktuarielidskrift, 31, 171-198.

Carlson, T. (1962), "Negative Binomial Rationale," Proceedings, Casualty Actuarial Society, 49, 177-183.

Daboni, L. (1974), "Some Models of Inference in Risk Theory from the Bayesian Viewpoint," ASTIN Bulletin, 8, 38-56.

Dunsmore, T. R. (1976), "Asymptotic Prediction Analysis," Biometrika 63, 627-630.

Hewitt, C. C. (1960), "The Negative Binomial Applied to the Canadian Merit Rating Plan for Individual Automobile Risk," Proceedings, Casualty Actuarial Society, 47, 55-65.

Lejeune, M. and Faulkenberry, G. B. (1982), "A Simple Predictive Density Function." Journal, American Statistical Association 77, 654-657.

Morgan, I. M. (1983), Credibility Theory under the Collective Risk Model, Ph.D. Dissertation, University of Wisconsin-Madison.

Simon, L. J. (1960), "The Negative Binomial and Poisson Distributions Compared," Proceedings, Casualty Actuarial Society 47, 20-24.

Tierney, L. and Kadane, J. B. (1984), "Accurate Approximations for Posterior Moments and Marginals," Tech Report No. 431, University of Minnesota, School of Statistics.