

Student's t-Copula 적합을 통한 Heavy Tail형 SCM 수요 데이터의 모델링 및 분석

김태성*, 이현수*
산업공학부, 국립 금오공과대학교*

Forecasting Modeling of Heavy Tail Typed Demand using Student's t-Copula Fitting in Supply Chain Management

Taesusung Kim*, Hyunsoo Lee*

School of Industrial Engineering, Kumoh National Institute of Technology*

요약 SCM의 관리 포인트가 공급중심에서 수요관리 중심으로 옮겨짐에 따라, 정확한 수요 예측을 위한 많은 기법들이 제시되어 왔다. 이 중 변수간의 인과관계 분석을 통한 수요예측이 많이 이루어지고 있음에도 불구하고, 연관된 변수들 간의 상관구조는 상관계수에 의존하였고, 이는 예측의 정확성을 저하시키는 요인으로 작용하였다. 본 논문에서는 기존 방법의 문제점들을 보완하며, SCM에서 발생하는 Heavy Tail형 데이터의 상관구조를 정밀하게 모델링할 수 있는 방법을 제시한다. 상관구조를 파악할 수 있는 프레임워크인 코플라 함수 중에서 Student's t-코플라 함수를 통하여 수요 예측모형을 수립하고, 관련 파라미터를 추정하는 기법을 실험과 함께 제시하였다. 이를 통해, 수요예측에 필요한 변수들 간의 상관구조 파악이 보다 명확해지며, 이는 SCM상의 체적효과의 완화로 이어져, 안정된 공급 사슬 네트워크의 관리에 기여할 것으로 기대된다.

주제어 : 코플라, 상관 구조, 인과 분석, 수요 예측, 헤비 테일 분포, 공급 사슬 네트워크

Abstract As the demand-oriented management has been getting important in Supply Chain Management (SCM), various forecasting methods have been suggested including regression analyses. However, dependency structures among variables have been captured by a correlation coefficient, only. It results in inaccurate demand predictions. This paper suggests a new and effective forecasting modeling framework using student's t-copula function. In order to show overall modeling procedures framework, heavy tail typed numerical data and its copula estimations are provided. The suggested methodology can contribute to decrease the bullwhip effect and to stabilize volatile environment in a supply chain network.

Key Words : Copula, Dependency Structure, Casual Analysis, Demand Forecasting, Heavy Tail Distribution, Supply Chain Network

* 이 연구는 금오공과대학교학술연구비에 의하여 지원된 논문임

Received 14 August 2013, Revised 9 September 2013

Accepted 20 September 2013

Corresponding Author: Hyunsoo Lee(Kumoh National Institute of Technology)

Email: hsl@kumoh.ac.k

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

정확한 수요 예측 (Demand Forecasting)은 기업의 핵심 경쟁력 중 하나인 SCM (Supply Chain Management/공급 가치사슬 경영)의 성공을 좌우하는 가장 중요한 요인으로 자리 잡아 왔다. 90년대 중반부터 2000년대 초반까지 SCM 시스템을 구축했던 다국적 기업 (Global Company)들의 목표는 공급 중심의 관점에서 투입 자원 및 재고를 최소화하며, 목적지까지의 총 기간을 줄이는데 필요한 모니터링 및 관련 최적화 알고리즘의 구현이었다. 그러나, 해외법인 및 공급/수요 다변화 등으로 인하여 생성된 대형화 (Large-scale) [1][2]와, 보다 복잡한 불확실성 (Complicate Uncertainties)을 내재하는 SCM 네트워크 [3]의 등장은 네트워크 상의 부분적인 최적화가 아니라, 불확실성을 일으키는 보다 근본적인 원인에 대한 시스템적 접근을 요구하게 되었다. 이러한 경향은 SCM 네트워크에서 채찍효과 (Bullwhip Effect)를 일으키는 근본적인 요인인 수요 예측의 정확성 및 정밀성에 대한 요구와 관련된 제반 시스템의 구축으로 이어져, 2000년대 중반을 거쳐 현재까지 수요 예측 및 관리 중심의 SCM으로 그 관리 초점이 옮겨지게 되었다.

현재 SCM에서 실시하는 수요예측은 시간구간 (Time Horizon) 및 예측 주체 (Forecasting Hierarchy) 등으로 세부적으로 나뉘어 관리된다. 일반적으로 수요 예측의 객체는 품목 (SKU, Stock Keeping Unit)으로 설정되며, 시간 구간에 따라 단기 예측, 중기 예측, 장기 예측으로 나뉘고, 예측의 주체에 대한 분류로 마케팅 예측 (Marketing Forecasting), 법인 예측 (Subsidiary Forecasting), 영업 예측 (Salesman Forecasting), 고객 예측 (Customer/Vendor Forecasting) 등으로 나뉘어 진행된다. 시간 구간에 대한 구분은 SKU에 해당하는 제품이나 부품의 생명주기 (Product Lifecycle) 및 관리 시점으로 나뉘어 설정된다. 과거에는 단기 예측구간의 중요성이 보다 중요하게 여겨졌으나, 현재에는 중기 예측과 장기 예측이 전략적으로 그 정확성이 요구되고 있는데, 이는 주로 SKU를 구성하는 핵심 자재나 부품의 사전 구매 및 준비로 사용될 수 있기 때문이다. 수요 예측 주체별 경향은, VMI (Vendor Managed Inventory)와 같은 협업 SCM 기법의 발달로 인하여 고객 예측이 중시되는 방향으로 나아가고 있다.

본 연구에서는 이 중 정확한 수요예측을 위한 정량적 예측기법에서 인과관계 규명을 통한 예측기법향상에 초점을 둔다.

일반적인 수요 예측 시스템은 수요 예측 주기 및 예측 계층에 따른 예측량 (Forecasting Amount) 및 예측가 (Forecasting Price)를 그 입력으로 사용한다. 이때, 전형적인 수요 예측 시스템은 예측 주체로 하여금 예측의 참조치 (Reference)를 제시하여 예측의 편이성 등을 제고하게 하는데, 이때 사용하는 참조 데이터로는 다음과 같은 것들이 있다.

- Type I : 과거 실적 데이터 (Historical Data) 및 전처리 (Pre-processing) 데이터
- Type II : 시계열 분석을 통한 전처리 데이터
- Type III : 인과 및 상관관계 분석을 통한 통계 처리 데이터
- Type IV : 다양하고 혼합된 데이터 마이닝을 통한 예측 데이터

수요 예측의 참조 데이터로서 과거 실적 데이터 (Type I)는 아직도 많은 기업에서 사용되고 있지만, 불확실성 및 복잡도의 증가는 기존의 수요 양상 (Pattern)과 다른 유형을 보여주게 되고, 이를 통해 Type II나 Type III, Type IV의 활용으로 변화되고 있다. 각 예측 참조치의 유형에 따른 비교는 표 1에서 기술하고 있다.

(Table 1) Comparison among forecasting methods as forecasting references

Type	Main methods	Advantage	Limitation
I	Historical Data /Smoothing methods	Direct usage with less-preprocessing	Weakness in exceptional cases
II	MA/AR/ARMA /ARIMA/NARCH/etc	Reflection of Time-related factors	Weakness in exceptional cases
III	Regression Analysis /Bayesian Networks	Reflection of factors with casual relationships	Errors from nonlinear models
IV	PCA/LDA/etc	Capturing ability of unnoticed patterns	Not-general method

표 1에서 보여주는 바와 같이, 예측 데이터가 전형적

인 시계열 데이터임에도 불구하고 계절성 (Seasonal Effect) 및 추세 (Trend)로서 설명할 수 없는 잡음 (Noise)이 크게 작용하기 때문에 Type II 형 참조데이터 보다는 Type III형 데이터가 실제로는 많이 사용되고 있다. 빅데이터 (Big Data) 및 데이터 마이닝 기법의 발전으로 Type IV형 예측 참조치에 대한 연구 또한 많이 진행되고 있는데, 이는 세부 예측항목마다 일반적으로 적용할 수 없다는 일관성에 약점을 가진다. 이러한 경향은 많은 예측 시스템들이 참조데이터로서 Type III형 데이터를 제시하는 것에서 나타난다. Oracle사의 SCM수요 예측 엔진인 Demantra [4]나 SAS사의 수요 예측 프로그램 [5] 은 수요예측 주체들과의 컨설팅 및 관련 데이터의 분석을 통하여 나타나는 인과관계를 지닌 요소를 통한 회귀식을 도출하여 이를 예측에 활용하거나, 여기에 베이저안 통계를 통해 정확성을 높이는 방식을 채택하고 있다. 그러나, 이러한 방식은 일반적인 회귀방정식 모형이 보여주듯이 변수간의 상관관계가 단순한 선형이나 계승적 비선형 관계 (Polynomial Nonlinear Regression Equation)의 모델링을 통해, 실제로 존재하는 상관관계의 손실 및 왜곡의 오류를 유발시킬 수 있고, 이는 최종적으로 왜곡된 예측 참조치 생성으로 이어져 왔다.

이에 본 논문에서는 코플라 (Copula)를 사용하여, SCM의 정확한 수요예측을 위하여 파악된 인과 관계 (Casual Relationship)를 가진 변수들의 상관관계 왜곡을 최소화하고, 이를 통하여 정확한 수요 예측 모형이 수립될 수 있도록 하는 방안을 제시한다.

이를 위하여 2장에서는 변수간의 상관관계 도출 방안 및 코플라에 대하여 조명하고, 3장에서는 연구의 유효성 검증을 위하여 사용된 SCM 예측데이터 및 관련 인자를 설명하며, 4장과 5장에서는 이에 대한 코플라 파라미터 (Parameter) 추정 및 모형 생성 기법을 제시한다.

2. 연구 배경 및 문헌 연구

수요예측을 위한 회귀방정식은 종속변수와 독립변수들 간의 상관관계를 분석하여 상관도가 높은 변수들과 관련된 교호작용의 상관도가 높은 변수식들을 추출하고 이를 조합하여 구성된다. 이때 사용되는 상관관계는 주로 피어슨 상관계수 (Pearson Correlation Coefficient)로

서 측정되는데, 두 변수 X (독립변수)와 Y (종속변수)간의 상관계수는 식 (1)에 의하여 정의된다.

$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E(X - \mu_X)(Y - \mu_Y)}{\sigma_X \sigma_Y} \quad (\text{eq. 1})$$

이를 통하여 정의된 상관계수는 -1과 1사이의 값을 가지는데, 이를 통하여 종속변수 Y의 추정을 위하여 독립변수 X를 회귀방정식에 구성할 것인지를 통계적으로 판단하게 된다. 기존의 회귀식을 이용한 SCM 수요예측에서의 문제점은 크게 두 가지로 요약될 수 있는데, 이는 다음과 같다.

- 상관도의 단일 값 추정으로 다차원의 상관관계 유실
- 독립변수들과 종속변수간의 비선형식 생성의 난이점

첫 번째 문제점인 상관도의 단일 값 추정은 식 (2)가 보여주듯이, 변수들 간의 상관관계가 다차원에서 정의되는데 불구하고, 하나의 단일값으로 변환되어 실제 변수들이 가지는 상관 맵핑이 유실되는 데에서 기인한다.

$$R \times R \rightarrow R^0 \quad (\text{eq. 2})$$

식 (1)에서 보여주는 피어슨 상관계수 역시 두 변수간의 관계를 선형으로 측정된 값으로서, X와 Y의 세부 범위에서 존재하는 상관관계가 유실되는 단점을 지닌다. 이러한 단점은 X, Y변수의 샘플링 데이터를 rank로 정렬하여 처리한 X'과 Y'의 상관관계인 스피어만 상관계수 (Spearman Correlation Coefficient / 식 (3))나 일관성을 측정하는 상관도인 크론바흐 알파계수 (Cronbach Alpha Coefficient) 등에서도 동일하게 나타난다. 즉, 두 변수간의 맵핑 관계가 단일 값으로 변환되어, 다차원의 관계가 유실되는 한계를 지니게 된다. 이러한 문제점은 Yen and Johnston [6]를 포함한 다양한 연구에서 지적되어 왔다.

$$\rho'_{X'Y'} = \frac{E(X' - \mu_{X'})(Y' - \mu_{Y'})}{\sigma_{X'} \sigma_{Y'}} \quad (\text{eq. 3})$$

상관관계를 나타내는 통계치의 또 다른 문제점의 하나는, 최종적으로 구성하고자 하는 독립변수와 종속변수와의 관계식 구성이 효율적이지 않다는 데에 있다. 즉, 종속변수 Y와 X가 상관관계가 있다고 가정될 때, 관계식을 구성하기 위하여 식 (4)와 같은 다항식을 세워서, 오차를 줄이는 각 n차 항의 계수들을 구하는 방법은 변수 간 비대칭/비선형 관계를 가지고 있을 때 적절하지 못하며, 효율적이지 않음을 보여주고 있다.

$$Y = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 \quad (\text{eq. 4})$$

이러한 문제점을 극복하기 위하여, 변수들 간의 상관관계를 단일한 값이 아니라, 변수들 간 상관구조의 차원을 유지 (식 (5))하면서 모형화하는 프레임워크들이 제시되었는데, 그 중의 하나가 코플라(Copula)이다.

$$R \times \dots \times R \rightarrow R^n \quad (\text{eq. 5})$$

코플라는 변수들 간의 결합 분포 (Joint Distribution)을 이용하여 상관관계를 모델링하는 틀로서, 1973년 Sklar [7]에 의하여 도입된 이래로, 경제, 사회 및 다양한 응용범위에서 사용되었고, 2000년대 초의 금융 파생상품 설계의 핵심 도구로서 자리 잡아왔다. 코플라는 확률분포함수인 C.D.F. (= F(x))로부터 유도될 수 있다 (식 (6)~(8)).

$$\Pr(X \leq x) = F(x) \geq y \quad (\text{eq. 6})$$

$$F(y)^{-1} = \{X : F(x) \geq y\} \quad (\text{eq. 7})$$

$$\Pr(F(y)^{-1} \leq x) = F(x) \quad (\text{eq. 8})$$

이때, 식 (8)은 식 (9)로 변환되며, 단일 분포를 따르는 변수에서 다변수 확률변수들로 확장 / 적용 (Multivariate cases) 될 때에, 식 (10)과 같이 코플라 c(·)의 정의를 이끌어낸다.

$$F(x) = F(F(y)^{-1}) \quad (\text{eq. 9})$$

$$F(F(y_1)^{-1}, F(y_2)^{-1}, \dots, F(y_d)^{-1}) = c(y_1, y_2, \dots, y_d) \quad (\text{eq. 10})$$

코플라는 변수들 간의 상관구조를 결합분포를 사용하여 정의함으로써, 기존의 상관계수형 통계치들이 가지는 단점들의 한계를 극복하고, 보다 상세한 상관관계를 모델링할 수 있다는 장점을 지니고 있다. 보다 자세한 특징, 관련 Theorems 및 응용은 Trivedi and Zimmer [8], Mai and Scherer [9]를 포함한 많은 연구에서 참고할 수 있다.

본 연구에서는 코플라를 사용하여 SCM의 수요 예측 모델의 모형 수립 및 분석을 통해 인과관계에 있는 변수들간의 상관관계를 명확히 모델링 하고, 보다 정확한 인과 모형이 만들어 질 수 있도록 한다.

3. 예측 데이터 및 인과 변수 설정

코플라가 상관구조를 모델링할 수 있는 우수한 도구임에도 불구하고, SCM의 수요예측 분야에서 이를 활용한 접근은 아직 성숙단계에 접어들지 않고 있다. Fridman, et al. [10]를 비롯한 많은 연구가 경제, 경영 데이터의 분석모형으로 코플라를 활용한 반면, SCM부문의로의 적용 및 응용은 Kaki, et al. [11]을 포함한 소수에 아직 머무르고 있으며, 특히 수요예측에서의 모델링 및 분석에 대한 연구는 시작점에 있다.

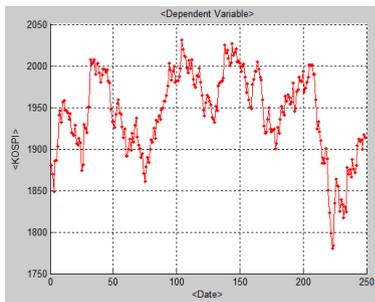
이에 본 연구에서는 SCM의 수요예측 관점에서 예측 및 실적데이터 (종속변수)와 인과관계를 갖는 변수 (독립변수)를 잡아 그 상관관계를 코플라를 통하여 모델링하고, 관련 파라미터를 적합시켜 분석한다. 이를 위하여 실제 SCM 데이터와 같은 정도의 변동성 (Volatility)를 보여주는 추가 데이터를 종속변수로 삼고, 인과관계가 있다고 판단되는 환율 (Won/Dollar Exchange Rate)을 독립변수로 삼아, 기존의 상관계수 방법과의 비교와 함께, 코플라를 통하여 상관구조를 명확하게 조명토록 한다. 표 2는 본 논문의 분석을 위하여 사용한 데이터, 종속변수 및 독립변수를 설명한다.

<Table 2> Dependent variable and independent variable for identifying dependent structure

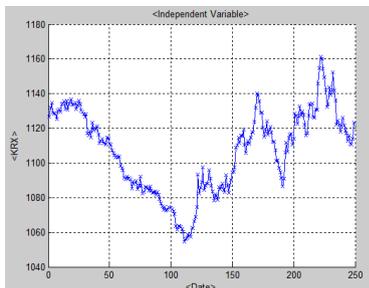
Variable	Dependent variable (Y)	Independent variable (X)
Data	A SKU's long-term forecasting data	Data with a casual effect
Processed example	Korean Stock Price	Korea Won-dollar exchange rate

본 연구에서 수요 예측 데이터를 주가로서 병치하여 설명하는 높은 것은, 주가 데이터 역시 수요와 공급 및 다양한 변수에 의하여 결정되는 데이터로서, 일반적인 회귀 식 및 시계열에 의한 방법으로는 설명이 어렵다는 데 기인한다.

연구의 유효성 증명을 위하여 사용한 종속변수 (Y)의 데이터로서 한국거래소 (<http://www.krx.co.kr>)의 2012년 8월 1일부터 2013년 7월 31일까지의 과거 1년간 KOSPI 종합주가지수 데이터를 사용하였다. 이는 실제 SCM부문에 장기 수요예측 데이터로서 대체될 수 있다. 그림 1 (a)는 종속변수의 데이터 플롯을 나타낸다. 마찬가지로, 인과관계로 있다고 판단되는 독립변수로는 같은 기간의 원/달러 환율을 사용하였다. 실제 SCM 수요예측에서는 사용된 종속변수 및 독립변수 대신에 관련된 SKU의 실적 및 인과관계가 파악된 변수값들이 사용된다.



(a) Data plot of a dependent variable

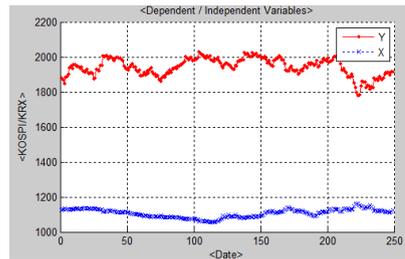


(b) Data plot of an independent variable

[Fig. 1] Data plots of dependent / independent variables

설정된 데이터를 통하여 피어슨 상관계수 (식 (1))를 구해보면 두 변수 (X, Y)사이의 피어슨 상관계수는 -0.5344로서 음의 상관관계를 가짐을 알 수 있다. 그러나, 절대값이 1에 가까울수록, 상관도가 높고, 0에 가까울수

록 상관도가 낮은 사실로 볼 때, 두 변수 사이의 피어슨 상관계수는 변수 X를 종속변수 Y를 위한 유효한 회귀 변수로서 사용할 것인지에 대한 추가적인 테스트를 필요로 한다. 그림 2는 같은 일자에 플롯된 두 변수들의 데이터를 보여준다.



[Fig. 2] Data plots of dependent and independent variables in the same graph

2장에서도 기술하였듯이, 기존의 상관계수는 두 변수들간의 상관구조를 정확하게 모델링할 수 없으므로, 이로 인하여 구축되는 인과모형 또한 한계를 드러내게 된다. 이를 극복하기 위하여 다음 장에서는 코플라를 사용하여 정확한 상관관계를 파악하고 모형화하는 방법을 제시한다.

4. Heavy Tail형 데이터의 Student's t-코플라 적합 모델

코플라 함수의 종류로는 모수형 (Parametric) 코플라와 비모수형 (Nonparametric) 코플라로 나누어 지며, 표준적인 모수형 코플라 함수에는 Product형, Farlie-Gumbel-Morgenstern 형, Student's t형, Gaussian형, Frank형, Ali-Mikhail-Haq형 등으로 분류된다. 자세한 정의 및 관련 파라미터는 Trivedi and Zimmer [8]을 따른다. 이 중, 3장에서 제시된 데이터의 상관구조를 위한 코플라 모델로는 모수형 (Parametric) 코플라 모델 중에서 Student's t형 코플라를 사용한다. 이는 데이터가 보여주는 Heavy Tail형 분포 (그림 3)를 설명하기에 적합한 모수형 코플라가 Student's t형 코플라이기 때문이다. Student's t-copula는 식 (11)과 같이 정의된다.

$$c(y_1, y_2; \theta_1, \theta_2) = \int_{-\infty}^{F_{\theta_1}^{-1}(y_1)} \int_{-\infty}^{F_{\theta_2}^{-1}(y_2)} \frac{1}{2\pi \sqrt{1-\theta_2^2}} \times \left\{ 1 + \frac{s^2 - 2\theta_2 st + t^2}{\nu(1-\theta_2^2)} \right\}^{-(\theta_1+2)/2} ds dt \quad (\text{eq. 11})$$

식 (11)에서 보는바와 같이 Student's t-코플라는 두 개의 파라미터 (θ_1, θ_2)를 가지는데, 이 중 θ_1 은 자유도 (degree of freedom)로서 heavy tail의 모양을 좌우하는 shape 파라미터이고, θ_2 는 추정 상관계수로서 두 데이터 사이의 상관 구조를 기술하는 파라미터로 사용한다.

이러한 파라미터들을 추정하기 위한 방법에는 여러 가지가 있는데, 본 논문에서는 Weib [12]의 방법을 변형한 Maximum Likelihood 방법을 통하여 두 변수를 추정한다. 이를 위하여 데이터를 통하여 식 (12)과 같이 Maximum Likelihood 함수를 추정하고, 식 (13)를 풀어내기 위한 비선형 최적화 프로그래밍을 통하여 각 파라미터를 추정한다.

$$L_y(\theta) = \sum_{i=1}^n \log c(y_{i,1}, y_{i,2} | \theta) \quad (\text{eq. 12})$$

where, $\theta = \{\theta_1, \theta_2\}$

$$\hat{\theta}(y_1, y_2) \cong \arg \max_{\theta} \sum_{i=1}^n \log c(y_{i,1}, y_{i,2} | \theta) \quad (\text{eq. 13})$$

식 (14)를 통하여 추정된 파라미터 및 이를 통한 Student's t-코플라 모델은 Heavy tail을 가진 데이터의 독립변수와 종속변수의 상관구조를 기술하는 SCM 예측 도구로 활용되게 된다.

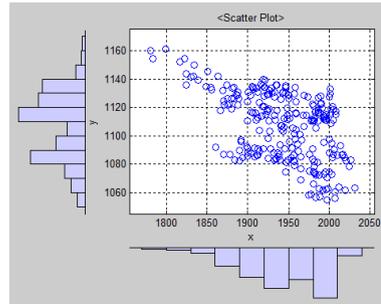
5. 수치해석 및 예측 모델 수립

본 장에서는 4장에서 기술한 방법을 통하여 3장에서 제시된 데이터에 맞는 Student's t-코플라 파라미터들을 적합하고, 이를 토대로 코플라 분포함수와 밀도함수를 수립하는 방법을 보여준다.

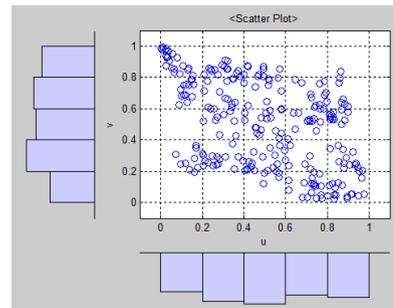
먼저, 두 데이터가 Heavy tail을 가짐을 확인하기 위하

여 상관관계를 기술하는 산점도를 확인하도록 한다 (그림 3 (a)).

이제, 그림 3 (b)와 같이 코플라 스페이스에서 변환된 데이터를 식 (12)과 (13)를 이용하여 자유도 (θ_1)와 상관 계수 (θ_2)를 추정한다. 이를 통하여 추출된 파라미터들의 값은 표 3에서 보여준다.



(a) Scatter plot with the data in original space



(b) Scatter plot with the data in copula space

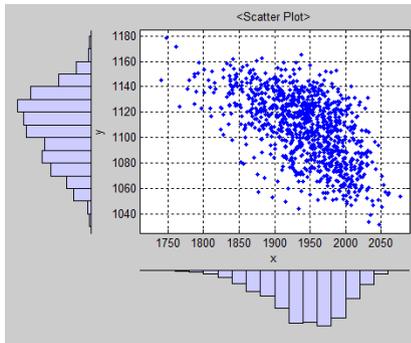
[Fig. 3] Scatter plot with the data in the section 3

[Table 3] Estimated parameter values in the Student's t-copula

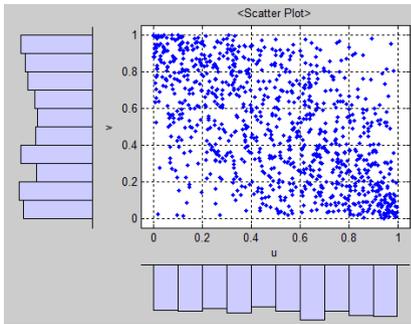
Parameter	Degree of Freedom	Correlation Coefficient
Value	18.4658	-0.5545

추정된 파라미터들을 통하여 구성된 Student's t-코플라 함수를 통해 생성한 데이터는 코플라 스페이스와 원래의 차원에서 각각 그림 4의 (a)와 (b)와 같이 생성되어 보다 명확한 상관구조를 보여주게 된다.

분포함수 (C.D.F)의 최종 모습은 그림 5와 같다. 그림 5에서 볼 수 있듯이 상관구조가 변수들 간의 맵핑 차원과 동일하게 파악되어, 보다 정밀한 수요예측이 이루어 질 수 있도록 한다.



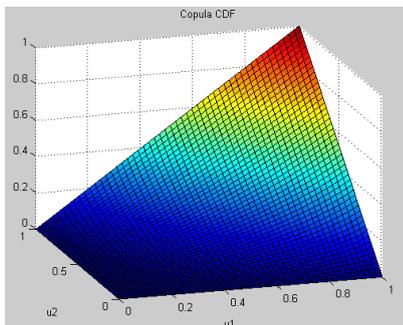
(a) Scatter plot with the generated data in copula space



(b) Scatter plot with the generated data in original space

[Fig. 4] Scatter plots with newly generated data

이러한 과정을 통하여 생성되는 Student's t-코플라



[Fig. 5] Estimated Student's t-copula cumulative density function

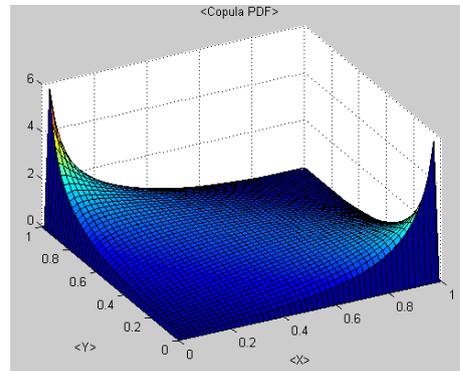
최종적으로 코플라 밀도함수 (copula probability density function)는 식 (14)를 통하여 그림 6과 같이 추출된다.

이렇게 형성된, 코플라 함수는 독립변수와 종속변수와

의 상관구조를 명확하게 나타내며, 이를 통해 기존의 상관계수를 통한 방법 등에서 파악되지 않았던 연관관계를 정밀하게 추정하고, 보다 정확한 수요예측을 위한 기본 분석 자료로 사용된다.

$$c'(y_1, y_2) = \frac{\partial}{\partial y_1 \partial y_2} c(y_1, y_2) \tag{eq. 14}$$

$$= \lim_{\Delta y_1, \Delta y_2 \rightarrow 0} c(y_1, y_2) - c(y_1 - \Delta y_1, y_2 - \Delta y_1)$$



[Fig. 6] Estimated Student's t-copula probability density function

6. 결론 및 차후 연구방향

본 연구에서는 코플라를 사용하여 SCM의 수요와 인과관계가 있는 독립변수와의 정확한 예측 모델을 수립하기 위한 프레임워크 및 적합 추정과 함께 수치 실험을 수행하였다. 제시된 기법은 Student's t-copula를 사용하여, 인과 관계를 통한 회귀방정식이 SCM의 수요예측 모델로서 유용함에도 불구하고, 상관계수의 단점 및 비선형 모델 생성의 난이성 등으로 인하여 예측의 정밀성이 떨어지는 한계를 극복하였으며, 통상적으로 SCM에서 발생하는 Heavy Tail 데이터 또한 모델링 할 수 있다는 장점을 지닌다. 이러한 장점은 실제 수요예측에 필요한 많은 변수들의 고려 시 다변수 코플라 적합을 통하여 정밀한 상관관계의 분석 및 특정 변수에 초점을 두었을 때, 높은 교호작용을 가지는 변수들의 추출에 이용될 수 있다.

이러한 기법의 사용은, 수요 관리 중심의 현 SCM 기본 관리방향과 부합하며, 인과관계가 있는 변수들 간의 상관구조를 세밀하게 묘사함으로써, 차후 다양한 데이터

마이닝 기법의 전처리 프로세스나 자료로서도 활용가능하다.

본 논문에서는 SCM의 수요예측 분야에서 아직 도입 단계에 있는 코플라의 적용 및 활용방안을 데이터 분석을 통하여 제시하였다. 실제 SCM의 수요 예측 데이터는 많은 다 변수들 간의 상관관계 뿐 아니라, 시계열을 포함한 복합적 예측기법을 활용하여 생성되기 때문에 이를 고려한 다변수 코플라 모델링과 효율적인 파라미터 추정 방법이 차후 연구방향으로 제시되며, 기업가치와 SCM의 상관관계를 규명하는 연구 [13] 와 같은 수요예측 이외의 영역에서도 이러한 기법의 활용이 기대된다.

ACKNOWLEDGMENTS

This paper was supported by Research Fund, Kumoh National Institute of Technology)

REFERENCES

- [1] H. Lee and A. Banerjee, An executable modeling and simulation framework for NP-Complete problems and different forms of uncertainties, *International Journal of Services Operations and Informatics*, Vol. 7, No. 1, pp. 1-17, 2012.
- [2] T. Kim and H. Lee, Topology modeling and dynamic control framework for NP-complete problem embedded supply chain network, *The 8th International Conference on Logistics and Service*, Tokyo, Japan, August 05-07, pp. 73-78, 2013.
- [3] H. Lee and A. Banerjee, Design, simulation and analysis of logistics with uncertainties, *International Conference on Value Chain Sustainability*, Louisville, Kentucky, October 19-21, pp. 1-6, 2009.
- [4] DOI : <http://www.oracle.com/u54s/products/applications/057028.pdf>
- [5] DOI :http://www.sas.com/resources/whitepaper/wp_8085.pdf
- [6] DOI :<http://www.jps.anl.gov/vol.2/3-Correlation.pdf>
- [7] A. Sklar, Random variables, joint distributions, and copulas, *kybernetika*, Vol. 9, pp. 49-460, 1973.
- [8] P. K. Trivedi and D. M. Zimmer, *Copula modeling: an introduction for practitioners*, NOW, 2005.
- [9] J. Mai and M. Scherer, *Simulating copulas, stochastic models, sampling algorithms, and applications*, Imperial College Press, 2012.
- [10] C. Friedman, J. Huang, Y. Zhang and W. Cao, Imputation via copula and transformation methods, with applications to financial and economic data, *International Journal of Data Analysis Techniques and Strategies*, Vol. 4, No. 4, pp. 315-339, 2012.
- [11] A. Kaki, A. Salo and S. Talluri, Scenario-based modeling of interdependent demand and supply uncertainties, *IEEE Transactions on Engineering Management*, Vol 1., pp. 1-13, 2013.
- [12] G.N.F. Weib, Copula parameter estimation: numerical considerations and implications for risk management, *The Journal of Risk*, Vol 13, No. 1, pp. 17-53, 2010.
- [13] Y. Kim, G. Jung, H. Lee, S. Kim, S. Jang and T. Kim, A study on the relationship between SCM and corporate value, *The Journal of Digital Policy and Management*, Vol 11. No. 2, pp. 91-99, 2013.

김태성(Kim, Taesung)



- 1993년 12월 : New Jersey Institute of Technology (공학석사)
- 2000년 12월 : Louisiana State University (공학박사)
- 2003년 3월 ~ 현재 : 금오공과대학교 산업공학부 교수
- 관심분야 : OR, 의사결정, SCM
- E-Mail : tkim@kumoh.ac.kr

이현수(Lee, Hyunsoo)



- 2002년 2월 : POSTECH (공학석사)
- 2010년 8월 : Texas A&M University (공학박사)
- 2011년 9월 ~ 현재 : 금오공과대학교 산업공학과 조교수
- 관심분야: 지능형 가상시스템 설계, 비선형 최적화, SCM

· E-Mail : hsl@kumoh.ac.kr