UNIVERSITY OF THE PHILIPPINES
SCHOOL OF STATISTICS

COLLOQUIUM ON THE STATISTICAL SCIENCES

THE SCHOOL OF STATISTICS
University of the Philippines

invites all students and faculty to a lecture on

"Likelihood analysis of Gaussian copula distributions (GCDs) via a parameter-expanded Monte Carlo EM (PX-MCEM) algorithm"

To be given by:

Dr. Alexander R. de Leon

Visiting Professor, UP School of Statistics
Associate Professor, Department of Mathematics & Statistics,
University of Calgary

21 April 2017, Friday
4:00 PM to 6:00 PM
Colloquium Room
2nd floor UP School of Statistics Building

Summary

Mixed discrete and continuous data arise in a variety of settings. In this talk, we adopt a general model based on the Gaussian copula we term Gaussian copula distributions (GCDs; Jiryaie et al., 2016, Journal of Statistical Computation & Simulation) to model the joint distribution of a mixture of (possibly non-Gaussian) continuous and discrete (e.g., binary, categorical) random variables, where the latter are described by thresholded continuous latent variables. An attractive feature of GCDs is their use of Gaussian copulas to separately model dependencies between non-Gaussian variables, thereby preserving the variables' distinct marginal properties. We adopt an efficient approach to maximum likelihood estimation for GCDs via a parameter-expanded Monte Carlo EM (PX-MCEM) algorithm (Liu et al., 1998, Biometrika). By using parameter expansion, we not only avoid the direct evaluation of the likelihood function, which involves computing multivariate Gaussian truncated orthant probabilities, but we also improve the computational efficiency of the algorithm. A Monte Carlo approach to evaluating the moments of truncated multivariate Gaussian distributions via the R package tmvtnorm avoids the tedious analytical calculations in the E-step. Another advantage of the PX-MCEM algorithm is that it has a computationally tractable M-step that merely entails the use of a root-solver (e.g., function multiroot in R package rootSolve); standard errors of estimates are calculated by bootstrap. Based on simulated data and real data from a breast cancer study at the University of Calgary, we illustrate the application of GCDs and its likelihood analysis in practice.

This is joint work with Mingchen Ren (University of Calgary) and Ying Yan (University of Calgary).