1. Estadística

Coordinadora:

- Nora Serdyukova, Departamento de Estadística, Universidad de Concepción, Concepción.

Índice

Estudio de simulación comparando el alpha de Cronbach y el alpha Ordinal

Sergio Contreras Espinoza[*]
Departamento de Estadística
Universidad del Bío-Bío
Concepción, Chile

Resumen

El objetivo principal de este estudio fue analizar las diferencias en la estimacin de la confiabilidad de un instrumento usando los estimadores alpha de Cronbach y alpha Ordinal en distintos contextos simulados, estos fueron diferentes nmeros de tems, opciones de respuesta, tamaos muestrales y valores de confiabilidad, tanto para casos simtricos como asimtricos. Para menos de 6 opciones de respuesta, alpha Ordinal es un mejor estimador que alpha de Cronbach en todos los casos simulados. Respecto del efecto de la asimetra de los datos, sta perjudica la estimacin de alpha de Cronbach, no teniendo un efecto tan grave en el caso de la estimacin utilizando alpha Ordinal.

Trabajo realizado en conjunto con:
M. Pereira[1]Departamento de Estadística, Universidad del Bío-Bío, Concepción, Chile.
F. Novoa-Muoz[2]Departamento de Estadística, Universidad del Bío-Bío, Concepción, Chile.

Referencias

[1] Contreras, S., & Novoa-Muoz, F., *Ventajas del alfa Ordinal respecto al alfa de Cronbach, ilustrado con la encuesta AUDIT-OMS*, Revista Panamericana de Salud Pblica. (2018) En prensa.

[2] Zumbo B. D., Gadermann A. M., & Zeisser C., *Ordinal versions of coefficients alpha and theta for Likertrating scales*, Journal of Modern Applied Statistical Methods. (2007) 6(2)4.

[*]Parcialmente financiado por Grupo de Matemática Aplicada GI 172409/C de la Universidad del Bío-Bío, e-mail: scontre@ubiobio.cl
[1]Parcialmente financiado por Grupo de Matemática Aplicada GI 172409/C de la Universidad del Bío-Bío, e-mail: mpereira@ubiobio.cl
[2]Parcialmente financiado por Grupo de Matemática Aplicada GI 172409/C de la Universidad del Bío-Bío, e-mail: fnovoa@ubiobio.cl

Estimation and Prediction using generalized Cauchy Covariance model under fixed domain asymptotics

Tarik Faouzi*
Department of statistics
University of Bio-Bio
Concepción, Chile

Resumen

We study estimation and prediction of Gaussian random fields with covariance models belonging to the generalized Cauchy (GC) class, under fixed domain asymptotics. Gaussian random field with this kind of covariance provide separate characterization of fractal dimension and long range dependence [LiTe09], an appealing feature in many physical, biological or geological systems.

The results of the paper are classified into three parts: first, we characterize the equivalence of two Gaussian measures with GC covariance function and we provide sufficient conditions for the equivalence of two Gaussian measures with Matérn (MT) and GC covariance functions and two Gaussian Measures with Generalized Wendland (GW) and GC covariance functions [Stein:1999, Zhang:2004].

In the second part, we establish strong consistency and asymptotic distribution of the maximum likelihood estimator of the microergodic parameter associated to GC covariance model, under fixed domain asymptotics. The third part elucidate the consequences of our results inf terms of prediction under fixed domain asymptotics.

Our findings are illustrated through a simulation study: the first compares the finite sample behavior of the maximum likelihood estimation of the microergodic parameter of the GC model with the given asymptotic distribution. We then compare the finite-sample behavior of the prediction and its associated mean square error when the true model is GC and the prediction is performed using the true model and a misspecified GW model.

Keywords: fractal dimension; long memory; microergodic parameter, maximum likelihood

Joint work with:
Moreno Bevilacqua[1], Departmentof Statistics, University of Valparaiso, Valparaiso, Chile.

*Partially supported by DIUBB 170308 3/I Chile. Supported by GI 172409/C e-mail: tfaouzi@ubiobio.cl
[1]Partially supported by FONDECYT grant 1160280 Chile, e-mail: moreno.bevilacqua@uv.cl

A stochastic methodology for risk assessment of a large earthquake when a long time has elapsed

Raúl Fierro[*]
Instituto de Matemáticas
Universidad de Valparaíso
Valparaíso, Chile

Resumen

We propose a stochastic methodology for risk assessment of a large earthquake when a long time has elapsed from the last large seismic event. We state an approximate probability distribution for the occurrence time of the next large earthquake, by knowing that the last large seismic event occurred a long time ago. We prove that, under reasonable conditions, such a distribution is exponential with a rate depending on the asymptotic slope of the cumulative intensity function corresponding to a non-homogeneous Poisson process. As it is not possible to obtain an empirical cumulative distribution function of the waiting time for the next large earthquake, an estimator of its cumulative distribution function based on existing data is derived. We conduct a simulation study for detecting in what scenario the proposed methodology performs well. Finally, a real-world data analysis is carried out to illustrate its potential applications.

Joint work with:
Víctor Leiva[1], Escuela de Ingeniería, Pontificia Universidad de Valparaíso, Valparaíso, Chile.

[*]Partially supported by FONDECYT No. 1160868, e-mail: raul.fierro@uv.cl
[1]Partially supported by FONDECYT No. 1160868, e-mail: victorleivasanchez@gmail.com

Trapezoidal Distributions and Trapezoidal beta Distribution

Jorge I. Figueroa-Zúñiga, Sebástian Niklitschek*
Departamento de Estadística
Universidad de Concepcíon
Concepción, Chile

Resumen

Beta Models have shown to be a powerful tool for the analysis and modeling of limited-range continuous variables. This mainly because of the great flexibility of its density function, which can cover a wide range of different shapes (symmetric, asymmetric, unimodal, bimodal). Recently Hahn (2008) proposed a new model for proportions more robust which considers equivalent tail-area events in both ends, namely "Rectangular beta Model". However, in many cases in the practice, the data are bounded and tail-area events occur on either end or at both ends independently. In such cases, it misses an extension of the beta and Rectangular beta Model that allows to capture this tail-behavior, for this reason in this work we propose the "Trapezoidal beta Model".

The focus of this paper is centered on the development of the "Trapezoidal beta Model", which is a generalization of both, the "beta Model."and the "Rectangular beta Model", and introduce a strategy for generating Trapezoidal distributions from any bounded density function. We study these distributions functions and we derive some fundamental properties such as the moments, moment generating function and characteristic function. To the "Trapezoidal beta Model"we will rewrite the model conveniently as a mixture model and we solve the problem of the parameter estimation using the EM algorithm. Finally, we report results of an application to a real data set. Model fitting comparison with several alternative models indicates that the model proposed presents the best fit and so it can be quite useful in real applications.

Key words: Beta distribution; Trapezoidal distribution; Maximum likelihood; EM algorithm.

*Acknowledges funding support by grant VRID N^o342/17, from the Universidad de Concepción, e-mail: jfigueroaz@udec.cl

Test de Bondad de ajustes para la distribución Hermite biinvariante

Pablo Gonzlez*
Departamento de Estadstica
Universidad del Bo-Bo
Chilln, Chile

Resumen

Los datos de conteo pueden aparecer bajo diferentes circunstancias. En un marco univariante, la distribucin Poisson es la que con mayor frecuencia ha sido empleada para modelar tales datos. En la prctica, los datos de conteo bivariantes surgen en varias disciplinas diferentes y la distribucin Poisson bivariante (DPB), siendo una generalizacin de la distribucin Poisson, Un caso particular de la DPB es la distribucin de hermite bivariante (DHB)juega un rol importante en modelarlos, siempre que dichos datos presenten una correlacin no negativa.

La distribucin de Hermite es cero-inflada respecto a la distribucin de Poisson, lo cual significa que la probabilidad del cero en la distribucin de Hermite es siempre superior a la probabilidad del cero de la Poisson, cuando ambas distribuciones tienen la misma esperanza. Tambin permite modelizar datos en los que aparece el fenmeno de sobredispersin, esto ocurre cuando la varianza es mayor que la esperanza, siendo esta sobredispersin moderada.

Contrastar la bondad de ajuste de las observaciones dadas con las hiptesis distribucionales asumidas es un aspecto crucial del anlisis de datos. Por esta razn, en este trabajo pretendemos proponer y estudiamos un test de bondad de ajuste para la DHB. Este test se basara en la funcin generatriz de probabilidades, el cual sea consistente contra alternativas fijas.

Trabajo realizado en conjunto con:
Francisco Novoa[1], Departamento de Estadstica , Universidad de Bo-Bo, Chilln, Chile.

Referencias

[1] Kocherlakota, S. y Kocherlakota, K., *Bivariate Discrete Distributions* ,New York: Marcel Dekker, (1992). 245-258.

[2] Johnson, N. L., Kotz, S. y Balakrishnan, N. *Discrete Multivariate Distributions.* (1997).. John Wiley and Sons.

*Parcialmente financiado por XXX, e-mail: pgonzalez@ubiobio.cl
[1]Parcialmente financiado por Grupo de Investigacin Matemtica Aplicada GI 172409/C de la Universidad del Bo-Bo, Concepcin, Chile, e-mail: fnovoa@ubiobio.cl

Monitoring urban environmental pollution by bivariate control charts: a new methodology and a case study in Santiago, Chile

Carolina Marchant*
Departamento de Matemática, Física y Estadística
Universidad Católica del Maule
Talca, Chile

Resumen

Particulate matter (PM) pollution is one of the main global urban environmental problems, affecting human health and life quality. Santiago, Chile, is one the most polluted cities around the world in terms of PM2.5 and PM10, because of a combination of anthropogenic, meteorological and topographic factors. Monitoring environmental risk is useful for preventing adverse effects on human health in highly polluted cities. In this work, we propose a new methodology based on bivariate quality control charts of Hotelling type using an asymmetric distribution for monitoring PM pollution in Santiago, Chile. This methodology is robust to extreme values of PM levels and constructed for both Phases I and II of control charts. We estimate the corresponding parameters with the maximum likelihood method and use parametric bootstrapping to obtain the distribution of the adapted Hotelling statistic. In addition, we consider the Mahalanobis distance to detect bivariate extreme values and use it to assess the adequacy of the distributional assumption. A simulation study is carried out for evaluating the performance of the methodology proposed in Phases I and II. A case study with PM pollution real data of Santiago, Chile, is provided. This study shows that the proposed methodology is useful for alerting early episodes of extreme air pollution, thus preventing adverse effects on human health. The obtained results are in agreement with the critical episodes reported by the Chilean health authority. [1]

Joint work with:
Víctor Leiva[1], Escuela de Ingeniería Industrial, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile.

Referencias

[1] Marchant C, Leiva V, Cysneiros FJA, Liu S (2018) Robust multivariate control charts based on Birnbaum-Saunders distributions. *Journal of Statistical Computation and Simulation* 88(1):182-202.

*e-mail: cmarchant@ucm.cl
[1]Partially supported by FONDECYT 1160868, e-mail: victorleivsanchez@gmail.com

A simulation-based study on Bayesian estimator for the skew Brownian motion.

Manuel J.P. Barahona, Laura Rifo, Maritza Seplveda, S. Torres [*]
Facultad de Ciencias
Universidad del Bo-Bo
Concepcin
Chile.

Resumen

In analyzing temporal data set from a continuous variable, diffusion processes can be suitable under certain conditions on increments distributions. We are interested in processes were a semi-permeable barrier splits the state space, producing a diffusion on each side. In this case, when Brownian motion seems to be a suitable model for the diffusion, the skew Brownian process appears as the natural choice.

In this work, we make some Bayesian inferences on the parameters involved in this class of processes, like evidence measures, hypotheses testing, and posterior and predictive estimates. The asymptotic behavior of the Bayes estimators for the skewness parameter is discussed. The analysis proposed is illustrated with an example on Dow Jones Industrial Average. We validate the inferences obtained through a simulated data set. Those examples show the suitability of the Bayesian approach to deal with temporal data.

Keywords: skew Brownian motion, full Bayesian significance test.

[*]e-mail: mpereira@ubiobio.cl

Estimation of the Regression Function With a Covariance Dependence Structure and Random Times

Tania Roa Rojas[*]
CIMFAV
Universidad de Valparaíso
Valparaíso, Chile

Resumen

In this work, we compute the least square estimator for the drift parameter in a regression model driven by fractional Brownian motion. For two random sampling consistency is shown, also a simulation study is provided in order to show the performance of the proposed method. Through the years and different research areas, we have noticed that some events, it does not occur at evenly spaced, because of this, we are will propose the estimation of a simple model but sampled at random times. Besides this, we will work with two different random times: irregular observations and renewal process.

Joint work with:
Héctor Araya[1], CIMFAV, Universidad de Valparaíso, Valparaíso, Chile.
Natalia Bahamonde[2], Instituto de Estadística, Pontificia Univerisdad Católica de Valparaíso, Valparaíso, Chile.

Soledad Torres[3], CIMFAV, Universidad de Valparaíso, Valparaíso, Chile.

Referencias

[1] Jose A. Vilar, *Kernel Estimation of the Regression Function with Random sampling Times*, Statistics & probability letters **50(1)**, (2000). 63-73.

[2] MATHAL, A. M.; MOSCHOPOULOS, P. G., *A form of multivariate gamma distribution*, Annals of the Institute of Statistical Mathematics **vol. 44, no 1**, (1992). 97-106.

[*]e-mail: tania.roa@postgrado.uv.cl
[1]Partially supported by Beca CONICYT-PCHA / Doctorado Nacional / 2016-21160138, e-mail: hector.araya@postgrado.uv.cl
[2]Partially supported by FONDECYT Grant 1160527, e-mail: natalia.bahamonde@pucv.cl
[3]Partially supported by FONDECYT Grant 1171335, e-mail: soledad.torres@uv.cl

On a Concordance Correlation Coefficient for Spatial Data

Ronny Vallejos*
Departmento de Matemática
Universidad Técnica Federico Santa María
Valparaíso, Chile

Resumen

In this work we define a spatial concordance coefficient for second-order stationary processes. This problem has been widely addressed in a non-spatial context, but here we consider a coefficient that for a fixed spatial lag allows one to compare two spatial sequences along a 45° line. The suggested index is a generalization of Lin's coefficient [1], which evaluates the agreement between two continuous variables by measuring the variation from a 45° line through the origin. The main properties of Lin's coefficient are preserved for the new coefficient. The proposed index is explored for the bivariate Matérn and Wendland covariance functions. The asymptotic normality of the maximum likelihood estimator of the spatial concordance coefficient for an increasing domain sampling framework is established for the Wendland covariance function following the guidelines given in [2]. Monte Carlo simulations are used to gain additional insights into the asymptotic properties for finite sample sizes. The results will be illustrated by real data examples to see how our method works in practice

Joint work with:
Aaron Ellison[1], Harvard Forest, Harvard University, Petersham, USA.
Andrew Richardson[2], School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, USA.
Javier Pérez[3], Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile.

Referencias

[1] LIN, L. I-K, *A concordance correlation coefficient to evaluate reproducibility.* Biometrics **45**, (1989). 255–268.

[2] MARDIA, K. V., MARSHALL, T. J, *Maximum likelihood estimation of models for residual covariance in spatial regression.* Biometrika **71**, (1984). 135–146.

*Partially supported by AC3E, FB-0008, e-mail: ronny.vallejos@usm.cl
[1]e-mail: aellison@fas.harvard.edu
[2]e-mail: Andrew.Richardson@nau.edu
[3]e-mail: javier.perezo@alumnos.usm.cl