International Journal of Applied Mathematics & Statistics,
Int. J. Appl. Math. Stat.; Vol. 27; Issue No. 3; Year 2012,
ISSN 0973-1377 (Print), ISSN 0973-7545 (Online)

Mathematical Modelling for Claim Severities using Normal and *t* Copulas

Y. RESTI[1], **N. ISMAIL**[2] and **S. H. JAAMAN**[2]

[1]Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sriwijaya,
Jl. Raya Palembang-Prabumulih Km.32, Inderalaya 30662, Sumatera Selatan, Indonesia
Email : farras_yani@yahoo.com

[2]Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor Darul Ehsan, Malaysia;
Email : ni@ukm.my, shj@ukm.my

ABSTRACT

This paper proposes the normal and the t copulas for modeling the dependence of claim severities for the Malaysian motor insurance data. The claim severities are fitted using two different assumptions; the independent assumption where the claim types are assumed to be independent, and the dependent assumption where the dependence between claim types are modeled by the normal and the t copulas. The result indicates that the t copula is an improvement over the normal copula, whereas the normal copula is better than the independent model.

Keywords: Claim severities, dependence, normal copula, t copula.

Mathematics Subject Classification: 62J12, 62G99

Computing Classification System: I.4

1. INTRODUCTION

In property and liability insurance, the process of establishing premium rates involves the estimation of two crucial elements, the claim frequency and the claim severity. Claim frequency can be defined as the claim count per exposure unit, whereas claim severity can be defined as the average claim cost per claim. Statistical estimates of claim frequency and claim severity are often calculated through the process of grouping risks with similar risk characteristics, which is known as risk classification.

Based on the actuarial literature, the Poisson regression model has been considered as a standard method for fitting the claim count data (Aitkin et al. 1990, Renshaw 1994). In other areas, several models have been recommended for handling overdispersion in count data, namely the quasi-Poisson (McCullagh & Nelder 1989, Brockman & Wright 1992), the negative binomial (Cameron & Trivedi 1986, Lawless 1987), and the generalized Poisson (Consul & Famoye 1992, Wang & Famoye 1997, Ismail & Jemain 2007) regression models.

Several regression models have also been suggested for estimating the claim severities, or the average claim costs per claim. Since it is well established that the claim severity distributions generally have positive support and are positively skewed, the gamma (McCullagh & Nelder 1989, Brockman & Wright 1992, Renshaw 1994, Ismail & Jemain 2006), and the inverse Gaussian regression models (Venter 2007, Resti et al. 2010) have been used by past researchers.

In insurance practice, claim frequencies and claim severities may give rise to multiple types. As an example, the claims for motor insurance may be categorized into several types such as the Own Damage (OD), the Third Party Property Damage (TPPD) and the Third Party Bodily Injury (TPBI). A standard method for estimating the risk premium is by estimating the claim frequencies and the claim severities independently for each claim type (Brockman & Wright 1992, Renshaw 1994, Haberman & Renshaw 1996). However, the modeling of risk premium via the estimation of claim frequency and claim severity independently for each claim type assumes that the joint behavior of two or more random variables is based upon an independent assumption. In practice, the independent assumption between claim types may not be true for all cases. One may question the impact of the dependence of claim types, such as, if there is an error in one type of claim, another type of claim may be affected, and the overall risk premium may also be affected.

As an alternative, a copula model can be used for handling the dependence of claim severities. A copula model expresses the joint distribution of two or more random variables by separating the joint distribution into two contributions; the marginal distributions of the individual variables, and the interdependency of the probabilities of the individual variables. The advantage of using a copula model is that each marginal distribution can be specified in the isolation of others, and then joined by the copula.

Copula models have been applied in several areas such as finance, insurance and environmental studies. In the actuarial and insurance literatures, Frees and Valdez (1998) and Klugman and Parsa (1999) applied the copula model for the claim sizes and the allocated loss adjusted expenses, Frees and Wang (2005) handles the serial time dependence through the *t*-copula by assuming the marginal distribution for claim severity data follows the generalized linear model (GLM), Pettere and Kollo (2006) used the copula model for the outstanding claim reserves, and Frees and Wang (2006) modeled the time dependence for count data by using the elliptical copulas. For more details, introduction to copulas can be found in Frees and Valdez (1998).

The objective of this article is to propose two types of copula models, the normal copula and the *t* copula, for accommodating the dependence of claim severities in the Malaysian motor insurance data. The estimation of risk premium is implemented by having three stages of modeling; the modeling of claim frequency by using the negative binomial regression, the modeling of claim type by applying the multinomial regression, and the modeling of claim severity by fitting the trivariate distribution with the gamma regression model for the marginals, and the normal and the *t* copulas for the dependence of claim severities.

2. METHODOLOGY

Consider an insurance claim data where $i = 1,2,...,n$ is the rating class and $j = 1,2,..$ is the claim type. For each i, we denote Y_i as the random variable for the claim count and e_i as the exposure measured in a car-year unit.

For each i and j, we denote M_{ij} as the random variable for the claim type and C_{ij} as the random variable for the claim severity equivalent to the average claim cost. The claim count for each i and j is denoted by y_{ij}. The rating factors are represented by the explanatory variables, $x_{ij1}, x_{ij2}, ..., x_{ijk}$. Table 1 shows an example of an insurance claim data with three claim types and four explanatory variables.

When a claim is made, it is possible to have one type of claim only, or a combination of two claim types, or a combination of three claim types. The joint distribution of the dependent variables, (Y_i, M_{ij}, C_{ij}), can be written as,

$$f(y_i, m_{ij}, c_{ij}) = \Pr(Y_i = y_i)\Pr(M_{ij} = m_{ij} \mid n_i) f(c_{ij} \mid n_i, m_{ij}),$$

which encompasses three components; the claim frequency where $\Pr(Y_i = y_i)$ is the probability of y_i claim count, the claim type where $\Pr(M_{ij} = m_{ij} \mid y_i)$ is the conditional probability of m_{ij} claim type given y_i claim count, and the claim severity where $f(c_{ij} \mid y_i, m_{ij})$ is the conditional density of c_{ij} claim severity given y_i claim count and m_{ij} claim type.

2.1 Claim Frequency Model

The dataset for modeling the claim frequencies are (Y_i, e_i), where Y_i denotes the claim count and e_i is the exposure. The claim count is represented by both paid and estimate of outstanding, and the claim frequency is equal to the claim count divided by the exposure, $Y_i e_i^{-1}$. If Y_i is distributed as a negative binomial, the probability mass function (p.m.f.) is,

$$\Pr(Y_i = y_i \mid \mu_i, a) = \frac{\Gamma(y_i + a^{-1})}{y_i!\Gamma(a^{-1})}\left(\frac{\mu_i}{\mu_i + a^{-1}}\right)^{y_i}\left(\frac{a^{-1}}{\mu_i + a^{-1}}\right)^{a^{-1}}, \quad y_i = 0,1,2,...$$

with mean $E(Y_i) = \mu_i$ and variance $Var(Y_i) = (1 + a\mu_i)\mu_i$, where a denotes the dispersion parameter. The covariates can be incorporated through a log link function, $\mu_i = e_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is the vector of regression parameters and $\mathbf{x}_i$ is the vector of explanatory variables. The maximum likelihood estimates of $\boldsymbol{\beta}$ and a can be obtained by maximizing the log likelihood with respect to $\boldsymbol{\beta}$ and a.

Table 1: An example of insurance claim data (three claim types and four explanatory variables)

Rating class	Claim count	Exposure	Claim type $j=1$			Claim type $j=2$			Claim type $j=3$			Explanatory variables x_{ijk}			
i	Y_i	e_i	M_{i1}	Claim severity C_{i1}	Claim count y_{i1}	M_{i2}	Claim severity C_{i2}	Claim count y_{i2}	M_{i3}	Claim severity C_{i3}	Claim count y_{i3}	x_{ij1}	x_{ij2}	x_{ij3}	x_{ij4}
1	Y_1	e_1	M_{11}	C_{11}	y_{11}	M_{12}	C_{12}	y_{12}	M_{13}	C_{13}	y_{13}	x_{1j1}	x_{1j2}	x_{1j3}	x_{1j4}
2	Y_2	e_2	M_{21}	C_{21}	y_{21}	M_{22}	C_{22}	y_{22}	M_{23}	C_{23}	y_{23}	x_{2j1}	x_{2j2}	x_{2j3}	x_{2j4}
3	Y_3	e_3	M_{32}	C_{31}	y_{31}	M_{32}	C_{32}	y_{32}	M_{33}	C_{33}	y_{33}	x_{3j1}	x_{3j2}	x_{3j3}	x_{3j4}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

2.2 Claim Type Model

The dataset for modeling the claim type are M_{ij}, where M_{ij} denotes the claim type. The multinomial probability of a claim of type m is,

$$\pi_m = \Pr(M_{ij} = m) \,,$$

where $\sum_m \pi_m = 1$. For each claim type, or for each j, the covariates can be included through a logit function, $\ln\left(\frac{\pi_m}{\pi_r}\right) = \mathbf{x_i^T \beta}$, where $\mathbf{x}_i$ is the vector of explanatory variables, $\boldsymbol{\beta}$ is the vector of regression parameters and r is the base level. The maximum likelihood estimates of $\boldsymbol{\beta}$ can be obtained by maximizing the log likelihood with respect to $\boldsymbol{\beta}$.

As an example, consider a data where three claim types are observed; type 1, type 2 and type 3. Therefore, $M = 1,2,...,7$, where $M = 1$ refers to the claims of type 1 only, $M = 2$ refers to the claims of type 2 only, $M = 3$ refers to the claims of type 3 only, $M = 4$ refers to the claims of both type 1 and type 2, $M = 5$ refers to the claims of both type 1 and type 3, $M = 6$ refers to the claims of both type 2 and type 3, and $M = 7$ refers to the claims of all three types.

2.3 Claim Severity Model

The dataset for modeling the claim severities are (C_{ij}, y_{ij}), where C_{ij} denotes the claim severity or the average claim cost per claim and y_{ij} is the claim count. The claim severity, which is already adjusted and trended for inflation, is represented by both paid and estimate of outstanding. The total claim cost is equal to the product of the claim count and the average claim cost, $y_{ij}C_{ij}$. For each claim type, or for each j, if C_i is distributed as a gamma, the probability density function (p.d.f.) is,

$$f(c_i \mid \mu_i, v) = \frac{1}{c_i \Gamma(\omega)} \left(\frac{c_i \omega}{\mu_i}\right)^{\omega} \exp\left(-\frac{c_i \omega}{\mu_i}\right), \quad c_i > 0 \,,$$

with mean $E(C_i) = \mu_i$ and variance $Var(C_i) = \omega^{-1}\mu_i^2 = \sigma^2 \mu_i^2$, where $\omega = \sigma^{-2}$ denotes the scale parameter. For a case where the variance vary within the classes and depend on the claim count, the scale parameter can be written as $\omega_i = y_i \omega = y_i \sigma^{-2}$ so that $Var(C_i) = \omega^{-1} y_i^{-1} \mu_i^2 = \sigma^2 y_i^{-1} \mu_i^2$. The covariates can be included via a log link function, $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is the vector of regression parameters and $\mathbf{x}_i$ is the vector of explanatory variables. The maximum likelihood estimates of $\boldsymbol{\beta}$ and ω can be obtained by maximizing the log likelihood with respect to $\boldsymbol{\beta}$ and ω.

2.4 Normal Copula Model

Consider a two-dimensional cumulative distribution function (c.d.f), F. The idea of Sklar's Theorem is to represent the c.d.f., F, in two parts; the marginal c.d.f., F_i, and the copula c.d.f, H, which describes the form of dependence in the distribution. Both F_i and H are connected by the c.d.f.

$$F(c_1,c_2) = H(F_1(c_1),F_2(c_2)) = H(u_1,u_2),$$

where U_1 and U_2 are the standard uniform random variables. By differentiation, the corresponding probability distribution function (p.d.f.), f, is given by,

$$f(c_1,c_2) = f_1(c_1)f_2(c_2)h(u_1,u_2),$$

where f_i is the marginal p.d.f. and h the copula p.d.f.

We will fit two types of elliptical copula; the normal and the *t*. The elliptical copula is,

$$H(u_1,u_2) = F(F_1^{-1}(u_1),F_2^{-1}(u_2)),$$

and the copula of a normal joint c.d.f. is called the normal copula. If ρ is the correlation parameter, the p.d.f. of a normal copula is given by,

$$h(u_1,u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{(\rho F_1^{-1}(u_1))^2 + (\rho F_2^{-1}(u_2))^2 - 2\rho F_1^{-1}(u_1)F_2^{-1}(u_2)}{1-\rho^2}\right)\right).$$

2.5 t-Copula Model

The copula of a student-*t* joint c.d.f. is called the *t* copula. Let $T_{v,\rho}$ be the student-*t* joint c.d.f. with degrees of freedom v and correlation ρ, and T_v^{-1} be the inverse of the student-*t* univariate c.d.f. with degrees of freedom v. The *t*-copula is defined as,

$$H(u_1,u_2) = T_v(T_v^{-1}(u_1),T_v^{-1}(u_2)),$$

and the p.d.f. is given by,

$$h(u_1,u_2) = \frac{1}{2}\left(\frac{\Gamma\left(\frac{v}{1}\right)}{\Gamma\left(\frac{v+1}{1}\right)}\right)^2 \frac{v}{\sqrt{1-\rho^2}} \frac{\left(\left(1+\frac{(T_v^{-1}(u_1))^2}{v}\right)\left(1+\frac{(T_v^{-1}(u_2))^2}{v}\right)\right)^{\frac{v+1}{2}}}{\left(1+\frac{(T_v^{-1}(u_1))^2+(T_v^{-1}(u_2))^2-2\rho T_v^{-1}(u_1)T_v^{-1}(u_2)}{v\sqrt{1-\rho^2}}\right)^{\frac{v+2}{2}}}$$

2.6 Maximum Likelihood Estimation

Consider an insurance claim data where the claims are categorized into three types, and the data is fitted to a trivariate distribution with gamma marginals and a normal copula. The joint p.d.f can be written as,

$$f_{i123}(c_{i1},c_{i2},c_{i3}) = f_{i1}(c_{i1})f_{i2}(c_{i2})f_{i3}(c_{i3})h(F_{i1}(c_{i1}),F_{i2}(c_{i2}),F_{i3}(c_{i3})) ,$$

where f_{ij} is the p.d.f. associated with the *i*th rating class and *j*th claim type, and h is the p.d.f. of the trivariate normal copula. The log likelihood is,

$$\log L = \sum_i \log f_{i1} + \sum_i \log f_{i2} + \sum_i \log f_{i3} + \sum_i \log h ,$$

where each marginal can be specified in the isolation of others and joined by the copula.

If the gamma regression model is used for the marginal claim severity, the covariates can be incorporated by writing the mean as $\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, and the claim count can be included by writing the scale parameter as $\omega_i = y_i \omega = y_i \sigma^{-2}$ in the fitting procedure. The maximum likelihood estimates of the marginal parameters, $\boldsymbol{\beta}$ and ω, and the copula parameter, ρ, can be obtained by maximizing the log likelihood with respect to $\boldsymbol{\beta}$, ω and ρ.

2.7 Risk Premium

A common method for estimating the risk premium of the ith rating class, r_i, is by estimating the claim frequencies and the claim severities independently for each claim type (Brockman & Wright 1992, Renshaw 1994, Haberman & Renshaw 1996),

$$\hat{r}_i = \hat{f}_{i1}\hat{c}_{i1} + \hat{f}_{i2}\hat{c}_{i2} + \hat{f}_{i3}\hat{c}_{i3} ,$$

where $\hat{f}_{ij} = e_{ij}^{-1} E(Y_{ij})$ denotes the estimate of frequency for the jth claim type, and $\hat{c}_{ij} = E(C_{ij})$ the estimate of claim severity (average claim cost) for the jth claim type. If the claim types are assumed to be independent, the risk premium can be written as,

$$\hat{r}_i = \hat{f}_i(\hat{p}_{i1}\hat{c}_{i1} + \hat{p}_{i2}\hat{c}_{i2} + \hat{p}_{i3}\hat{c}_{i3}) ,$$

where $\hat{p}_{ij} = E(M_{ij})$ is the estimate of probability for the jth claim type.

However, if the claim types are assumed to be dependent, the risk premium can be rewritten as,

$$\hat{r}_i = \hat{f}_i(\hat{p}_{i1}\hat{c}_{i1} + \hat{p}_{i2}\hat{c}_{i2} + \hat{p}_{i3}\hat{c}_{i3} + \hat{p}_{i12}\hat{c}_{i12} + \hat{p}_{i13}\hat{c}_{i13} + \hat{p}_{i23}\hat{c}_{i23} + \hat{p}_{i123}\hat{c}_{i123}) .$$

3. RESULT AND DISCUSSION

The copula models are fitted to the Malaysian motor insurance claims experience. The database, which is supplied by Insurance Services Malaysia (ISM) Berhad, provides information on the private car insurance portfolios of ten general insurance companies in 2000-2003. The sample data contains 572,627 policies with 52,522 claims which can be categorized into three types; the Own Damage (OD), the Third Party Property Damage (TPPD) and the Third Party Bodily Injury (TPBI). Four rating factors are considered, and they are the scope of coverage, the vehicle make, the vehicle cubic capacity and the vehicle age. The rating factors and the rating classes are shown in Table 2.

Table 2: Rating factors and rating classes

No	Rating factors	Rating classes
1	Coverage	Comprehensive
		Non-comprehensive
2	Vehicle make	Local
		Foreign
3	Vehicle cubic capacity (cc)	0-1000 cc
		1001-1300 cc
		1301-1500 cc
		1501-1800 cc
		1801+ cc
4	Vehicle age	0-1 year
		2-3 year
		4-5 year
		6-7 year
		8+ year

The result of fitting the negative binomial regression model to the claim frequencies is shown in Table 3, whereas the result of fitting the multinomial logit regression model to the claim types is shown in Table 4. Based on the results in Table 3, all rating factors (scope of coverage, vehicle make, vehicle cc and vehicle age) are significant. In particular, the risk of claim frequency is higher for comprehensive coverage, local vehicle, vehicle with cubic capacity greater than 1300, and vehicle aging more than one year. The results in Table 4 indicate that the TPBI-OD-TPPD claim has the largest probability, whereas the OD claim has the smallest. Therefore, the OD claim is used as the base level.

Table 3: Negative Binomial regression model for claim frequency

Regression parameter	Negative binomial		
	est.	s.e.	*p*-value
Intercept	-9.53	0.28	0.00
Non-comprehensive	-0.88	0.27	0.00
Foreign	-1.02	0.24	0.00
0-1000 cc	-2.46	0.34	0.00
1001-1300 cc	-2.09	0.31	0.00
0-1 yr	-0.79	0.27	0.00

Table 4: Multinomial logit regression model for claim type

Combination	Type	Intercept		
		est.	s.e.	*p*-value
1	TPBI	2.03	0.08	0.00
3	TPPD	1.39	0.06	0.00
4	TPBI-OD	2.18	0.08	0.00
5	TPBI-TPPD	2.59	0.10	0.00
6	OD-TPPD	1.70	0.07	0.00
7	OD-TPPD-TPBI	2.77	0.11	0.00

The trivariate normal and *t* copulas with unstructured correlation matrix are fitted to the claim severities, and the results are shown in Table 5. The estimates obtained from fitting the marginal claim severities are used as initial values for estimating the parameters in the copula model. The log likelihood indicates that the *t* copula is an improvement over the normal copula, whereas the normal copula is better than the independent model.

Based on the log likelihood, the AIC and the BIC, the *t* copula is the best model for accommodating the dependence between the TPBI, the OD and the TPPD claim severities. The significant rating factors for the TPBI claims are the scope of coverage (non-comprehensive), the vehicle cc (0-1000cc, 1001-1300cc) and the vehicle age (0-1yr), the significant rating factors for the OD claims are the scope of coverage (non-comprehensive) and the vehicle cc (0-1000cc, 1001-1300cc, 1801+cc), whereas the significant rating factors for the TPPD claims are the scope of coverage (non-comprehensive), the vehicle cc (0-1000cc, 1001-1300cc) and the vehicle age (2-3yr). Based on the *t* copula model, the risk of severity is higher for comprehensive coverage and vehicle with cubic capacity below 1000. For the TPBI claim, the severity is higher for vehicle aging more than one year, whereas for the TPPD claim, the severity is higher for vehicle aging 2-3 years.

Table 5: Normal and *t* copulas for claim severity

Claim type	Parameter	Normal copula	t copula
		est. (s.e.)	est. (s.e.)
TPBI	Intercept	8.16 (0.30)	8.02 (0.12)
	Non-comprehensive	-0.86 (0.55)	-0.80 (0.16)
	0-1000 cc	1.28 (0.53)	1.46 (0.20)
	1001-1300 cc	1.10 (0.24)	1.26 (0.23)
	0-1 yr	-0.85 (0.27)	-1.04 (0.14)
	Scale	0.64 (0.21)	0.49 (0.05)
OD	Intercept	7.61 (0.28)	7.42 (0.07)
	Non-comprehensive	-1.32 (0.15)	-0.59 (0.06)
	0-1000 cc	1.33 (0.36)	1.12 (0.10)
	1001-1300 cc	1.07 (0.33)	0.79 (0.09)
	1801+ cc	0.53 (0.32)	0.23 (0.16)
	Scale	0.67 (0.84)	0.46 (0.00)
TPPD	Intercept	7.01 (0.16)	6.54 (0.10)
	Non-comprehensive	-0.77 (0.19)	-0.57 (0.12)
	0-1000 cc	0.84 (0.15)	1.11 (0.10)
	1001-1300 cc	0.86 (0.24)	1.05 (0.21)
	2-3yr	0.35 (0.16)	0.40 (0.22)
	Scale	0.25 (1.42)	0.19 (0.14)
Copula parameter		ρ_{12} =0.68	ρ_{12} =0.63
		ρ_{13} =0.46	ρ_{13} =0.45
		ρ_{23} =0.32	ρ_{23} =0.37
Log-likelihood		-146.16	-101.90
AIC		316.32	227.80
BIC		347.58	259.06

4. CONCLUSION

This paper accommodates the dependence between the claim severities using the normal copula and the *t*-copula. In particular, the copula models are fitted to the Malaysian motor insurance claims experience based on the data supplied by Insurance Services Malaysia (ISM) Berhad, which provides information on the private car insurance portfolios of ten general insurance companies in 2000-2003.

The main advantage of using the copula model is that each marginal distribution can be specified independently based on the distribution of the individual variable and then joined by the copula which considers the dependence of the variables. The results of fitting the independent regression models and the copula models indicate that the dependence of claim severities is significant. Based on the log likelihood, the AIC and the BIC, the *t* copula is the best model for accommodating the dependence between the TPBI, the OD and the TPPD claim severities.

Acknowledgment
The authors gratefully acknowledge the financial support received in the form of a research grant (Science Fund: 06-01-02-SF0482) from the Ministry of Science, Technology and Innovation (MOSTI), Malaysia. The authors are also pleasured to thank Insurance Services Malaysia Berhad (ISM) for supplying the data.

5. REFERENCES

Aitkin, M., Anderson, D., Francis, B. & Hinde, J., 1990. *Statistical modelling in GLIM*. New York, Oxford University Press.

Brockman, M.H. & Wright, T.S., 1992. Statistical motor rating: making effective use of your data. *Journal of the Institute of Actuaries* 119(3): 457-543.

Cameron, A.C. & Trivedi, P.K., 1986. Econometric models based on count data: comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*. 1: 29-53.

Consul, P.C. & Famoye, F., 1992. Generalized Poisson regression model. *Communication Statistics (Theory & Methodology)* 2(1): 89-109.

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

Frees, E.W. & Valdez, E.A., 1998. Understanding relationships using copulas. *North American Actuarial Journal* 2(1): 1-25.

Frees, E.W. & Wang, P. 2005. Credibility using copulas. *North American Actuarial Journal* 9(2): 31-48.

Frees, E.W. & Wang, P., 2006. Copula credibility for aggregate loss models. *Insurance: Mathematics and Economics* 8: 360-373.

Frees, E.W. & Valdez, E.A., 2008. Hierarchical insurance claims modeling. *Journal of the American Statistical Association* 103(484): 1457–1469.

Haberman, S. & Renshaw, A.E., 1996. Generalized linear models and actuarial science. *The Statistician* 45(4): 407-436.

Ismail, N. & Jemain, A.A., 2007. Handling overdispersion with negative binomial and generalized Poisson regression models. *Casualty Actuarial Society Forum* Winter: 103-158.

Ismail, N. & Jemain, A.A., 2006. A comparison of risk classification methods for claim severity data. *Journal of Modern Applied Statistical Methods* 5(2): 513-528.

Klugman, S.A. & R. Parsa, 1999. Goodness-of-fit procedure for multivariate truncated data. *Insurance: Mathematics and Economics* 24: 139-148.

Lawless, J.F., 1987. Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics* 15(3): 209-225.

Lemaire, J., 1985. *Automobile insurance actuarial models*. Boston, Kluwer Academic Publishers Group.

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

Nelsen, R.B., 2006. *An Introduction to Copulas*, 2nd Edition. New York, Springer.

Pettere, G. & Kollo, T., 2006. Modelling claim size in time via copulas. ICA Papers. Department of Mathematics, Statistics, and Philosophy, University of Tampere, Vammala. http://papers.ica2006.com/1A.html.

Renshaw, A.E., 1994. Modelling the claims process in the presence of covariates. *ASTIN Bulletin* 24(2): 265-285.

Resti, Y., Ismail, N. & Jaaman, S.H., 2010. Handling the dependence of claim severities with copula models. *Journal of Mathematics and Statistics* 6(2): 136-142.

Venter, G.G., 2007. GLM beyond the exponential family with loss reserve application. *ASTIN Bulletin* 37(2): 345-364.

Wang, W. & Famoye, F., 1997. Modeling household fertility decisions with generalized Poisson regression. *Journal of Population Economics* 10: 273-283.