

GOF DAYS 2012

Sevilla, 17 noviembre 2.012

<http://mat.uab.es/~acabana/gofdays.htm>

El propósito de estas jornadas es compartir nuestro entusiasmo por las diversas áreas de la bondad de ajuste y reunir a especialistas para intercambiar desarrollos recientes y fomentar el trabajo conjunto. Las jornadas están abiertas a alumnos del Programa de Doctorado “Matemáticas” de la Universidad de Sevilla.

PROGRAMA

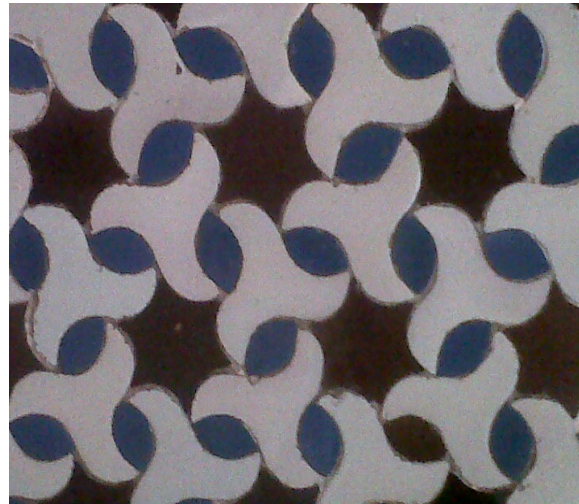
- 9:00** *Conditional Stochastic Dominance* **M.A. Delgado**
9:30 *Two-Sample Stochastic Dominance Testing Uniform in Covariates* **J.C. Escanciano**
10:00 *A nonparametric ANOVA-type test for regression curves based on characteristic functions* **J.C. Pardo-Fernández**
10:30 Discusión. Modera F. López Blazquez.
11:30 *Recortes Imparciales y Similitud de Muestras* **J.A. Cuesta-Albertos**
12:00 *Bondad de Ajuste para datos Sesgados* **J. Ojeda**
12:30 *Nonparametric testing in regression with functional data* **C. Sánchez-Sellero**
13:00 Discusión. Modera J.A. Cristóbal.
15:30 *Una nueva metodología en tests de bondad de ajuste para las innovaciones en modelos AR(1)* **P. Puig**
16:00 *New goodness-of-fit diagnostics for discrete response dynamic models* **C. Velasco**
16:30 *El proceso característico empírico de los residuos en modelos GARCH* **M.D. Jiménez Gamero**
17:00 Discusión. Modera A. Cabaña.
18:00 *Empirical likelihood approach using Phi-divergences* **N. Martín**
18:30 *Contrastes de bondad de ajuste basados en la correlación máxima de Hoeffding* **A. Grané**
19:00 Discusión. Modera L. Pardo Llorente.

LUGAR: Aula 2.4 de la Facultad de Matemáticas de la US (<http://www.matematicas.us.es/index.php>).

FINANCIAN: Grupo FQM-153, Junta de Andalucía, Proyecto MTM2008-00018, Programa de Doctorado “Matemáticas” de la Universidad de Sevilla (<http://www.imus.us.es/doctorado>).



Programa de Doctorado
MATEMÁTICAS



GOF DAYS 2012

Sevilla, 17 noviembre 2.012

PROGRAMA
RESÚMENES



FACULTAD DE MATEMÁTICAS. UNIVERSIDAD DE SEVILLA

Objetivo

El propósito de estas jornadas es compartir nuestro entusiasmo por las diversas áreas de la bondad de ajuste y reunir a especialistas para intercambiar desarrollos recientes y fomentar el trabajo conjunto. Las jornadas están abiertas a alumnos del Programa de Doctorado “Matemáticas” de la Universidad de Sevilla.

Organización

- Alejandra Cabaña Nigro, Universitat Autònoma de Barcelona
- María Dolores Jiménez Gamero, Universidad de Sevilla

Financiación

El taller está generosamente respaldado por

- Grupo FQM-153, Junta de Andalucía.
- Proyecto MTM2008-00018.
- Programa de Doctorado “Matemáticas” de la Universidad de Sevilla.

Lugar de celebración

El taller tendrá lugar en el aula 2.4 de la Facultad de Matemáticas de la Universidad de Sevilla, sita en la segunda planta. La Facultad de Matemáticas se encuentra a unos 10-15 minutos andando desde la residencia universitaria Rector Estanislao del Campo, donde tenéis resevado alojamiento todos aquellos que lo solicitásteis. Nos gustaría agradecer a la Facultad de Matemáticas de la Universidad de Sevilla el habernos cedido gratuitamente el espacio donde se celebrará el evento.

Algunas direcciones de interés

- GOF DAYS 2012 <http://mat.uab.es/acabana/gofdays.htm>
- Facultad de Matemáticas de la Universidad de Sevilla <http://www.matematicas.us.es/index.php>
- Residencia Universitaria Rector Estanislao del Campo <http://www.rusevilla.com/>
- Programa de Doctorado “Matemáticas” de la Universidad de Sevilla <http://www.imus.us.es/doctorado>.

PROGRAMA

9:00 *Conditional Stochastic Dominance*

M.A. Delgado, J.C. Escanciano

9:30 *Two-Sample Stochastic Dominance Testing Uniform in Covariates*

J.C. Escanciano, M.A. Delgado

10:00 *A nonparametric ANOVA-type test for regression curves based on characteristic functions*

J.C. Pardo-Fernández, M.D. Jiménez-Gamero, A. El Ghouch

10:30 Discusión. Modera F. López Blazquez.

11:00 Café

11:30 *Recortes Imparciales y Similitud de Muestras*

J.A. Cuesta-Albertos, P. Álvarez-Esteban, E. del Barrio y C. Matrán

12:00 *Bondad de Ajuste para datos Sesgados*

J. Ojeda, J.A. Cristóbal, J.T. Alcalá

12:30 *Nonparametric testing in regression with functional data*

C. Sánchez-Sellero, V. Patilea, M. Saumard

13:00 Discusión. Modera J.A. Cristóbal.

13:30 Almuerzo

15:30 *Una nueva metodología en tests de bondad de ajuste para las innovaciones en modelos $AR(1)$*

P. Puig, D. Moriña, J. Valero

16:00 *New goodness-of-fit diagnostics for discrete response dynamic models*

I. Kheifets, **C. Velasco**

16:30 *El proceso característico empírico de los residuos en modelos GARCH*

M.D. Jiménez Gamero

17:00 Discusión. Modera A. Cabaña.

17:30 Café

18:00 *Empirical likelihood approach using Phi-divergences*

N. Martín, L. Pardo Llorente

18:30 *Contrastes de bondad de ajuste basados en la correlación máxima de Hoeffding*

A. Grané

19:00 Discusión. Modera L. Pardo Llorente.

RESÚMENES

Por orden alfabético del ponente (en negrita)

Recortes Imparciales y Similitud de Muestras

J.A. Cuesta-Albertos¹, P. Álvarez-Esteban², E. del Barrio² y C. Matrán²

¹Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria

²Departamento de Estadística e Investigación Operativa, Universidad de Valladolid

La primera parte de este trabajo puede incluirse en el campo de los tests de ajuste entre dos muestras. Los trabajos en este campo suelen centrarse en la obtención de tests para contrastar la igualdad de distribución.

El punto de vista que adoptamos aquí es algo más general: Nuestra hipótesis nula establece que ambas muestras provienen de versiones contaminadas (a lo sumo en una proporción α) de la misma distribución. Llamaremos α -*similares* a dos distribuciones que verifiquen esta hipótesis nula.

Nuestra propuesta para resolver este problema utiliza los llamados *recortes imparciales*. Demostraremos que este problema está relacionado con el cálculo de las distancias mínimas entre conjuntos de *probabilidades recortadas*. Sucede que un recorte superior al nivel de similitud entre las distribuciones de partida ocasiona la aparición de un *sobreaajuste* que hace que las muestras sobrerrecortadas estén más próximas de lo esperado por efecto del azar. Este fenómeno, combinado con un procedimiento de remuestreo, permite estimar el nivel de similitud entre las dos distribuciones de partida.

En la segunda parte, consideramos el problema de k muestras. El objetivo será saber si la mayor parte de cada muestra se ajusta a un patrón general determinado por las muestras restantes. Aplicaremos los resultados al análisis de la existencia de diferencias significativas entre las calificaciones otorgadas por los diferentes correctores en cierta asignatura dentro de las pruebas de Selectividad.

Palabras clave: Similaridad, recortes imparciales, sobreaajuste.

Conditional Stochastic Dominance

M.A. Delgado¹, J.C. Escanciano²

¹Departamento de Economía, Universidad Carlos III de Madrid

²Department of Economics, Indiana University

This article proposes bootstrap-based stochastic dominance tests for nonparametric conditional distributions and their moments. We exploit the fact that a conditional distribution dominates the other if and only if the difference between the marginal joint distributions is monotonic in the explanatory variable at each value of the dependent variable. The proposed test statistic compares restricted and unrestricted estimators of the difference between the joint distributions, and it can be implemented under minimal smoothness requirements on the underlying nonparametric curves and without resorting to smooth estimation. The finite sample properties of the proposed test is examined by means of a Monte Carlo study. We report an application to studying the impact on post-intervention earnings of the National Supported Work Demonstration, a randomized labor training program carried out in the 1970s.

Palabras clave: Nonparametric testing, Conditional stochastic dominance, Conditional inequality restrictions, Least concave majorant, Treatment effects.

Two-Sample Stochastic Dominance Testing Uniform in Covariates

J.C. Escanciano¹, M.A. Delgado²

¹Department of Economics, Indiana University

²Departamento de Economía, Universidad Carlos III de Madrid

This article proposes two-sample tests for the hypothesis of stochastic dominance controlling for covariates. The concept of stochastic order of random variables defined on different parent populations is considered in conditional terms. That is, we test stochastic order of conditional distributions uniformly in the explanatory variables. The integrated difference between the conditional distributions in the two populations with respect to a covariate matching measure forms a basis for the test. The test statistic is the distance between the restricted estimator of this integrated difference, under the stochastic dominance assumption, and its unrestricted counterpart. The asymptotic distribution of the test statistic is non-pivotal and the test is implemented with the assistance of a bootstrap technique. We provide the asymptotic justification of the test under minimal regularity conditions, the asymptotic behavior under certain local alternatives, a study of its finite sample performance by means of Monte Carlo simulations, and an application to examining differences in wages distributions across different racial groups using US Current Population Surveys data.

Palabras clave: Nonparametric testing, Two sample tests, Stochastic dominance, Conditional distributions, Conditional moments, Covariate matching, Least concave majorant.

Contrastes de bondad de ajuste basados en la correlación máxima de Hoeffding

A. Grané

Departamento de Estadística, Universidad Carlos III de Madrid

En Fortiana y Grané (2003) se propone el estadístico Q_n para contrastar la bondad de ajuste de una muestra a una distribución completamente especificada (o equivalentemente, un contraste de uniformidad en $[0, 1]$). Dicho estadístico se obtiene a partir de la correlación máxima de Hoeffding y está muy relacionado con el estadístico de Gini. Los autores obtienen las distribuciones exacta y asintótica y estudian su potencia para un amplio conjunto de alternativas. Grané y Tchirina (2012) estudian la eficiencia asintótica de Bahadur para Q_n , obteniendo los dominios de optimalidad asintótica. Grané (2012) desarrolla el contraste basado en Q_n para muestras censuradas, obteniendo las distribuciones exacta y asintótica bajo ciertos esquemas de censura.

Palabras clave: L -estadísticos, bondad de ajuste, correlación máxima de Hoeffding, muestras censuradas.

Referencias

- J. Fortiana, A. Grané (2003) Goodness-of-fit tests based on maximum correlations and their orthogonal decompositions, *Journal of the Royal Statistical Society B*, 65, 115–126.
- A. Grané (2012) Exact goodness-of-fit tests for censored data. *Annals of the Institute of Statistical Mathematics*, 64, 1187–1203.
- A. Grané, A. Tchirina (2012) Asymptotic properties of a goodness-of-fit tests based on maximum correlations, *Statistics*, DOI:10.1080/02331888.2011.588709.

El proceso característico empírico de los residuos en modelos GARCH

M.D. Jiménez Gamero

Departamento de Estadística e Investigación Operativa, Universidad de Sevilla

Empirical processes play a prominent role in statistics since statistical procedures often involve functionals of them. In certain settings, such as linear models or time series, some variables of interest, such as the innovations, cannot be directly observed and the inference is based on the residuals, whose calculation involves the estimation of certain parameters in the model. GARCH models belong to this class. This work is devoted to studying the characteristic empirical process of the residuals in these models. Specifically, it is shown the uniform convergence to the population characteristic function of the innovations in compact sets. The weak convergence of this empirical process, suitably normalized, is also studied. The limit depends on the population characteristic function of the innovations, the equation defining the model and the parameter estimators employed to calculate the residuals. Applications of the obtained results for testing symmetry and goodness of fit to the law of the innovations are given.

Palabras clave: GARCH; residuals; empirical characteristic function; symmetry; goodness-of-fit.

Empirical likelihood approach using Phi-divergences

N. Martín¹, L. Pardo Llorente²

Departamento de Estadística, Universidad Carlos III de Madrid

Departamento de Estadística e Investigación Operativa I, Universidad Complutense de Madrid

The main purpose of this paper is to introduce a new family of empirical test statistics for testing goodness-of-fit. This new family is based on a distance between two probability vectors, namely, the probability vector obtained by maximizing the empirical likelihood subject to the unbiased estimating functions for the problem under consideration and the probability vector obtained under the simple null hypothesis. The considered distance is the phi-divergence measure. The asymptotic distribution is then derived for this family of test statistics. A simulation study is carried out to compare its performance with respect to the empirical likelihood ratio test when confidence intervals are constructed based on them with small sample sizes. The “empirical modified likelihood ratio test statistic” is proposed as an alternative of the empirical likelihood ratio test-statistic for the problem of goodness-of-fit.

Palabras clave: Empirical likelihood, empirical phi-divergence test statistic, phi-divergence measures, power function.

Bondad de Ajuste para datos Sesgados

J. Ojeda, J.A. Cristóbal, J.T. Alcalá

Departamento de Métodos Estadísticos, Universidad de Zaragoza

El Sesgo por Selección es un problema habitual en estudios de tipo observacional. El hecho de que no sea posible controlar la forma en que se registran los datos supone un problema no sólo a la hora de desarrollar estimadores, sino al realizar inferencia.

En esta charla se presentan y discuten algunos Contrastes de Bondad de Ajuste para la función de regresión cuando los datos están sesgados por selección. Se discuten además algunos esquemas de remuestreo que han sido utilizados para implementar los contrastes citados. Frente a la corrección del sesgo, la estrategia utilizada en este caso para solventar el problema que conlleva el sesgo en los datos es la de la compensación o preservación de la media. Dicha estrategia presenta algunas implicaciones particulares en los métodos de remuestreo utilizados que también se abordarán.

Palabras clave: Sesgo por selección, Estadística Noparamétrica, Procesos Empíricos.

Referencias

Rao, C. R. (1997) *Statistic and True. Putting chance to work*. World Scientific Publishing.

Vardi, Y. (1985) Empirical distributions in selection bias models, *Ann. Statist.*, 13, 178–205.

Cristóbal, José A. and Alcalá, José T. (2001) An overview of nonparametric contributions to the problem of functional estimation from biased data, *Test*, 2, 309–332.

A nonparametric ANOVA-type test for regression curves based on characteristic functions

J.C. Pardo-Fernández¹, M.D. Jiménez-Gamero², A. El Ghouch³

¹Departamento de Estadística e Investigación Operativa, Universidade de Vigo

²Departamento de Estadística e Investigación Operativa, Universidad de Sevilla

³ISBA, Université catholique de Louvain

In this talk we will present a new procedure to test for the equality of k regression curves in a fully nonparametric context. The test is based on the comparison of two empirical estimators of the characteristic functions of the regression residuals in each population. The asymptotic behaviour of the test statistic is studied in detail. It is shown that under the null hypothesis the distribution of the test statistic converges to a combination of χ_1^2 random variables. Under certain restrictions on the populations, the asymptotic null distribution of the test statistic is χ_{k-1}^2 . The practical performance of the test based on the asymptotic null distribution is investigated by means of simulations.

Palabras clave: comparison of regression curves, empirical characteristic function, regression residuals.

Una nueva metodología en tests de bondad de ajuste para las innovaciones en modelos AR(1)

P. Puig¹, D. Morriña¹, J. Valero²

¹Departament de Matemàtiques, Universitat Autònoma de Barcelona

²Escola Superior d'Agricultura de Barcelona, Universitat Politècnica de Catalunya

Consideremos la serie Y_t , que satisface un modelo AR(1), $Y_t = \alpha Y_{t-1} + W_t$, $|\alpha| < 1$, donde las innovaciones W_t , iid, siguen una distribución continua arbitraria, no necesariamente Gaussiana, con media μ y varianza σ^2 . Sea la serie transformada mediante exponenciación, $X_t = e^{Y_t}$. McKenzie (1982), demuestra que la estructura correlacional de X_t y de Y_t es la misma, si y sólo si la distribución marginal de Y_t es Gamma. Generalizando este resultado podemos caracterizar la distribución marginal de Y_t en términos de la función de autocorrelación (ACF) de X_t . En particular, demostramos el siguiente resultado:

Teorema: La ACF de X_t es de la forma $\rho_k = (e^{\frac{\sigma^2 \alpha^k}{1-\alpha^2}} - 1) / (e^{\frac{\sigma^2}{1-\alpha^2}} - 1)$, si y sólo si las innovaciones W_t de la serie Y_t son Gaussianas.

Estimando la función de autocorrelación de X_t a partir de los datos y utilizando la metodología presentada en Anderson et al. (2004), construimos tests de bondad de ajuste para la distribución de las innovaciones W_t .

Palabras clave: caracterización de distribuciones, función de autocorrelación.

Referencias

E. McKenzie (1982). Product autoregression: a time-series characterization of the gamma distribution. *Journal of applied probability*, **19**, 463–468.

T.W. Anderson, R.A. Lockhart, M.A. Stephens (2004). An omnibus test for the time series model AR(1). *Journal of Econometrics*, **118**, 111–127.

Nonparametric testing in regression with functional data

C. Sánchez-Sellero¹, V. Patilea², M. Saumard³

¹Departamento de Estadística e I.O., Univ. de Santiago de Compostela, Spain

²CREST (Ensaï) & IRMAR, France

³INSA-IRMAR, France

En esta exposición se mostrarán dos trabajos recientes sobre contrastes no paramétricos en modelos de regresión con datos funcionales. Se empieza considerando covariable funcional y respuesta escalar. En ese caso, la covariable funcional se proyecta sobre una base de funciones de $L^2[0, 1]$. Se toma un número finito de coeficientes, que puede crecer con el tamaño muestral. El test se basa en una técnica de contraste con covariable finito-dimensional, aplicada sobre los coeficientes de la covariable funcional en la base. Cuando la variable respuesta también es funcional, se emplean los productos escalares de las respuestas en el núcleo del U-estadístico de contraste.

Palabras clave: regresión con datos funcionales, bondad de ajuste.

Referencias

V. Patilea, C. Sánchez-Sellero, M. Saumard (2012). Projection-based nonparametric goodness-of-fit testing with functional covariates. arXiv:1205.5578 [math.ST]

V. Patilea, C. Sánchez-Sellero, M. Saumard (2012). Nonparametric testing for no-effect with functional responses and functional covariates. arXiv:1209.2085 [math.ST]

New goodness-of-fit diagnostics for discrete response dynamic models

I. Kheifets¹, C. Velasco²

¹Department of Economics, New Economic School, Moscow

²Departamento de Economía, Universidad Carlos III de Madrid

This paper proposes new specification tests for dynamic models with discrete responses. In particular, we test the specification of the conditional distribution of multinomial and count data, which is key to apply efficient maximum likelihood methods, to obtain consistent estimates of partial effects and to construct appropriate predictions of the probability of future events. The traditional approach is based on a random continuation transformation of discrete data which leads to continuous uniform iid series under the true conditional distribution, see e.g. Kheifets and Velasco (2012). Then, standard specification techniques can be applied to the transformed series, but the extra random noise involved in the continuation process may affect the power properties of these methods. We investigate in this paper an alternative estimate of a continuous cumulative distribution function (cdf) based only on discrete data, proposed by Czado, Gneiting and Held (2009), which can be compared directly to a $[0, 1]$ uniform cdf. We analyze the asymptotic properties of goodness-of-fit tests based on this new approach and explore the properties in finite samples of a bootstrap algorithm to approximate the critical values of test statistics, which are model and parameter estimation dependent. We find that in many relevant cases our new approach performs much better than random-continuation counterparts.

Palabras clave: Specification tests, count data, dynamic discrete choice models, conditional probability integral transform.

Referencias

- I. Kheifets and C. Velasco (2012) Model Adequacy Checks for Discrete Choice Dynamic Models. En: X. Chen, N.R. Swanson, *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis. Essays in Honor of Halbert L. White Jr.*, pp. 363–382, Springer, New York.
- C. Czado, T. Gneiting, L. Held (2009) Predictive model assessment for count data. *Biometrics*, 65, 1254–1261.