ACCOUNTING FOR INDIVIDUAL OVER-DISPERSION IN A BONUS-MALUS AUTOMOBILE INSURANCE SYSTEM

By

MENG SHENGWANG, YUAN WEI

Department of Statistics
Renmin University, Beijing, China

AND

G.A. WHITMORE

Faculty of Management
McGill University, Montreal, Canada

ABSTRACT

Individual automobile insurance claims are characterized by over-dispersion relative to the Poisson model. In addition, claim propensities vary among individuals in any insurance portfolio. This paper presents a model which takes account of both characteristics. The model employs the negative-binomial distribution as the distribution for individual-level claims and a Pareto distribution as the distribution for claim propensities within the portfolio. The paper shows that the resulting model is tractable and has a number of attractive properties which make it suitable for this application. The fit of the model to actual claim numbers for automobile third party liability insurance is examined and found acceptable. Bayes theorem is then applied to this model to calculate illustrative optimal premiums under the Bonus-Malus System (BMS).

1. INTRODUCTION

The Poisson distribution has a long history in insurance as a model for claim counts for individuals. Actual experience shows, however, that the distribution of claim counts in repeated observations for a *single individual* tend to have greater dispersion than can be accounted for by the Poisson model. This characteristic is certainly observed in the field of automobile insurance. It also has been known since the earliest studies of insurance that the propensity to make claims differs among individuals. In automobile insurance, these differing propensities are explained by a variety of

ASTIN BULLETIN, Vol. 29, No. 2, 1999, pp 327-337

personality, health, physical and environmental risk factors for individual policyholders that are not accounted for by premium-related risk adjustments.

In the next section, we propose the negative-binomial distribution as a model for individual automobile insurance claims that can account for over-dispersion relative to the Poisson model. We then take account of variation in individual claim propensity by assuming that the mean number of claims is distributed across individual policyholders according to a Pareto distribution. The section examines the properties of this mixture model and shows that it is both mathematically tractable and suitable for this application in a number of respects. Section 3 presents a demonstration of the fit of the model to actual claim numbers for automobile third party liability insurance. In Section 4, Bayes theorem is applied to this model to calculate illustrative optimal premiums under the Bonus-Malus System (BMS).

2. Negative Binomial-Pareto Model

We assume that the number of claims K of an individual policyholder in a given time period follows a negative binomial probability distribution $\mathrm{NB}(\mu, r)$ with probability function

$$p(k|\mu) = \frac{\Gamma(r+k)}{\Gamma(r)k!}\frac{r^r\mu^k}{(r+\mu)^{r+k}} \quad \mathit{for} \quad k = 0,\ 1,\ \dots \tag{2.1}$$

Parameter $\mu > 0$ is the mean number of claims in the period, i.e., $E(K|\mu) = \mu$, and, hence, measures the individual's claim propensity. Parameter $r > 0$ is an unknown constant that is assumed to be the same value for all individuals. The variance of the number of claims K is given by

$$\sigma^2(K|\mu) = \mu\left(1 + \frac{\mu}{r}\right) . \tag{2.2}$$

The quantity μ/r in (2.2) determines the extent of over-dispersion in the negative binomial model relative to the Poisson model. This quantity is proportional to the mean number of claims μ and inversely proportional to the parameter r. Thus, the negative binomial model implies that over-dispersion increases with the mean number of claims μ. Moreover, parameter r governs the responsiveness of over-dispersion to this mean number. The larger is r, the smaller is the degree of over-dispersion. In the limit, as r tends to ∞, $p(k|\mu)$ in (2.1) tends to a Poisson distribution with mean μ.

The probability distribution of the mean parameter μ among policyholders in the insurance portfolio is assumed to be a Pareto distribution with the following density function.

$$f(\mu) = \frac{\Gamma(s\zeta + sr + 1)}{\Gamma(s\zeta)\Gamma(sr+1)} \frac{r^{sr+1}\mu^{s\zeta-1}}{(r+\mu)^{s\zeta+sr+1}} \tag{2.3}$$

We denote this Pareto distribution by Pareto (ζ, r, s). Parameter $\zeta > 0$ is the mean of the distribution, i.e., $E(\mu) = \zeta$. Parameter r is the same parameter that appears in the negative binomial distribution in (2.1). Parameter $s > 0$ measures the homogeneity of claim propensities among individual policyholders, with larger values of s implying more uniform propensities. As s increases, the density function $f(\mu)$ becomes more concentrated and, in the limit, tends to a degenerate distribution centered on ζ. This parameterization of $f(\mu)$ is somewhat elaborate but facilitates interpretation of the portfolio effect, as we shall explain shortly. We choose a Pareto distribution for several reasons. It is conjugate to the negative binomial distribution which makes it mathematically tractable. It is right skewed and unimodal which makes it suitable for describing the variation that is typically found in the mean claim parameter of individual policyholders. Finally, it has considerable mathematical flexibility for fitting different distribution patterns. This three-parameter version of the Pareto distribution is sometimes called the *generalized* Pareto distribution. Klugman et al. (1998:574) give some properties of this distribution.

The marginal distribution of K, the number of claims in the period for a randomly chosen policyholder from the insurance portfolio, is obtained from the distributions in (2.1) and (2.3) by integrating over the mean parameter μ as follows.

$$p(k) = \int_0^\infty p(k|\mu)f(\mu)d\mu = \frac{\Gamma(s\zeta + sr + 1)\Gamma(r + sr + 1)\Gamma(s\zeta + k)\Gamma(r + k)}{\Gamma(s\zeta)\Gamma(sr+1)\Gamma(r)\Gamma(r + sr + s\zeta + k + 1)k!} \tag{2.4}$$

We refer to this marginal distribution as a *negative binomial-Pareto distribution* and denote it by NBP(ζ, r, s). We note for later that the probabilities for this distribution are readily computed recursively using the following properties.

$$p(0) = \frac{\Gamma(s\zeta + sr + 1)\Gamma(sr + r + 1)}{\Gamma(sr+1)\Gamma(s\zeta + sr + r + 1)} \tag{2.5}$$

$$\frac{p(k+1)}{p(k)} = \frac{(r+k)(s\zeta + k)}{(k+1)(s\zeta + sr + r + k + 1)} \quad for \quad k = 0,\ 1,\ \ldots \tag{2.6}$$

The NBP model has characteristics that are consistent with those of traditional models for insurance claims but it also has new features that allow it to capture important aspects of real insurance portfolios. Moreover, it is a tractable model that lends itself to useful operational interpretation. We now look at the NBP distribution more closely.

1. It is known that a negative binomial distribution may be derived from a Poisson distribution by letting the Poisson mean parameter have a gamma distribution. It is this characterization, in fact, which has been used in the insurance literature to justify the use of the negative binomial distribution as a model that accounts for variation in claim propensity. Here we use this same rationale for explaining individual over-dispersion. The negative binomial distribution in (2.1) is a gamma mixture of Poisson distributions where the gamma mean parameter is μ and its shape parameter is r.
2. We have already noted that parameter r controls the extent of over-dispersion of the individual claim distribution. As r approaches ∞, the individual claim distribution approaches a Poisson distribution. In addition, however, as r increases, the Pareto distribution for μ in (2.3) approaches a gamma distribution with mean parameter ζ and shape parameter $s\zeta$. It follows therefore that, as r approaches ∞, the NBP distribution in (2.4) approaches a negative binomial distribution $\mathrm{NB}(\zeta, s\zeta)$.
3. The NBP distribution also approaches a negative binomial distribution $\mathrm{NB}(\zeta, r)$ as s approaches ∞, i.e., as claim propensities become more uniform in the portfolio. Thus, the negative binomial distribution for the number of claims is a special case of the NBP model under two different scenarios – when there is no over-dispersion or when there is no variation in claim propensity.
4. If both r and s approach infinity, the NBP model reduces to the simple Poisson model (with mean parameter ζ).
5. The NBP model uses conjugacy to extend the Poisson model in two aspects. The negative binomial model for individual claims in (2.1) follows from the conjugacy of the Poisson and gamma distributions. The NBP model in (2.4) follows from the conjugacy of the negative binomial and Pareto distributions.

Admittedly, the mathematical convenience of conjugacy is no guarantee that individual over-dispersion and variation in claim propensity have precisely the distributional forms implied by the NBP model. The negative binomial and Pareto distributions, however, are plausible models for these phenomena and are likely to capture their influence to a good approximation.

Now we wish to examine how the NBP model responds to claim experience. Consider a policyholder, drawn randomly from the insurance portfolio, who is observed to have the sequence of claims k_1, ..., k_t over t periods. We assume that each claim number is drawn independently from the negative binomial distribution (2.1) for that policyholder. The density function $f(\mu)$ in (2.3) serves as a prior distribution for the mean number of claims for the policyholder. We use Bayes theorem to derive the posterior distribution for the policyholder's mean number of claims, given this claim record, as follows.

$$f(\mu|k_1, \ldots, k_t) \propto f(\mu) \prod_{i=1}^{t} p(k_i|\mu) \tag{2.7}$$

Letting $c = \sum_{i=1}^{t} k_i$, the posterior distribution $f(\mu|k_1, \ldots, k_t)$ remains a Pareto distribution but now with the updated parameters Pareto(ζ_t, r, s_t), where

$$\zeta_t = \frac{s\zeta + c}{s + t} , \tag{2.8a}$$

$$s_t = s + t . \tag{2.8b}$$

Observe that the parameter r continues to remain fixed. The result in (2.8b) shows that parameter s_t serves as a time counter with s being the implicit time count implied by the prior density $f(\mu)$. The parameter ζ_t serves as the current estimate of the claim rate for the policyholder with ζ being the implicit claim rate implied by the prior density $f(\mu)$.

For a policyholder with the claim record $k_1, \ldots, k_t$, we are interested in the probability distribution for the claim number in the next period, namely, $K = K_{t+1}$. It follows from (2.4) and (2.7) that the predictive distribution for K, given c and t, is of the form

$$K|c, t \sim NBP(\zeta_t, r, s_t) . \tag{2.9}$$

Under a BMS, the policyholder's premium for period $t + 1$ would be set with reference to the parameters of this distribution, which reflect the policyholder's claim record. This distribution is readily computed using (2.5) and (2.6), with an appropriate substitution of the updated parameters. We note for later reference that the mean and variance of the number of claims under this predictive distribution are given by

$$E(K|c, t) = \zeta_t , \tag{2.10a}$$

$$\sigma^2(K|c, t) = \zeta_t + \frac{(s_t + 1)\zeta_t^2 + (r + 1)\zeta_t}{rs_t - 1} . \tag{2.10b}$$

3. Fitting the Model to Automobile Insurance Data

It would be ideal to have access to claim data at the individual level for multiple time periods ($t > 1$) in order to validate the negative binomial model for individual claim experience and the Pareto model for the distribution of claim propensity for the insurance portfolio. However, we do not have access to detailed claims data of this type. We do have access to published claim numbers for the automobile third party liability insurance portfolio of a Belgian insurance company for a recent year (Lemaire, 1995, page 25). The data and the fitted NBP distribution appear in Table 1. The frequency $n(k)$ represents the number of policyholders in the portfolio for whom the insurer experienced k claims. The total number of policyholders is given by $n = \sum_k n(k)$.

TABLE 1

Observed and fitted claim numbers for the NBP distribution

Claim number k	*Observed Frequency* $n(k)$	*Fitted Frequency* $n\hat{p}(k)$
0	96978	96980.0
1	9240	9235.9
2	704	702.1
3	43	51.8
4	9	3.9
> 4	0	0.3
Total	106974	106974.0

The claim distribution in Table 1 corresponds to the marginal distribution $p(k)$ in (2.4) and, hence, can be used to estimate the parameters ζ, r and s of the NBP model on the assumption that it is a valid model. We have estimated the parameters by maximum likelihood using the *ml* procedure in the statistical package STATA and obtained the estimates $\hat{\zeta} = 0.1011$, $\hat{r} = 3.736$ and $\hat{s} = 36.93$. The standard errors for the parameter estimates provided in the maximum likelihood output show that r and s are not estimated with great accuracy from the marginal distribution $p(k)$. The fitted values in Table 1 are computed using these parameter estimates.

The estimation results give some useful insights into the claim distribution pattern. Given the estimate of r, we can gauge the extent of over-dispersion (relative to the Poisson model) for the claim distribution of an individual policyholder with any given mean claim rate μ. The estimated mean of the fitted NBP distribution in Table 1 is $\hat{\zeta} = 0.1011$ claim per year. For a policyholder having this average claim rate, we see from (2.2) that $\mu/\hat{r} = 0.1011/3.736 = 0.027$. Thus, the variance is inflated by about 3 percent relative to a Poisson model for this kind of policyholder. This is a small effect. For a policyholder with a mean claim rate of $\mu = 0.5$ per year (a very extreme case), the over-dispersion amounts to $\mu/\hat{r} = 0.5/3.736 = 0.13$, which is still only a modest effect. The magnitude of over-dispersion for other automobile insurance portfolios remains an outstanding empirical question that we will examine in subsequent research. Another interesting result that can be inferred from the NBP parameter estimates is the extent of variation in the mean number of claims among policyholders. The coefficient of variation of the density function $f(\mu)$ in (2.3) (expressed as a fraction) is given by

$$C.V. = \left(\frac{\zeta + r}{\zeta(rs-1)}\right)^{\frac{1}{2}}. \tag{3.1}$$

From our parameter estimates, C.V. is estimated to be 0.53 or 53%. This estimate suggests that there is substantial variation among individual policyholders in terms of claim propensity. We also note that since $\hat{s}\hat{\zeta} = 3.731 > 1$, it follows that the mode of the density function $f(\mu)$ lies away from zero.

We have also compared the fit of the NBP model with two competing mixture models that have been proposed in the literature for claim distributions, namely, the pure negative binomial model and the Poisson-inverse Gaussian model. As we noted earlier, the negative binomial model is a special case of the NBP model in which either $r = \infty$ or $s = \infty$. The fitted distributions for the pure negative binomial and Poisson-inverse Gaussian models are taken from Lemaire (1995). Table 2 shows chi-square goodness-of-fit statistics for the three models, subject to pooling categories so that all expected frequencies $n\hat{p}(k)$ are 2 or more and at least 80% are 5 or more. The chi-square statistics are based on maximum likelihood estimates derived from the full frequency distribution. The table shows the chi-square statistic for each model, as well as its degrees of freedom. All of the chi-square statistics are significant at the 5% level, indicating that all models show a lack of fit. We comment shortly on why this may be so. The Poisson-inverse Gaussian model has the smallest chi-square statistic but by a small margin. There is little basis for choosing among the models in terms of their fit to the observed frequencies.

TABLE 2

COMPARISON OF FIT OF THREE MODELS

Model	*Chi-square Statistic*
Negative Binomial-Pareto	$\chi^2 = 6.74$, $df = 1$
Negative Binomial	$\chi^2 = 9.15$, $df = 2$
Poisson-Inverse Gaussian	$\chi^2 = 6.25$, $df = 2$

There is one feature of the data set in Table 1 that may explain the apparent lack of fit. We note that the frequency counts $n(3) = 43$, $n(4) = 9$ and $\sum_{k=5}^{\infty} n(k) = 0$ in Table 1 exhibit an unnatural pattern. The pattern suggests that perhaps administrative or other actions by the insurer (including warnings or suspensions of coverage) may have prevented claim numbers in excess of 2 from following their 'natural' statistical pattern. The sudden drop to zero of frequency counts for 5 or more claims is surprising. We cannot verify our suspicions about this pattern. We note, however, that a distortion of frequency counts would expose the measures of fit to suspect frequencies $n(k)$ for $k > 2$. This effect may explain the lack of fit of all models. The possible distortion of larger frequency counts argues for parameter estimation based on a truncated version of the NBP distribution but there are too few degrees of freedom here to allow the truncated version to be tested for fit.

4. Application of Model in Setting Automobile Insurance Premiums

The expected value principle for calculating an insurance premium sets the premium equal to $P(c, t) = (1 + \rho)E(K|c, t)$ where $E(K|c, t)$ is the expected number of claims in period $t + 1$, conditional on the policyholder having made a total of c claims in an earlier experience period of length t. The constant ρ is a multiplicative loading factor for the premium. For the NBP model, the expected value $E(K|c, t)$ is given in (2.10a). Setting a new policyholder's premium to 100, i.e., P(0, 0) = 100, we can calculate the relative premiums $100P(c, t)/P(0, 0)$ for various experience conditions (c, t) under the BMS rule. These relative premiums are shown in Table 3. Tables 4 and 5 show comparable relative premiums for the pure negative binomial and Poisson-inverse Gaussian models taken from Lemaire (1995, pages 165 and 169).

TABLE 3

BMS RELATIVE PREMIUMS FOR THE NBP MODEL BASED ON THE EXPECTED VALUE PRINCIPLE

Years	*Total claims c*						
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**
0	100						
1	97.36	123.45	149.53	175.61	201.69	227.78	253.86
2	94.86	120.28	145.69	171.10	196.51	221.93	247.34
3	92.49	117.26	142.04	166.82	191.59	216.37	241.14
4	90.23	114.40	138.57	162.74	186.91	211.08	235.25
5	88.08	111.67	135.26	158.86	182.45	206.05	229.64

The NBP model has several other noteworthy features in the context of BMS automobile insurance pricing.

1. Tables 3, 4 and 5 reveal one striking characteristic, namely, that the NBP model provides for more moderate (i.e., lower) relative premiums for policyholders with some claim experience. This characteristic can be explained by the fact that individual claims follow a Poisson process in both the pure negative binomial and Poisson-inverse Gaussian models. The NBP model, in contrast, assumes that individual claim experience will be over-dispersed relative to a Poisson model and the over-dispersion is larger for policyholders with larger mean claim rates. Therefore, extreme individual claim counts are more likely under the NBP model than under one based on a Poisson mixture. More moderate relative premiums are the result.

TABLE 4

BMS RELATIVE PREMIUMS FOR THE NEGATIVE BINOMIAL MODEL BASED ON THE EXPECTED VALUE PRINCIPLE

Years	*Total claims c*						
t	0	1	2	3	4	5	6
0	100						
1	94.07	152.69	211.31	269.92	328.54	387.16	445.77
2	88.81	144.15	199.49	254.83	310.16	365.50	420.84
3	84.11	136.51	188.92	214.33	297.73	346.14	398.54
4	79.88	129.65	179.42	229.19	278.96	328.73	378.49

TABLE 5

BMS RELATIVE PREMIUMS FOR THE POISSON-INVERSE GAUSSIAN MODEL BASED ON THE EXPECTED VALUE PRINCIPLE

Years	*Total claims c*						
t	0	1	2	3	4	5	6
0	100						
1	94.24	149.58	225.39	316.09	415.46	519.41	625.81
2	89.37	139.14	206.71	287.49	376.17	469.16	564.49
3	85.19	130.41	191.31	264.03	344.02	428.07	514.37
4	81.55	122.98	178.37	244.44	317.23	393.85	472.64

2. BMS pricing is financially balanced because it can be shown for the NBP model that

$$E(K) = E_{\{k_1, \ldots, k_t\}}[E_K(K|k_1, \ldots, k_t)] \,. \tag{4.1}$$

3. As an insurer gains experience with a policyholder (more precisely, as t increases), the posterior Pareto distribution for μ in (2.7) becomes concentrated around its mean value ζ_t which, in turn, approaches c/t as t increases. Hence, for large t, the posterior estimate of the policyholder's mean claim rate will be approximately c/t, the empirical claim rate.
4. If the variance principle is used for setting an insurance premium in the NBP context, the policyholder with claim experience (c, t) will pay the following premium

$$P(c, t) = E(K|c, t) + \lambda\sigma^2(K|c, t) \,, \tag{4.2}$$

where λ is the safety parameter. Tables 6 and 7 show BMS relative premiums using this pricing principle in conjunction with the NBP model. The values of λ used for the computations are 0.235 and 1.88, respectively, which correspond to safety loadings of 25% and 200% of the net premium. The latter represents an extreme assumption. Even so, however, the tables show that the relative premiums differ little from those in Table 3 that were calculated under the expected value principle. Thus, the choice between these two premium calculation methods makes little difference under the NBP model.

TABLE 6

BMS RELATIVE PREMIUMS FOR THE NBP MODEL BASED ON THE VARIANCE PRINCIPLE WITH λ = 0.235

Years	*Total claims c*						
t	0	1	2	3	4	5	6
0	100						
1	97.33	123.58	149.89	176.28	202.74	229.27	255.87
2	94.80	120.36	145.99	171.68	197.44	223.27	249.17
3	92.40	117.31	142.28	167.32	192.42	217.58	242.81
4	90.12	114.41	138.76	163.17	187.64	212.17	236.77
5	87.95	111.65	135.41	159.22	183.10	207.03	231.02

TABLE 7

BMS RELATIVE PREMIUMS FOR THE NBP MODEL BASED ON THE VARIANCE PRINCIPLE WITH $\lambda = 1.88$

Years	*Total claims c*						
t	0	1	2	3	4	5	6
0	100						
1	97.26	123.88	150.74	177.84	205.17	232.75	260.56
2	94.67	120.56	146.69	173.04	199.61	226.41	253.44
3	92.21	117.42	142.85	168.49	194.34	220.41	246.70
4	89.88	114.44	139.20	164.17	189.34	214.72	240.30
5	87.66	111.60	135.74	160.07	187.60	209.32	234.23

ACKNOWLEDGEMENTS

The authors thank two anonymous referees for suggestions that have greatly improved the presentation and deepened the results of this paper. We are grateful to the Centre for International Management Studies of McGill University and the Natural Sciences and Engineering Research Council of Canada for financial support for this research.

CITED REFERENCES

KLUGMAN, Stuart, Harry H. PANJER and Gordon E. WILLMOT (1998). *Loss Models: From Data to Decisions*, John Wiley & Sons, Inc.

LEMAIRE, Jean (1995). *Bonus-Malus Systems for Automobile Insurance*, Kluwer Academic Publishers.

YAN, Maosong (1989). *Bayes Risk Decisions*, Qinghua University Publishing House.

Professor G.A. WHITMORE
Faculty of Management
McGill University
1001 Sherbrooke Street West
Montreal
Quebec H3A 1G5
Canada
e-mail: whitmore@management.mcgill.ca
tel: 514-398-4049
fax: 514-398-3876