

Maximum Entropy Empirical Likelihood Methods Based on Bivariate Laplace Transforms and Moment Generating Functions

Andrew Luong

École d'Actuariat, Université Laval, Québec, Canada

Email: Andrew.Luong@act.ulaval.ca

How to cite this paper: Luong, A. (2018) Maximum Entropy Empirical Likelihood Methods Based on Bivariate Laplace Transforms and Moment Generating Functions. *Open Journal of Statistics*, 8, 264-283. <https://doi.org/10.4236/ojs.2018.82017>

Received: February 14, 2018

Accepted: April 7, 2018

Published: April 10, 2018

Copyright © 2018 by author and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Maximum Entropy Empirical Likelihood (MEEL) methods are extended to bivariate distributions with closed form expressions for their bivariate Laplace transforms (BLT) or moment generating functions (BMGF) without closed form expressions for their bivariate density functions which make the implementation of the likelihood methods difficult. These distributions are often encountered in joint modeling in actuarial science and finance. Moment conditions to implement MEEL methods are given and a bivariate Laplace transform power mixture (BLTPM) is also introduced, the new operator generalizes the existing univariate one in the literature. Many new bivariate distributions including infinitely divisible (ID) distributions with closed form expressions for their BLT can be created using this operator and MEEL methods can also be applied to these bivariate distributions.

Keywords

Bivariate Power Mixture Operator, Laplace Transform Copulas, Chi-Square Test Statistics, Bivariate Normality, Infinite Divisibility, Empirical Likelihood

1. Introduction

Bivariate distributions are useful for joint modelling and naturally fitting these distributions is a necessity for pricing in insurance and finance. For example, in finance fitting a bivariate jump diffusion model to joint returns data allowed us to price a basket option accordingly, see Ruijter and Oosterlee [1] (p. B658) for a bivariate jump diffusion model. The authors advocated a two dimensions cosine method for pricing basket option. In actuarial sciences, one might consider modeling two claim amounts from an accident, *i.e.*, the amount from body

damage and the amount from material damage which are incurred simultaneously in a car accident, see Partrat [2] (p. 225). Joint survival analysis is also useful for lifetime study, see Crowder [3] (pp. 121-136).

In order to find suitable bivariate parametric models, compound methods and copulas are often used. New bivariate distributions created using the traditional distribution or survival copulas are often continuous with closed form distribution functions or density functions so that methods based on likelihood functions can be applied for statistical inferences. It is also well known that many useful bivariate infinitely divisible distributions do not have closed form density function nor distribution function yet very useful for applications.

For the univariate case, the LT power mixture (LTPM) operator as introduced by Abate and Whitt [4] (pp. 92-93) for the univariate set up for creating new distribution using Laplace transforms (LT) can be used to generate many distributions with closed form LT. Furthermore, it can also be used to generate distribution which is infinitely divisible. In this paper we shall generalize the LTPM to a bivariate version, the BLTPM operator and show that the BLTPM operator can be used to generate new bivariate distributions with closed form BLT. Furthermore, the traditional survival copulas such as the Clayton copula, see Shih and Louis [5] can be used as a LT copula if the property of complete monotonicity of two specific related functions are satisfied. For another class of survival copulas, see Crowder [3] (pp. 121-138). Consequently, some distribution or survival function copulas which are defined using a generator based on a LT of a nonnegative random variable can be used to generate new distribution with prescribed marginals specified by their marginal LTs. A similar bivariate PM operator based on distribution functions or survival functions, the BDSPM operator has been introduced by Marshall and Olkin [6] (pp. 834-836) to create new distributions functions and with frailty induced distributions, it also related to a class of distribution or survival Copula functions defined with a generator. The BLTPM operator can be used to generate bivariate infinite divisible distributions with closed form BLTs. It appears to be useful to have bivariate infinite divisible distributions for joint modeling as they are related to corresponding bivariate Lévy processes with stationary and independent increments. These types of processes are useful as they often lead to elegant results in risk theory in actuarial sciences and in finance.

It appears natural to extent inferences based on bivariate maximum entropy empirical likelihood (MEEL) to distributions with closed form bivariate LT (BLT) or closed form bivariate moment generating function (BMGF) which are similar to the univariate case as given by Luong [7]. Beside only the BLT or BMGF is needed without asking the explicit expression of the Bivariate distribution function or density function, MEEL methods appear to be practical and the methods also offer simultaneously a goodness of fit test statistics which follows a chi-square distribution which is relatively simple to implement and appear to be useful for practitioners. Along the same vein of univariate MEEL methods, bivariate MEEL methods of estimation and tests are based on constraints specified by moments

conditions which can be identified with elements of a basis where the true score functions are projected as in the univariate set up, see Luong [7] (pp. 463-465); these moment conditions which define elements of the base can be extracted from the model BLT or BMGF.

It is also worth to mention that many bivariate distributions with a singular component in a domain such as the Bivariate exponential distribution as introduced by Marshall and Olkin [8] (p. 36) or bivariate phase type distribution as given by Assaf *et al.* [9] (p. 692) do have a closed form bivariate LT so that model testing for these distributions with singular components in the domain can be handled by bivariate MEEL methods. This motivates us to extend MEEL methods to bivariate distributions and the paper is organized as follows.

In Section 2 some examples from actuarial science and finance are given to illustrate the usefulness of bivariate models without closed form density functions but with closed form BLTs. The BLTPM operator is introduced in Section 3 and it is shown that it can be used to create bivariate ID distribution with closed form BLT. MEEL methods are introduced in Section 3 with the proposal of two bases to generate moment conditions. The elements of the bases are based on BLT or BMGF and do not need the density functions explicitly. These bases are proposed to balance efficiencies and numerical tractability as a base with a large number of elements tend to create numerical difficulties on implementation of the MEEL methods. The asymptotic properties which already appear in the literature are restated emphasizing bases and projections of the score functions. We also discuss numerical implementation of the MEEL methods using penalty function approach as given by Luong [7]. Section 5 illustrates the use of the methods by conducting a limited simulation study with a bivariate compound Poisson model as proposed by Partrat [2] (pp. 220-223). With the range of parameters as considered, we observe that the MEEL estimators are two to four time more efficient than MM estimators and at the same time provide a chi-square goodness of fit test statistics for model selection. A base of nine elements is chosen in the study and there is no major numerical difficulty encountered on implementation using the penalty function approach. The methods avoid the use of the covariance matrix of the elements of the base as required for the optimum generalized moment methods (GMM) as discussed in Luong [7], the expressions for this optimum covariance matrix can be complicated.

In the next section, we shall give some examples to illustrate that in many situations we end up working with bivariate distributions with closed form BLT or closed form BMGF but without a closed form for distributions functions or density functions.

2. Some Examples

Example 1

The following bivariate distribution introduced by Partrat [2] (pp. 220-222) is nonnegative and continuous in general but it is discrete at the origin where there is

a mass assigned to it. It can be created from compounding using a bivariate Poisson distribution and has closed form BLT and suitable for modelling two claim amounts from two types of claim in actuarial sciences. The vector of claim amount will have two components and each component can be expressed as a random sum as in the univariate case but the two components will be correlated due to the two numbers of claims given by the vector $N = (N_1, N_2)'$ are dependent and follow a bivariate Poisson distribution. More precisely, the claim amount is a vector $X = (Y, Z)'$. Each of the components of X can be expressed each as a random sum. Note that $Y = \sum_{i=1}^{N_1} U_i$, the random variables $\{U_i, i = 1, 2, \dots\}$ are independent and identically distributed as $U \geq 0$ which is a nonnegative continuous random variable with LT, $L_U(s)$. Similarly, $Z = \sum_{i=1}^{N_2} V_i$, the random variables $\{V_i, i = 1, 2, \dots\}$ are independent and identically distributed as $V \geq 0$ which is a nonnegative continuous random variable with LT, $L_V(t)$. The random variables $\{U_i, i = 1, 2, \dots\}$, $\{V_i, i = 1, 2, \dots\}$, $\{N_1, N_2\}$ are assumed to be independent and with other related independence assumptions as given by H1-H7 in Partrat [2] (pp. 220-221) leads to the BLT of the vector X expressible as $L_X(s, t) = g(L_U(s), L_V(t))$, see Partrat [2] (p. 221).

The bivariate probability generating function (BPGF) for $N = (N_1, N_2)'$ is denoted by $g(s, t)$ and if $N = (N_1, N_2)'$ follows a bivariate Poisson distribution then its PGF is given as $g(s, t) = e^{\phi_1(s-1) + \phi_2(t-1) + \phi_{12}(st-1)}$, see Partrat [2] (p. 222).

The parameters $\phi_1, \phi_2, \phi_{12}$ are all nonnegative. This is the only bivariate Poisson distribution which is infinitely divisible (ID), see Dwass and Teicher [10]. Now if U and V are also ID then with property of the bivariate Poisson PGF, it can be seen that

$$\left[g(L_U(s), L_V(t)) \right]^n$$

is a BLT for each positive integer n , using the remarks given by Abate and Whitt [4] based on results given by Feller [11] (p. 176, 449). Consequently, we can conclude that X is also ID. As BLT is closely related to BMGF, a similar conclusion can be made using BMGF instead of BLT. Also in many circumstances by inspecting the BLT or BMGF which has a closed form expression, we can infer the property of infinite divisibility of the distribution without having to use more sophisticated methods such as the one based on the Lévy-Khintchine representation, see Dwass and Teicher [10] for this representation which is general but very technical to characterize ID distribution and Lévy process.

Clearly, we can differentiate $L_X(s, t)$ or using a conditioning argument with the random sums representations to obtain the first two moments of the vector X and these first two moments will be included in the set of moment conditions for MEEL methods developed subsequently in section 4. The following example gives a model which is used in finance.

Example 2

Ruijter and Oosterlee [1] (p. B658) consider joint modelling of returns of two

assets in a period of time. The return random vector is $\mathbf{X} = (Y, Z)'$ and is representable as

$$\mathbf{X} = \mathbf{B} + \sum_{i=1}^N \mathbf{J}_i,$$

with \mathbf{B} following a bivariate normal distribution, $\mathbf{B} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ and the covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$, the random vector \mathbf{B} is independent of the random sum $\sum_{i=1}^N \mathbf{J}_i$, the \mathbf{J}_i 's are independent and identically distributed as \mathbf{J} which follows a bivariate normal distribution, $\mathbf{J} \sim N(\boldsymbol{\mu}^J, \boldsymbol{\Sigma}^J)$ with

$$\boldsymbol{\mu}^J = (\mu_1^J, \mu_2^J)', \quad \boldsymbol{\Sigma}^J = \begin{pmatrix} \sigma_{11}^J & \sigma_{12}^J \\ \sigma_{12}^J & \sigma_{22}^J \end{pmatrix}$$

This is a jump-diffusion model with the diffusion part given by \mathbf{B} and the jump part by the random sum $\sum_{i=1}^N \mathbf{J}_i$. Also, the \mathbf{J}_i 's, \mathbf{B} , \mathbf{Z} and \mathbf{N} are independent with \mathbf{N} following a Poisson distribution with parameter λ . Comparing to a bivariate normal model, the bivariate jump diffusion model has an extra jump component. Ruijter and Oosterlee [1] (p. B658) show that by letting $\mathbf{s} = (s, t)'$, the BMGF for \mathbf{W} is given by

$$M_{\mathbf{X}}(s, t) = e^{s'\boldsymbol{\mu} + \frac{1}{2}s'\boldsymbol{\Sigma}s} \cdot \exp\left(\lambda \left(e^{s'\boldsymbol{\mu}^J + \frac{1}{2}s'\boldsymbol{\Sigma}^J s} - 1 \right)\right), \quad -\infty < s, t < \infty.$$

The domain of $M_{\mathbf{X}}(s, t)$ is the entire plane due to the use of the Poisson and normal distributions. Clearly, this is an ID BMGF and the corresponding jump diffusion process is a bivariate Lévy process. Also, the first two moments of the vector \mathbf{X} can be extracted from the BMGF. For this model, MEEL methods can be used for estimation and model testing. The class of bivariate normal mean variance mixture as described by McNeil *et al* [12] (pp. 77-78) is another example where bivariate MEEL methods might be suitable.

Furthermore, MEEL methods can also be used for testing the null composite hypothesis which specifies that the parametric model fits the data. The test statistics based on MEEL methods can be constructed in a unified way and obtained simultaneously with the estimation procedures. This feature is useful for doing applied works and the test statistics is less complicated than statistics based on the Mahalanobis distance, see McAssey [13], Muldhokar *et al*. [14] for statistics based on such a distance. The unique asymptotic chi-square distribution that the statistics follows across the composite hypothesis make them suitable statistics to replace the traditional chi-square test statistics as proposed by Moore and Stubblebine [15] which require closed form bivariate density functions. Model testing procedure is easy to implement if it is based on a statistics with a unique asymptotic distribution for all $\boldsymbol{\beta}_0 \in \Omega$, the parameter space is denoted by Ω . The main requirement is the model BLT has a closed form expression.

These properties of bivariate MEEL methods are similar to properties of univariate MEEL methods as discussed by Luong [7] and will be further discussed in section 4. Note that asymptotic theory for empirical likelihood methods as developed by Qin and Lawless [16] (p302-308) is not restricted to univariate distribution and the same can be said for the maximum entropy version. In fact, asymptotic theory for empirical likelihoods is well established for model with multivariate observations which are independent and identically distributed using a set of fixed moment conditions. The example 2 as given by Qin and Lawless [16] (pp. 311-312) is an example of the use of empirical likelihood methods for a bivariate model. In practice, we would like to choose the moment conditions so that high efficiencies can be achieved and handle the procedures numerically. We focus on these points for a class of distributions with closed form BLT or BMGF and do not emphasize asymptotic theory in the paper.

3. Distributions Created Using the BLTPM Operator

3.1. The BLTPM Operator

The Laplace transform power mixture operator (LTPM) has been introduced in the literature for creating univariate non-negative distribution and can be viewed as a continuous type of compounding operator and it makes use of LT of a distribution and the LT of a mixing random variable to create a new distribution specified by its LT. It is due to Abate and Whitt [4] (pp. 92-93). They use the acronym PM but to emphasize the explicitly use of LTs with this operator we shall use the acronym LTPM and to distinguish it with a similar power mixture operator introduced by Marshall and Olkin [6] (pp. 834-836) which make explicitly use of distributions instead of LTs. Abate and Whitt [4] also point out that a new ID can be obtained if we start with an ID distribution and the mixing distribution is non-negative continuous and ID. They also give an elegant and insightful interpretation when instead of ID distribution we can consider the corresponding Lévy process behind. The new ID distribution is related to a new Lévy process constructed from a Lévy process subject to a change time process induced by the mixing random variable and since it is ID, the change time process is also a Lévy process. To summarize, we can say that starting with a Lévy process and subject this process to another non-negative Lévy process used as a change of time process then a new Lévy process can be created or equivalently a new ID non-negative distribution is created which is the distribution of the increments of the newly created Lévy process. Observe that the LTPM operator can be used to generate new distributions in general and with further conditions imposed on the procedures, the new distribution can be ID. Consequently, it appears to be natural to generalize this operator to a bivariate version to create bivariate distributions using two univariate LTs and the LT of a mixing random variable, this is the bivariate power mixture operator (BLTPM). Clearly, it is natural extension of the LTPM operator which is used to create univariate distribution. This operator is similar to another power mixture operator which makes use of distributions or

survival functions instead of LTs given by Marshall and Olkin [6] (p. 835), we shall use the acronym BDSPPM for this operator.

For illustrations, some examples of new bivariate distribution specified by BLT and created using the BLTPM operator will be given for illustrations in section 3.2. The BLTPM operator can also be used to create new distribution with one marginal distribution which is discrete and the other marginal being continuous. These types of bivariate distributions appear to be useful for actuarial science and possibly for other fields as well.

Often, we need to know whether a function can be considered as LT of some univariate random variable, the notion of completely monotonicity of a function is useful for characterizing a function to be a proper LT of a random variable and can be found in Feller [11] (pp. 439-440) but it is reproduced here for completeness as Definition and Theorem below.

Definition

A function $\phi(s)$ with the domain given by $s \geq 0$ is completely monotone (CM) if it possesses derivatives $\phi^{(n)}$ of all order and $(-1)^n \phi^{(n)}(s) \geq 0$.

Theorem

A function $\phi(s)$ with the domain given by $s \geq 0$ is the LT of a random variable if and only if it is CM.

Now if we assume that the LT of a random variable X is $\hat{f}(s)$ so that $\hat{f}^\alpha(s)$ is a proper LT, also suppose that the LT of another random variable Y is $\hat{g}(t)$ so that $\hat{g}^\alpha(t)$ is a proper LT and α is a non-negative random variable with distribution function H and LT $\hat{h}(u)$. Furthermore, we assume that conditioning on a realized value α , the LT $\hat{f}^\alpha(s)$ which is the LT of a random variable denoted by $X^{(\alpha)}$ and the LT $\hat{g}^\alpha(t)$ which is the LT of a random variable denoted by $Y^{(\alpha)}$ are independent but when integrate out with respect to the distribution of α will give a new bivariate LT for a vector $X = (Y, Z)'$,

$$L_X(s, t) = \int_0^\infty \hat{f}^\alpha(s) \cdot \hat{g}^\alpha(t) dH(\alpha). \quad (1)$$

Using the LT of α , it can be re-expressed as

$$L_X(s, t) = \hat{h}\left(-\log(\hat{f}(s)) - \log(\hat{g}(t))\right). \quad (2)$$

Expression (2) gives a definition of the BLTPM operator and it generalizes expression (6.1) given by Abate and Whitt [4] (p. 92) for the univariate LTPM operator. Similarly, if MGFs are used instead and they are given respectively by $\widehat{f}_M(s)$, $\widehat{g}_M(t)$ and $\widehat{h}_M(u)$, a new bivariate distribution with BMGF can be created similarly.

$$M_X(s, t) = \widehat{h}_M\left(\log(\widehat{f}_M(s)) + \log(\widehat{g}_M(t))\right). \quad (3)$$

Observe that if $X^{(\alpha)}, Y^{(\alpha)}$ are ID and $\alpha \geq 0$ is continuous and ID then the new bivariate distribution created is also ID. We can interpret this property using Lévy processes, as two independent Lévy processes subject to the same time change Lévy process is a bivariate Lévy process. This generalizes the argument used for the

univariate case given by Abate and Whitt [4] (p. 92).

A form of bivariate PM operator, the BSDPM has been used in the literature but with the use of distribution functions or survival functions, see the seminal works of Marshall and Olkin [6] and a class of survival or distribution copula functions are also obtained. This class is defined using a generator which is based on the LT of some nonnegative random variable which forms a subclass of the Archimedean class and will be further discussed in section 3.2. Due to the analogy between the BLTPM operator and BDSPM operator, these distribution or survival copula functions subject to some conditions are also LT copulas. For the Archimedean class, see Klugman *et al.* [17] (pp. 187-213).

In section 3.2, we shall see that additional requirements are needed when working with LT's instead of distribution functions or survival functions for the use of these copulas traditionally used to create new bivariate distribution function using prescribed marginal distributions.

This is due to the conditions $e^{-\alpha\Phi^{-1}(L_1(s))}$ and $e^{-\alpha\Phi^{-1}(L_2(t))}$, $\alpha > 0$ are proper LT are more difficult to verify than the corresponding conditions, $e^{-\alpha\Phi^{-1}(F_1(x))}$ and $e^{-\alpha\Phi^{-1}(F_2(y))}$ are proper distribution functions with $F_1(x)$ and $F_2(y)$ being distribution functions when the BDSPM operator is used, see Marshall and Olkin [6] (p. 834).

We shall give a few examples below to illustrate the use of the BLTPM operator to create bivariate ID distribution using the LT of a gamma mixing random variable. The same procedure can be used with other mixing random variables such as the inverse Gaussian random variable or the exponential mixture inverse Gaussian (EMIG) random variable. The EMIG distribution as studied by Abate and Whitt [4] (pp. 95-96) is also ID, continuous and nonnegative. This distribution has not received much attention in the statistical literature despite it is used frequently in queueing theory.

3.2. Some Examples

Example 3

For modeling two type of claims in insurance for one period of time, we might want to construct a joint continuous model using the BLTPM operator with the mixing random variable following a gamma distribution with LT

$\hat{h}(u) = (1+u)^{-\tau}$, $\tau > 0$, $\hat{f}(s)$ and $\hat{g}(t)$ are LT of inverse Gaussian distributions with LTs given respectively by

$$\hat{f}(s) = \exp\left(\frac{\theta_1}{\mu_1} \left(1 - \sqrt{1 + \frac{2\mu_1^2 s}{\theta_1}}\right)\right) \text{ and } \hat{g}(t) = \exp\left(\frac{\theta_2}{\mu_2} \left(1 - \sqrt{1 + \frac{2\mu_2^2 t}{\theta_2}}\right)\right),$$

$\theta_i, \mu_i > 0, i = 1, 2$ see Klugman *et al.* [18] (p. 472) for LT of an inverse Gaussian distribution.

Apply the BLTBPM operator using these LTs, this will give a new bivariate continuous distribution for a random vector $X = (Y, Z)'$ which is ID and nonnegative with its BLT given by

$$L_X(s, t) = \left(1 - \frac{\theta_1}{\mu_1} \left(1 - \sqrt{1 + \frac{2\mu_1^2 s}{\theta_1}} \right) - \frac{\theta_2}{\mu_2} \left(1 - \sqrt{1 + \frac{2\mu_2^2 t}{\theta_2}} \right) \right)^{-\tau}.$$

From $L_X(s, t)$, the Pearson product moment correlation coefficient can be obtained and the coefficient can be used to study the dependence between Y and Z . Other dependence measures can also be used to study associativity and dependence, see Chapter 4 of the book by Balakrishnan and Lai [19] (pp. 141-173) and for total positivity dependence, see Barlow and Proschan [20] (pp. 142-150) but we do not go into details into these dependence studies for distributions created using the BLTPM operator as the main focus of the paper are on statistical inference methods using BLTs or BMGFs.

Example 4

Bivariate mixed models with one marginal for counts and the other one for claim amounts appear to be useful in actuarial science. These models can also be constructed using the BLTPM operator. For example, let the random vector $\mathbf{X} = (Y, Z)'$ with Y being discrete and Z being continuous and nonnegative. We might specify $\hat{h}(u) = (1+u)^{-\tau}$, $\tau > 0$, $\hat{f}(s) = \exp(\lambda(e^{-s} - 1))$ which is the LT of a Poisson random variable with $\lambda > 0$, let and

$$\hat{g}(t) = \exp\left(\frac{\theta}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 t}{\theta}} \right)\right), \quad \theta \text{ and } \mu > 0 \text{ which is the LT of an inverse}$$

Gaussian distribution.

The BLT of the random vector $\mathbf{X} = (Y, Z)'$ created using the BLTPM operator is

$$L_X(s, t) = \left(1 - \lambda(e^{-s} - 1) - \frac{\theta}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 t}{\theta}} \right) \right)^{-\tau}.$$

Setting $t = 0$ we obtain the marginal LT of Y which is given by

$L_Y(s) = \left(1 - \lambda(e^{-s} - 1) \right)^{-\tau}$. One can recognize that this is a LT of a negative binomial distribution and for this distribution, the corresponding probability generating function (PGF) $P_Y(s) = (1 - \lambda(s - 1))^{-\tau}$, see Klugman *et al.* [17] (p. 83) for the PGF of a negative binomial distribution.

The marginal $L_Z(t) = \left(1 - \frac{\theta}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 t}{\theta}} \right) \right)^{-\tau}$ is the LT of a continuous random variable.

Example 5

In this example we shall obtain a bivariate negative binomial distribution using the BLTPM operator. Let $\hat{h}(u) = (1+u)^{-\tau}$, $\tau > 0$, $\hat{f}(s) = \exp(\lambda_1(e^{-s} - 1))$ and $\hat{g}(t) = \exp(\lambda_2(e^{-s} - 1))$, λ_1 and $\lambda_2 > 0$. Apply the BLTPM operator this gives a joint distribution for $\mathbf{X} = (Y, Z)'$ with BLT given by

$$L_X(s, t) = \left(1 - \lambda_1(e^{-s} - 1) - \lambda_2(e^{-s} - 1) \right)^{-\tau}.$$

This is a BLT of a bivariate negative binomial distribution which is also ID. In

the literature, this distribution has been constructed using various methods and by many authors, see Mardia [21] (p. 84). We can see that the BLTPM operator is very useful for constructing bivariate distribution with the property being ID. This property is useful in finance and actuarial sciences as the corresponding bivariate process is a bivariate Lévy process with stationary and independent increments. This property often leads to nice results in risk theory and pricing formulas in finance. Simulation of sample paths from these processes are also simplified using this property.

We briefly mention on how to simulate from bivariate distribution with a specified $L_X(s, t)$ created by the BLTPM operator. The procedures are similar to the procedures described in section 5 and given by Marshall and Olkin [6] (p. 840) but the inverse transform method might not be necessary explicitly.

For the inverse transform method, see Ross [22] (pp. 69-70).

The main steps are:

1) Generate an observation for the mixing random variable α from the distribution H which has LT $\hat{h}(u)$.

2) Use the observed value α , generate an observation $X^{(\alpha)}$ from a distribution with LT $\hat{f}^\alpha(s)$. Often from the expression of $\hat{f}^\alpha(s)$, we have a procedure to simulate from this distribution and the inverse method might not be needed explicitly at this step.

3) Use the same observed value α , simulate another observation $Y^{(\alpha)}$ which is independent of $X^{(\alpha)}$ obtained from 2) from a distribution with LT $\hat{g}^\alpha(t)$.

The pair of observations $(X^{(\alpha)}, Y^{(\alpha)})'$ obtained will follow a bivariate distribution with BLT as specified by $L_X(s, t)$.

3.3. LT Copulas

Observe that the BLTPM operator given by expression (2) is similar to the BDSPM operator given by expression 2.2 in Marshall and Olkin [6] (p834) where a class of distribution or survival copulas can be obtained, it is considered below. Without loss of generality we consider distribution copulas which are related to the BDSPM operator. By using the specified marginal distributions F_1 and F_2 , a new bivariate distribution $F(x, y)$ can be constructed with these two specified marginals which is given by

$$F(x, y) = \phi(\phi^{-1}(F_1(x)) + \phi^{-1}(F_2(y))), \quad (4)$$

ϕ is a univariate LT and its inverse given by ϕ^{-1} .

If we let $\phi = (1 + s)^\tau$ which is a LT of a gamma distribution, its inverse is given by $\phi^{-1} = s^{\frac{1}{\tau}} - 1$, then we will have the classical distribution or survival distribution Clayton copula as shown in example 4.2 given by Marshall and Olkin [6] (p. 837). For other copulas of the form given by expression (4), see Shih and Louis [5].

For new bivariate distribution constructed using the BLTPM, we restrict the attention to nonnegative distributions. Note that if the marginal distributions are

specified respectively using LTs and given by $L_1(s)$ and $L_2(t)$ and subject to the two functions $e^{-\alpha\Phi^{-1}(L_1(s))}$ and $e^{-\alpha\Phi^{-1}(L_2(t))}$ are proper LT, a BLT of a new distribution with marginal LTs given respectively by $L_1(s)$ and $L_2(t)$ can also be obtained as

$$L_X(s, t) = \Phi(\Phi^{-1}(L_1(s)) + \Phi^{-1}(L_2(t))), \quad (5)$$

which is similar to the procedure as given by expression (4) using distribution or survival functions. This also means that some of the distribution or survival copulas of the class given by expression (4) can also be used as LT copulas. But the drawback of this approach using LT copula here is even by specifying the marginals $L_1(s)$ and $L_2(t)$ being LT of ID distributions, $L_X(s, t)$ constructed using expression (5) might not necessary be a BLT of a bivariate ID distribution.

It is also more difficult to simulate from a BLT $L_X(s, t)$ obtained using the LT copula as given by expression (5). Despite the algorithm used to generate a pair of observations from a specified BLT procedure is already given as above but when apply the procedures here we encounter the difficulty on how to simulate from distributions with LTs given by $K_1(s) = e^{-\alpha\Phi^{-1}(L_1(s))}$ and $K_2(t) = e^{-\alpha\Phi^{-1}(L_2(t))}$ as we might not be able to recognize these distributions. The inverse transform method can be applied by using the approximate quantile function based on moment generating functions $M_1(s) = e^{-\alpha\Phi^{-1}(L_1(-s))}$ and $M_2(s) = e^{-\alpha\Phi^{-1}(L_2(-t))}$ to obtain a simulated sample with some accuracy. The approximate quantile functions based on moment generating functions can be obtained explicitly using the saddle point technique and it is given by Arevelillo [23]. If these univariate MGFs do not exist, these LTs can be converted to characteristic functions and the numerical methods based on the inverse method and the approximate quantile functions as described by Glasserman and Liu [24] (pp. 1613-1615) can also be used to generate a pair of observations with some degree of accuracies.

There is a vast literature on survival or distribution copula, we just mention a few here. For a good general review on distribution or survival copula models with emphasis on goodness-of-fit test statistics, see Genest *et al.* [25] and for a good review of inference methods based on survival copulas used in actuarial science, see Klugman *et al.* [17] (pp. 187-213), Frees and Valdez [26].

As mentioned earlier, asymptotic theory for Maximum entropy empirical likelihood (MEEL) and empirical likelihood (EL) for models where we have X_1, \dots, X_n a sample of multivariate observations independent and identically distributed (iid) as X which follows a d-variate multivariate distribution $F_{\boldsymbol{\beta}}, \boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ are well established but assuming a set of moment conditions or constraints has been chosen, see Qin and Lawless [16], Schennach [27], Owen [28], Mittelhammer *et al.* [29]; also see discussions in section 3.2 by Luong [7] (pp. 471-472).

For applications, the question on how to choose moment conditions or equivalently bases so that MEEL methods have high efficiencies is a relevant one and we would like to address mainly this issue here for models with closed form

BLTs or BMGFs as introduced in previous sections. The moment conditions are viewed as constraints and can be identified with elements of a chosen basis used to project the true score functions as in the univariate case, see Luong [7] (pp. 461-467). Beside estimation, empirical likelihood type methods also give a chi-square test statistics for model testing which is useful for practical applications. We shall emphasize moment conditions and focus on how to choose constraints for models with closed form BLTs or BMGFs in the next section as asymptotic theory is similar to the univariate case and already discussed in section 3.2 as given by Luong [7].

4. MEEL Methods

Let X_1, \dots, X_n be sample of bivariate observations and they are independent and identically distributed as $X = (Y, Z)'$ which follows a bivariate distribution $F_{\beta}, \beta = (\beta_1, \dots, \beta_p)'$, F_{β} has no closed form but its BLT or BMGF has a closed form expression and given respectively by $L_{\beta}(s, t)$ or $M_{\beta}(s, t)$. The true vector of parameters is denoted by $\beta_0 \in \Omega$. The parameter space Ω is assumed to be compact.

Suppose that we can identify k constraints of the form $g_1(x; \beta), \dots, g_k(x; \beta)$ where these functions $g_i(x; \beta), i = 1, \dots, k$ are unbiased estimating functions with the property $E_{\beta}(g_i(x; \beta)) = 0, i = 1, \dots, k$, these functions also form a basis $B = \{g_1(x; \beta), \dots, g_k(x; \beta)\}$ used to project the true score functions and MEEL estimators can be viewed as equivalent to quaslikelihood estimators based on the projected score functions, see Luong [7] (pp. 466-468). Therefore, the constraints are linked to a basis and the ability to obtain projected score functions close to the true score functions for some restricted but useful parameter space will make MEEL estimators have high efficiencies in practice.

4.1. Choice of Bases

In this section, we shall recommend a base when using a model BLT to extract moment conditions which can be used as a guideline for forming other bases if needed for applying bivariate MEEL methods.

We shall define the first five g_i 's as follows which make use of the first two moments of X , *i.e.*,

$$\begin{aligned} g_1(x; \beta) &= y - E_{\beta}(y), \quad g_2(x; \beta) = z - E_{\beta}(z) \\ g_3(x; \beta) &= (y - E_{\beta}(y))^2 - V_{\beta}(y), \quad g_4(x; \beta) = (z - E_{\beta}(z))^2 - V_{\beta}(z), \quad (6) \\ g_5(x; \beta) &= (y - E_{\beta}(y))(z - E_{\beta}(z)) - Cov_{\beta}(y, z). \end{aligned}$$

The variances of y and z are given respectively by $V_{\beta}(y)$ and $V_{\beta}(z)$, the covariance between y and z is denoted by $Cov_{\beta}(y, z)$. Subsequently we will pick the g_i 's directly using the model BLT $L_{\beta}(s, t)$. We shall choose first four points of the BLT $L_{\beta}(s, t)$, $(s_1, t_1), (s_2, t_2), (s_3, t_3), (s_4, t_4)$ which belong to a circle of radius $r_1 = 0.01$ in the nonnegative quadrant and centered at the origin, *i.e.*, let

$$(s_1, t_1) = r_1 (\cos(0), \sin(0)), \quad (s_2, t_2) = r_1 \left(\cos\left(\frac{\pi}{6}\right), \sin\left(\frac{\pi}{6}\right) \right)$$

$$(s_3, t_3) = r_1 \left(\cos\left(\frac{2\pi}{6}\right), \sin\left(\frac{2\pi}{6}\right) \right), (s_4, t_4) = r_1 \left(\cos\left(\frac{3\pi}{6}\right), \sin\left(\frac{3\pi}{6}\right) \right)$$

which lead to define

$$g_6(x; \beta) = e^{s_1 y + t_1 z} - L_\beta(s_1, t_1), \quad g_7(x; \beta) = e^{s_2 y + t_2 z} - L_\beta(s_2, t_2),$$

$$g_8(x; \beta) = e^{s_3 y + t_3 z} - L_\beta(s_3, t_3), \quad g_9(x; \beta) = e^{s_4 y + t_4 z} - L_\beta(s_4, t_4).$$

We might want to add another 4 points on the part of the circle of radius $r_2 = 0.02$ of the nonnegative quadrant and centered at the origin by letting

$$(s_5, t_5) = r_2 (\cos(0), \sin(0)), \quad (s_6, t_6) = r_2 \left(\cos\left(\frac{\pi}{6}\right), \sin\left(\frac{\pi}{6}\right) \right)$$

$$(s_7, t_7) = r_1 \left(\cos\left(\frac{2\pi}{6}\right), \sin\left(\frac{2\pi}{6}\right) \right), \quad (s_8, t_8) = r_2 \left(\cos\left(\frac{3\pi}{6}\right), \sin\left(\frac{3\pi}{6}\right) \right)$$

which leads to define

$$g_{10}(x; \beta) = e^{s_5 y + t_5 z} - L_\beta(s_5, t_5), \quad g_{11}(x; \beta) = e^{s_6 y + t_6 z} - L_\beta(s_6, t_6),$$

$$g_{12}(x; \beta) = e^{s_7 y + t_7 z} - L_\beta(s_7, t_7), \quad g_{13}(x; \beta) = e^{s_8 y + t_8 z} - L_\beta(s_8, t_8).$$

If needed, pick another 4 points on the circle of radius $r_3 = 0.03$ of the nonnegative quadrant centered at the origin by letting

$$(s_9, t_9) = r_3 (\cos(0), \sin(0)), \quad (s_{10}, t_{10}) = r_3 \left(\cos\left(\frac{\pi}{6}\right), \sin\left(\frac{\pi}{6}\right) \right)$$

$$(s_{11}, t_{11}) = r_3 \left(\cos\left(\frac{2\pi}{6}\right), \sin\left(\frac{2\pi}{6}\right) \right), \quad (s_{12}, t_{12}) = r_3 \left(\cos\left(\frac{3\pi}{6}\right), \sin\left(\frac{3\pi}{6}\right) \right)$$

which leads to define

$$g_{14}(x; \beta) = e^{s_9 y + t_9 z} - L_\beta(s_9, t_9), \quad g_{15}(x; \beta) = e^{s_{10} y + t_{10} z} - L_\beta(s_{10}, t_{10}),$$

$$g_{16}(x; \beta) = e^{s_{11} y + t_{11} z} - L_\beta(s_{11}, t_{11}), \quad g_{17}(x; \beta) = e^{s_{12} y + t_{12} z} - L_\beta(s_{12}, t_{12}).$$

With all these g_i 's, the basis can be formed and given by

$$B_{LT} = \{g_1(x; \beta), \dots, g_k(x; \beta)\}, k = 17. \tag{7}$$

Therefore, if the number of parameters m in the model, $m < k$, bivariate MEEL methods can be used for estimation and model testing using this basis to generate constraints. Note that the choice of $g_i(x; \beta), i = 1, \dots, 17$ do not put restrictions on the parameter space as the model BLT is well defined on the nonnegative quadrant for all values of the vector $s = (s, t)'$. This might not be the case with the use of the BMGF $M_\beta(s, t)$ where the domain of s might depend on β . We have to be ensured that if restrictions are imposed, the restricted parameter space is all we need for applications. The choice of points are based on the intuitive ground that the BLT contains more information at points near the origin and obviously there is some arbitrariness on the choice of points or equivalently moment conditions for using MEEL methods.

If we use moment conditions based on the BMGF such as estimation for the

jump-diffusion model as given by example 2, in general we might want to keep the first five $g_i(x; \beta)$'s as given by expression (6) for BLT and add 12 additional g_j 's with chosen points on the circle of radius $r = 0.01$ centered at the origin without restricting on the first quadrant and define

$$(s_j, t_j) = r \left(\cos\left(\frac{j\pi}{6}\right), \sin\left(\frac{j\pi}{6}\right) \right), j = 0, 1, 2, \dots, 11 \text{ which leads to define}$$

$$g_6(x; \beta) = e^{s_0 z_1 + t_0 z_2} - M_\beta(s_0, t_0),$$

$$g_7(x; \beta) = e^{s_1 z_1 + t_1 z_2} - M_\beta(s_1, t_1),$$

$$\vdots$$

$$g_{17}(x; \beta) = e^{s_{11} z_1 + t_{11} z_2} - M_\beta(s_{11}, t_{11}).$$

Consequently, the base

$$B_{MGF} = \{g_1(x; \beta), \dots, g_5(x; \beta), g_6(x; \beta), \dots, g_{17}(x; \beta)\} \tag{8}$$

which makes use of the model BMGF $M_\beta(s, t)$ can be used for generating constraints.

Note that the BMGF for the jump-diffusion model as given by example 2 is well defined for all points $s = (s, t)'$. Note that with a basis of $k = 17$ elements, it will allow the number of parameters $p < 17$. It appears that such a base gives a good balance between efficiency and numerical tractability for the MEEL methods and can be used as a guideline to form other bases if necessary in practice.

4.2. Estimation and Model Testing

Asymptotic properties for Multivariate MEEL methods and for EL methods for models with independent and identically distributed of vector of observations are well established, see Qin and Lawless, Mittelhammer *et al.* [29] (p. 323), Schennach [27]. Now if we use the constraints as defined using a basis $B = \{g_1, \dots, g_k\}$, MEEL estimators are obtained as the vector $\hat{\beta}$ which maximizes the entropy or minimizes

$$\sum_{i=1}^n \pi_i^*(\beta) \ln \pi_i^*(\beta) \tag{9}$$

with

$$\pi_i^* = \frac{\exp\left(-\sum_{j=1}^k \lambda_j(\beta) \left(\sum_{i=1}^n g_j(x_i; \beta)\right)\right)}{\sum_{i=1}^n \exp\left(-\sum_{j=1}^k \lambda_j(\beta) \left(\sum_{i=1}^n g_j(x_i; \beta)\right)\right)}, i = 1, \dots, n \tag{10}$$

subject to

$$\sum_{i=1}^n \pi_i^*(\lambda, \beta) (g_j(x_i; \beta)) = 0, j = 1, \dots, k. \tag{11}$$

The λ_j 's depend on β and are only implicitly defined using expression (11).

Using standard regularity conditions as indicated in the references just listed, also see section 3.2 given by Luong [7] (pp. 471-472), we then have consistency and asymptotic normality for the MEEL estimators given by $\hat{\beta}$, *i.e.*, $\hat{\beta} \xrightarrow{p} \beta_0$, β_0 is the vector of the true parameters,

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{p} N(0, \Omega),$$

$$\Omega = \left[E \left[\frac{\partial g(x, \beta)}{\partial \beta} \Big|_{\beta=\beta_0} \right] \left(E \left[g(x, \beta) g(x, \beta)' \right] \Big|_{\beta=\beta_0} \right)^{-1} E \left[\frac{\partial g(x, \beta)}{\partial \beta'} \Big|_{\beta=\beta_0} \right] \right]^{-1}, \quad (12)$$

$$g(x, \beta) = (g_1(x, \beta), \dots, g_k(x, \beta))', \quad \Sigma(\beta_0) = E \left[g(x, \beta) g(x, \beta)' \right] \Big|_{\beta=\beta_0}.$$

An estimator $\hat{\Omega}$ for Ω is given below,

$$\hat{\Omega} = \left[\left[\sum_{i=1}^n \hat{\pi}_i \frac{\partial g(x_i, \beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} \right] \left(\sum_{i=1}^n \hat{\pi}_i g(x_i, \beta) g(x_i, \beta)' \Big|_{\beta=\hat{\beta}} \right)^{-1} \right. \\ \left. \cdot \left[\sum_{i=1}^n \hat{\pi}_i \frac{\partial g(x_i, \beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} \right] \right]^{-1}$$

$\hat{\pi}_i = \pi_i^*(\hat{\beta}), i = 1, \dots, n$. If we let $\hat{\pi}_i = \frac{1}{n}, i = 1, \dots, n$, we have another consistent

estimator for Ω . The asymptotic results are very similar to the ones for the univariate case as given by Luong [7] but for completeness they are reproduced to make the paper easier to follow.

For model testing, it is viewed as testing the validity of the moment conditions, *i.e.*, testing the null hypothesis $H_0 : E_{\beta} (g_j(x)), j = 1, \dots, k$ just as in the univariate case, the expectations are under the true parametric model.

The following test statistics given below is a chi-square test statistics which has an asymptotic chi-square distribution with $r = k - p$ degree of freedom, *i.e.*,

$$2nKL(\pi^*(\hat{\beta}), p_n) = 2n \left(\sum_{i=1}^n \pi_i^*(\hat{\beta}) \left[\log(\pi_i^*(\hat{\beta})) - \log\left(\frac{1}{n}\right) \right] \right) \xrightarrow{L} \chi^2(k - p). \quad (13)$$

This test might be useful for testing bivariate normality in empirical finance and it is based on the MEEL estimators $\hat{\beta}$ and since the basis used spans the true score functions of the bivariate normal model which makes $\hat{\beta}$ as efficient as the maximum likelihood estimator $\hat{\beta}_{ML}$, as the projected score functions are identical to the score functions. Furthermore, the asymptotic chi-square distributions are practical for the use of these test statistics and we do not need intensive simulations to implement model testing procedures as in other methods. Also, MEEL methods can be applied for model with a singular part in the domain provided that the model BLT has a closed form expression.

4.3. Numerical Implementations

As for the univariate case, a penalty approach can be used which means that we can perform unconstrained minimization using the following objective function with respect to $\lambda = (\lambda_1, \dots, \lambda_k)'$ and β_1, \dots, β_p ,

$$\sum_{i=1}^n \pi_i^*(\lambda, \beta) \ln \pi_i^*(\lambda, \beta) \\ + \frac{K}{2} \left[\left(\sum_{i=1}^n \pi_i^*(\lambda, \beta) [g_1(x_i, \beta)] \right)^2 + \dots + \left(\sum_{i=1}^n \pi_i^*(\lambda, \beta) [g_k(x_i, \beta)] \right)^2 \right] \quad (14)$$

The penalty constant K is a large positive value, direct search algorithms which are derivative free are recommended and these direct search algorithms are also more stable.

It is important to note that only a local minimizer is found each time using these algorithms, some strategies are needed to identify the global minimizer. Often, we might need a starting vector being close to the vector of the estimators to initialize the algorithm, this is important when working with real data and bivariate models might have many parameters. We might want to adopt the following strategy to look for a good starting vector. Using the two submodels based on the two marginal distributions of the model separately we can fit using univariate MEEL methods using $\{y_i, i=1, \dots, n\}$ and $\{z_i, i=1, \dots, n\}$ as described in Luong [7] to obtain respectively $\widehat{\beta}_1^{(0)}$ the vector which estimates the vector of parameters β_1 of the first marginal model and similarly obtain $\widehat{\beta}_2^{(0)}$ which estimates the vector of parameters β_2 of the second marginal model. The vector of parameters of the bivariate model is β and β can be partitioned into three components, *i.e.*,

$\beta = (\beta_1, \beta_2, \beta_3)'$. Now if we apply MEEL methods with the bivariate model but fixing

$\beta_1 = \widehat{\beta}_1^{(0)}, \beta_2 = \widehat{\beta}_2^{(0)}$ and only estimate β_3 , the vector which estimates β_3 is denoted by $\widehat{\beta}_3^{(0)}$. We can construct the following starting vector

$$\widehat{\beta}^{(0)} = \left(\widehat{\beta}_1^{(0)}, \widehat{\beta}_2^{(0)}, \widehat{\beta}_3^{(0)} \right)', \quad \widehat{\lambda}^{(0)} = (0, \dots, 0)$$

and fit the bivariate model using penalty method by minimizing jointly with respect to the vector λ and the vector β the objective function as given by expression (14).

5. A Simulation Study

In this section we perform a limited simulation study for model given by example 1. We compare the performance of the Method of moment (MM) estimators with the MEEL estimators using the following basis B_{MEEL} with $k=9$ elements which is formed using the first 9 elements of the basis given by expression (7) as the model only has five parameters given by the vector

$\beta = (\phi_1, \phi_2, \phi_{12}, \beta_1, \beta_2)'$ and these parameters will be introduced below. Using the model in example 1 by specifying an exponential distribution for U with

density function $f(u; \beta_1) = \frac{1}{\beta_1} e^{-\frac{u}{\beta_1}}, \beta_1 > 0$ and specifying another exponential

distribution for V with density function $f(v; \beta_2) = \frac{1}{\beta_2} e^{-\frac{v}{\beta_2}}, \beta_2 > 0$. The random

vector $N = (N_1, N_2)'$ follows a bivariate Poisson distribution with parameters $\phi_1, \phi_2, \phi_{12}$ and these parameters all all positive. We have n independent and identically distributed observations

$X_i = (Y_i, Z_i)', i=1, 2, \dots, n$. They have the same distributions as $X = (Y, Z)'$ and using the model as specified, the following first two moments of X can be obtained, they are also given by Partrat [2] (p. 221),

$$E(Y) = \phi_1\beta_1, E(Z) = \phi_2\beta_2, V(Y) = 2\phi_1\beta_1^2, V(Z) = 2\phi_2\beta_2^2, Cov(Y, Z) = \phi_{12}\beta_1\beta_2 \quad (15)$$

where $E(\cdot), V(\cdot)$ and $Cov(\cdot)$ denote respectively the mean, the variance and covariance of the expression inside the parenthesis. Now if we replace $E(Y), E(Z), V(Y), V(Z)$ and $Cov(Y, Z)$ by their sample counterparts which are given by $\bar{Y}, \bar{Z}, s_Y^2, s_Z^2$ and s_{YZ} in expression (15) and by solving the system of equation, it will give us the MM estimators given by the vector

$$\tilde{\beta} = (\tilde{\phi}_1, \tilde{\phi}_2, \tilde{\phi}_{12}, \tilde{\beta}_1, \tilde{\beta}_2)' \quad \text{with}$$

$$\tilde{\beta}_1 = \frac{s_Y^2}{2\bar{Y}}, \tilde{\beta}_2 = \frac{s_Z^2}{2\bar{Z}}, \tilde{\phi}_1 = \frac{\bar{Y}}{\beta_1}, \tilde{\phi}_2 = \frac{\bar{Z}}{\beta_2}, \tilde{\phi}_{12} = \frac{s_{YZ}}{\beta_1\beta_2}$$

For MEEL methods we use the basis $B_{MEEL} = \{g_1, \dots, g_9\}$ and we only keep the first 9 elements of the basis B as given by expression (7).

The model BLT is

$$L_X(s, t) = \exp\left(\phi_1\left(\frac{1}{1+\beta_1s} - 1\right) + \phi_2\left(\frac{1}{1+\beta_2t} - 1\right) + \phi_{12}\left(\frac{1}{1+\beta_1s} \frac{1}{1+\beta_2t} - 1\right)\right),$$

see Partrat [2] (p. 221) and example 1 in section (2). For efficiency, observe that the quasi-score functions of the MM estimators belong to the linear space spanned by g_1, \dots, g_5 and based on asymptotic theory MEEL estimators will be more efficient than MM estimators for large samples already as g_1, \dots, g_5 belong to the basis used for MEEL methods, see Luong [7] for quasi-score functions related to MEEL estimation. Another advantage of the MEEL methods over methods of moments (MM) is that a chi-square statistics with four degree of freedom is available for model testing as the methods provide a unified approach to estimation and model testing.

We conduct a limited simulation study using this model, to simulate an observation from the bivariate Poisson distribution we use a trivariate reduction technique which is based on the following representation in distribution. We simulate first an observation R_{12} from a Poisson distribution with parameter λ_{12} then simulate an independent observation R_1 which follows a Poisson distribution with parameter λ_1 and another observation which is independent of the first two observations which follows a Poisson distribution with parameter λ_2 . Finally, we construct the vector of observations obtained by simulations as $N = (N_1, N_2)'$ with its components given by $N_1 = R_1 + R_{12}$ with $N_2 = R_2 + R_{12}$ and the equalities hold in distribution. The vector $N = (N_1, N_2)'$ will follow a bivariate Poisson distribution with parameters $\phi_1 = \lambda_1 + \lambda_{12}$, $\phi_2 = \lambda_2 + \lambda_{12}$, $\phi_{12} = \lambda_{12}$, see Johnson *et al.* [30] (pp. 124-126) for this representation. To simulate an observation $X_i = (Y_i, Z_i), i = 1, \dots, n$ we can use the random sum representations of Y_i and Z_i and therefore it suffices to simulate from exponential distributions and perform the appropriate summations.

We use $M = 100$ samples and each of the sample is of size $n = 1000$. For the range of parameters we fix $\beta_1 = \beta_2$ and let β_1 varies from 1, 2, ..., 10. For the parameters of the bivariate Poisson distribution we fix $\phi_{12} = 0.1$, $\phi_1 = \phi_2$ and

let $\phi_1 = 1.1, 3.1, 4.1, 5.1, 6.1, 8.1, 10.1$. Similarly, we fix $\beta_1 = \beta_2$ and let β_1 varies from 1, 2, ..., 10 and we fix $\phi_{12} = 0.2$, $\phi_1 = \phi_2$ and let $\phi_1 = 1.2, 3.2, 4.2, 5.2, 6.2, 8.2, 10.2$.

The mean square errors (*MSE*) are estimated using simulated samples. The mean square error of an estimator $\hat{\pi}$ for π_0 is defined as

$MSE(\hat{\pi}) = E(\hat{\pi} - \pi_0)^2$. We use the following criterion for comparison between the efficiencies of MEEL estimators versus MM estimators

$$ARE = \frac{MSE(\hat{\phi}_1) + MSE(\hat{\phi}_2) + MSE(\hat{\phi}_{12}) + MSE(\hat{\beta}_1) + MSE(\hat{\beta}_2)}{MSE(\tilde{\phi}_1) + MSE(\tilde{\phi}_2) + MSE(\tilde{\phi}_{12}) + MSE(\tilde{\beta}_1) + MSE(\tilde{\beta}_2)}$$

The results are displayed in **Table 1** where we observe that overall the MEEL estimators are two to four time more efficient than MM estimators with the *ARE* estimated using simulated samples. We also observe that each individual estimate *MSE* is smaller for MEEL estimators than their counterpart for MM estimators. Furthermore, we also have a chi-square goodness-of-fit statistics for model selection for a model without closed form density function. This model testing statistics can be useful for applications in general. Despite the study is limited as

Table 1. Overall relative efficiency of MEEL estimators versus MM estimators.

$$ARE = \frac{MSE(\hat{\phi}_1) + MSE(\hat{\phi}_2) + MSE(\hat{\phi}_{12}) + MSE(\hat{\beta}_1) + MSE(\hat{\beta}_2)}{MSE(\tilde{\phi}_1) + MSE(\tilde{\phi}_2) + MSE(\tilde{\phi}_{12}) + MSE(\tilde{\beta}_1) + MSE(\tilde{\beta}_2)}$$

($\phi_{12} = 0.1, \phi_1 = \phi_2, \beta_1 = \beta_2$)

ϕ_i/β_i	1	2	3	4	5	6	8	10
1.1	0.5335	0.4046	0.3099	0.3170	0.2127	0.1783	0.1514	0.1941
3.1	0.5033	0.4405	0.4408	0.4868	0.4508	0.3161	0.2532	0.2392
4.1	0.3655	0.4629	0.4102	0.3506	0.2924	0.3418	0.2921	0.2859
5.1	0.4405	0.4374	0.3374	0.3913	0.3246	0.3063	0.3119	0.2887
6.1	0.3717	0.3741	0.4499	0.3476	0.3316	0.3269	0.2876	0.3398
8.1	0.3920	0.3778	0.3682	0.3296	0.2768	0.3011	0.2322	0.2613
10.1	0.3221	0.3597	0.2950	0.3032	0.3414	0.2836	0.2186	0.2562
($\phi_{12} = 0.2, \phi_1 = \phi_2, \beta_1 = \beta_2$)								
ϕ_i/β_i	1	2	3	4	5	6	8	10
1.2	0.4735	0.5555	0.4253	0.3800	0.4123	0.4077	0.4708	0.7260
3.2	0.4654	0.5085	0.3709	0.3865	0.4153	0.4300	0.3662	0.4414
4.2	0.3899	0.4755	0.3437	0.2910	0.3252	0.3526	0.3929	0.2642
5.2	0.3817	0.4329	0.3574	0.3301	0.2530	0.3645	0.3737	0.4066
6.2	0.3910	0.3654	0.4597	0.3235	0.3505	0.2529	0.3301	0.3078
8.2	0.5020	0.3809	0.3976	0.3671	0.3014	0.3129	0.2254	0.2506
10.2	0.4410	0.4415	0.3799	0.2931	0.3213	0.2506	0.2382	0.2700

we do not have large computing resources but it does confirm the asymptotic results on efficiencies of MEEL estimators. More works need to be done for assessing the efficiencies of MEEL methods by using various parametric models especially in finite samples.

6. Conclusion

MEEL estimators using the proposed moment conditions or bases appear to have the potential of higher efficiencies than MM estimators in general. The methods also provide chi-square goodness-of-fit test statistics for model testing. These features make the methods useful for inferences for bivariate distributions with closed form BLTs or BMGFs without closed form for the bivariate density functions. For these distributions, the implementation of likelihood methods might be difficult.

Acknowledgements

The helpful and constructive comments of a referee which lead to an improvement of the presentation of the paper and support from the editorial staffs of Open Journal of Statistics to process the paper are all gratefully acknowledged.

References

- [1] Ruijter, M.J. and Oosterlee, C.W. (2012) Two Dimensional Fourier Cosine Series Expansion Method for Pricing Financial Options. *SIAM Journal of Scientific Computing*, **34**, B642-B671. <https://doi.org/10.1137/120862053>
- [2] Partrat, C. (1995) Compound Model for Two Dependent Kinds of Claims. *Insurance: Mathematics and Economics*, **15**, 219-231. [https://doi.org/10.1016/0167-6687\(94\)90796-X](https://doi.org/10.1016/0167-6687(94)90796-X)
- [3] Crowder, M. (2012) Multivariate Survival Analysis and Competing Risks. CRC Press, Boca Raton. <https://doi.org/10.1201/b11893>
- [4] Abate, J. and Whitt, W. (1996) An Operational Calculus for Probability Distributions. *Advanced in Applied Probability*, **28**, 75-113. <https://doi.org/10.2307/1427914>
- [5] Shih, J.H. and Louis, T.A. (1995) Inferences of the Association Parameter in Copula Models for Bivariate Survival Data. *Biometrics*, **51**, 1384-1399. <https://doi.org/10.2307/2533269>
- [6] Marshall, A.W. and Olkin, I. (1988) Families of Multivariate Distributions. *Journal of the American Statistical Association*, **83**, 834-841. <https://doi.org/10.1080/01621459.1988.10478671>
- [7] Luong, A. (2017) Maximum Entropy Empirical Likelihood Methods Based on Laplace Transforms for Nonnegative Continuous Distribution with Actuarial Applications. *Open Journal of Statistics*, **7**, 459-482. <https://doi.org/10.4236/ojs.2017.73033>
- [8] Marshall, A.W. and Olkin, I. (1967) A Multivariate Exponential Distribution. *Journal of the American Statistical Association*, **62**, 30-44. <https://doi.org/10.1080/01621459.1967.10482885>
- [9] Assaf, D., Langberg, N.A., Savits, T.H. and Shaked, M. (1984) Multivariate Phase Type Distributions. *Operations Research*, **32**, 688-702.

- <https://doi.org/10.1287/opre.32.3.688>
- [10] Dwass, M. and Teicher, H. (1957) On Infinitely Divisible Random Vectors. *Annals of Mathematical Statistics*, **28**, 461-470. <https://doi.org/10.1214/aoms/1177706974>
- [11] Feller, W. (1971) An Introduction to Probability and Its Applications. Volume 2, Wiley, New York.
- [12] McNeil, A.J., Frey, R. and Embrechts, P. (2005) Quantitative Risk Management: Concepts, Techniques, Tools. Princeton University Press, Princeton.
- [13] McAssey, M.P. (2012) An Empirical Goodness-of-fit Test for Multivariate Distributions. *Journal of Applied Statistics*, **40**, 1120-1131. <https://doi.org/10.1080/02664763.2013.780160>
- [14] Mudholkar, G.S., McDermott, M. and Srivastava, D.K. (1992) A Test of p Variate Normality. *Biometrika*, **79**, 850-854. <https://doi.org/10.1093/biomet/79.4.850>
- [15] Moore, D.S. and Stubblebine, J.B. (1981) Chi-Square Tests for Multivariate Normality with Applications to Common Stock Prices. *Communications in Statistics, Theory and Methods A*, **10**, 713-738. <https://doi.org/10.1080/03610928108828070>
- [16] Qin, J. and Lawless, J. (1994) Empirical Likelihood and General Estimating Equations. *Annals of Statistics*, **22**, 300-325. <https://doi.org/10.1214/aos/1176325370>
- [17] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2013) Loss Models: Further Topics. Wiley, New York. <https://doi.org/10.1002/9781118787106>
- [18] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2012) Loss Models: From Data to Decisions. Fourth Edition, Wiley, New York.
- [19] Balakrishnan, N. and Lai, C.-D. (2009) Continuous Bivariate Distributions. Second Edition, Springer, New York, Chapter 4.
- [20] Barlow, R.E. and Proschan, F. (1981) Statistical Theory of Reliability and Life Testing. Holt, Rinehart and Winston, New York.
- [21] Mardia, K.V. (1970) Families of Bivariate Distributions. Griffin, London.
- [22] Ross, S. (2013) Simulation. Fifth Edition, Elsevier, New York.
- [23] Arevalillo, J.M. (2003) Inverting a Saddlepoint Approximation. *Statistics and Probability Letters*, **61**, 421-428. [https://doi.org/10.1016/S0167-7152\(02\)00402-9](https://doi.org/10.1016/S0167-7152(02)00402-9)
- [24] Glasserman, P. and Liu, Z. (2010) Sensitivity Estimates from Characteristic Functions. *Operations Research*, **58**, 1611-1623. <https://doi.org/10.1287/opre.1100.0837>
- [25] Genest, C., Rémillard, B. and Beaudoin, D. (2009) Goodness-of-Fit Tests for Copulas: A Review and Power Study. *Insurance: Mathematics and Economics*, **44**, 199-213. <https://doi.org/10.1016/j.insmatheco.2007.10.005>
- [26] Frees, E.W. and Valdez, E.A. (1998) Understanding Relationships Using Copulas. *North American Actuarial Journal*, **2**, 1-25. <https://doi.org/10.1080/10920277.1998.10595667>
- [27] Schennach, S. (2007) Point Estimation with Exponential tilted Likelihood. *Annals of Statistics*, **35**, 634-672. <https://doi.org/10.1214/009053606000001208>
- [28] Owen, A. (2001) Empirical Likelihood. Chapman and Hall, New York. <https://doi.org/10.1201/9781420036152>
- [29] Mittelhammer, R.C., Judge, G.G. and Miller, D.J. (2000) Econometrics Foundations. Cambridge University Press, Cambridge.
- [30] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997) Discrete Multivariate Distributions. Wiley, New York.