ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 3

A PARAMETRIC MODEL FOR HEALTH CARE CLAIMS

BY

JOHN P. KMETIC

Department of Actuarial and Management Sciences
Faculty of Management
University of Manitoba
Winnipeg, Manitoba

September, 1993

Abstract

This paper develops a parametric model for health care insurance claims using currently available software to fit parametric models to the data. Model selection is based on the Chi-squared goodness-of-fit test and a comparison of the empirical limited expected value function to parametric limited expected value functions. We make use of a copula function to model the bivariate portion of the data.

Acknowledgments

Throughout the duration of this project I have received much assistance and guidance from several sources. I would like to take this opportunity to thank them all.

I gratefully acknowledge the donor of the data, without them this project would not have been possible. I would also like to thank Rajesh Barnwal for some of our late evening discussions, they provided a good testing ground for some of my ideas. Last but not least, much appreciation is extended to Jacques Carriere for the use of his office and for his guidance and input throughout each stage of the paper.

John Kmetic

List of Tables

List of Abbreviations

AB	Alberta
BC	British Columbia
cdf	cumulative distribution function
LEV	limited expected value
MB	Manitoba
ON	Ontario
pdf	probability density function
PQ	Quebec

Contents

1. INTRODUCTION

Health care insurance consists of many different types of coverage under one umbrella. Some examples of the types of events covered are drugs, dental, ambulance, chiropractor, and the list goes on. Until now the data has been analyzed without the use of parametric models. The purpose of this paper will be to attempt to fit a parametric model to such data. We were attempting to find a practical model, unfortunately for the model we present here it may not always be possible to find the moments of the distribution. A simplifying assumption of independence is suggested which would allow, if they exist, the calculation of the moments.

We were also interested in testing some of the tools available to fit parametric models. In particular we used Klugman's [5] FIT software for these purposes. Appendix I contains the distributions which were available to choose from in the software. The software will graph the data, graph the fitted distributions, perform a Chi-squared goodness-of-fit test, calculate the empirical limited expected value (LEV) function , and calculate the model LEV function. A quasi-Newton-Raphson method is employed to calculate the parameter estimates and an alternative method, called the simplex method, can be used when the quasi-Newton-Raphson method fails. Typically one would try the quasi-Newton-Raphson method after using the alternative method if the estimates have converged.

Finally once we have the parametric model, we try to use it to learn some more about the dependence structure of the data.

During the development of this paper it became obvious that we could manipulate the grouping of the data to get a good Chi-squared p-value. Some attention is given to this problem but no solution is presented as it is beyond the scope of this paper.

2. DATA CHARACTERISTICS

The data was supplied by an anonymous donor on two diskettes, one for males and one for females. The data were for 1991 health care claims for small employers of 5 to 35 people with claims recorded for the insured only. We assumed only one claim per year per category could occur. Each data record included the date of birth of the insured, province of residence, and the claims made on the policy categorized by claim type. There were 23 different types of claims that could be made, including a miscellaneous medical supplies category. The claim amounts recorded were those covered by the insured's policy before any deductibles and co-insurance. This paper will concentrate on the female's claim data.

Inspection of the data revealed that the Drugs category was by far the most common claim type, accounting for more claims than the other categories combined. The remaining categories were very sparse, and so it was decided to combine these categories in any further analysis. Lastly, we did not consider age to be a factor, as a plot of age versus claim amount did not reveal any discernible pattern. Adding age as a variable would also create other problems, such as deciding on the grouping and it would also make the model less parsimonious. We decided to group by province, since each province has a different benefit system. Differing levels of tax and cost of services by province would also be a factor. Note that we eliminated Saskatchewan and the Atlantic provinces as the data was very sparse for these areas. We are left with female data for 1991, grouped by province, with claims categorized by Drugs and Other.

2.1 Number of Claims

A check of the independence of a claim in Drugs versus a claim in Other revealed that the two categories were indeed dependent.

Contingency tables for Claim versus No Claim for Drugs by Other are given in Tables 1 to 5 for each province.

TABLE 1

Claim Contingency Table for British Columbia

Drugs	Other		
	No Claim	Claim	Total
No Claim	3	85	88
Claim	721	224	945
Total	724	309	1033

TABLE 2

Claim Contingency Table for Alberta

Drugs	Other		
	No Claim	Claim	Total
No Claim	3	12	15
Claim	671	56	727
Total	674	68	742

TABLE 3

Claim Contingency Table for Manitoba

Drugs	Other		
	No Claim	Claim	Total
No Claim	1	82	83
Claim	524	124	648
Total	525	206	731

TABLE 4

Claim Contingency Table for Ontario

Drugs	Other		
	No Claim	Claim	Total
No Claim	8	140	148
Claim	2969	373	3342
Total	2977	513	3490

TABLE 5

Claim Contingency Table for Quebec

Drugs	Other		
	No Claim	Claim	Total
No Claim	8	209	217
Claim	1469	550	759
Total	1477	759	2236

In all cases a Chi-square test of independence of the rows and columns of the contingency table failed with a p-value of 0. All calculations were done using BMDP [1] Program 4F. Thus the occurrence/nonoccurrence of a claim in Drugs is not independent of the occurrence/nonoccurrence of a claim in Other. This implies that we have four probabilities to consider:

1. No claim occurs in either category.
2. A claim occurs in Drugs but not in Other.
3. A claim occurs in Other but not in Drugs.
4. A claim occurs in both categories, Drugs and Other.

We can now see that there will be four components to the probability density function of this data; a mass at (0,0), a marginal distribution for Drugs conditional on no claim occurring in Other, a marginal distribution for Other conditional on no claim occurring in Drugs, and a bivariate distribution for which both Drugs and Other are strictly positive.

2.2 Claim Severity

Number (4) above is the bivariate component of the distribution. We performed a nonparametric test of independence of the claim severity, by taking all those insureds for which a claim occurred in both the Drugs category and the Other category. The Spearman Rank Correlation Coefficients are calculated for each province in Table 6.

TABLE 6

Spearman's ρ by Province

Province	$\hat{\rho}$	p-value
British Columbia	0.1302	0.0510
Alberta	0.2523	0.0583
Manitoba	0.0511	0.5716
Ontario	-0.0799	0.1230
Quebec	0.1494	0.0004

The calculations were done using BMDP [1] Program 3D. The p-value is under the following hypothesis:

H_0: The severity of Drugs is independent of the severity of Other.

vs.

H_A: The severity of Drugs is not independent of the severity of Other.

We can see that only Manitoba has a marginally high p-value, the remaining provinces have evidence to reject the hypothesis. Thus we conclude that the distribution of the claim severity of Drugs and Other, conditional on both being positive, is dependent.

3. MODEL

3.1 Likelihood Equation and Copula Function

In section 2 it was found that there would be four components for each province making up the distribution of this data.

1. Let P_{00} = Probability of no claims in either category, and let N_{00} = the number of such claims.
2. Let P_{01} = Probability of no claim in Drugs and a claim in Other, and let N_{01} = the number of such claims.
3. Let P_{10} = Probability of a claim in Drugs and no claim in Other, and let N_{10} = the number of such claims.
4. Let P_{11} = Probability of a claim in Drugs and a claim in Other, and let N_{11} = the number of such claims.

Note that $\sum\sum P_{ij} = 1$ and let $\sum\sum N_{ij} = N$.

Further let X = Drugs and Y = Other, and denote the marginal probability density function (pdf) of X as $f_X(x)$ and the marginal cumulative distribution function (cdf) of X as $F_X(x)$. Similarly let the marginal pdf and marginal cdf of Y be denoted as $f_Y(y)$ and $F_Y(y)$ respectively. Let the joint pdf and cdf of X and Y be denoted as $f^*_{XY}(x,y)$ and $F^*_{XY}(x,y)$ respectively. The asterix indicates that these two functions are conditional on both X and Y being strictly positive.

The complete information likelihood equation is given by:

$$L = P_{00}^{N_{00}} \{P_{10}^{N_{10}} \prod_{i \in D_{10}} f_X(x_i, \theta_x)\}\{P_{01}^{N_{01}} \prod_{i \in D_{01}} f_Y(y_i, \theta_y)\}\{P_{11}^{N_{11}} \prod_{i \in D_{11}} f^*_{XY}(x_i, y_i, \theta_{xy})\}$$

Using the fact that the P_{ij}'s sum to one, and working with the log-likelihood we find that the maximum likelihood estimate of P_{ij} is

$$\hat{P}_{ij} = \frac{N_{ij}}{N} \qquad (1)$$

We chose to group the observations, since there was a lot of clustering present in the data. In section 3.5 we will discuss the grouping in more detail. Since we grouped the data we are actually working with the grouped likelihood which is given by the product of the following three equations.

$$L(\theta_x)=\prod_{i=1}^{m_x}[F_X(c_i,\theta_x)-F_X(c_{i-1},\theta_x)]^{f_i} \tag{2}$$

$$L(\theta_y)=\prod_{j=1}^{m_y}[F_Y(c_j,\theta_y)-F_Y(c_{j-1},\theta_y)]^{g_j} \tag{3}$$

$$L(\theta_{xy})=\prod_{k=1}^{m_{xy}}[F_{XY}(c_k,d_k,\theta_{xy})-F_{XY}(c_{k-1},d_{k-1},\theta_{xy})]^{h_k} \tag{4}$$

Note that we have m_x, m_y, and m_{xy} groupings each with f_i, g_j, and h_k observations. We can easily use Klugman's [5] FIT software to find the parameter estimates for each of equations (2) and (3) above, but equation (4) presents a more difficult problem. One problem is the choice of distribution, and the other problem is finding the parameter estimates for the chosen bivariate distribution. Unlike the univariate case where it is easy to fit many distributions, it is computationally expensive to do this in the bivariate case. To avoid these problems we decided to use a Copula function to model the dependence structure. Avoiding the theoretical details, we find that the Copula function $C(u,v)$ can be used to model the dependence structure of a Bivariate distribution by making use of the marginal cumulative distribution functions of the two variables. If $H(x,y)$ is the bivariate cdf of a continuous distribution with continuous marginals $F(x)=H(x,\infty)$ and $G(y)=H(\infty,y)$ then there exists a unique distribution function $C(u,v)$ such that $H(x,y) = C(F(x),G(y))$.

Since no test exists for testing one-parameter families of copulas we used an ad hoc method using the relationships between Spearman's ρ and Kendall's τ, derived in Carriere [2], for given one-parameter families of copulas. For the Morgenstern family we find that $\tau=g(\rho)$ with

$$g(\rho)=\frac{2}{3}\rho$$

For the mixed Frechet copula, we find that $\tau=g(\rho)$ with

$$g(\rho)=\frac{(2\rho+\rho^{5/3})}{3}$$

For the normal copula, we find that $\tau=g(\rho)$ with

$$g(\rho)=\frac{2}{\pi}\arcsin\{2\sin\{\rho\pi/6\}\}$$

Using $\hat{\rho}$ we calculated the functions $g(\hat{\rho})$ for each of the copulas and compared them to $\hat{\tau}$ in Table 7.

TABLE 7

Ad-hoc Copula Test

			$g(\hat{\rho})$		
Province	$\hat{\rho}$	$\hat{\tau}$	$\frac{2}{3}\hat{\rho}$	$\frac{2\hat{\rho}+\hat{\rho}^{5/3}}{3}$	$\frac{2}{\pi}\arcsin\{2\sin\{\hat{\rho}\pi/6\}\}$
BC	0.1302	0.0876	0.0868	0.0979	0.0870
AB	0.2523	0.1723	0.1682	0.2018	0.1697
MB	0.0511	0.0328	0.0341	0.0364	0.0341
ON	-0.0799	-0.0503	-0.0533	-0.0582	-0.0533
PQ	0.1494	0.1013	0.0996	0.1136	0.0999

We can see that both the Morgenstern copula and the normal copula closely match the function $\hat{\tau}=g(\hat{\rho})$. Both $\hat{\rho}$ and $\hat{\tau}$ were calculated using BMDP [1] Program 3S. Note that the values for $\hat{\rho}$ in Table 7 are the same as those given earlier in Table 6 of section 2.2. We will work with the Morgenstern copula, as it has a closed form for the distribution function. Mardia [6] gives the distribution function of the one parameter Morgenstern copula as

$$H(x,y)=F(x)G(y)[1+\alpha\{1-F(x)\}\{1-G(y)\}] \qquad |\alpha|\leq 1 \tag{5}$$

For continuous random variables, the density is given by

$$h(x,y)=f(x)g(y)[1+\alpha\{2F(x)-1\}\{2G(y)-1\}] \qquad |\alpha|\leq 1 \tag{6}$$

Following Carriere [2], we will re-parametrize with $\alpha=3\rho$, so that we require $|\rho|\leq 1/3$. Note that in Table 7 $|\hat{\rho}|\leq 1/3$ for every province.

Note that there are many theoretical questions left unanswered regarding the use of the copula function in place of the maximum likelihood estimator. We may want to know about the efficiency and consistency of the estimator. These questions will not be addressed in this paper. Note also that we grouped the data, and so we used estimates from the grouped likelihood rather than the complete information likelihood.

The pdf of our model is given by

$$h(x,y)=\begin{cases} P_{00} & x=0, y=0 \\ P_{10} f_X(x) & x>0, y=0 \\ P_{01} f_Y(y) & x=0, y>0 \\ P_{11} \dfrac{\partial^2 C(F_X^*(x), F_Y^*(y))}{\partial x \partial y} & x>0, y>0 \end{cases} \tag{7}$$

where the last partial derivative is given by (6) with $F(x)$ and $G(y)$ replaced by $F_X^*(x)$ and $F_Y^*(y)$ respectively.

We are now ready to give estimates of f_X, f_Y, f_X^*, and f_Y^*.

3.2 Estimates for the Number of Claims.

Using equation (1) given in section 3.1, and the contingency tables given in section 2.1, we can find the maximum likelihood estimates of P_{00}, P_{10}, P_{01}, and P_{11}. The estimates for each province are summarized in Table 8.

TABLE 8

Maximum Likelihood Estimates for the Number of Claims

Province	P_{00}	P_{10}	P_{01}	P_{11}
British Columbia	0.00290	0.69797	0.08228	0.21684
Alberta	0.00404	0.90431	0.01617	0.07547
Manitoba	0.00137	0.71683	0.11218	0.16963
Ontario	0.00229	0.85072	0.04011	0.10688
Quebec	0.00358	0.65698	0.09347	0.24597

We can see that in all five provinces P_{00} the estimated probability of no claims in either Drugs or Other, is very small. It is also apparent that a claim in Drugs and no claim in Other is the dominant form of a claim.

3.3 The Marginal Distributions.

When we solve the 2 systems of grouped likelihood equations defined by (2) and (3) in 3.1, we are finding the marginal distributions f_X and f_Y for Drugs and Other respectively. The fitting was done using Klugman's [5] FIT software. The software is able to accept ungrouped data of up to 3,000 observations or grouped data. Following Hogg and Klugman [3], we will use the Limited Expected Value (LEV) function to compare the best models.

The distribution, estimated parameter values, and corresponding p-values for British Columbia (BC) - Drugs are given in Table 9.

TABLE 9

Marginal Models for British Columbia - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.139$	$\hat{\lambda}=102.5$		0.8990
Weibull	$\hat{\tau}=1.073$	$\hat{\lambda}=119.7$		0.8444
Burr	$\hat{\alpha}=9.666$	$\hat{\lambda}=823.0$	$\hat{\gamma}=1.142$	0.8560
Inverse Burr	$\hat{\tau}=0.3221$	$\hat{\lambda}=182.0$	$\hat{\gamma}=2.706$	0.5972
Generalized Pareto	$\hat{\alpha}=21.79$	$\hat{\lambda}=2{,}028$	$\hat{\tau}=1.204$	0.8548
Transformed Gamma	$\hat{\alpha}=1.454$	$\hat{\lambda}=72.75$	$\hat{\tau}=0.8732$	0.8516
Transformed Beta	$\hat{\alpha}=9.938$	$\hat{\lambda}=846.9$	$\hat{\gamma}=1.135$	0.7217
	$\hat{\tau}=1.008$			

We can eliminate the Inverse Burr and Transformed Beta distributions from further consideration, as their p-values are relatively low compared to the other distributions in Table 3. The Transformed Gamma model can also be eliminated from further discussion, since it does not provide a better fit than the two parameter Gamma model. The Limited Expected Value (LEV) function for each or the four remaining models, as well as the empirical LEV function, are given in Table 10.

TABLE 10

Empirical and Model LEV Functions for BC - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Burr	Generalized Pareto
48	40.58	40.48	40.29	40.59	40.66
95	67.54	67.09	66.88	67.18	67.22
142	85.12	84.60	84.47	84.54	84.58
189	96.56	96.02	95.96	95.81	95.86
236	104.0	103.4	103.4	103.1	103.2
283	108.9	108.2	108.2	107.9	108.0
330	112.1	111.3	111.2	111.1	111.1
∞	121.0	116.7	116.4	117.7	117.5

It is clear that all 4 models are equally good, with the Burr and the Generalized Pareto having the best fit in the tail of the LEV function. We select the Gamma model, since it has the best p-value and also provided a good fit to the LEV function.

TABLE 11

Marginal Models for Alberta - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.050$	$\hat{\lambda}=111.1$		0.9785
Weibull	$\hat{\tau}=1.031$	$\hat{\lambda}=118.9$		0.9766
Transformed Gamma	$\hat{\alpha}=1.005$	$\hat{\lambda}=117.7$	$\hat{\tau}=1.027$	0.9529

The distributions and their parameter values for Alberta (AB) - Drugs, are given in Table 11. We were only able to find three models that would fit the data, so all three are compared via their LEV functions in Table 12.

TABLE 12

Empirical and Model LEV Functions for AB - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Transformed Gamma
32	28.35	28.25	28.26	28.22
57	45.71	45.61	45.68	45.58
82	59.64	59.61	59.77	59.60
108	71.39	71.28	71.56	71.29
133	80.57	80.27	80.64	80.29
158	87.93	87.49	87.96	87.52
184	93.99	93.49	94.04	93.53
209	98.62	98.09	98.72	98.13
234	102.3	101.8	102.5	101.8
∞	117.6	116.7	117.4	116.4

It is apparent in Table 12 that all three distributions fit the empirical LEV function quite well. The Weibull distribution has the best fit of all three but the Gamma is also very good. We select either the Gamma or Weibull distributions as the best models, eliminating the Transformed Gamma on the basis of parsimony.

In Table 13, the distributions and parameter estimates, and p-values are given for Manitoba (MB) - Drugs.

TABLE 13

Marginal Models for Manitoba - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.266$	$\hat{\lambda}=94.13$		0.6698
Weibull	$\hat{\tau}=1.151$	$\hat{\lambda}=119.3$		0.8446
Inverse Burr	$\hat{\tau}=0.3120$	$\hat{\lambda}=180.9$	$\hat{\gamma}=2.967$	0.9654
Transformed Gamma	$\hat{\alpha}=0.5630$	$\hat{\lambda}=184.6$	$\hat{\tau}=1.727$	0.9930

The LEV functions of the four distributions are compared in Table 14.

TABLE 14

Empirical and Model LEV Functions for MB - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Inverse Burr	Transformed Gamma
25	23.01	23.21	23.18	22.92	22.97
48	40.81	41.00	41.00	40.72	40.75
73	56.84	56.75	56.84	56.77	56.75
95	68.35	67.97	68.15	68.29	68.25
130	82.48	81.71	81.95	82.28	82.28
160	91.16	90.35	90.56	90.85	90.89
∞	110.5	115.4	113.5	118.9	109.3

The Transformed Gamma model provides the best fit in the tail of the LEV function, the next best fit is the Weibull model. It appears that the best two parameter choice is the Weibull model and the best three parameter choice, and the overall best choice is the Transformed Gamma model.

The parameter estimates and p-values for Ontario (ON) - Drugs are given in Table 15.

TABLE 15

Marginal Models for Ontario - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.143$	$\hat{\lambda}=126.0$		0.9486
Weibull	$\hat{\tau}=1.082$	$\hat{\lambda}=147.6$		0.8567
Burr	$\hat{\alpha}=13.18$	$\hat{\lambda}=1,412$	$\hat{\gamma}=1.121$	0.8741
Generalized Pareto	$\hat{\alpha}=48.06$	$\hat{\lambda}=5,848$	$\hat{\tau}=1.164$	0.8900
Transformed Gamma	$\hat{\alpha}=1.257$	$\hat{\lambda}=111.1$	$\hat{\tau}=0.9449$	0.8951
Transformed Beta	$\hat{\alpha}=728.3$	$\hat{\lambda}=116,400$	$\hat{\gamma}=0.9485$	0.7381
	$\hat{\tau}=1.251$			

We can eliminate the Transformed Gamma as it does not provide a better fit than the simpler two parameter Gamma model. We can also eliminate the Transformed Beta distribution, citing parsimony and the relatively low p-value. The remaining distributions are compared via their LEV functions in Table 16.

TABLE 16

Empirical and Model LEV Functions for ON - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Burr	Generalized Pareto
50	43.39	43.40	43.26	43.41	43.45
100	74.69	74.45	74.31	74.45	74.48
145	94.65	94.33	94.27	94.29	94.32
190	109.1	108.6	108.6	108.5	108.6
235	119.4	118.9	118.9	118.8	118.8
280	126.8	126.2	126.2	126.1	126.1
∞	145.9	144.0	143.2	145.2	144.7

All four distributions provide a very good fit to the empirical LEV function. The Burr model provides the best fit in the tail by a marginal amount. We select the Gamma model since it had the highest p-value and very closely matches the empirical LEV function.

The parameter estimates and p-values for Quebec (PQ) - Drugs are given in Table 17.

TABLE 17

Marginal Models for Quebec - Drugs

Distribution	Parameter Values			χ^2 p-value
Log-logistic	$\hat{\gamma}=2.177$	$\hat{\lambda}=89.48$		0.9822
Inverse Gamma	$\hat{\alpha}=2.455$	$\hat{\lambda}=191.4$		0.9788
Burr	$\hat{\alpha}=1.175$	$\hat{\lambda}=99.43$	$\hat{\gamma}=1.990$	0.9930
Inverse Burr	$\hat{\tau}=0.7103$	$\hat{\lambda}=110.8$	$\hat{\gamma}=2.292$	0.9933
Generalized Pareto	$\hat{\alpha}=2.662$	$\hat{\lambda}=18.92$	$\hat{\tau}=11.39$	0.9853
Inverse Transformed Gamma	$\hat{\alpha}=3.538$	$\hat{\lambda}=392.9$	$\hat{\tau}=0.7890$	0.9822
Transformed Beta	$\hat{\alpha}=0.9914$	$\hat{\lambda}=111.3$	$\hat{\gamma}=2.310$	0.9706
	$\hat{\tau}=0.6981$			

We can eliminate the Transformed Beta for the sake of parsimony, and we can also eliminate the Inverse Transformed Gamma since it does not provide a better fit than the Inverse Gamma. The five remaining distributions are compared in Table 18.

TABLE 18

Empirical and Model LEV Functions for PQ - Drugs

Upper Limit	Empirical LEV	Log-logistic	Inverse Gamma	Burr	Inverse Burr	Generalized Pareto
95	69.66	73.93	76.75	73.10	72.26	75.62
210	104.4	103.7	106.2	103.2	102.5	105.5
325	115.5	113.9	116.4	113.4	112.7	115.7
435	120.2	118.5	121.0	118.0	117.2	120.3
545	122.9	121.2	123.7	120.5	119.7	122.9
650	124.5	122.8	125.3	122.1	121.3	124.4
760	125.6	124.1	126.5	123.2	122.4	125.5
∞	128.8	130.2	131.5	128.0	127.4	129.6

We select the Log-logistic as the best model, since it is the best two parameter model and fits the empirical LEV as well as the Burr model, which has the highest p-value and best fits the empirical LEV function.

The distribution, estimated parameter values, and corresponding p-values for British Columbia (BC) - Other are given in Table 19.

TABLE 19

Marginal Models for British Columbia - Other

Distribution	Parameter Values			χ^2 p-value
Lognormal	$\hat{\mu}=3.879$	$\hat{\sigma}=1.417$		0.8720
Inverse Gaussian	$\hat{\mu}=121.2$	$\hat{\lambda}=37.59$		0.7939
Pareto	$\hat{\alpha}=1.638$	$\hat{\lambda}=94.08$		0.8406
Gamma	$\hat{\alpha}=0.4625$	$\hat{\lambda}=233.4$		0.8669
Weibull	$\hat{\tau}=0.6491$	$\hat{\lambda}=81.79$		0.9780
Burr	$\hat{\tau}=7.075$	$\hat{\lambda}=1{,}107$	$\hat{\gamma}=0.7209$	0.9679
Inverse Burr	$\hat{\tau}=0.2612$	$\hat{\lambda}=196.3$	$\hat{\gamma}=1.802$	0.9859
Generalized Pareto	$\hat{\alpha}=2.970$	$\hat{\lambda}=389.7$	$\hat{\tau}=0.6019$	0.9747
Transformed Gamma	$\hat{\alpha}=1.787$	$\hat{\lambda}=21.42$	$\hat{\tau}=0.4851$	0.9619
Transformed Beta	$\hat{\alpha}=0.7167$	$\hat{\lambda}=175.6$	$\hat{\gamma}=2.251$	0.9615
	$\hat{\tau}=0.1967$			

We can eliminate the Lognormal, Inverse Gaussian, Pareto, Gamma, Inverse Burr and the Transformed Beta models since they all have a very poor fit in the tail of the LEV function. The remaining models are compared in Table 20.

TABLE 20

Empirical and Model LEV Functions for BC - Other

Upper Limit	Empirical LEV	Weibull	Burr	Generalized Pareto	Transformed Gamma
50	37.14	32.77	33.30	32.98	33.43
100	57.36	52.42	52.88	52.60	52.99
150	70.87	65.91	66.15	65.89	66.27
195	79.54	74.82	74.86	74.61	75.03
245	86.90	82.33	82.21	81.94	82.42
295	92.48	88.05	87.86	87.56	88.10
345	96.85	92.49	92.30	91.98	92.55
450	103.6	99.01	98.98	98.65	99.22
∞	115.3	111.9	116.9	119.1	114.8

We select the Weibull model as the best two parameter model, and the Transformed Gamma model as the best three parameter model and overall best model. The Transformed Gamma model has the best fit to the LEV of all the models and a very high p-value.

The parameter estimates and the model p-values for Alberta (AB) - Other are given in Table 21.

TABLE 21

Marginal Models for Alberta - Other

Distribution	Parameter Values			χ^2 p-value
Lognormal	$\hat{\mu}=4.887$	$\hat{\sigma}=1.128$		0.8889
Inverse Gaussian	$\hat{\mu}=246.0$	$\hat{\lambda}=129.1$		0.8190
Pareto	$\hat{\alpha}=16.30$	$\hat{\lambda}=3,175$		0.8795
Log-logistic	$\hat{\gamma}=1.437$	$\hat{\lambda}=132.8$		0.8521
Gamma	$\hat{\alpha}=0.9870$	$\hat{\lambda}=204.8$		0.8690
Weibull	$\hat{\tau}=0.9872$	$\hat{\lambda}=201.6$		0.8704
Burr	$\hat{\alpha}=2.335$	$\hat{\lambda}=336.8$	$\hat{\gamma}=1.171$	0.6765
Inverse Burr	$\hat{\tau}=0.5946$	$\hat{\lambda}=214.4$	$\hat{\gamma}=1.735$	0.6507
Generalized Pareto	$\hat{\alpha}=3.281$	$\hat{\lambda}=417.0$	$\hat{\tau}=1.274$	0.6831
Transformed Gamma	$\hat{\alpha}=4.938$	$\hat{\lambda}=3.122$	$\hat{\tau}=0.4053$	0.6991

The Lognormal, Inverse Gaussian, and Log-logistic all have a very poor fit in the tail of the LEV function. The Burr, Inverse Burr, Generalized Pareto and Transformed Gamma all have relatively low p-values and can also be eliminated. The remaining three models are compared in Table 22.

TABLE 22

Empirical and Model LEV Functions for AB - Other

Upper Limit	Empirical LEV	Pareto	Gamma	Weibull
75	63.42	62.33	62.48	62.41
145	104.1	102.7	103.1	103.0
220	134.4	133.0	133.6	133.5
295	155.9	154.2	154.7	154.6
∞	190.6	207.5	202.1	202.7

All three models over estimate the LEV function in the tail. The Pareto seems a little further out than the Weibull and the Gamma, so we suggest taking either the Gamma or the Weibull as the model of choice.

The parameter estimates and p-values for Manitoba (MB) - Other are given in Table 23.

TABLE 23

Marginal Models for Manitoba - Other

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.275$	$\hat{\lambda}=86.70$		0.9055
Weibull	$\hat{\tau}=1.161$	$\hat{\lambda}=115.4$		0.8519
Inverse Burr	$\hat{\tau}=0.5129$	$\hat{\lambda}=132.7$	$\hat{\gamma}=2.198$	0.6005
Generalized Pareto	$\hat{\alpha}=1.543$	$\hat{\lambda}=1,209$	$\hat{\tau}=1.344$	0.8121
Transformed Gamma	$\hat{\alpha}=1.851$	$\hat{\lambda}=48.07$	$\hat{\tau}=0.7961$	0.8327

The Inverse Burr has a relatively low p-value, and the Transformed Gamma does not give a much better fit than the two parameter Gamma model, so it is not considered further. The remaining models are compared via their LEV functions in Table 24.

TABLE 24

Empirical and Model LEV Functions for MB - Other

Upper Limit	Empirical LEV	Gamma	Weibull	Generalized Pareto
20	18.88	18.92	18.84	18.97
40	35.15	35.20	35.07	35.27
60	48.64	48.90	48.78	48.91
85	62.35	62.79	62.76	62.71
110	73.34	73.70	73.77	73.51
190	95.86	94.76	94.93	94.50
∞	112.0	110.5	109.5	112.6

The Generalized Pareto has the best fit in the tail of the LEV function, but by a marginal amount. We suggest either the Gamma model or the Weibull model, as both fit the empirical LEV function quite well and have high p-values.

The parameter estimates and p-values for Ontario (ON) - Other are given in Table 25 below.

TABLE 25

Marginal Models for Ontario - Other

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.246$	$\hat{\lambda}=200.6$		0.6862
Weibull	$\hat{\alpha}=1.146$	$\hat{\lambda}=261.1$		0.7635
Burr	$\hat{\alpha}=39.14$	$\hat{\lambda}=6,056$	$\hat{\gamma}=1.161$	0.6108
Inverse Burr	$\hat{\tau}=0.3275$	$\hat{\lambda}=383.9$	$\hat{\gamma}=2.905$	0.9778
Transformed Gamma	$\hat{\alpha}=0.9599$	$\hat{\lambda}=272.6$	$\hat{\tau}=1.175$	0.6072

We can eliminate the Burr and Transformed Gamma as they have relatively low p-values. The remaining distributions are compared via their LEV functions in Table 26.

TABLE 26

Empirical and Model LEV Functions for ON - Other

Upper Limit	Empirical LEV	Gamma	Weibull	Inverse Burr
55	50.66	51.10	50.92	50.56
180	136.5	135.1	135.2	135.8
290	183.4	180.3	180.9	181.8
380	207.0	204.0	204.9	205.2
480	223.8	221.1	222.0	221.7
620	237.8	235.1	235.6	235.1
∞	259.6	250.0	248.7	260.0

The Inverse Burr model has the best fit to the LEV function and the highest p-value. Although the Inverse Burr model has the highest p-value and provides the best fit to the empirical LEV function, the first two moments do not exist, making it a rather useless model. We suggest the Weibull model as the best choice.

The parameter estimates and p-values for Quebec (PQ) - Other are given in Table 27.

TABLE 27

Marginal Models for Quebec - Other

Distribution	Parameter Values			χ^2 p-value
Lognormal	$\hat{\mu}=4.979$	$\hat{\sigma}=0.9728$		0.7515
Generalized Pareto	$\hat{\alpha}=3.322$	$\hat{\lambda}=286.3$	$\hat{\tau}=1.878$	0.7613
Transformed Gamma	$\hat{\alpha}=17.65$	$\hat{\lambda}=0.001222$	$\hat{\tau}=0.2436$	0.8530

The LEV functions of all three distributions are compared in Table 28.

TABLE 28

Empirical and Model LEV Functions for PQ - Other

Upper Limit	Empirical LEV	Lognormal	Generalized Pareto	Transformed Gamma
58	52.84	54.43	53.79	53.99
116	94.07	95.25	94.80	94.88
175	124.3	124.9	125.0	125.0
233	145.9	146.1	146.6	146.7
290	161.9	161.7	162.4	162.6
360	176.8	176.0	176.7	177.2
600	206.6	203.4	203.2	203.7
∞	226.3	233.1	231.5	225.5

We select the Transformed Gamma as the best model since it has the highest p-value and the best fit to the LEV function. The best two parameter choice is the Lognormal model.

3.4 The Conditional Marginal Distributions.

Recall that we need the conditional marginal distribution functions $F_X^*(x)$ and $F_Y^*(y)$ for the copula function. We will use Klugman's [5] FIT software to find the conditional marginal densities. We again chose to group the data. In section 3.5 we will discuss the grouping of the data in more detail. Following Hogg and Klugman [3], we will also use the Limited Expected Value (LEV) function to compare the best models.

The distribution, estimated parameter values, and corresponding p-values for British Columbia (BC) - Drugs are given in Table 29.

TABLE 29

Conditional Marginal Models for British Columbia - Drugs

Distribution	Parameter Values			χ^2 p-value
Pareto	$\hat{\alpha}=339.2$	$\hat{\lambda}=48,520$		0.9144
Gamma	$\hat{\alpha}=1.004$	$\hat{\lambda}=142.7$		0.9150
Weibull	$\hat{\tau}=1.001$	$\hat{\lambda}=143.4$		0.9149
Burr	$\hat{\alpha}=29.56$	$\hat{\lambda}=3,905$	$\hat{\gamma}=1.018$	0.8281
Inverse Burr	$\hat{\tau}=0.3369$	$\hat{\lambda}=222.3$	$\hat{\gamma}=2.423$	0.7898
Generalized Pareto	$\hat{\alpha}=40.31$	$\hat{\lambda}=5,531$	$\hat{\tau}=1.024$	0.8282
Transformed Gamma	$\hat{\alpha}=1.095$	$\hat{\lambda}=127.8$	$\hat{\tau}=0.9488$	0.8285
Transformed Beta	$\hat{\alpha}=10.76$	$\hat{\lambda}=1,292$	$\hat{\gamma}=1.135$	0.6835
	$\hat{\tau}=0.8672$			

We can eliminate the Burr, Inverse Burr, and Transformed Beta distributions form further consideration, as their p-values are relatively low compared to the other distributions. The Transformed Gamma model and the Generalized Pareto model can also be eliminated from further discussion, since they do not provide a better fit than the two parameter models. The Limited Expected Value (LEV) function for each or the three remaining models as well as the empirical LEV function are given in Table 30.

TABLE 30

Empirical and Model LEV Functions for BC - Drugs

Upper Limit	Empirical LEV	Pareto	Gamma	Weibull
45	38.97	38.61	38.65	38.64
90	67.40	66.80	66.89	66.87
135	88.09	87.40	87.52	87.49
180	103.4	102.4	102.6	102.5
220	113.5	112.4	112.5	112.5
265	121.8	120.7	120.8	120.8
310	127.7	126.8	126.9	126.9
∞	145.7	143.4	143.3	143.3

It is clear that all 3 models are equally good. We recommend either the Gamma or the Weibull model, and because the Gamma model was selected for the unconditional marginal distribution, we will choose it as the model of choice.

TABLE 31
Conditional Marginal Models for Alberta - Drugs

Distribution	Parameter Values			χ^2 p-value
Lognormal	$\hat{\mu}=4.840$	$\hat{\sigma}=1.089$		0.9793
Log-logistic	$\hat{\gamma}=1.497$	$\hat{\lambda}=126.9$		0.9720
Gamma	$\hat{\alpha}=1.073$	$\hat{\lambda}=175.9$		0.9463
Weibull	$\hat{\tau}=1.036$	$\hat{\lambda}=191.2$		0.9403
Burr	$\hat{\alpha}=1.732$	$\hat{\lambda}=223.6$	$\hat{\gamma}=1.297$	0.9373
Inverse Burr	$\hat{\tau}=0.6722$	$\hat{\lambda}=182.1$	$\hat{\gamma}=1.731$	0.9311
Generalized Pareto	$\hat{\alpha}=2.944$	$\hat{\lambda}=290.4$	$\hat{\tau}=1.481$	0.9403
Transformed Gamma	$\hat{\alpha}=8.370$	$\hat{\lambda}=0.1976$	$\hat{\tau}=0.3216$	0.9435

The distributions and their parameter values for Alberta (AB) - Drugs, are given in Table 31. The Log-logistic, Burr, Inverse Burr, and Generalized Pareto can all be eliminated, as they have a very poor fit in the tail of the LEV function. The Transformed Gamma distribution can also be eliminated as it does not provide a better fit than the two parameter Gamma model. The three remaining distributions are compared in Table 32.

TABLE 32
Empirical and Model LEV Functions for AB - Drugs

Upper Limit	Empirical LEV	Lognormal	Gamma	Weibull
60	52.50	53.90	52.14	51.96
110	85.09	86.17	84.79	84.52
160	109.6	110.1	109.8	109.5
210	128.4	128.3	128.8	128.6
295	151.2	150.8	151.4	151.2
∞	182.8	228.7	188.8	188.5

All 3 distributions over estimate the empirical LEV function in the tail and we eliminate the Lognormal since it is the worst. We select either the Gamma or Weibull distributions as the best model. In Table 33, the distributions and parameter estimates, and p-values are given for Manitoba (MB) - Drugs.

TABLE 33

Conditional Marginal Models for Manitoba - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.360$	$\hat{\lambda}=71.34$		0.6871
Weibull	$\hat{\tau}=1.241$	$\hat{\lambda}=102.2$		0.8025
Inverse Burr	$\hat{\tau}=0.2176$	$\hat{\lambda}=169.9$	$\hat{\gamma}=4.252$	0.8811
Transformed Gamma	$\hat{\alpha}=0.4011$	$\hat{\lambda}=178.4$	$\hat{\tau}=2.405$	0.8789

The LEV functions of the four distributions are compared in Table 34.

TABLE 34

Empirical and Model LEV Functions for MB - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Inverse Burr	Transformed Gamma
15	14.27	14.42	14.40	14.17	14.21
27	24.53	24.83	24.83	24.44	24.50
42	35.90	36.31	36.40	36.01	36.05
59	47.15	47.37	47.62	47.49	47.46
80	59.00	58.50	58.94	59.35	59.21
96	66.48	65.37	65.93	66.76	66.54
117	74.36	72.64	73.28	74.51	74.26
136	79.72	77.81	78.45	79.83	79.60
168	85.79	84.23	84.73	85.90	85.73
∞	97.98	97.00	95.38	95.45	92.39

The Gamma model provides the best fit in the tail of the LEV function, but has the lowest p-value. We recommend the Weibull as the best two parameter model and the Transformed Gamma as the best 3 parameter model, to be consistent with the models selected in the case of the unconditional marginals.

The parameter estimates and p-values for Ontario (ON) - Drugs are given in Table 35.

TABLE 35

Conditional Marginal Models for Ontario - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.271$	$\hat{\lambda}=121.7$		0.9726
Weibull	$\hat{\tau}=1.171$	$\hat{\lambda}=161.1$		0.9663
Burr	$\hat{\alpha}=13.25$	$\hat{\lambda}=1,337$	$\hat{\gamma}=1.200$	0.9459
Inverse Burr	$\hat{\tau}=0.4713$	$\hat{\lambda}=197.5$	$\hat{\gamma}=2.306$	0.8905
Transformed Gamma	$\hat{\alpha}=1.274$	$\hat{\lambda}=121.4$	$\hat{\tau}=0.9985$	0.9515

We can eliminate the Transformed Gamma as it does not provide a better fit than the simpler two parameter Gamma model. We can also eliminate the Inverse Burr distribution since it has a very poor fit in the tail of the LEV function. The remaining distributions are compared via their LEV functions in Table 36.

TABLE 36

Empirical and Model LEV Functions for ON - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Burr
16	15.53	15.56	15.52	15.54
28	26.47	26.48	26.41	26.44
39	35.83	35.88	35.79	35.84
50	44.53	44.69	44.60	44.66
65	55.49	55.79	55.73	55.77
80	65.54	65.88	65.86	65.89
99	77.13	77.30	77.36	77.35
114	85.35	85.34	85.47	85.41
130	93.12	93.05	93.24	93.14
152	102.4	102.3	102.6	102.4
174	110.4	110.3	110.6	110.4
203	119.2	119.0	119.3	119.1
235	126.9	126.7	127.0	126.8
∞	151.8	154.7	152.5	155.1

All three distributions provide a very good fit to the empirical LEV function. The Weibull model provides the best fit in the tail by a marginal amount. We select either the Gamma or the Weibull models since both have similar p-value and very closely match the empirical LEV function.

The parameter estimates and p-values for Quebec (PQ) - Drugs are given in Table 37.

TABLE 37

Conditional Marginal Models for Quebec - Drugs

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.042$	$\hat{\lambda}=127.4$		0.9320
Weibull	$\hat{\tau}=1.024$	$\hat{\lambda}=134.0$		0.9293
Burr	$\hat{\alpha}=15.15$	$\hat{\lambda}=1{,}680$	$\hat{\gamma}=1.058$	0.9005
Inverse Burr	$\hat{\tau}=0.3476$	$\hat{\lambda}=200.2$	$\hat{\gamma}=2.475$	0.8937
Generalized Pareto	$\hat{\alpha}=2.836$	$\hat{\lambda}=3{,}412$	$\hat{\tau}=1.072$	0.8983
Transformed Gamma	$\hat{\alpha}=1.145$	$\hat{\lambda}=112.8$	$\hat{\tau}=0.9443$	0.8955

We can eliminate the Transformed Gamma, since it does not provide a better fit than the simpler model. The remaining models are compared in Table 38.

TABLE 38

Empirical and Model LEV Functions for PQ - Drugs

Upper Limit	Empirical LEV	Gamma	Weibull	Burr	Inverse Burr	Generalized Pareto
28	25.48	25.42	25.39	25.47	25.23	25.47
37	32.59	32.55	32.51	32.61	32.36	32.61
50	42.07	42.04	41.99	42.11	41.89	42.11
61	49.40	49.38	49.33	49.45	49.29	49.44
76	58.45	58.46	58.40	58.51	58.47	58.49
92	67.14	67.07	67.02	67.09	67.18	67.07
108	74.86	74.69	74.65	74.68	74.87	74.65
126	82.45	82.22	82.19	82.16	82.43	82.14
149	90.73	90.46	90.44	90.34	90.61	90.32
172	97.61	97.36	97.35	97.18	97.39	97.17
206	105.7	105.6	105.6	105.3	105.4	105.3
245	112.7	112.7	112.7	112.4	112.3	112.4
328	122.3	122.3	122.3	122.1	121.8	122.1
∞	136.4	132.8	132.7	134.0	143.1	133.7

All five models provide a good fit to the empirical LEV function. For the sake of parsimony we select either the Gamma or the Weibull models.

The distribution, estimated parameter values, and corresponding p-values for British Columbia (BC) - Other are given in Table 39.

TABLE 39

Conditional Marginal Models for British Columbia - Other

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.066$	$\hat{\lambda}=134.0$		0.6610
Weibull	$\hat{\tau}=1.031$	$\hat{\lambda}=144.5$		0.6364
Burr	$\hat{\tau}=4.348$	$\hat{\lambda}=463.7$	$\hat{\gamma}=1.131$	0.6070
Inverse Burr	$\hat{\tau}=0.5366$	$\hat{\lambda}=161.4$	$\hat{\gamma}=1.932$	0.4197
Generalized Pareto	$\hat{\alpha}=6.313$	$\hat{\lambda}=665.7$	$\hat{\tau}=1.190$	0.6290
Transformed Gamma	$\hat{\alpha}=2.065$	$\hat{\lambda}=42.16$	$\hat{\tau}=0.6645$	0.6648

We can eliminate the Inverse Burr due to the relatively low p-value. The Transformed Gamma is not much better than the two parameter Gamma so it can also be eliminated. The remaining models are compared in Table 40

TABLE 40

Empirical and Model LEV Functions for BC - Other

Upper Limit	Empirical LEV	Gamma	Weibull	Burr	Generalized Pareto
10	9.754	9.712	9.693	9.740	9.754
30	27.52	27.35	27.26	27.44	27.48
50	42.83	42.75	42.60	42.81	42.85
70	55.78	56.16	55.96	56.09	56.11
95	69.51	70.47	70.24	70.14	70.14
125	83.30	84.61	84.37	83.93	83.89
160	96.51	97.69	97.47	96.63	96.58
200	108.5	109.1	108.9	107.8	107.7
255	119.9	120.3	120.1	118.9	118.8
310	127.5	127.7	127.7	126.6	126.6
∞	145.5	142.8	142.7	150.0	149.2

We select either the Weibull model or the Gamma model, as they both fit the empirical LEV function quite well.

The parameter estimates and the model p-values for Alberta (AB) - Other are given in Table 41.

TABLE 41

Conditional Marginal Models for Alberta - Other

Distribution	Parameter Values			χ^2 p-value
Pareto	$\hat{\alpha}=6.349$	$\hat{\lambda}=1,165$		0.9500
Gamma	$\hat{\alpha}=0.8765$	$\hat{\lambda}=236.8$		0.9801
Weibull	$\hat{\tau}=0.9188$	$\hat{\lambda}=201.3$		0.9774
Inverse Burr	$\hat{\tau}=0.3348$	$\hat{\lambda}=328.1$	$\hat{\gamma}=2.246$	0.8665
Transformed Gamma	$\hat{\alpha}=0.7760$	$\hat{\lambda}=270.0$	$\hat{\tau}=1.086$	0.9142

The Pareto and Inverse Burr both have a relatively poor fit in the tail of the LEV function, and the Transformed Gamma is eliminated on the basis of parsimony. The remaining two models are compared in Table 42.

TABLE 42

Empirical and Model LEV Functions for AB - Other

Upper Limit	Empirical LEV	Gamma	Weibull
30	27.59	27.36	27.43
80	64.64	64.38	64.47
140	98.93	98.41	98.41
260	146.1	143.6	143.4
360	169.3	166.4	166.2
∞	201.4	207.5	209.5

Both models over estimate the LEV function in the tail. We suggest taking either the Gamma or the Weibull as the model of choice.

The parameter estimates and p-values for Manitoba (MB) - Other are given in Table 43.

TABLE 43

Conditional Marginal Models for Manitoba - Other

Distribution	Parameter Values			χ^2 p-value
Lognormal	$\hat{\mu}=4.238$	$\hat{\sigma}=1.265$		0.7723
Pareto	$\hat{\alpha}=3.550$	$\hat{\lambda}=347.5$		0.9595
Log-logistic	$\hat{\gamma}=1.331$	$\hat{\lambda}=71.00$		0.7972
Gamma	$\hat{\alpha}=0.8553$	$\hat{\lambda}=145.7$		0.7028
Weibull	$\hat{\tau}=0.8906$	$\hat{\lambda}=118.9$		0.7892
Burr	$\hat{\alpha}=2.226$	$\hat{\lambda}=186.0$	$\hat{\gamma}=1.096$	0.9426
Inverse Burr	$\hat{\tau}=0.5554$	$\hat{\lambda}=127.7$	$\hat{\gamma}=1.701$	0.9555
Generalized Pareto	$\hat{\alpha}=2.753$	$\hat{\lambda}=219.9$	$\hat{\tau}=1.128$	0.9387
Transformed Gamma	$\hat{\alpha}=3.798$	$\hat{\lambda}=3.875$	$\hat{\tau}=0.4216$	0.9086
Transformed Beta	$\hat{\alpha}=0.9127$	$\hat{\lambda}=123.9$	$\hat{\gamma}=1.799$	0.8819
	$\hat{\tau}=0.5179$			

The Log-logistic, Burr, Inverse Burr, and Transformed Beta can be eliminated as they have relatively poor fits in the tail of the LEV function. We can also eliminate the Gamma model as it has a relatively low p-value, and we can eliminate the Generalize Pareto as it does not give a better fit than the two parameter model. The remaining models are compared via their LEV functions in Table 44.

TABLE 44

Empirical and Model LEV Functions for MB - Other

Upper Limit	Empirical LEV	Pareto	Weibull	Burr	Transformed Gamma
15	14.03	13.92	13.81	14.06	14.11
30	26.19	25.94	25.78	26.23	26.25
55	43.13	42.59	42.57	42.99	42.91
80	56.63	55.93	56.27	56.30	56.18
120	73.08	72.32	73.32	72.46	72.45
145	80.85	80.28	81.64	80.25	80.36
330	115.2	111.5	112.9	111.0	111.6
∞	131.1	136.3	125.8	143.9	133.0

The Transformed Gamma has the best overall fit, but the Pareto and Weibull are also good. The Burr has the worst fit in the tail. We suggest the Pareto model as it has the highest p-value, but the Transformed Gamma is also a good model.

The parameter estimates and p-values for Ontario (ON) - Other are given in Table 45 below.

TABLE 45

Conditional Marginal Models for Ontario - Other

Distribution	Parameter Values			χ^2 p-value
Gamma	$\hat{\alpha}=1.463$	$\hat{\lambda}=172.9$		0.8667
Weibull	$\hat{\alpha}=1.234$	$\hat{\lambda}=268.3$		0.6879
Burr	$\hat{\alpha}=4.456$	$\hat{\lambda}=709.7$	$\hat{\gamma}=1.393$	0.6638
Generalized Pareto	$\hat{\alpha}=14.82$	$\hat{\lambda}=2,178$	$\hat{\tau}=1.628$	0.7090
Transformed Gamma	$\hat{\alpha}=2.469$	$\hat{\lambda}=71.62$	$\hat{\tau}=0.7555$	0.7312

The Gamma model has a much higher p-value than the other distributions, and also has as good a fit to the LEV function as any of the other models. We select the Gamma model, and in Table 46 , the empirical LEV and the model LEV are compared.

TABLE 46

Empirical and Model LEV Functions for ON - Other

Upper Limit	Empirical LEV	Gamma
135	111.8	113.5
262	176.9	177.6
393	214.1	214.1
520	233.7	232.9
∞	261.6	253.0

The Gamma model under estimates the LEV function in the tail, but overall provides a good fit.

The parameter estimates and p-values for Quebec (PQ) - Other are given in Table 47.

TABLE 47

Conditional Marginal Models for Quebec - Other

Distribution	Parameter Values			χ^2 p-value
Pareto	$\hat{\alpha}=166.6$	$\hat{\lambda}=35,810$		0.7070
Gamma	$\hat{\alpha}=1.043$	$\hat{\lambda}=206.3$		0.7656
Weibull	$\hat{\tau}=1.022$	$\hat{\lambda}=217.1$		0.7453
Burr	$\hat{\alpha}=5.968$	$\hat{\lambda}=992.1$	$\hat{\gamma}=1.102$	0.8039
Inverse Burr	$\hat{\tau}=0.4404$	$\hat{\lambda}=277.1$	$\hat{\gamma}=2.170$	0.8894
Generalized Pareto	$\hat{\alpha}=10.26$	$\hat{\lambda}=1,820$	$\hat{\tau}=1.119$	0.7782
Transformed Gamma	$\hat{\alpha}=1.324$	$\hat{\lambda}=147.5$	$\hat{\tau}=0.8606$	0.7292
Transformed Beta	$\hat{\alpha}=1.235$	$\hat{\lambda}=304.9$	$\hat{\gamma}=1.952$	0.8231
	$\hat{\tau}=0.4973$			

We can eliminate the Pareto as the p-value is relatively low. We can also eliminate the Inverse Burr and the Transformed Beta, as they both have poor fits in the tail of the LEV function. The Transformed Gamma is also eliminated , since it does not provide a better fit than the two parameter model. The LEV functions of the remaining distributions are compared in Table 48.

TABLE 48

Empirical and Model LEV Functions for PQ - Other

Upper Limit	Empirical LEV	Gamma	Weibull	Burr	Generalized Pareto
20	19.16	19.19	19.16	19.26	19.25
45	40.98	40.90	40.82	41.08	41.05
80	67.96	67.43	67.30	67.64	67.57
105	84.69	83.92	83.76	84.05	83.96
135	102.1	101.3	101.2	101.3	101.2
175	121.3	121.1	120.9	120.7	120.6
220	138.7	139.2	139.1	138.5	138.4
285	158.2	159.5	159.4	158.3	158.3
375	177.8	179.0	178.9	177.5	177.6
520	197.3	197.2	197.1	196.0	196.1
∞	221.4	215.3	215.2	221.5	219.9

All four models follow the LEV function quite closely, with the Burr being the best in the tail, followed by the Generalized Pareto. We select the Burr as the best model, and either the Gamma or the Weibull as the best two parameter model.

3.5 The Grouping

In sections 3.3 and 3.4, the Chi-square p-value of each model was given. In some cases it was surprising how many distributions we were able to fit to the grouped data. We were also able to manipulate the grouping to get a high p-value. Since the data has been discretisized it is subject to clustering, and one of the purposes of grouping data is to overcome clustering. The question is: how far can we go? Briefly consulting some of the research done in the area, it was found that the problem has not really been solved. Most of the literature we consulted usually deals with equiprobable classes, and specific families of alternatives. Here we are dealing with no specific alternative. We tried grouping the data into equiprobable intervals, with success in only 2 instances. Kallenberg et al [4] suggest that for alternatives with heavy tails, partitions with some smaller classes in the tails may lead to an increase in power. In general they recommend trying 5 partitions with probability .05, .3, .3, .3,.05 or 6 partitions with probability .05, .15, .3, .3, .15, .05. We tried this type of grouping with British Columbia, Alberta, Manitoba, and Ontario for the conditional models and only in British Columbia were we able to successfully fit any model.

4. CONCLUSIONS

The distributions that were fit to the data in sections 3.3 and 3.4 are summarized in Tables 49 – 52 below.

The summary of marginal models fitted to the data for Drugs by province is listed in Table 49.

TABLE 49

Summary of Marginal Models for Drugs

Province	Distribution	Parameter Values		
British Columbia	Gamma	$\hat{\alpha}=1.139$	$\hat{\lambda}=102.5$	
Alberta	Gamma	$\hat{\alpha}=1.050$	$\hat{\lambda}=111.1$	
	Weibull	$\hat{\tau}=1.031$	$\hat{\lambda}=118.9$	
Manitoba	Transformed Gamma	$\hat{\alpha}=0.5630$	$\hat{\lambda}=184.6$	$\hat{\tau}=1.727$
Ontario	Gamma	$\hat{\alpha}=1.143$	$\hat{\lambda}=126.0$	
Quebec	Log-logistic	$\hat{\gamma}=2.177$	$\hat{\lambda}=89.48$	

It appears that the Gamma model is the most common model.

The summary of marginal models fitted to the Other category by province is given in Table 50.

TABLE 50

Summary of Marginal Models for Other

Province	Distribution	Parameter Values		
British Columbia	Transformed Gamma	$\hat{\alpha}=1.787$	$\hat{\lambda}=21.42$	$\hat{\tau}=0.4851$
Alberta	Gamma	$\hat{\alpha}=0.9870$	$\hat{\lambda}=204.8$	
	Weibull	$\hat{\tau}=0.9872$	$\hat{\lambda}=201.6$	
Manitoba	Gamma	$\hat{\alpha}=1.275$	$\hat{\lambda}=86.70$	
	Weibull	$\hat{\tau}=1.161$	$\hat{\lambda}=115.4$	
Ontario	Weibull	$\hat{\tau}=1.146$	$\hat{\lambda}=261.1$	
Quebec	Transformed Gamma	$\hat{\alpha}=17.65$	$\hat{\lambda}=0.001222$	$\hat{\tau}=0.2436$

The Gamma is again a popular model.

The summary of conditional marginal models fitted to the data for Drugs by province is listed in Table 51.

TABLE 51

Summary of Conditional Marginal Models for Drugs

Province	Distribution	Parameter Values		
British Columbia	Gamma	$\hat{\alpha}=1.004$	$\hat{\lambda}=142.7$	
	Weibull	$\hat{\tau}=1.001$	$\hat{\lambda}=143.4$	
Alberta	Gamma	$\hat{\alpha}=1.073$	$\hat{\lambda}=175.9$	
	Weibull	$\hat{\tau}=1.036$	$\hat{\lambda}=191.2$	
Manitoba	Transformed Gamma	$\hat{\alpha}=0.4011$	$\hat{\lambda}=178.4$	$\hat{\tau}=2.405$
Ontario	Gamma	$\hat{\alpha}=1.271$	$\hat{\lambda}=121.7$	
	Weibull	$\hat{\tau}=1.171$	$\hat{\lambda}=161.1$	
Quebec	Gamma	$\hat{\alpha}=1.042$	$\hat{\lambda}=127.4$	
	Weibull	$\hat{\tau}=1.024$	$\hat{\lambda}=134.0$	

It appears that the Gamma and Weibull models are the most common models.

The summary of conditional marginal models fitted to the Other category by province is given in Table 52.

TABLE 52

Summary of Conditional Marginal Models for Other

Province	Distribution	Parameter Values		
British Columbia	Gamma	$\hat{\alpha}=1.066$	$\hat{\lambda}=134.0$	
	Weibull	$\hat{\tau}=1.031$	$\hat{\lambda}=144.5$	
Alberta	Gamma	$\hat{\alpha}=0.8765$	$\hat{\lambda}=236.8$	
	Weibull	$\hat{\tau}=0.9188$	$\hat{\lambda}=201.3$	
Manitoba	Pareto	$\hat{\alpha}=3.550$	$\hat{\lambda}=347.5$	
Ontario	Gamma	$\hat{\alpha}=1.463$	$\hat{\lambda}=172.9$	
Quebec	Burr	$\hat{\tau}=0.4404$	$\hat{\lambda}=277.1$	$\hat{\gamma}=2.170$

The Gamma is again a popular model.

Table 53 shows the maximum likelihood estimates for the number of claims in each province.

TABLE 53

Maximum Likelihood Estimates for the Number of Claims

Province	p_{00}	p_{10}	p_{01}	p_{11}
British Columbia	0.00290	0.69797	0.08228	0.21684
Alberta	0.00404	0.90431	0.01617	0.07547
Manitoba	0.00137	0.71683	0.11218	0.16963
Ontario	0.00229	0.85072	0.04011	0.10688
Quebec	0.00358	0.65698	0.09347	0.24597

We note that the probability of no claims for either category is very small, and that a claim in Drugs with no claim in Other is the dominant form of a claim. A claim in both categories is the next most popular form of a claim.

We were able to fit parametric models to the marginal and conditional marginal distributions of this data. Some questions about the chi-squared test were asked but not answered. Due to the severe clustering of the data it was sometimes necessary to reduce the number of groups to as few as 5 or 6. The question still remains, did we "over group" the data, and hide the true underlying distribution, or worse yet, lead ourselves to believe that there was indeed any such underlying distribution at all? Although Klugman's [5] FIT software has the capability to group the data, we feel that there is certainly a market for a more sophisticated package that could greatly reduce the amount of work required to group large data sets. Klugman's [5] FIT software was very fast, but as the author states, it provides very little error checking.

We did not address the theoretical implications of our model with regards to the copula function. As mentioned previously there are many unanswered questions about the efficiency and consistency of the model.

Referring back to equation (6) in section 3.1 we can see that, in general, it will not be easy to find the moments of the distribution. An assumption of independence would facilitate, in most cases, the calculation of the moments. Note that in either case it would not be difficult to calculate the cumulative distribution function.

In sections 3.3 and 3.4 we follow the lead of Hogg and Klugman [3] and use the limited expected value function to compare the best models we found. In most cases the LEV function did not supply a lot of new information, rather we found it useful in eliminating many of the models which were not listed in the tables.

Making a quick comparison between the models for the conditional marginals and the models for the marginals reveals that the Gamma and the Weibull are both very good models for this data. In most cases the Gamma model provides the best fit to the data. In all but a few cases, if the Gamma model fit well then the Weibull model would also fit well.

Bibliography

[1] BMDP Statistical Software Inc., Program Version 1987 (IBM/OS); Manual Edition 1990.

[2] Carriere, Jacques (Working Paper). "A Large-Sample Test for One-Parameter Families of Copulas" Faculty of Management, University of Manitoba, Winnipeg Manitoba.

[3] Hogg, Robert V., Klugman, Stuart A. (1984). *Loss Distributions*, John Wiley & Sons, Inc.

[4] Kallenberg, W. C. M., Oosterhoff, J., Schriever, B. F. (1985). "The Number of Classes in Chi-squared Goodness-of-Fit Tests" *Journal of the American Statistical Association*, Vol. 80, No. 392, pp. 959--968.

[5] Klugman, Stuart A. *FIT* (ver 3.0), January 1993. Programs for fitting parametric severity and frequency models to data and collective risk calculations using those models. Copyright 1992-3 by Stuart Klugman.

[6] Mardia, K. V. (1970). *Families of Bivariate Distributions*, Griffin, London.

Appendix I

The densities for seven of the fifteen distribution models used in Klugman's FIT [5] software are given in Table 54.

TABLE 54

Distributions

Distribution	Probability Density Function
Transformed Beta	$f(x)=\frac{\Gamma(\alpha+\tau)\gamma\lambda^{\gamma\alpha}x^{\gamma\tau-1}}{\Gamma(\alpha)\Gamma(\tau)(\lambda^{\gamma}+x^{\gamma})^{\alpha+\tau}}$
Generalized Pareto	$f(x)=\frac{\Gamma(\alpha+\tau)\lambda^{\alpha}x^{\tau-1}}{\Gamma(\alpha)\Gamma(\tau)(\lambda+x)^{\alpha+\tau}}$
Pareto	$f(x)=\frac{\alpha\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}$
Inverse Pareto	$f(x)=\frac{\tau\lambda x^{\tau-1}}{(\lambda+x)^{\tau+1}}$
Burr	$f(x)=\frac{\alpha\gamma\lambda^{\gamma\alpha}x^{\gamma-1}}{(\lambda^{\gamma}+x^{\gamma})^{\alpha+1}}$
Inverse Burr	$f(x)=\frac{\tau\gamma\lambda^{\gamma}x^{\gamma\tau-1}}{(\lambda^{\gamma}+x^{\gamma})^{\tau+1}}$
Log-logistic	$f(x)=\frac{\gamma\lambda^{\gamma}x^{\gamma-1}}{(\lambda^{\gamma}+x^{\gamma})^{2}}$

All of the densities in Table 1 are special cases of the Transformed Beta distribution. All parameters must take positive values and the support is always positive as well. All unused parameters in the less general cases are set equal to 1.

The remaining eight distributions used in Klugman's [5] FIT software are listed in Table 55.

TABLE 55

Distributions

Distribution	Probability Density Function
Transformed Gamma	$f(x)=\dfrac{\tau x^{\alpha\tau-1}e^{-(x/\lambda)^{\tau}}}{\lambda^{\alpha\tau}\Gamma(\alpha)}$
Inverse Transformed Gamma	$f(x)=\dfrac{\tau\lambda^{\alpha\tau}e^{-(\lambda/x)^{\tau}}}{x^{\alpha\tau+1}\Gamma(\alpha)}$
Gamma	$f(x)=\dfrac{x^{\alpha-1}e^{-(x/\lambda)}}{\lambda^{\alpha}\Gamma(\alpha)}$
Inverse Gamma	$f(x)=\dfrac{\lambda^{\alpha}e^{-(\lambda/x)}}{x^{\alpha+1}\Gamma(\alpha)}$
Weibull	$f(x)=\dfrac{\tau x^{\tau-1}e^{-(x/\lambda)^{\tau}}}{\lambda^{\tau}}$
Inverse Weibull	$f(x)=\dfrac{\tau\lambda^{\tau}e^{-(\lambda/x)^{\tau}}}{x^{\tau+1}}$
Lognormal	$f(x)=\dfrac{1}{x\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}$
Inverse Gaussian	$f(x)=\left(\dfrac{\lambda}{2\pi x^3}\right)^{1/2}\exp\left\{-\dfrac{\lambda(x-\mu)^2}{2x\mu^2}\right\}$

The first two distributions, the Transformed Gamma and the Inverse Transformed Gamma, are the general case of the 4 distributions which follow them. All parameters must take positive values, with the exception of the Lognormal distribution, where μ can be negative. Once again all support is positive.