Bridging Minimum Bias and Maximum Likelihood Methods through Weighted Equation

Noriszura Ismail and Abdul Aziz Jemain

Abstract

In classification ratemaking, the multiplicative and additive models derived by actuaries are based on two common methods; minimum bias and maximum likelihood. These models are already considered as established and standard, particularly in automobile and general liability insurance. This paper aims to identify the relationship between both methods by rewriting the equations of both minimum bias and maximum likelihood as a weighted equation. The weighted equation is in the form of a weighted difference between observed and fitted rates. The advantage of having the weighted equation is that the solution can be solved using regression model. Compared to the classical method introduced by Bailey and Simon (1960), the regression model provides an improved and simplified programming algorithm. In addition, the parameter estimates could also be rewritten as a weighted solution; for multiplicative model the solution can be written in the form of a weighted proportion whereas for additive model, the form is of a weighted difference. In this paper, the weighted equation will be applied on three types of classification ratemaking data; ship damage incidents data of McCullagh and Nelder (1989), data from Bailey and Simon (1960) on Canadian private automobile liability insurance and UK private car motor insurance data from Coutts (1984).

Keywords: Classification ratemaking; Minimum bias; Maximum likelihood; Multiplicative; Additive.

1. INTRODUCTION

In order to determine pure premium rates in casualty insurance, actuaries have to fulfil two requirements. First, they have to ensure that the insurer will receive premiums at a level adequate to cover losses and expenses. Next, they have to allocate premiums "fairly" between insureds, i.e., high risk insured should pay higher premium. For the first requirement, actuaries are required to adjust the overall level of premiums, taking into account short-term economic effects such as inflation, and other external factors such as government regulation, that can be dealt with minimum statistical analysis. However, for the second requirement, the relative premium levels need to be determined. At this stage, statistical modelling and actuarial judgement are important and actuaries can achieve this by using classification ratemaking.

The goal of classification ratemaking is to group homogeneous risks and charge each group a premium to commensurate with the expected average loss. Failure to achieve this goal may lead to adverse selection to insureds and economic losses to insurers. The risks may be categorized according to rating factors; for instance in auto insurance, driver's gender, claim experience, and location, or vehicle's make, capacity and year, can be considered as rating factors.

Among the pioneer studies of classification ratemaking, Bailey and Simon (1960) compared the systematic bias of multiplicative and additive models. Following their work, a few studies focusing and debating on additive and multiplicative models were published. Bailey (1963) compared multiplicative and additive models by producing two statistical criteria, namely the minimum chi-squares and the average absolute difference. Freifelder (1986) predicted the pattern of over and under estimation of multiplicative and additive models if true models are misspecified. Jee (1989) compared the predictive accuracy of multiplicative and additive models, and Holler *et al.* (1999) compared the initial values sensitivity of multiplicative and additive models.

In addition, researchers of classification ratemaking also suggested various statistical procedures to estimate the model parameters. Bailey and Simon (1960) suggested the minimum chi-squares, Bailey (1963) used the zero bias, Jung (1968) produced a heuristic method for minimum modified chi-squares, Ajne (1975) suggested the method of moments, Chamberlain (1980) used the weighted least squares, Coutts (1984) produced the method of orthogonal weighted least squares with logit transformation, Harrington (1986) suggested the maximum likelihood method for models with functional form, and Brockmann and Wright (1992) used the generalized linear models with Poisson error structure for claim frequency and Gamma error structure for claim severity. With the development of computing packages in the recent years, various statistical packages were also suggested and used, including GLIM by Brown (1988) and SAS by Holler *et al.* (1999) and Mildenhall (1999).

Based on the literature review, most researchers studied classification ratemaking in terms of two main perspectives; the models of multiplicative vs. additive, and the methods of minimum bias vs. maximum likelihood; using a variety of criteria, namely biasness, interaction terms, goodness of fit, initial value sensitivity and prediction accuracy. This paper differs such that it tries to bridge both methods via a weighted equation. This author believes that the weighted equation makes understanding the similarities and differences between both methods an easier task.

The objective of this paper is to bridge minimum bias and maximum likelihood methods by rewriting their equations as a weighted equation. The weighted equation can be written in the form of a weighted difference between observed and fitted rates. The advantage of having the weighted equation is that the solution can be solved using regression model. Compared to the classical method introduced by Bailey and Simon (1960), the regression model provides an improved and simplified programming algorithm. In addition, the parameter estimates could also be rewritten as a weighted solution; for multiplicative model it is in the form of a weighted proportion whereas for additive model, the form is of a weighted difference. In this paper, the weighted equation will be applied on three types of classification ratemaking data; ship damage incidents data of McCullagh and Nelder (1989), data from Bailey and Simon (1960) on Canadian private automobile liability insurance and UK private car motor insurance data from Coutts (1984).

Rewriting the equations of minimum bias and maximum likelihood as a weighted equation has its own advantages; the mathematical concept of the weighted equation is simpler, hence providing an easier understanding, particularly for insurance practitioners; the weighted equation allows the usage of regression model as an alternative programming algorithm to calculate the parameter estimates; the weighted equation provides a basic step to further understand the more complex distributions, primarily the distributions involving dispersion parameter; the weights of the parameter solution shows that each of multiplicative and additive models has similar solution; and finally, the weights of the parameter solution also shows that models with larger sample size and number of parameter have slower convergence.

2. CLASSIFICATION RATEMAKING

In casualty insurance, the risk premium, i.e., the premium excluding expenses, is equal to the product of claim frequency and severity. Classification ratemaking is the statistical procedure that classifies risks in claim frequency and severity models into groups of homogeneous risks, categorized by the rating factors. In this study, classification ratemaking is used to estimate claim frequency rates, expressed in terms of frequency per unit of exposure. For instance, the exposure unit used for auto insurance is based on a car-year unit.

Consider a regression model with n observations of claim frequency rates and p explanatory variables inclusive of intercept and dummy variables. Next, consider a data of frequency rates involving three rating factors, each respectively with three, two and three rating classes. Thus, this data has a total of $n = 18$ observed rates with $p = 6$ explanatory variables. In addition, let $\mathbf{c}$, $\mathbf{e}$ and $\mathbf{r}$ denote the vectors for claim counts, exposures and observed rates, respectively. Therefore, the observed rate for the *i*th rating class, $i = 1,2,...,18$, is equivalent to $r_i = {c_i}/{e_i}$.

Furthermore, let $\mathbf{X}$ be the matrix of explanatory variables with the *i*th row equivalent to vector $\mathbf{x}_\mathbf{i}^\mathbf{T}$, and $\boldsymbol{\beta}$ be the vector of regression parameters. If x_{ij}, $i = 1,2,...,18$, $j = 1,2,...,6$, is the *ij*th element of matrix $\mathbf{X}$, the value for x_{ij} is either one or zero. Table 1 summarize the regression model for the data.

Table 1. Data summary

i	c_i	e_i	$r_i = {c_i}/{e_i}$	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	x_{i6}
1	c_1	e_1	r_1	1	0	0	0	0	0
2	c_2	e_2	r_2	1	0	0	0	1	0
3	$\vdots$	$\vdots$	$\vdots$	1	0	0	0	0	1
4				1	0	0	1	0	0
5				1	0	0	1	1	0
6				1	0	0	1	0	1
7				1	1	0	0	0	0
8				1	1	0	0	1	0
9				1	1	0	0	0	1
10				1	1	0	1	0	0
11				1	1	0	1	1	0
12				1	1	0	1	0	1
13				1	0	1	0	0	0
14				1	0	1	0	1	0
15				1	0	1	0	0	1
16				1	0	1	1	0	0
17				1	0	1	1	1	0
18	c_{18}	e_{18}	r_{18}	1	0	1	1	0	1

Moreover, let $\mathbf{f}$, a function of $\mathbf{X}$ and $\boldsymbol{\beta}$, denotes the vector for fitted rates. For a multiplicative model, the *i*th fitted rate is equivalent to

$$f_i = \exp(\mathbf{x}_\mathbf{i}^\mathbf{T}\boldsymbol{\beta}),$$

which can also be written as

$$f_i = f_{i(-j)} \exp(\beta_j x_{ij}), \tag{1}$$

where $f_{i(-j)}$ is the *i*th multiplicative fitted rate without the *j*th effect. As for an additive model, the *i*th fitted rate is equal to

$$f_i = \mathbf{x_i^T}\boldsymbol{\beta},$$

which can also be written as

$$f_i = f_{i(-j)} + \beta_j x_{ij}, \tag{2}$$

where $f_{i(-j)}$ is the *i*th additive fitted rate without the *j*th effect. Thus, the objective of classification ratemaking is to have the fitted rates, f_i, be as close as possible to the observed rates, r_i, for all *i*.

3. MINIMUM BIAS

Bailey and Simon (1960) were among the pioneer researchers that consider bias in classification ratemaking and introduced the minimum bias method. They proposed a famous list of four criteria for an acceptable set of classification rates:

i. It should reproduce experience for each class and overall, i.e., be balanced for each class and overall.

ii. It should reflect the relative credibility of the various classes.

iii. It should provide minimum amount of departure from the raw data.

iv. It should produce a rate for each class of risks which is close enough to the experience so that the differences could reasonably be caused by chance.

3.1 Bailey Zero Bias

Bailey and Simon (1960) proposed a suitable test for Criterion (i) by calculating,

$$\frac{\sum_i e_i f_i}{\sum_i e_i r_i}, \tag{3}$$

for each *j* and total. A set of rates is balanced, i.e., zero bias, if equation (3) equals 1.00. Automatically, zero bias for each class implies zero bias overall.

From this test, Bailey (1963) derived a minimum bias model by setting the average difference between observed and fitted rates to be equal to zero. The zero bias equation for each *j* can be written in the form of a weighted difference between observed and fitted rates,

$$\sum_i w_i (r_i - f_i) = 0, \tag{4}$$

where w_i is equal to $e_i x_{ij}$.

Substituting (1) into (4), the zero bias equation for multiplicative model become

$$\sum_i e_i r_i x_{ij} = \sum_i e_i f_{i(-j)} \exp(\beta_j x_{ij}) x_{ij} .$$

Since x_{ij} is either one or zero, the solution for each *j* could be obtained and written in the form of a weighted proportion of observed over multiplicative fitted rates without the *j*th effect,

$$\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}} , \tag{5}$$

where v_i is the normalized weight of $\frac{z_i}{\sum_i z_i}$ and z_i is $e_i f_{i(-j)} x_{ij}$.

For additive model, the zero bias equation after substituting (2) into (4) is

$$\sum_i e_i (r_i - f_{i(-j)}) x_{ij} = \sum_i e_i (\beta_j x_{ij}) x_{ij} .$$

Again, since x_{ij} is either one or zero, the solution for each *j* could be obtained. However, for additive model, it is in term of a weighted difference between observed and additive fitted rates without the *j*th effect,

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}) , \tag{6}$$

where v_i is $\frac{z_i}{\sum_i z_i}$ and z_i is $e_i x_{ij}$.

3.2 Minimum Chi-Squares

Bailey and Simon (1960) also suggested the χ^2 statistics as an appropriate test for Criterion (iv),

$$\chi^2 = \sum_i \frac{e_i}{f_i} (r_i - f_i)^2 .$$

The same test is also suitable for Criterion (ii) and (iii) as well.

By minimizing the χ^2 statistics, another minimum bias model was derived. For each *j*, the minimum χ^2 equation could be written in the form of a weighted difference between observed and fitted rates,

$$\frac{\partial \chi^2}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0\,, \tag{7}$$

where w_i is $\dfrac{e_i (r_i + f_i)}{f_i^2} \dfrac{\partial f_i}{\partial \beta_j}$.

For multiplicative model,

$$\frac{\partial f_i}{\partial \beta_j} = f_i x_{ij}\,, \tag{8}$$

whereas in additive model,

$$\frac{\partial f_i}{\partial \beta_j} = x_{ij}\,. \tag{9}$$

If multiplicative model is chosen, by substituting (1) and (8) into (7), the parameter solution is equivalent to

$$\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}}\,, \tag{10}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $e_i (r_i + f_i) x_{ij}$.

For additive model, the parameter solution after substituting (2) and (9) into (7) is

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)})\,, \tag{11}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i (r_i + f_i)}{f_i^2} x_{ij}$.

4. MAXIMUM LIKELIHOOD

Assume that the *i*th claim frequency count, $c_i = e_i r_i$, comes from a distribution whose probability density function is $g(c_i; f_i)$. A maximum likelihood method maximizes the likelihood function,

$$L = \prod_i g(c_i; f_i),$$

or equivalently, the log likelihood function,

$$\ell = \log L = \sum_i \log\left(g(c_i; f_i)\right).$$

Thus, the parameter solution can be obtained by setting $\frac{\partial \ell}{\partial \beta_j} = 0$ for each *j*.

4.1 Normal Distribution

If $c_i = e_i r_i$ is assumed to follow Normal distribution with mean $e_i f_i$, $g(c_i; f_i)$ can be written as

$$g(c_i; f_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}\left(e_i r_i - e_i f_i\right)^2\right\}.$$

Hence, the likelihood equation for each *j* is equivalent to

$$\frac{\partial \ell}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0, \tag{12}$$

where w_i is $e_i^{\,2} \frac{\partial f_i}{\partial \beta_j}$.

Assuming multiplicative model, the solution after substituting (1) and (8) into (12) is

$$\exp(\beta_j) = \sum_i v_i \left(\frac{r_i}{f_{i(-j)}}\right), \tag{13}$$

where v_i is $\frac{z_i}{\sum_i z_i}$ and z_i is $e_i^2 f_{i(-j)}^2 x_{ij}$.

For additive model, by substituting (2) and (9) into (12), the parameter solution is equivalent to

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}), \tag{14}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $e_i^2 x_{ij}$.

4.2 Poisson Distribution

The same weighted equation could also be used to show that Poisson multiplicative is actually equivalent to zero bias multiplicative, derived by Bailey (1963). If $c_i = e_i r_i$ is assumed to have Poisson distribution with mean $e_i f_i$, the probability density function is

$$g(c_i; f_i) = \frac{\exp(-e_i f_i)(e_i f_i)^{e_i r_i}}{(e_i r_i)!}.$$

As a result, for each *j*, the likelihood equation is equal to

$$\frac{\partial \ell}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0, \tag{15}$$

where w_i is $\dfrac{e_i}{f_i} \dfrac{\partial f_i}{\partial \beta_j}$.

Substituting (1) and (8) into (15) for multiplicative model, the parameter solution can be written as

$$\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}},$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $e_i f_{i(-j)} x_{ij}$. This solution is equivalent to zero bias multiplicative shown by (5).

If additive model is chosen, by substituting (2) and (9) into (15), the parameter solution is equal to

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}), \tag{16}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i}{f_i} x_{ij}$.

4.3 Binomial Distribution

Assuming $c_i = e_i r_i$ comes from Binomial distribution with mean $e_i f_i$, $g(c_i; f_i)$ can be written as

$$g(c_i; f_i) = \binom{e_i}{c_i} f_i^{c_i} (1 - f_i)^{e_i - c_i} .$$

For each *j*, the likelihood equation is equivalent to

$$\frac{\partial \ell}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0 , \tag{17}$$

where w_i is $\dfrac{e_i}{f_i(1-f_i)} \dfrac{\partial f_i}{\partial \beta_j}$.

Using multiplicative model, the solution after substituting (1) and (8) into (17) is

$$\exp(\beta_j) = \sum_i v_i \left(\frac{r_i}{f_{i(-j)}} \right), \tag{18}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i f_{i(-j)}}{1 - f_i} x_{ij}$.

If additive model is chosen, by substituting (2) and (9) into (17), the solution can be written as

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}) , \tag{19}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i}{f_i(1-f_i)} x_{ij}$.

4.4 Negative Binomial Distribution

The advantage of using the weighted equation is that it can be used as an introductory step to understand the fitting procedure of a distribution with dispersion parameter. If the dependent variable, C_i, is a count with mean $E(C_i) = \mu_i$, a standard statistical procedure is to fit the data with Poisson distribution using multiplicative model. However, if overdispersion is detected in the data, i.e., $Var(C_i) > E(C_i)$, the parameter estimates for standard Poisson are still consistent, but inefficient. As an alternative, the standard overdispersion model is the Negative Binomial distribution with multiplicative model. If C_i

is distributed as Negative Binomial$(\mu_i; a)$, the probability density function is (Lawless, 1987),

$$g(c_i;\mu_i,a) = \frac{\Gamma(c_i + \frac{1}{a})}{c_i!\Gamma(\frac{1}{a})}\left(\frac{a\mu_i}{1+a\mu_i}\right)^{c_i}\left(\frac{1}{1+a\mu_i}\right)^{\frac{1}{a}},$$

and the mean and variance are

$$E(C_i) = \mu_i,$$

$$Var(C_i) = \mu_i(1+a\mu_i),$$

where a is the dispersion parameter. Since $a \geq 0$ and $\mu_i \geq 0$ for all *i*, the distribution allows for overdispersion.

For our classification ratemaking example, $c_i = e_i r_i$ and $\mu_i = e_i f_i$. Thus, the likelihood equation can also be written as

$$\frac{\partial \ell}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0, \qquad (20)$$

where w_i is $\dfrac{e_i}{f_i(1+ae_i f_i)}\dfrac{\partial f_i}{\partial \beta_j}$. Notice that the weight for Poisson (15) is a special case of the weight for Negative Binomial (20), when the dispersion parameter, a, is equal to zero.

Assuming multiplicative model, by substituting (1) and (8) into (20), the parameter solution is

$$\exp(\beta_j) = \sum_i v_i \left(\frac{r_i}{f_{i(-j)}}\right), \qquad (21)$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i f_{i(-j)}}{1+ae_i f_i} x_{ij}$.

For additive model, the parameter solution after substituting (2) and (9) into (20) is equal to

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}), \qquad (22)$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i}{f_i(1+ae_i f_i)} x_{ij}$.

4.5 Generalized Poisson Distribution

Another alternative for overdispersion is to use the Generalized Poisson distribution. The advantage of using Generalized Poisson distribution is that it can be used for both overdispersion, i.e., $Var(C_i) > E(C_i)$, as well as underdispersion, i.e., $Var(C_i) < E(C_i)$. If C_i is assumed to follow Generalized Poisson distribution, $g(c_i;\mu_i,a)$ can be written as (Wang and Famoye, 1997),

$$g(c_i;\mu_i,a) = \left(\frac{\mu_i}{1+a\mu_i}\right)^{c_i} \frac{(1+ac_i)^{c_i-1}}{c_i\,!} \exp\left(-\frac{\mu_i(1+ac_i)}{1+a\mu_i}\right),$$

with mean and variance,

$$E(C_i) = \mu_i,$$

$$Var(C_i) = \mu_i(1+a\mu_i)^2.$$

Since $a \ge 0$ or $a \le 0$, the distribution allows for either overdispersion or underdispersion. Assuming $c_i = e_i r_i$ and $\mu_i = e_i f_i$, the likelihood equation is

$$\frac{\partial \ell}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0, \tag{23}$$

where w_i is $\dfrac{e_i}{f_i(1+ae_i f_i)^2}\dfrac{\partial f_i}{\partial \beta_j}$. Again, the weight for Poisson (15) is a special case of the weight for Generalized Poisson (23), when the dispersion parameter, a, is equal to zero.

Substituting (1) and (8) into (23) for multiplicative model, the parameter solution is

$$\exp(\beta_j) = \sum_i v_i \left(\frac{r_i}{f_{i(-j)}}\right), \tag{24}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i f_{i(-j)}}{(1+ae_i f_i)^2} x_{ij}$.

For additive model, by substituting (2) and (9) into (23), the parameter solution can be written as

$$\beta_j = \sum_i v_i (r_i - f_{i(-j)}), \tag{25}$$

where v_i is $\frac{z_i}{\sum_i z_i}$ and z_i is $\frac{e_i}{f_i(1+ae_i f_i)^2} x_{ij}$.

5. OTHER MODELS

5.1 Least Squares

The same weighted equation could also be extended to other error functions as well. Define the sum squares error as (Brown, 1988),

$$S = \sum_i \frac{(e_i r_i - e_i f_i)^2}{e_i} = \sum_i e_i (r_i - f_i)^2 .$$

So, the least squares equation can be written as

$$\frac{\partial S}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0 , \tag{26}$$

where w_i is $e_i \frac{\partial f_i}{\partial \beta_j}$.

Substituting (1) and (8) into (26) for multiplicative model, the parameter solution is

$$\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}} , \tag{27}$$

where v_i is $\frac{z_i}{\sum_i z_i}$ and z_i is $e_i f_{i(-j)}^2 x_{ij}$.

Extending this equation to least squares with additive model, it can be shown that least squares additive is equivalent to zero bias additive, derived by Bailey (1963). The parameter solution after substituting (2) and (9) into (26) is equivalent to

$$\beta_j = \sum_i v_i (r_i - f_{(-j)}) ,$$

where $v_i = \frac{z_i}{\sum_i z_i}$ and $z_i = e_i x_{ij}$. This solution is equivalent to the zero bias additive shown by (6).

5.2 Minimum Modified Chi-Squares

If the function of error is a modified χ^2 statistics which is defined as,

$$\chi^2_{\text{mod}} = \sum_i \frac{e_i}{r_i}(r_i - f_i)^2,$$

the equation for minimum modified χ^2 is equivalent to

$$\frac{\partial \chi^2_{\text{mod}}}{\partial \beta_j} = \sum_i w_i (r_i - f_i) = 0, \tag{28}$$

where w_i is $\dfrac{e_i}{r_i}\dfrac{\partial f_i}{\partial \beta_j}$.

For multiplicative model, by substituting (1) and (8) into (28), the parameter solution can be written as

$$\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}}, \tag{29}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $z_i = \dfrac{e_i f^2_{i(-j)}}{r_i} x_{ij}$.

Substituting (2) and (9) into (28) for additive model, the parameter solution is

$$\beta_j = \sum_i v_i (r_i - f_{(-j)}), \tag{30}$$

where v_i is $\dfrac{z_i}{\sum_i z_i}$ and z_i is $\dfrac{e_i}{r_i} x_{ij}$.

Table 2 summarizes the weighted equations and parameter solutions for all of the models discussed above. From the table, the following conclusions can be made:

i. For additive models, the zero bias and least squares are equivalent.

ii. For multiplicative models, the zero bias and Poisson are equal.

iii. The weighted equation, which is in the form of a weighted difference between observed and fitted rates, show that all models are similar; each model is distinguished only by its weight.

iv. The weights in the parameter solutions show that each of multiplicative and additive models is expected to produce similar parameter estimates.

6. MODEL PROGRAMMING

6.1 Classical Method

The classical iterative method for finding parameter solutions was first introduced by Bailey and Simon (1960). This method solves the parameter individually for each *j*. In the first iteration, vector of initial values, $\boldsymbol{\beta}^{(0)}$, are needed to calculate the vector of next parameter estimates, $\boldsymbol{\beta}^{(1)}$. The process of iteration is then repeated until all solutions converge.

Table 2. Summary of weighted equations and parameter solutions

Models	w_i for weighted equation, $\sum_i w_i(r_i - f_i) = 0$	z_i for multiplicative parameter solutions, $\exp(\beta_j) = \sum_i v_i \dfrac{r_i}{f_{i(-j)}}$, $v_i = \dfrac{z_i}{\sum_i z_i}$	z_i for additive parameter solutions, $\beta_j = \sum_i v_i(r_i - f_{i(-j)})$, $v_i = \dfrac{z_i}{\sum_i z_i}$
Zero bias	$w_i = e_i x_{ij}$	$z_i = e_i f_{i(-j)} x_{ij}$	$z_i = e_i x_{ij}$
Poisson	$w_i = \dfrac{e_i}{f_i}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = e_i f_{i(-j)} x_{ij}$	$z_i = \dfrac{e_i}{f_i} x_{ij}$
Least Squares	$w_i = e_i \dfrac{\partial f_i}{\partial \beta_j}$	$z_i = e_i f_{i(-j)}{}^2 x_{ij}$	$z_i = e_i x_{ij}$
Minimum χ^2	$w_i = \dfrac{e_i(r_i + f_i)}{f_i^2}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = e_i(r_i + f_i) x_{ij}$	$z_i = \dfrac{e_i(r_i + f_i)}{f_i^2} x_{ij}$
Normal	$w_i = e_i{}^2 \dfrac{\partial f_i}{\partial \beta_j}$	$z_i = e_i{}^2 f_{i(-j)}{}^2 x_{ij}$	$z_i = e_i{}^2 x_{ij}$
Binomial	$w_i = \dfrac{e_i}{f_i(1-f_i)}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = \dfrac{e_i f_{i(-j)}}{1-f_i} x_{ij}$	$z_i = \dfrac{e_i}{f_i(1-f_i)} x_{ij}$
Negative Binomial	$w_i = \dfrac{e_i}{f_i(1+ae_i f_i)}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = \dfrac{e_i f_{i(-j)}}{1+ae_i f_i} x_{ij}$	$z_i = \dfrac{e_i}{f_i(1+ae_i f_i)} x_{ij}$
Generalized Poisson	$w_i = \dfrac{e_i}{f_i(1+ae_i f_i)^2}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = \dfrac{e_i f_{i(-j)}}{(1+ae_i f_i)^2} x_{ij}$	$z_i = \dfrac{e_i}{f_i(1+ae_i f_i)^2} x_{ij}$
Minimum modified χ^2	$w_i = \dfrac{e_i}{r_i}\dfrac{\partial f_i}{\partial \beta_j}$	$z_i = \dfrac{e_i f_{i(-j)}{}^2}{r_i} x_{ij}$	$z_i = \dfrac{e_i}{r_i} x_{ij}$

An example for the parameter solution of zero bias multiplicative, $\exp(\beta_j) = \sum_i v_i \frac{r_i}{f_{i(-j)}}$, where v_i is $\frac{z_i}{\sum_i z_i}$ and z_i is $e_i f_{i(-j)} x_{ij}$, is discussed here.

Let $f_{i(-j)}$ denotes the the *i*th row of $\mathbf{f_{(-j)}}$, the vector of multiplicative fitted rates without the *j*th effect. For multiplicative model, $\mathbf{f_{(-j)}} = \exp(\mathbf{X_{(-j)}} \boldsymbol{\beta}_{\mathbf{(-j)}})$, where $\mathbf{X_{(-j)}}$ denotes the matrix of explanatory variables without the *j*th column and $\boldsymbol{\beta}_{\mathbf{(-j)}}$ the vector of regression parameters without the *j*th row.

Moreover, let $\mathbf{x_j}$ denotes the vector equivalent to the *j*th column of matrix **X**. Thus, x_{ij} is equal to the *i*th row of vector $\mathbf{x_j}$. Further, let z_i, the *i*th row of vector $\mathbf{z}$, equal to the product of e_i, $f_{i(-j)}$ and x_{ij}. Therefore, v_i, the *i*th row of vector $\mathbf{v}$, is equivalent to the proportion of z_i over sum of z_i for all *i*.

For multiplicative models, the same programming can be used if z_i is written as

$$z_i = e_i^{\,b} r_i^{\,d} f_{i(-j)}^{\;\;g} (1-f_i)^h (r_i + f_i)^k (1 + a e_i f_i)^l \; x_{ij} .$$

For example, in zero bias multiplicative, $a = 0, b = 1, d = 0, g = 1, h = 0, k = 0$, and $l = 0$. Similarly, z_i for additive model is of the form,

$$z_i = e_i^b r_i^d f_i^g (1-f_i)^h (r_i + f_i)^k (1 + a e_i f_i)^l \; x_{ij} .$$

For instance, in zero bias additive, $a = 0, b = 1, d = 0, g = 0, h = 0, k = 0$, and $l = 0$.

Examples of S-PLUS programming for both multiplicative and additive models are shown in Appendix A. The same programming can be used since z_i can be written in a functional form of a, b, d, g, h, k and l. Note that for minimum modified χ^2, both multiplicative and additive models contain the observed rate, r_i, as the denominator in z_i. Thus, to avoid a "division by zero", it is suggested that a small constant is added to r_i in the programming.

6.2 Regression Model

In regression model, the estimates for $\beta_j, j = 1, 2, \ldots, p$, can be found by minimizing,

$$\sum_i w_i (r_i - f_i(\boldsymbol{\beta}))^2 ,$$

or equivalently, they are the solution of,

$$\sum_i w_i (r_i - f_i(\boldsymbol{\beta})) \frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_j} = 0 ,$$

for each *j*. This equation is similar to the weighted equation derived for classification rates discussed previously. Hence, the parameter solutions for classification ratemaking are allowed to be solved using a regression model.

By using Taylor series approximation, it can be shown that (Venables and Ripley, 1997),

$$\boldsymbol{\beta}^{(1)} = (\mathbf{Z}^{(0)\mathrm{T}} \mathbf{W}^{(0)} \mathbf{Z}^{(0)})^{-1} \mathbf{Z}^{(0)\mathrm{T}} \mathbf{W}^{(0)} (\mathbf{r} - \mathbf{s}^{(0)}) ,$$

where,

$\mathbf{Z}^{(0)} = (n \times p)$ matrix whose *ij*th element is equal to $\left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_j}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}}$

$\mathbf{W}^{(0)} = (n \times n)$ diagonal matrix of weight, evaluated at $\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)}$

$\mathbf{s}^{(0)}$ = vector where the *i*th row is equal to $f_i(\boldsymbol{\beta}^{(0)}) - \sum_{j=1}^{p} \beta_j^{(0)} z_{ij}^{(0)}$

In the first iteration, the vector of initial values, $\boldsymbol{\beta}^{(0)}$, are needed to calculate $\boldsymbol{\beta}^{(1)}$. The process of iteration is then repeated until the solution converges. Since the parameter estimates are represented by vector $\boldsymbol{\beta}$, the regression model solves them simultaneously, thus providing a faster convergence compared to the classical approach.

Consider an additive model where the *ij*th element of matrix $\mathbf{Z}^{(0)}$ is equal to $\left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_j}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}} = x_{ij}$, which is free of $\boldsymbol{\beta}^{(0)}$. Since x_{ij} is the *ij*th element of matrix $\mathbf{X}$ and both matrices have the same dimension, $\mathbf{Z}^{(0)} = \mathbf{X}$ and $\mathbf{s}^{(0)} = \mathbf{f}(\boldsymbol{\beta}^{(0)}) - \mathbf{X}\boldsymbol{\beta}^{(0)} = \mathbf{0}$.

For example, the weighted equation for least squares (26) is equivalent to

$$\sum_i e_i (r_i - f_i) \frac{\partial f_i}{\partial \beta_j} = 0 .$$

Here, the *i*th diagonal element of matrix $\mathbf{W}^{(0)}$ is e_i, which is also free of $\boldsymbol{\beta}^{(0)}$. Therefore, for additive model, the vector of parameter estimates for least squares is

$$\boldsymbol{\beta}^{(0)} = \boldsymbol{\beta} = (\mathbf{X}^{\mathrm{T}} \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^{\mathrm{T}} \mathbf{W} \mathbf{r} ,$$

which is equivalent to the normal equation in linear regression model, thus allowing the solution to be solved without any iteration.

However, if multiplicative model is assumed, the *ij*th element of matrix $\mathbf{Z}^{(0)}$ is $\left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_j}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}} = f_i(\boldsymbol{\beta}^{(0)})x_{ij}$, or equivalently $\mathbf{Z}^{(0)} = \mathbf{F}^{(0)}\mathbf{X}$, where $\mathbf{F}^{(0)}$ is the diagonal matrix whose *i*th diagonal element is $f_i(\boldsymbol{\beta}^{(0)})$. For this reason, vector $\mathbf{s}_{(0)}$ can also be written as

$$\mathbf{s}^{(0)} = \mathbf{f}(\boldsymbol{\beta}^{(0)}) - \mathbf{F}^{(0)}\mathbf{X}\boldsymbol{\beta}^{(0)}.$$

For all models discussed previously, the same programming can be used if the *i*th diagonal element of weight matrix is written as,

$$e_i^b r_i^d f_i^g (1-f_i)^h (r_i + f_i)^k (1+ae_i f_i)^l .$$

The simplest form is the weight for least squares whereby $a = 0, b = 1, d = 0, g = 0, h = 0, k = 0$ and $l = 0$.

Examples of S-PLUS programming for both multiplicative and additive models are shown in Appendix B. The same programming can be used since the weight can be written in a functional form of a, b, d, g, h, k and l. Note that for minimum modified χ^2, the weight, w_i, contain the observed rate, r_i, as the denominator. Thus, to avoid a "division by zero", it is suggested that a small constant is added to r_i in the programming.

7. EXAMPLES

Consider three types of classification ratemaking data; ship damage incidents data of McCullagh and Nelder (1989), data from Bailey and Simon (1960) on Canadian private automobile liability insurance, and UK private car motor insurance data from Coutts (1984). These data are also available and can be accessed from the Internet in the following websites; http://sunsite.univie.ac.at/statlib/datasets/ships for McCullagh and Nelder (1989) data, http://www.casact.org/library/astin/vol1no4/192.pdf for the data of Bailey and Simon (1960), and http://www.actuaries.org.uk/files/pdf/library/JIA-111/0087-0148.pdf for Coutts (1984) data.

For ship damage incidents data, the number of damage incidents and exposure for each class are available. The risk of damage was associated with three rating factors; ship type, year of construction and period of operation, involving a total of 40 classes, including 6 classes with zero exposure. For Canadian private automobile liability insurance data, the number of claims incurred and exposure for each class are available. Two rating factors are considered; class and merit ratings, involving a total of 20 classes. Finally, for UK private car motor insurance data, the incurred claim count and exposure for each class are available. Four rating factors are considered; coverage, vehicle age, vehicle group and policyholder age, involving a total of 120 classes.

Bailey and Simon (1960) also suggested the average absolute difference as a suitable test for Criterion (iii),

$$\frac{\sum_i e_i \left| r_i - f_i \right|}{\sum_i e_i r_i}.$$

Therefore, the χ^2 statistics, a test for Criterion (iv), and the average absolute difference, a test for Criterion (iii), will be calculated for all models. Table 3, Table 4 and Table 5 show the parameter estimates, χ^2 statistics and average absolute difference for the models discussed above.

Table 3. Parameters, χ^2 and absolute difference for ship data

Parameters & bias measures		Multiplicative models					
		Zero bias / Poisson	Least squares	Minimum χ^2	Normal	Binomial	Minimum modified χ^2
Intercept	$\exp(\beta_1)$	0.002	0.002	0.002	0.002	0.002	0.002
Ship type *B*	$\exp(\beta_2)$	0.581	0.563	0.568	0.588	0.581	0.593
Ship type *C*	$\exp(\beta_3)$	0.503	0.436	0.781	0.317	0.503	0.231
Ship type *D*	$\exp(\beta_4)$	0.927	1.087	1.113	0.926	0.927	0.652
Ship type *E*	$\exp(\beta_5)$	1.385	1.384	1.575	1.123	1.385	1.113
Const. yr 65-69	$\exp(\beta_6)$	2.008	2.071	2.040	2.038	2.008	1.938
Const. yr 70-74	$\exp(\beta_7)$	2.267	2.157	2.242	2.395	2.267	2.242
Const. yr 75-79	$\exp(\beta_8)$	1.574	1.368	1.584	1.767	1.573	1.576
Operation yr 75-79	$\exp(\beta_9)$	1.469	1.437	1.443	1.447	1.469	1.544
χ^2		42.275	45.211	36.393	59.567	42.277	85.18
absolute difference		0.187	0.194	0.209	0.165	0.187	0.169

Parameters (10^{-3}) & bias measures		Additive models					
		Zero bias/ Least squares	Poisson	Minimum χ^2	Normal	Binomial	Minimum Modified χ^2
Intercept	β_1	2.665	2.430	2.477	0.962	2.431	2.472
Ship type *B*	β_2	-1.821	-1.565	-1.619	-0.104	-1.566	-1.596
Ship type *C*	β_3	-2.149	-1.730	-1.026	-1.081	-1.731	-2.793
Ship type *D*	β_4	-0.376	-0.651	0.433	0.868	-0.650	-1.682
Ship type *E*	β_5	1.756	2.019	2.932	2.650	2.017	0.849
Const. yr 65-69	β_6	1.094	1.055	1.097	1.108	1.055	1.003
Const. yr 70-74	β_7	1.536	1.604	1.661	1.657	1.603	1.323
Const. yr 75-79	β_8	0.453	0.714	0.747	0.968	0.713	0.665
Operation yr 75-79	β_9	0.837	0.791	0.778	0.770	0.792	0.846
χ^2		41.063	40.042	35.591	50.044	40.040	71.436
absolute difference		0.168	0.160	0.178	0.179	0.160	0.154

Table 4. Parameters, χ^2 and absolute difference for Canadian data

Parameters & bias measures		Multiplicative models					
		Zero bias / Poisson	Least squares	Minimum χ^2	Normal	Binomial	Minimum modified χ^2
Intercept	$\exp(\beta_1)$	0.080	0.081	0.080	0.079	0.080	0.080
Class 2	$\exp(\beta_2)$	1.350	1.330	1.351	1.392	1.347	1.347
Class 3	$\exp(\beta_3)$	1.599	1.586	1.598	1.628	1.597	1.599
Class 4	$\exp(\beta_4)$	1.692	1.660	1.697	1.742	1.686	1.682
Class 5	$\exp(\beta_5)$	1.241	1.223	1.242	1.286	1.238	1.238
Merit rating X	$\exp(\beta_6)$	1.313	1.307	1.312	1.334	1.312	1.314
Merit rating Y	$\exp(\beta_7)$	1.427	1.405	1.428	1.483	1.423	1.423
Merit rating B	$\exp(\beta_8)$	1.637	1.611	1.640	1.705	1.632	1.633
χ^2		577.826	625.268	577.037	754.403	580.754	583.899
absolute difference		0.028	0.032	0.028	0.020	0.028	0.028

Parameters (10^{-2}) & bias measures		Additive models					
		Zero bias/ Least squares	Poisson	Minimum χ^2	Normal	Binomial	Minimum Modified χ^2
Intercept	β_1	7.878	7.877	7.876	7.875	7.877	7.878
Class 2	β_2	3.080	3.126	3.129	3.207	3.120	3.121
Class 3	β_3	5.296	5.242	5.248	5.081	5.252	5.232
Class 4	β_4	6.489	6.529	6.531	6.637	6.521	6.523
Class 5	β_5	2.100	2.167	2.174	2.323	2.158	2.152
Merit rating X	β_6	2.793	2.757	2.760	2.697	2.762	2.751
Merit rating Y	β_7	3.827	3.858	3.861	3.938	3.853	3.850
Merit rating B	β_8	5.884	5.878	5.881	5.896	5.879	5.870
χ^2		97.829	95.926	95.904	108.302	95.970	96.100
absolute difference		0.008	0.007	0.007	0.005	0.007	0.007

Table 5. Parameters, χ^2 and absolute difference for UK data

Parameters & bias measures		Multiplicative models					
		Zero bias / Poisson	Least squares	Minimum χ^2	Normal	Binomial	Minimum modified χ^2
Intercept	$\exp(\beta_1)$	0.243	0.241	0.249	0.242	0.242	0.227
Coverage N.Comp	$\exp(\beta_2)$	0.756	0.748	0.759	0.823	0.755	0.750
Veh. age 4-7	$\exp(\beta_3)$	0.804	0.817	0.803	0.827	0.806	0.806
Veh. age 8+	$\exp(\beta_4)$	0.643	0.649	0.642	0.618	0.645	0.647
Veh. group B	$\exp(\beta_5)$	1.139	1.123	1.133	1.133	1.137	1.151
Veh. group C	$\exp(\beta_6)$	1.238	1.218	1.231	1.268	1.235	1.251
Veh. group D	$\exp(\beta_7)$	1.605	1.574	1.600	1.614	1.600	1.615
P/H age 21-24	$\exp(\beta_8)$	0.846	0.838	0.841	0.813	0.843	0.864
P/H age 25-29	$\exp(\beta_9)$	0.639	0.647	0.631	0.622	0.641	0.667
P/H age 30-34	$\exp(\beta_{10})$	0.574	0.591	0.567	0.563	0.578	0.596
P/H age 35+	$\exp(\beta_{11})$	0.514	0.524	0.505	0.509	0.516	0.541
χ^2		107.049	109.491	106.487	122.232	107.290	112.873
absolute difference		0.063	0.065	0.063	0.063	0.063	0.064

Table 5, continued. Parameters, χ^2 and absolute difference for UK data

Parameters (10^{-2}) & bias measures		Additive models					
		Zero bias/ Least squares	Poisson	Minimum χ^2	Normal	Binomial	Minimum Modified χ^2
Intercept	β_1	21.864	21.387	21.732	14.444	21.458	20.643
Coverage N.Comp	β_2	-2.878	-2.500	-2.445	-1.451	-2.538	-2.603
Veh. age 4-7	β_3	-3.253	-3.397	-3.393	-2.773	-3.380	-3.391
Veh. age 8+	β_4	-5.443	-5.463	-5.504	-5.608	-5.457	-5.379
Veh. group B	β_5	1.208	1.269	1.271	1.144	1.261	1.266
Veh. group C	β_6	2.162	2.168	2.167	2.518	2.165	2.175
Veh. group D	β_7	6.241	6.193	6.273	6.557	6.204	6.020
P/H age 21-24	β_8	-1.883	-1.736	-1.864	4.035	-1.762	-1.413
P/H age 25-29	β_9	-5.767	-5.528	-5.789	0.933	-5.567	-4.962
P/H age 30-34	β_{10}	-7.096	-6.937	-7.206	-0.396	-6.967	-6.385
P/H age 35+	β_{11}	-8.437	-7.968	-8.282	-1.398	-8.035	-7.302
χ^2		133.859	127.935	127.252	234.493	127.902	133.542
absolute difference		0.072	0.072	0.072	0.077	0.072	0.074

Several conclusions can be made regarding the programming and results of parameter estimates from the fitted models:

i. The classical approach and regression model give equivalent parameter estimates.

ii. The regression model has faster convergence.

iii. The additive models are more sensitive to initial values.

iv. Each of multiplicative and additive models produced similar parameter estimates.

8. CONCLUSIONS

This paper bridged the minimum bias and maximum likelihood methods for both additive and multiplicative models via a weighted equation. The equations for both minimum bias and maximum likelihood can be rewritten as a weighted equation, in the form of a weighted difference between observed and fitted rates. The parameter estimates could also be rewritten as a weighted solution; for multiplicative model it is in the form of a weighted proportion whereas for additive model, the form is of a weighted difference.

Applying the weighted equation for maximum likelihood and minimum bias equations has several advantages; the weighted equation is mathematically and conceptually simpler, the weighted equation also allows the usage of regression model, and finally, the weighted equation provides an initial understanding of the fitting procedure for distribution with overdispersion parameter. In addition, the weights of the parameter solutions for both multiplicative and additive models show that they have similar estimates.

Acknowledgment

The authors gratefully acknowledged the financial support received in the form of a research grant (IRPA RMK8: 09-02-02-0112-EA274) from the Ministry of Science, Technology and Innovation (MOSTI), Malaysia.

The authors are also pleasured to thank the reviewer for his interest in the paper and for his valuable comments and suggestions.

APPENDIX A

S-Plus classical programming for multiplicative model

```
Classic.multi <- function(data,b,d,g,h,k,iter=200)
{
    X <- as.matrix(data[,-(1:2)])
    count <- as.vector(data[,1])
    exposure <- as.vector(data[,2])
    rate <- count/exposure
    parameter <- dim(X)[2]
    new.beta <- rep(c(0.5), dim(X)[2])
    for (i in 1:iter)
    {
        for (j in 1:parameter)
        {
            beta <- new.beta
            fitted <- as.vector(exp(X%*%log(beta)))
            fitted.noj <- as.vector(exp(X[,-j]%*%log(beta[-j])))
            z <- as.vector(exposure^b*(rate+0.5/exposure)^d*fitted.noj^g*
                (1-fitted)^h*(rate+fitted)^k*X[,j])
            v <- as.vector(z/sum(z))
            new.beta[j] <- as.vector(sum(v*(rate/fitted.noj)))
        }
        if (all(abs(new.beta-beta)<0.0000001))
            break
    }
    fitted <- as.vector(exp(X%*%log(beta)))
    chi.square <- sum((exposure*(rate-fitted)^2)/fitted)
    abs.difference <- sum(exposure*abs(rate-fitted))/sum(exposure*rate)
```

```
    list( beta=round(beta,4), chi.square=round(chi.square,3),
            absolute.difference=round(abs.difference,3))
}
```

S-Plus classical programming for additive model

```
Classic.add <- function(data,b,d,g,h,k,iter=300)
{
    X <- as.matrix(data[,-(1:2)])
    count <- as.vector(data[,1])
    exposure <- as.vector(data[,2])
    rate <- count/exposure
    parameter <- dim(X)[2]
    new.beta <- rep(c(0.000001), dim(X)[2])
    for (i in 1:iter)
    {
        for (j in 1:parameter)
        {
            beta <- new.beta
            fitted <- as.vector(X%*%beta)
            fitted.noj <- as.vector(X[,-j]%*%beta[-j])
            z <- as.vector(exposure^b*(rate+0.5/exposure)^d*fitted^g*
                (fitted*(1-fitted))^h*(rate+fitted)^k*X[,j])
            v <- as.vector(z/sum(z))
            new.beta[j] <- as.vector(sum(v*(rate-fitted.noj)))
        }
        if (all(abs(new.beta-beta)<0.0000001))      break
    }
    fitted <- as.vector(X%*%beta)
    chi.square <- sum((exposure*(rate-fitted)^2)/fitted)
    abs.difference <- sum(exposure*abs(rate-fitted))/sum(exposure*rate)
    list( beta=round(beta,6), chi.square=round(chi.square,3),
            absolute.difference=round(abs.difference,3))
}
```

APPENDIX B

S-Plus regression programming for multiplicative model

```
Regression.multi <- function(data,b,d,g,h,k,iter=20)
{
    X <- as.matrix(data[, -(1:2)])
    count <- as.vector(data[,1])
    exposure <- as.vector(data[,2])
    rate <- count/exposure
    new.beta <- rep(c(1),dim(X)[2])
    for (i in 1:iter)
    {
        beta <- new.beta
        fitted <- as.vector(exp(X%*%beta))
        Z <- diag(fitted)%*%X
```

```
        W <-    diag(exposure^b*(rate+0.5/exposure)^d*fitted^g*(1-fitted)^h*
                (rate+fitted)^k)
        r.s <- rate-fitted+as.vector(Z%*%beta)
        new.beta <- as.vector(solve(t(Z)%*%W%*%Z)%*%t(Z)%*%W%*%r.s)
        if (all(abs(new.beta-beta)<0.0000001))
            break
    }
    fitted <- as.vector(exp(X%*%beta))
    chi.square <- sum((exposure*(rate-fitted)^2)/fitted)
    abs.difference <- sum(exposure*abs(rate-fitted))/sum(exposure*rate)
    list( beta=round(exp(beta),4), chi.square=round(chi.square,3),
            absolute.difference=round(abs.difference,3))
}
```

S-Plus regression programming for additive model

```
Regression.add <- function(data,b,d,g,h,k,iter=20)
{
    X <- as.matrix(data[, -(1:2)])
    count <- as.vector(data[,1])
    exposure <- as.vector(data[,2])
    rate <- count/exposure
    new.beta <- rep(c(0.000001),dim(X)[2])
    for (i in 1:iter)
    {
        beta <- new.beta
        fitted <- as.vector(X%*%beta)
        W <- diag(exposure^b*(rate+0.5/exposure)^d*fitted^g*(fitted*(1-fitted))^h*
                    (rate+fitted)^k)
        r.s <- rate-fitted+as.vector(X%*%beta)
        new.beta <- as.vector(solve(t(X)%*%W%*%X)%*%t(X)%*%W%*%r.s)
        if (all(abs(new.beta-beta)<0.0000001))
            break
    }
    fitted <- as.vector(X%*%beta)
    chi.square <- sum((exposure*(rate-fitted)^2)/fitted)
    abs.difference <- sum(exposure*abs(rate-fitted))/sum(exposure*rate)
    list( beta=round(exp(beta),6), chi.square=round(chi.square,3),
            absolute.difference=round(abs.difference,3))
}
```

9. REFERENCES

Ajne, B., "A note on the multiplicative ratemaking model", *ASTIN Bulletin*, (1975), Vol. 8, No. 2, 144-153.

Bailey, R.A., "Insurance rates with minimum bias", *Proceedings of the Casualty Actuarial Society*, (1963), Vol. 50, No. 93, 4-14.

Bailey, R.A. and L.J., Simon, "Two studies in automobile insurance ratemaking", *ASTIN Bulletin*, (1960), Vol. 4, No. 1, 192-217.

Brockmann, M.J. and T.S., Wright, "Statistical motor rating: Making effective use of your data", *Journal of the Institute of Actuaries*, (1992), Vol. 119, No. 3, 457-543.

Brown, R.L., "Minimum bias with generalized linear models", *Proceedings of the Casualty Actuarial Society*, (1988), Vol. 75, No. 143, 187-217.

Chamberlain, C., 1980, "Relativity pricing through analysis of variance", *Discussion Paper Program of the Casualty Actuarial Society,* (1980), 4-24.

Coutts, S.M., "Motor insurance rating, an actuarial approach", *Journal of the Institute of Actuaries*, (1984), Vol. 111, 87-148.

Freifelder, L., "Estimation of classification factor relativities: A modelling approach", *Journal of Risk and Insurance,* (1986), Vol. 53, 135-143.

Harrington, S.E., 1986, "Estimation and testing for functional form in pure premium regression models", *ASTIN Bulletin*, (1986), Vol. 16, 31-43.

Holler, K.D., D., Sommer, and G., Trahair, "Something old, something new in classification ratemaking with a novel use of GLMs for credit insurance", *Forums of the Casualty Actuarial Society,* (1999), Winter, 31-84.

Jee, B., "A comparative analysis of alternative pure premium models in the automobile risk classification system", *Journal of risk and insurance*, (1989), Vol. 56, 434-459.

Jung, J., "On automobile insurance ratemaking", *ASTIN Bulletin*, (1968), Vol. 5, No. 1, 41-48.

Lawless, J.F., "Negative binomial and mixed Poisson regression", *The Canadian Journal of Statistics*, (1987), Vol. 15, No. 3, 209-225.

McCullagh, P., and J.A., Nelder, *Generalized Linear Models (2nd Edition)*, (1989), Chapman and Hall, London.

Mildenhall, S.J., 1999, "A systematic relationship between minimum bias and generalized linear models", *Proceedings of the Casualty Actuarial Society,* (1999), Vol. 86. No. 164, 393-487.

S-PLUS 2000, *Guide to Statistics, Vol 1*, MathSoft. Inc., Seattle.

Venables, W.N., and B.D., Ripley, *Modern Applied Statistics with S-PLUS (2nd Edition)*, (1997), Springer-Verlag, New York.

Wang, W., and F., Famoye, F., "Modeling household fertility decisions with generalized Poisson regression", *Journal of Population Economics*, (1997), Vol. 10, 273-283.

Biographies of Authors

Noriszura Ismail is a lecturer in National University of Malaysia since July 1993, teaching Actuarial Science courses. She received her MSc. (1993) and BSc. (1991) in Actuarial Science from University of Iowa, and is now currently pursuing her PhD in Statistics. She has presented several papers in Seminars, including those locally and in South East Asia, and published several papers in local and Asian Journals.

Abdul Aziz Jemain is an Associate Professor in National University of Malaysia, teaching in Statistics Department since 1982. He received his MSc. (1982) in Medical Statistics from London School of Hygiene and Tropical Medicine and PhD (1989) in Statistics from University of Reading. He has written several articles, including local and international Proceedings and Journals, and co-authors several local books.