

# REVISTA ESTADISTICA

Vol. 62 - NRO. 178 – junio 2010

## CONTENIDO – CONTENTS

### **MODELOS DE ESPACIO ESTADO BASADOS EN LA DISTRIBUCION NORMAL INVERSA GAUSSIANA: UNA APLICACION AL ANALISIS DE SERIES DE TIEMPO DE LA ECONOMIA VENEZOLANA**

SABA INFANTE

*Departamento de Matemáticas, Facyt, Universidad de Carabobo, Venezuela*  
sinfante@uc.edu.ve

JAVIER ROJAS

*Universidad Simón Bolívar, Venezuela*  
jjrojaspy@yahoo.es

ARACELIS HERNANDEZ

*Departamento de Matemáticas, Facyt, Universidad de Carabobo, Venezuela*  
arhernan@uc.edu.ve

VIRGINIA CARTAYA

*Oficina de Investigaciones Económicas, Banco Central de Venezuela*  
vcartaya@bcv.org.ve

### **RESUMEN**

El presente trabajo tiene como primer objetivo mostrar como funcionan los algoritmos computacionales: muestreador de Gibbs, filtro de Kalman, filtro de Kalman extendido y filtro de partículas con remuestreo, en el modelaje de series de tiempo. Un segundo objetivo que se propone es comparar el uso de estos algoritmos para modelar los estados desconocidos del sistema económico de Venezuela utilizando la distribución Normal Inversa Gaussiana (NIG), caso especial de las distribuciones Hiperbólicas generalizadas de Barndorff-Nielsen (1978). En el estudio se consideran modelos con estructuras lineales y no lineales, con errores Gaussianos y no Gaussianos; se analizan las series del Producto Interno Bruto (PIB) petrolero, y no petrolero; y la tasa de cambio Dólar a Bolívar de la economía Venezolana; también se realizó un estudio de simulación. Se demuestra que el estimador de la media a posteriori de los estados desconocidos del sistema se ajustan bien a las series verdaderas bajo los algoritmos propuestos cuando el modelo es lineal, independientemente si los errores son Gaussianos o no Gaussianos; mientras que cuando el modelo es no

lineal (caso simulado), el filtro de partículas con remuestreo es el que mejor se comporta. Se estimaron las densidades a posteriori de los estados y se obtuvo densidades bimodales y trimodales, este resultado es una señal de que las series analizadas tienen un comportamiento no lineal, no estacionario y son volátiles. Además, se simularon muestras para comparar la eficiencia de los filtros y se utilizó como medida de bondad de ajuste la raíz cuadrada del error cuadrático medio empírico (RCECME). En este sentido, si el modelo es lineal todos los algoritmos tienen la misma eficiencia debido a que la RCECME de los valores estimados son pequeños, mientras que cuando el modelo es no lineal hay una marcada diferencia de la RCECME de valores estimados entre el filtro de Kalman extendido con respecto al muestreador de Gibbs y el filtro de partículas con remuestreo.

### ***Palabras clave***

Métodos de Monte Carlo; distribución normal inversa Gaussiana; modelos espacio estado.

### **ABSTRACT**

The present article has as a first objective to show how the following computational algorithms can work in time series modeling: the Gibbs sampler, the Kalman filter, the extended Kalman filter and a particle filter with resampling. A second objective is to compare the use of these algorithms to model the unknown states of the Venezuelan economic system using the Normal Inverse Gaussian distribution (NIG), a special case of the generalized Hyperbolic distribution of Barndorff-Nielsen (1978). The study considers models with linear and nonlinear structures, and with Gaussian and non-Gaussian errors to analyze the series of the Gross Internal Product (PIB) for oil and non-oil and the series of the US Dollar to Bolivar Exchange rate of the Venezuelan economy. A simulation study was also done. We show that the estimator of the posterior mean of the unknown state adjusts well to the real series under the proposed algorithms when the model is linear, independently of whether the errors are Gaussian or non-Gaussian. When the model is non-linear (simulated case), the particle filter with resampling has the best behavior. The posterior density of the states was estimated obtaining bimodal and trimodal density showing that the analyzed series have nonlinear behavior, are non stationary and are volatile. In addition, samples were simulated to compare the efficiency of the filters using the empirical root mean squared error (RCECME) as goodness of fit measure. In this sense, if the model is linear, all the algorithms have the same efficiency because the RCECME of the estimated values are small, whereas when the model is nonlinear there is a marked difference in the RCECME of the estimated values for the extended Kalman filter, the Gibbs filter and the particle filter with resampling.

### ***Key words***

Monte Carlo methods; normal inverse Gaussian distribution; space state models.

## **APLICACION DE LOS MODELOS DE ESPACIO DE ESTADOS PARA LA ESTIMACION DE UN INDICADOR AGREGADO DE LA ACTIVIDAD ECONOMICA PROVINCIAL**

MÉNDEZ, FERNANDA

*Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística.  
Facultad de Ciencias Económicas y Estadística. Universidad Nacional de Rosario  
nandixx@hotmail.com*

### **RESUMEN**

En este trabajo se aplica la metodología propuesta por Stock y Watson, para estimar un indicador agregado de la actividad económica santafesina. El procedimiento postula un modelo probabilístico de espacio de estados que se utiliza para predecir un proceso latente y estimar así un indicador agregado coincidente de la actividad económica. Se fundamenta en la hipótesis de que los comovimientos observados en las series indicadoras son capturados por una única variable no observable, común a todas ellas, denominada el “estado de la economía”. Se plantean varios modelos y la estimación máximo verosímil de los parámetros del modelo se logra representando el sistema en la forma de un modelos espacio de estados y aplicando el filtro de Kalman. También se analiza si el indicador agregado coincidente estimado mediante el modelo que se selecciona es consistente con la evolución económica de la provincia en el período considerado.

### ***Palabras clave***

Indicador agregado de la actividad económica regional, modelos de espacio de estados, filtro de Kalman, modelo de Stock y Watson.

### **ABSTRACT**

This paper applies the methodology developed by Stock and Watson to estimate an aggregate indicator of Santa Fe's economic activity. The procedure postulates a probabilistic state space model that is used to predict a latent process and therefore to estimate an aggregate coincident indicator of economic activity. It is based on the hypothesis that the comovements in the indicators series are captured by a single common, unobserved variable called the "state of the economy". Various models are presented and the maximum likelihood estimation of the parameters of the model is accomplished by representing the system in state space form and using the Kalman filter. Also it is analyzed if the aggregate coincident indicator estimated is consistent with knowledge of Santa Fe's economic history.

### ***Key words***

Aggregated indicator of regional economic activity, state space models, Kalman filter, Stock-Watson model.

## **COMPARACION INTERTEMPORAL DEL INDICE DE MARGINACION EN MEXICO USANDO EL MODELO DE COMPONENTES PRINCIPALES COMUNES EN VECTORES DEPENDIENTES**

HUMBERTO SOTO DE LA ROSA  
*Comisión Económica para América Latina y el Caribe*  
*Presidente Masaryk 29, CP 11570, México, DF*  
humberto.soto@cepal.org

### **RESUMEN**

El Análisis de Componentes Principales Comunes en Vectores Dependientes (CPCD) permite determinar la significancia estadística de las variaciones intertemporales de un índice sintético (multidimensional), lo cual tiene muchas aplicaciones, por ejemplo, en la evaluación del desempeño de políticas públicas destinadas a impulsar un aspecto del desarrollo económico o social de una población, que puede ser medido a partir de un índice sintético. Se compilan los principales resultados teóricos desarrollados para el modelo CPCD, generalización del Análisis de Componentes Principales (ACP) ampliamente utilizado en la construcción de índices sintéticos, y se presentan los resultados de su aplicación para analizar la evolución de la Marginación

en México entre 1990 y 2000, utilizando como base el índice sintético de Marginación que el Consejo Nacional de Población mexicano ha construido desde 1990 usando el modelo ACP, concluyendo que la marginación en México ha tenido una reducción estadísticamente significativa en el periodo de análisis.

**Palabras clave**

Componentes principales comunes; vectores dependientes; índices sintéticos; marginación; comparabilidad intertemporal.

**ABSTRACT**

Principal Components Analysis for Dependent Random Vectors (CPCD) allows determining the statistical significance of intertemporal variations of a synthetic index (multidimensional). This has several applications, for example, on impact assessment of public policies stimulating a particular issue of social or economic development of a population, which can be measured using a synthetic index. The principal theoretical results developed for the CPCD model, generalization of the Principal Components Analysis (PCA) are compiled, and the results of an application analyzing the evolution of marginalization in Mexico between 1990 and 2000 are presented, using as baseline the synthetic marginalization index generated by the Mexican Population Council since 1990 with the PCA model, concluding that marginalization in Mexico has had a statistically significant reduction on the reference period.

**Key words**

Common principal components; dependent vectors; syntetic indexes; marginalization; intertemporal comparability.

**MODELOS MIXTOS GENERALIZADOS PARA EL ESTUDIO  
DE LA ASOCIACIÓN ENTRE ALGUNAS VARIABLES  
SOCIOECONÓMICAS Y LAS TASAS DE INCIDENCIA DE  
CÁNCER EN LOCALIDADES DE CÓRDOBA, ARGENTINA**

MARGARITA DIAZ

*Instituto de Estadística y Demografía, Facultad de Ciencias Económicas  
Universidad Nacional de Córdoba, Argentina  
mdiaz@eco.unc.edu.ar*

FERNANDO GARCIA

*Instituto de Estadística y Demografía, Facultad de Ciencias Económicas  
Universidad Nacional de Córdoba, Argentina  
walfer@coop5.com.ar*

PATRICIA CARO

*Instituto de Estadística y Demografía, Facultad de Ciencias Económicas  
Universidad Nacional de Córdoba, Argentina  
pacaro@eco.unc.edu.ar*

MARIA DEL PILAR DIAZ

*Cátedra de Bioestadística, Escuela de Nutrición, Facultad de Ciencias Médicas  
Universidad Nacional de Córdoba, Argentina  
pdiaz@fcm.unc.edu.ar*

## RESUMEN

El objetivo de este trabajo es construir Modelos Lineales Mixtos Generalizados (GLMM) que permitan cuantificar el efecto de diversos factores socio- económicos en la incidencia del cáncer en la provincia de Córdoba, Argentina, usando la base de datos de la casuística de tumores incidentes para el año 2004. Se incluyeron como covariables las necesidades básicas insatisfechas, la falta de obra social, la desocupación y la falta de instrucción; mientras que, la variabilidad regional fue incorporada mediante una zonificación de las localidades (zona). Se construyeron diferentes modelos de regresión Poisson, para cada sexo y tipo de tumor, incorporando coeficientes aleatorios para zona, a fin de capturar la asociación espacial posible entre las respuestas de las unidades de cada uno de esos agrupamientos. Los GLMM estimados para explicar el patrón espacial de incidencia de los diferentes cánceres identificaron diversos factores de riesgo y lograron definir regiones vulnerables y no vulnerables asociadas a dicho patrón.

### *Palabras clave*

Efecto aleatorio, Poisson, distribución espacial, factores socio-económicos, Argentina, cáncer.

## ABSTRACT

The aim of this work was to build Generalized linear mixed models (GLMM) to quantify the socio-economic factor effects on the cancer incidence pattern in Córdoba province, Argentina. For that, the 2004-database containing the tumorincidence series was used. Several characteristics, like basic economic needs, social security lack, unemployment and non-education status were included into the model, whereas the spatial variability was incorporated by zone stratification. To capture the spatial association among responses within the zone classes, Poisson regression models with random coefficients were fitted for each sex and type of tumor. The proposed GLMMs explained the spatial pattern of incidence of various cancers and identified socio-economic risk factors, getting to identify vulnerable and no vulnerable regions associated with that pattern.

### *Key words*

Random effect, Poisson, spatial distribution, socio-economic factors, Argentina, cancer.