

## Handling the Dependence of Claim Severities with Copula Models

<sup>1</sup>Yulia Resti, <sup>2</sup>Noriszura Ismail and <sup>2</sup>Saiful Hafizah Jaaman

<sup>1</sup>Department of Mathematics, Faculty of Mathematics and Science, University of Sriwijaya,  
Jl. Raya Palembang-Prabumulih Km.32 Inderalaya 30662 Sumatera Selatan, Indonesia

<sup>2</sup>Actuarial Science Program, School of Mathematical Sciences, Faculty of Science and Technology,  
University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

---

**Abstract: Problem statement:** Several studies have been carried out on the modeling of claim severity data in actuarial literature as well as in insurance practice. Since it is well established that the claim cost distributions generally have positive support and are positively skewed, the regression models of Gamma and Lognormal have been used by practitioners for modeling claim severities. However, the fitting of claim severities via regression models assumes that the claim types are independent. **Approach:** In this study, independent assumption between claim types will be investigated as we will consider three types of Malaysian motor insurance claims namely Third Party Body Injury (TPBI), Third Party Property Damage (TPPD) and Own Damage (OD) and applied the normal, t, Frank and Clayton copulas for modeling dependence structures between these claim types. **Results:** The AIC and BIC indicated that the Clayton is the best copula for modeling dependence between TPBI and OD claims and between TPPD and OD claims, whereas the t-copula is the best copula for modeling dependence between TPBI and TPPD claims. **Conclusion:** This study modeled the dependence between insurance claim types using copulas on the Malaysian motor insurance claim severity data. The main advantage of using copula is that each marginal distribution can be specified independently based on the distribution of individual variable and then joined by the copula which takes into account the dependence between these variables. Based on the results, the estimated of copula parameter for claim severities indicate that the dependence between claim types is significant.

**Key words:** Claim severity, claim types, dependence, copula

---

### INTRODUCTION

The pricing of premium for fire, motor and workmen compensation insurances in Malaysia is governed by their respective tariffs formulated by General Insurance Association of Malaysia (PIAM). The main objective of tariffs is to guarantee the premium rate to be at least at the level required by the Malaysian government, ensuring the price competition among local insurers to be above the market's economic level. However, one of the effects caused by the world economic crisis in 1997 is the process of liberalization spreading gradually in most financial sectors in Malaysia, including non-life insurance sector. Therefore, a thorough and comprehensive preparation towards the development of a more matured and open insurance market should be undertaken by the sector and regulatory concerned. In achieving this target, one of the main tasks that should be given serious attention

is the determination of appropriate premium rate, especially in low premium and high volume non-life insurance businesses, which can be accomplished via statistical modeling.

Statistical modeling of premium rate requires two crucial estimates; the probabilities associated with the occurrence of insured events namely claim frequency and the magnitude of such events namely claim severity.

Claim frequency can be defined as the number of claims per exposure unit, whereas claim severity is the average claim cost per claim. Based on the actuarial literature, statistical estimates of claim frequency and severity are often calculated through the process of grouping risks with similar risk characteristics for the purpose of establishing "fair" premium price, known as risk classification.

Several studies have been carried out on the modeling of claim severity in actuarial literature as well

---

**Corresponding Author:** Noriszura Ismail, Actuarial Science Program, School of Mathematical Sciences,  
Faculty of Science and Technology, University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor,  
Malaysia

as in insurance practice. Since it is well established that the claim cost distributions generally have positive support and are positively skewed, the distributions of Gamma and Lognormal have been used by practitioners for modeling claim severities. For examples, McCullagh and Nelder (1989) fit the UK own damage costs for privately owned and comprehensively insured vehicles using Gamma regression model by assuming the coefficient of variation is constant within classes and the mean is incorporated in the model via an inverse link function, Brockman and Wright (1992) fit the UK own damage costs for comprehensive motor policies also to the Gamma regression via a log link function, Renshaw (1994) fit the UK motor insurance claim severity also to the Gamma regression via a log link function and Ismail and Jemain (2009) fit the Gamma and Inverse Gaussian regressions via the log, linear and inverse link functions to the Malaysian motor claim costs data. As a comparison, several actuarial studies also reported claim severity results from Normal distribution via Box-Cox transformation and one such example can be found in Harrington (1986) who fitted two types of motor insurance data, the UK and the Massachusetts data.

However, the fitting of claim severities through regression models assumes that the claim types are independent. In this study, such assumption will be investigated as we will consider three types of Malaysian motor insurance claims namely Third Party Body Injury (TPBI), Third Party Property Damage (TPPD) and Own Damage (OD) and applied copula for modeling the dependence structures between these claim types. In other words, instead of implementing a traditional univariate claim analysis, we will perform a bivariate analysis for claim severity data, taking into account the possibility of damage in an accident which resulted in more than one claim types and taking into consideration the impact of dependence of one claim type on another claim type incurred out of the same accident.

Copula model expresses the joint distribution of two or more random variables by separating the joint distribution into two contributions; the marginal distributions of individual variables and the interdependency of probabilities of individual variables. An advantage of copula is that each marginal distribution can be specified in 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 actuarial literature, Frees and Valdez (1998) and Klugman and Parsa (1999) applied copulas for modeling claim sizes and allocated loss adjusted expenses, Frees and Wang (2005) handles

serial time dependences through t-copula by assuming the marginal distribution for claim severity data follows a Generalized Linear Model (GLM) and Frees and Wang (2006) model time dependencies for count data by using elliptical copulas. Introductions to copulas can be found in Frees and Valdez (1998).

The objective of this study is to model the dependence between insurance claim types using copula. The copula models are applied on the Malaysian motor insurance claim severity data which is divided into three types namely TPBI, TPPD and OD. Specifically, two stages of fitting will be involved. First, the TPBI, TPPD and OD claim severities are fitted independently to the regression models of Gamma and Inverse Gaussian. Then, for investigating the dependence between claim types, the TPBI, TPPD and OD claim severities are fitted, compared and tested on the Normal and t-copulas which belong to the elliptical families and the Clayton and Frank copulas which belong to the Archimedean families.

### MATERIALS AND METHODS

**Gamma regression model:** Let  $C_i$  be the random variable for claim severity or equivalently the claim cost for the  $i$ -th risk class,  $i = 1, 2, \dots, n$ , where  $n$  denotes the number of risk classes. If  $C_i$  follows a gamma distribution, the probability density function (pdf) is:

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

with mean,  $E(C_i) = \mu_i$  and variance,  $\text{Var}(C_i) = v^{-1}\mu_i^2$ , where  $v$  denotes the index parameter. To incorporate covariates and to ensure non-negativity, the mean is included in the regression model via a log link function,  $\log(\mu_i) = x_i^T \beta$ , where  $x_i$  denotes the vector of explanatory variables and  $\beta$  the vector of regression parameters. The regression parameters,  $\beta$  and the index parameter,  $v$ , can be fitted using maximum likelihood procedure.

**Inverse Gaussian regression model:** If  $C_i$  is distributed as Inverse Gaussian distribution, the pdf is:

$$f(c_i; \mu_i, \sigma) = \frac{1}{\sigma\sqrt{2\pi c_i^3}} \exp\left[-\frac{1}{2c_i} \left(\frac{c_i - \mu_i}{\mu_i \sigma}\right)^2\right], \quad c_i > 0 \quad (2)$$

with mean,  $E(C_i) = \mu_i$  and variance,  $\text{Var}(C_i) = \sigma^2\mu_i^3$ , where  $\sigma$  denotes the scale parameter. The mean is also

included in the regression model via a log link function,  $\log(\mu_i) = x_i^T \beta$  and similar to the Gamma regression, the regression parameters,  $\beta$  and the scale parameter,  $\sigma$ , for the Inverse Gaussian can be fitted using maximum likelihood procedure.

**Normal copula:** The idea of Sklar's Theorem for a two-dimensional cumulative distribution function (cdf),  $F$ , is to represent the function into two parts; the marginal cdf,  $F_i$  and the copula,  $H$ , which describes the form of dependence in the distribution. Both  $F_i$  and  $H$  are connected by the cdf:

$$F(c_1, c_2) = H(F_1(c_1), F_2(c_2)) = H(u_1, u_2) \tag{3}$$

where,  $U_1$  and  $U_2$  denote the standard uniform random variables.

By differentiation, the corresponding probability distribution function (pdf) is given by,

$$f(c_1, c_2) = f_1(c_1) f_2(c_2) h(u_1, u_2) \tag{4}$$

where,  $f_i$  is the marginal pdf and  $h$  the copula pdf.

We will fit two families of copula; the elliptical and the Archimedean. An elliptical copula corresponds to an elliptical distribution by the Sklar's Theorem. Let  $F$  be the multivariate cdf of an elliptical distribution, whereas let  $F_i$  be the cdf of the  $i$ th margin and  $F_i^{-1}$ ,  $i=1,2$ , be the inverse function (or the quantile function). The elliptical copula is:

$$H(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) \tag{5}$$

The copula of a normal joint cdf is called normal copula. If  $\rho$  is the correlation parameter, the pdf of a normal copula is given by:

$$h(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2} \left[ \frac{(\rho F^{-1}(u_1))^2 + (\rho F^{-1}(u_2))^2 - 2\rho F^{-1}(u_1)F^{-1}(u_2)}{1-\rho^2} \right] \right\} \tag{6}$$

The likelihood function depends on the association of copula and marginals. As an example, if the observed severity arise from the first and second claim types where the marginal density functions respectively are  $f_1(c_1)$  and  $f_2(c_2)$ , the contribution to the likelihood can be written as:

$$L_{12} = f_{12}(c_1, c_2) = f_1(c_1) f_2(c_2) h(u_1, u_2) \tag{7}$$

The marginal parameters and copula parameter (correlation parameter) can be obtained using maximum likelihood procedure.

**t-copula:** The copula of a student t joint cdf is called t-copula. If  $\rho$  is the correlation parameter and  $v$  is the degrees of freedom, the pdf is given by:

$$h(u_1, u_2) = \frac{1}{2} \left[ \frac{\Gamma\left(\frac{v}{1}\right)}{\Gamma\left(\frac{v+1}{1}\right)} \right]^2 \left\{ \frac{v}{\sqrt{1-\rho^2}} \frac{\left( \left( 1 + \frac{(F_1^{-1}(u_1))^2}{v} \right)^{\frac{v+1}{2}} \right) \left( 1 + \frac{(F_1^{-1}(u_2))^2}{v} \right)^{\frac{v+1}{2}}}{\left( (F_1^{-1}(u_1))^2 + (F_1^{-1}(u_2))^2 - \frac{v+2}{2} \right)^{\frac{v+2}{2}}} \right\} \tag{8}$$

Similar to the normal copula, the marginal parameters and copula parameter (correlation parameter) for t-copula can be obtained using maximum likelihood procedure.

**Clayton copula:** An Archimedean copula is constructed through a generator,  $\phi$ , as:

$$H(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)) \tag{9}$$

where,  $\phi^{-1}$  is the inverse of the generator and  $U_1$  and  $U_2$  are standard uniform random variables. A generator uniquely determines an Archimedean copula.

The generator of Clayton copula with space parameter,  $\alpha$ , is given by:

$$\phi(u) = u^{-\alpha} - 1 \tag{10}$$

and the inverse is:

$$\phi^{-1}(u) = (u + 1)^{-1/\alpha} \tag{11}$$

From the generator,  $\phi$ , and the inverse,  $\phi^{-1}$ , the copula of Clayton can be obtained:

$$H(u_1, u_2) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha} \tag{12}$$

and the pdf is:

$$h(u_1, u_2) = \alpha^2 \frac{\Gamma\left(\frac{1}{\alpha} + 2\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \cdot (u_1^{-\alpha-1})(u_2^{-\alpha-1})(u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-\frac{1}{\alpha}-2} \quad (13)$$

The marginal parameters and copula parameter (space parameter) can be obtained using maximum likelihood procedure.

**Frank copula:** The generator of Frank copula with space parameter,  $\alpha$ , is given by:

$$\varphi(u) = \ln\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha}}\right) \quad (14)$$

and the inverse is:

$$\varphi^{-1}(u) = -\frac{1}{\alpha} \ln(1 + e^u(e^{-\alpha} - 1)) \quad (15)$$

From the generator,  $\varphi$  and the inverse,  $\varphi^{-1}$ , the copula of Frank can be obtained:

$$H(u_1, u_2) = -\frac{1}{\alpha} \ln\left[1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)(e^{-\alpha} - 1)}{(e^{-\alpha} - 1)^2}\right] \quad (16)$$

and the pdf is:

$$h(u_1, u_2) = \frac{-\alpha(e^{-\alpha u_1} + 1)(e^{-\alpha u_2} + 1)}{(e^{-\alpha} - 1) \left((e^{-\alpha} - 1) + e^{-\alpha u_1} e^{-\alpha u_2}\right)^2} \quad (17)$$

Similar to the Clayton copula, the marginal parameters and copula parameter (space parameter) for Frank copula can be obtained using maximum likelihood procedure.

## RESULTS

The database for the Malaysian claim severities, which is supplied by Insurance Services Malaysia Berhad (ISM), provides information on private car insurance portfolios in years 2000-2003. The sample data contains 572,627 policies with 52,522 (9.17%) claims which can be categorized into three claim types; OD, TPPD and TPBI. The risk of each claim is associated with four rating factors namely scope of coverage, vehicle make, vehicle cubic capacity and vehicle year and the rating factors and classes are shown in Table 1. The best regression models for gamma and inverse Gaussian, each for OD, TPPD and TPBI claims, are presented in Table 2-4.

Table 1: Rating factors and rating classes

Rating factors	Rating classes
Coverage	Comprehensive Non-comprehensive
Vehicle make	Local Foreign
Vehicle cubic capacity (cc)	0-1000 cc 1001-1300 cc 1301-1500 cc 1501-1800 cc 1801+cc
Vehicle age	0-1 year 2-3 year 4-5 year 6-7 year 8+year

Table 2: Fitted Gamma and inverse Gaussian for OD severities

Parameter	Gamma regression			Inverse Gaussian regression		
	est.	s.e	p-value	est.	s.e	p-value
Intercept	7.42	0.07	0.00	7.50000	0.06	0.00
$\beta_1$ (non-comprehensive)	-0.48	0.08	0.00	-0.73000	0.08	0.00
$\beta_2$ (0-1000 cc)	1.04	0.11	0.00	1.30000	0.13	0.00
$\beta_3$ (1001-1300 cc)	0.76	0.11	0.00	0.97000	0.11	0.00
$\beta_4$ (1801+ cc)	0.65	0.11	0.00	0.59000	0.09	0.00
Index/scale	$v = 0.16$	0.02	-	$\sigma = 0.00006$	0.00	-
Log likelihood	-793.86			-776.11000		
AIC	1599.72			1564.22000		
BIC	1615.23			1579.73000		

Table 3: Fitted gamma and inverse Gaussian for TPPD severities

Parameter	Gamma regression inverse			Gaussian regression		
	est.	s.e	p-value	est.	s.e	p-value
Intercept	6.54	0.10	0.00	6.510	0.14	0.00
$\beta_1$ (non-comprehensive)	-0.55	0.13	0.00	-0.670	0.17	0.00
$\beta_2$ (0-1000 cc)	1.12	0.17	0.00	1.270	0.31	0.00
$\beta_3$ (1001-1300 cc)	1.12	0.17	0.00	1.170	0.30	0.00
$\beta_4$ (2-3 year)	0.41	0.17	0.01	-	-	-
Index/scale	$v = 0.41$	0.06	-	$\sigma = 0.001$	0.00	-
Log likelihood	-729.51			-739.938		
AIC	1471.01			1491.877		
BIC	1486.40			1507.263		

**Table 4:** Fitted Gamma and inverse Gaussian for TPBI severities

Parameter	Gamma regression			Inverse Gaussian regression		
	est.	s.e	p-value	est.	s.e	p-value
Intercept	8.020	0.12	0.000	10.530	0.21	0.00
$\beta_1$ (Non-comprehensive)	-1.010	0.15	0.000	-	-	-
$\beta_3$ (0-1000 cc)	1.440	0.20	0.000	0.940	0.33	0.00
$\beta_4$ (1001-1300 cc)	1.240	0.19	0.000	-	-	-
$\beta_7$ (0-1 year)	-1.140	0.19	0.000	-2.050	0.67	0.00
$\beta_8$ (2-3 year)	-	-	-	1.450	0.27	0.00
Index/scale	$\nu = 0.500$	0.07	-	$\sigma = 0.001$	0.00	-
Log likelihood	-843.841		-892.653			
AIC	1699.683		1799.306			
BIC	1715.131		1817.329			

**Table 5:** Fitted copulas for TPBI and OD severities

Claim type	Parameter	Normal copula	t-copula	Frank copula	Clayton copula
		est. (s.e.)	est. (s.e.)	est. (s.e.)	est. (s.e.)
TPBI	Intercept	8.02 (0.25)	7.90 (0.15)	7.21 (0.19)	7.33 (0.12)
	$\beta_1$ (Non-comprehensive)	-0.90 (0.13)	-0.85 (0.13)	-1.01 (0.14)	-0.82 (0.12)
	$\beta_3$ (0-1000 cc)	1.43 (0.15)	1.42 (0.15)	1.47 (0.15)	1.56 (0.16)
	$\beta_4$ (1001-1300 cc)	1.35 (0.64)	1.35 (0.45)	1.49 (0.45)	1.08 (0.42)
	$\beta_7$ (0-1 yr)	-1.34 (0.78)	-1.31 (0.71)	-1.17 (0.49)	-1.09 (0.41)
	Index, $\nu$	0.46 (0.49)	0.49 (0.45)	0.42 (0.61)	0.46 (0.45)
OD	Intercept	6.59 (0.16)	6.68 (0.15)	6.47 (0.39)	7.29 (0.24)
	$\beta_1$ (Non-comprehensive)	-0.65 (0.51)	-0.64 (0.44)	-0.68 (0.28)	-0.71 (0.12)
	$\beta_3$ (0-1000 cc)	1.24 (0.51)	1.25 (0.47)	1.34 (0.64)	1.38 (0.42)
	$\beta_4$ (1001-1300 cc)	1.07 (0.92)	1.05 (0.21)	1.10 (0.76)	1.06 (0.19)
	$\beta_6$ (1801+ cc)	0.88 (0.94)	0.89 (0.25)	0.90 (0.66)	0.84 (0.17)
	Scale, $\sigma$	0.00005 (0.51)	0.00006 (0.30)	0.00005 (0.32)	0.00005 (0.60)
Copula parameter		$\rho = 0.50$	$\rho = 0.59$	$\alpha = 0.31$	$\alpha = .75$
Log-likelihood		-918.51	-911.48	-928.30	-903.41
AIC		1861.02	1846.96	1880.60	1830.82
BIC		1892.28	1878.22	1911.86	1862.08

**Table 6:** Fitted copulas for TPBI and TPPD severities

Claim type	Parameter	Normal copula	t-copula	Frank copula	Clayton copula
		est. (s.e.)	est. (s.e.)	est. (s.e.)	est. (s.e.)
TPBI	Intercept	8.02 (0.29)	8.02 (0.12)	8.02 (0.12)	8.02 (0.12)
	$\beta_1$ (Non-comprehensive)	-0.93 (0.03)	-0.90 (0.15)	-0.93 (0.15)	-0.94 (0.15)
	$\beta_3$ (0-1000 cc)	1.33 (0.27)	1.35 (0.20)	1.51 (0.22)	1.54 (0.20)
	$\beta_4$ (1001-1300 cc)	1.22 (0.06)	1.22 (0.19)	1.39 (0.19)	1.36 (0.19)
	$\beta_7$ (0-1 yr)	-0.89 (0.22)	-0.85 (0.19)	-1.03 (0.22)	-1.03 (0.19)
	Index, $\nu$	0.50 (0.40)	0.52 (0.07)	0.43 (0.07)	0.43 (0.07)
TPPD	Intercept	6.54 (0.34)	6.54 (0.10)	6.54 (0.12)	6.54 (0.10)
	$\beta_1$ (Non-comprehensive)	-0.45 (0.91)	-0.41 (0.13)	-0.57 (0.13)	-0.35 (0.14)
	$\beta_3$ (0-1000 cc)	1.11 (0.55)	1.16 (0.17)	1.20 (0.16)	1.31 (0.17)
	$\beta_4$ (1001-1300 cc)	1.19 (0.19)	1.21 (0.17)	1.43 (0.13)	1.02 (0.12)
	$\beta_8$ (2-3 yr)	0.71 (0.80)	0.64 (0.17)	0.67 (0.17)	0.32 (0.17)
	Index, $\nu$	0.35 (0.16)	0.36 (0.06)	0.31 (0.06)	0.38 (0.05)
Copula parameter		$\rho = 0.54$	$\rho = 0.57$	$\alpha = 0.36$	$\alpha = 0.75$
Log-likelihood		-399.98	-395.03	-419.79	-401.88
AIC		823.96	814.06	863.58	827.76
BIC		855.22	845.32	894.84	859.02

The dependence between claim types is investigated by fitting the marginal distribution first, followed by fitting the copula models, so that the parameter estimates obtained from fitting the marginals can be used as initial values for estimating the parameters in copulas. In particular, the marginal distribution for the OD, TPPD and TPBI severities

are the inverse Gaussian, Gamma and Gamma regression models respectively. As for the copula models, the claim severities are fitted to the normal, t, Frank and Clayton copulas. Table 5-7 summarized the results of fitting copula models respectively to the TPBI-OD claims, TPBI-TPPD claims and TPPD-OD claims.

**Table 7:** Fitted copulas for TPPD and OD severities

Claim type	Parameter	Normal copula est. (s.e.)	t-copula est. (s.e.)	Frank copula est. (s.e.)	Clayton copula est. (s.e.)
TPPD	Intercept	6.20 (0.09)	6.35 (0.09)	8.12 (0.28)	6.71 (0.09)
	$\beta_1$ (NonComprehensive)	-0.48 (0.24)	-0.41 (0.20)	-0.57 (0.42)	-0.35 (0.59)
	$\beta_3$ (0-1000 cc)	1.08 (0.73)	1.08 (0.25)	1.05 (0.32)	1.23 (0.30)
	$\beta_4$ (1001-1300 cc)	1.15 (0.31)	1.16 (0.15)	1.32 (0.29)	0.92 (0.19)
	$\beta_8$ (2-3 yr)	0.74 (0.29)	0.67 (0.24)	0.66 (0.22)	0.35 (0.24)
	Index, v	0.36 (0.47)	0.37 (0.22)	0.35 (0.11)	0.40 (0.38)
OD	Intercept	7.45 (0.25)	7.08 (0.13)	6.82 (1.303)	7.15 (0.47)
	$\beta_1$ (NonComprehensive)	-0.65 (0.01)	-0.62 (0.08)	-0.77 (0.080)	-0.71 (0.07)
	$\beta_3$ (0-1000 cc)	1.26 (0.44)	1.20 (0.41)	1.33 (0.76)	1.29 (0.41)
	$\beta_4$ (1001-1300 cc)	1.03 (0.28)	0.97 (0.17)	1.04 (0.33)	1.03 (0.13)
	$\beta_6$ (1801+ cc)	0.92 (0.72)	0.87 (0.70)	0.86 (0.23)	0.84 (0.12)
	Scale, $\sigma$	0.00005 (0.52)	0.00006 (0.25)	0.00005 (0.31)	0.00005 (0.20)
Copula parameter		$\rho = 0.50$	$\rho = 0.58$	$\alpha = 0.35$	$\alpha = 0.77$
Log-likelihood		-865.15	-860.76	-864.76	-850.30
AIC		1754.29	1745.52	1753.52	1724.60
BIC		1785.55	1776.79	1784.79	1755.87

**DISCUSSION**

Based on the results in **Table 2**, the rating factors for scope of coverage and vehicle cubic capacity are significant for OD claim severities. In particular, the risks for non-comprehensive coverage and vehicle with 1301-1800 c.c. are lower compared to others. The log likelihood, AIC and BIC shows that the inverse Gaussian is a better model compared to the Gamma.

The results in **Table 3** illustrated that the rating factors for scope of coverage, vehicle cubic capacity and vehicle year are significant for TPPD claim severities. Specifically, non-comprehensive coverage, vehicle with more than 1300 c.c. and vehicle aged 0-1 and 4+ years have lower risks. Based on the log likelihood, AIC and BIC, the Gamma is a better model compared to the inverse Gaussian.

Similar to the TPPD claim severities, the significant rating factors for TPBI claim severities, as shown in **Table 4**, are scope of coverage, vehicle cubic capacity and vehicle year. Comparison based on the log likelihood, AIC and BIC shows that the Gamma is a better model compared to the inverse Gaussian. For Gamma model, the risks for non-comprehensive coverage, vehicle with 1301+ c.c. and vehicle aged 0-1 year are lower compared to others.

The estimates of copula parameter (correlation coefficient,  $\rho$ , or space,  $\alpha$ ) for TPBI-OD severities, TPBI-TPPD severities and TPPD-OD severities shown in **Table 5-7** indicate that the dependence between claim types is significant. In particular, the log likelihood in **Table 5-7** shows that the t-copula is a better model compared to the normal copula for elliptical family, whereas the Clayton copula is an improvement over the Frank copula for Archimedean family. Based on the AIC and BIC, the Clayton copula

is the best model for accommodating the dependence between TPBI and OD claim severities and between TPPD and OD claim severities, whereas the t-copula is the best distribution for modeling the dependence between TPBI and TPPD claim severities.

**CONCLUSION**

This study models the dependence between insurance claim types using copulas on the Malaysian motor insurance claim severity data which were divided into three types; TPBI, TPPD and OD. Four types of copulas namely normal, t, Frank and Clayton are fitted to the severity data. One main advantage of using copula is that each marginal distribution can be specified independently based on the distribution of individual variable and then joined by the copula which takes into account the dependence between these variables. The marginal distribution selected for the TPBI, TPPD and OD claim severities respectively are the Gamma, Gamma and inverse Gaussian regression models. Based on the log likelihood, the t-copula is superior than the Normal and the Clayton copula is better than the Frank for all TPBI-OD, TPBI-TPPD and TPPD-OD claim severities. The AIC and BIC indicate that the Clayton is the best copula for modeling TPBI-OD and TPPD-OD severities, whereas the t-copula is the best copula for modeling TPBI-TPPD 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 pleased to thank

Insurance Services Malaysia Berhad (ISM) for supplying the data.

#### REFERENCES

- Brockman, M.H. and T.S. Wright, 1992. Statistical motor rating: Making effective use of your data. *J. Inst. Actuar.*, 119: 457-543.
- Frees, E.W. and E.A. Valdez, 1998. Understanding relationship using copulas. *North Am. Actuar. J.*, 2: 1-25.
- Frees, E.W. and P. Wang, 2005. Credibility using copulas. *North Am. Actuarial J.*, 9: 31-48.
- Frees, E.W. and P. Wang, 2006. Copula credibility for aggregate loss models. *Insur.: Math. Econ.*, 38: 360-373.
- Harrington, S.E., 1986. Estimation and testing for functional form in pure premium regression models. *ASTIN Bull.*, 16: 31-43.
- Ismail, N. and A.A. Jemain, 2009. Comparison of minimum bias and maximum likelihood methods for claim severity. *Casualty Actuar. Soc. E-Forum*, Winter, 243-275.
- Klugman, S.A. and R. Parsa, 1999. Goodness-of-fit procedure for multivariate truncated data. *Insurance: Math. Econ.*, 24: 139-148.
- McCullagh, P. and J.A. Nelder, 1989. *Generalized Linear Models*. 2nd Edn., Chapman and Hall, pp: 296-300.
- Renshaw, A.E., 1994. Modeling the claims process in the presence of covariates. *ASTIN Bull.*, 24: 265-285.