

Número especial en  
aplicaciones en la industria y estadística industrial

**Editor**

B. Piedad Urdinola, Ph.D.  
UNIVERSIDAD NACIONAL DE COLOMBIA, BOGOTÁ, COLOMBIA

**Editor invitado**

Jorge Luis Romeu, Ph.D.  
SYRACUSE UNIVERSITY, NEW YORK, USA

**Comité invitado**

Alberto Luceña, Ph.D.  
UNIVERSIDAD DE CANTABRIA, CANTABRIA, ESPAÑA

José Alberto Vargas, Ph.D.  
UNIVERSIDAD NACIONAL DE COLOMBIA, BOGOTÁ, COLOMBIA

Manuel Presedo, Ph.D.  
UNIVERSIDADE DA CORUÑA, LA CORUÑA, ESPAÑA

Sergio Yañez, Ph.D.(c)  
UNIVERSIDAD NACIONAL DE COLOMBIA, MEDELLÍN, COLOMBIA

Victor Aguirre, Ph.D.  
INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO, ITAM, MÉXICO





Revista Colombiana de Estadística	Bogotá	Vol. 34	Nº 2
ISSN 0120 - 1751	COLOMBIA	junio-2011	Págs. 211-402

## Contenido

**Javier Tarrío & Salvador Naya**

*Influencia de la adición de nano y microsílíce en la estabilidad térmica de una resina epoxi. Aplicaciones del ANOVA funcional* .....211-230

**Sergio Yáñez, Hugo Brango, Mario C. Jaramillo & Carlos M. Lopera**

*Comparación entre riesgos competitivos vía el estimador cópula-gráfico* .. 231-248

**Carlos M. Lopera, & Eva Cristina Manotas**

*Aplicación del análisis de datos recurrentes sobre interruptores FL245 en Interconexión Eléctrica S.A.*.....249-266

**Enrique Raúl Villa-Diharce & Pedro Enrique Monjardin**

*Análisis bivariado de confiabilidad basado en cópulas* ..... 267-285

**Rubén Darío Guevara & José Alberto Vargas**

*Process Capability Analysis Plot for a Product with Bilateral Specifications* ..... 287-301

**John Cook, Jairo Fúquene & Luis Pericchi**

*Skeptical and Optimistic Robust Priors for Clinical Trials* ..... 333-345

**Carlos G. Maté**

*A Multivariate Analysis Approach to Forecasts Combination. Application to Foreign Exchange (FX) Markets* .....347-375

**Jorge Andrés Perdomo & Darrell Hueth**

*Funciones de producción, análisis de economías a escala y eficiencia técnica en el eje cafetero colombiano: una aproximación con frontera estocástica* ..... 377-402



# Editorial

## Número especial de la *Revista Colombiana de Estadística* en aplicaciones en la industria y estadística industrial

JORGE LUIS ROMEU<sup>1,2,a</sup>, B. PIEDAD URDINOLA<sup>3,b</sup>

<sup>1</sup>DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING, SYRACUSE UNIVERSITY, NEW YORK, UNITED STATES

<sup>2</sup>PROYECTO JUÁREZ LINCOLN MARTÍ,

<sup>3</sup>DEPARTAMENTO DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE COLOMBIA, BOGOTÁ, COLOMBIA

---

Durante 2010, escribimos en los Boletines Electrónicos del Proyecto Juárez Lincoln Martí de Educación Internacional (<http://web.cortland.edu/matresearch>), varios artículos sobre la necesidad de tener, en Iberoamérica, más publicaciones que trataran sobre estadística aplicada en español. La Revista Colombiana de Estadística, a través de su editora ejecutiva de ese momento, recogió este llamado. Y así nació el Número especial de aplicaciones en la industria y estadística industrial, que hoy ponemos en sus manos.

La importancia de la estadística en el mundo moderno es indiscutible. Se utiliza en la agricultura, en la industria, el comercio y las ciencias sociales, entre otras áreas de importancia económica, que tanto necesitan desarrollarse en nuestra región. Por este motivo, es necesario reforzar el estudio de las aplicaciones, además del estudio teórico de la estadística, en donde todavía tenemos un gran déficit. Similar a este déficit, es el bajo dominio en la lengua predominante en las ciencias de nuestro tiempo, -el inglés-, por parte de la mayoría de los estudiantes latinoamericanos.

El presente Número especial de la RCE pretende ayudar a llenar este vacío, presentando ocho excelentes trabajos de estadística aplicada, escritos en español, o con un extenso resumen en español, para los dos trabajos escritos en inglés. Recibimos 20 trabajos, de autores de más de ocho países de Iberoamérica, algunos de los cuales fueron pasados al número regular de RCE, dado su carácter teórico. Se escogieron los presentes, por su área de interés, y por su aplicación a la resolución de problemas prácticos en esas áreas.

Pensamos que muchos profesores de estadística aplicada (ingeniería, ciencias comerciales y agropecuarias, econometría, etc.) podrán utilizar este material en sus clases, y aun animar a alumnos a convertir sus tesis de maestría en artículos, y someterlos a futuros números de la RCE.

---

<sup>a</sup>Editor invitado número especial de la Revista Colombiana de Estadística.  
Profesor e investigador. E-mail: romeu@cortland.edu

<sup>b</sup>Editora número especial de la Revista Colombiana de Estadística.  
Profesora asociada. E-mail: bpurdinolac@bt.unal.edu.co

Pensamos también que con la publicación de trabajos aplicados en español se incremente el número potencial de autores y lectores, logrando así que los departamentos de estadística que actualmente prefieren que sus profesores publiquen en inglés, también consideren aceptables los trabajos publicados en español.

Por último, este número no habría salido sin la constante participación y el apoyo de los miembros del Comité Invitado, sin el cuidadoso y dedicado trabajo de los árbitros, que leyeron y relevaron las distintas versiones de los artículos, y la asistente editorial de la revista. Gracias,

# Editorial

## Special Issue of the *Colombian Journal of Statistics on Industrial Applications and Industrial Statistics*

JORGE LUIS ROMEU<sup>1,2,a</sup>, B. PIEDAD URDINOLA<sup>3,b</sup>

<sup>1</sup>DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING, SYRACUSE UNIVERSITY, NEW YORK, UNITED STATES

<sup>2</sup>PROYECTO JUÁREZ LINCOLN MARTÍ,

<sup>3</sup>DEPARTAMENTO DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE COLOMBIA, BOGOTÁ, COLOMBIA

---

During 2010, we wrote several articles in the Juarez Lincoln Marti International Education Project Electronic Bulletins, about the need for having, in Iberoamerica, more publications and more papers dealing with applied statistics, written in Spanish. The Colombian Journal of Statistics (RCE), through its Executive Editor at the time, responded to this call. And this is how the Special Issue on Applications in Industry and Industrial Statistics that we bring today to our readers, was born.

The importance of statistics in the modern world is unquestionable. It is used in agriculture, in industry, in business, and social sciences, among many areas of economic importance that are in need, and which require further development, in our region. For these reasons, to underline the study of applications, in addition to the study of theoretical statistics, is necessary. And this is an area where we still have a large deficit. In addition, many -if not most -of our Latin American students have a low level of English language understanding, which is the language in which most science articles are written today. Thus, they cannot read them, let alone write them.

The objective of this Special Issue of RCE is, precisely, to fill such void, by presenting eight excellent articles in applied statistics, written in Spanish, or with an extended summary in Spanish, for the two articles that are written in English. We received twenty submissions, written by authors from eight Ibero-American countries. Some of them were passed to the regular issue of RCE, given their theoretical slant. The remaining eight were selected for their areas of interest, and for their application techniques to solving practical problems in such areas.

We believe that many applied statistics professors in engineering, business, agriculture and econometrics, among others, may be able to use this material in their classes. And perhaps to encourage their masters students to convert their theses into articles, and to submit them to future issues of RCE.

---

<sup>a</sup>Invited Editor Special Issue of the Colombian Journal of Statistics.  
Professor and Research. E-mail: romeu@cortland.edu

<sup>b</sup>Editor in Chief Special Issue of the Colombian Journal of Statistics.  
Associate professor. E-mail: bpurdinolac@bt.unal.edu.co



We believe that, with the publication of applied papers in Spanish, we will help increase the number of potential authors and readers. And with this, to convince some statistics departments that currently require that their faculty publish in English, to also consider acceptable to have papers published in Spanish.

Finally, this issue would not exist without the constant participation and support of the members of the guest Committee, without the careful and dedicated readings of the reviewers, who went through several versions of each article, and without the help of the Editorial Assistant of this journal. Thank you.

## Influencia de la adición de nano y microsílíce en la estabilidad térmica de una resina epoxi. Aplicaciones del ANOVA funcional

### Influence of Nano and Micro Silica Addition on Thermal Stability of an Epoxy Resin. Use of Functional ANOVA

JAVIER TARRÍO<sup>a</sup>, SALVADOR NAYA<sup>b</sup>

DEPARTAMENTO DE MATEMÁTICAS, ESCUELA POLITÉCNICA SUPERIOR, UNIVERSIDAD DE A  
CORUÑA, FERROL, ESPAÑA

---

#### Resumen

El principal objeto de este trabajo ha sido el empleo de una nueva técnica que aúna el análisis funcional y el diseño de experimentos: ANOVA funcional para un factor tratamiento. Mediante esta herramienta se ha medido la influencia que tiene la adición de una mezcla de nano y micro partículas de humo de sílice, subproducto de la industria del silicio, en la degradación térmica de una resina epoxi. Para tal fin se ha realizado un diseño de experimentos con un factor tratamiento (cantidad de humo de sílice) a tres niveles diferentes. Los datos se han obtenido mediante el empleo del Análisis Termogravimétrico (TG), dando como resultado cinco curvas o trayectorias de degradación por nivel. El ANOVA funcional aprovecha toda la información de cada curva o dato funcional.

El empleo del ANOVA funcional a partir de los datos TG ha dado como resultado que la cantidad de humo de sílice influye significativamente en la estabilidad térmica del compuesto y de la resina epoxi involucrada en cada muestra. Estos hechos pueden ser un indicio de la interacción entre la fase orgánica e inorgánica del material.

**Palabras clave:** análisis de datos funcionales, ANOVA, diseño de experimentos.

#### Abstract

The main purpose of this work has been the use of a new technique that combines the functional analysis and design of experiments: a one way functional ANOVA. Using this tool is intended to measure the influence of the addition of a mixture of nano and micro particles of silica fume, by-product

---

<sup>a</sup>Profesor titular. E-mail: jtarrío@udc.es

<sup>b</sup>Profesor titular. E-mail: salva@udc.es

of silicon industry, on the thermal degradation of epoxy resin. To achieve this, it has been performed a design of experiments with a treatment factor (amount of fumed silica) at three different levels. The data were obtained using thermogravimetric analysis (TG), resulting in five curves or degradation pathways by level. The functional ANOVA uses all information of each curve or functional data.

The use of functional ANOVA from TG data has resulted in the amount of silica fume significantly affect the thermal stability of the compound and the epoxy resin involved in each sample. These facts may be indicative of the interaction between the organic and the inorganic phase.

**Key words:** ANOVA, Experimental design, Functional data analysis.

## 1. Introducción

En los últimos años se ha producido un gran avance en técnicas estadísticas asociadas al incremento y mejora de las aplicaciones informáticas. Algunas de estas técnicas han surgido para dar respuesta al problema que supone el trabajar con grandes bases de datos, que en ocasiones corresponden a observaciones de una variable aleatoria tomadas a lo largo de un intervalo continuo. Un ejemplo de estas nuevas técnicas es lo que se conoce como estadística o análisis de datos funcionales (FDA, sus siglas en inglés).

La estadística de datos funcionales aborda muchos de los problemas de la estadística clásica: descripción de variables, modelización, clasificación, inferencia, etc. Habría que mencionar la publicación del libro de Ramsay & Silverman (1997) como el pionero en la divulgación de este tipo de estudios, si bien muchas de las aplicaciones a problemas concretos en áreas tan dispares como la medicina o la ingeniería se presentan en una segunda obra de los mismos autores (Ramsay & Silverman 2005). Otros dos textos importantes para el acercamiento a este estudio son, por un lado, el libro de de Ferraty & Vieu (2006), obra referencia en el enfoque no paramétrico del FDA, y el más reciente y generalista de Ferraty & Romain (2010).

En este trabajo se propone la aplicación de algunas de estas nuevas técnicas del campo de la estadística de datos funcionales en el estudio de las propiedades de un nuevo material compuesto para la industria: las resinas poliméricas epoxi con refuerzo de nano y micropartículas de sílice. El material en cuestión es un nanocompuesto de matriz polimérica (resina epoxi) y carga inorgánica (partículas de humo de sílice).

Las resinas epoxi son un polímero termoestable ampliamente utilizado. Algunas de las ventajas que exhiben las resinas epoxi frente a otros tipos de resinas poliméricas son, entre otras, su baja contracción de curado, la práctica ausencia de volátiles durante el mismo, resistencia mecánica y durabilidad, adhesión, resistencia química y aislamiento eléctrico. Además, estas resinas son compatibles con un gran número de materiales, por lo que son ampliamente utilizadas como matriz de materiales compuestos. De hecho, aunque su uso está de alguna forma limitado debido a su alta rigidez, provocada por una densa estructura de entrecruzamien-

to, sus propiedades mecánicas pueden mejorarse mediante la adición de partículas inorgánicas (Harsch, Karger-Kocsis & Holst 2007). La forma, volumen, tamaño, características superficiales y dispersión de las partículas dentro de la matriz condicionan altamente las propiedades mecánicas del material compuesto resultante (Lee & Lichtenhan 1999, Mehta, Mirabella, Bafna & Rufener 2004, Hsiue, Chen & Liu 2000, Shao-Yun, Xi-Qiao, Bernd & Yiu-Wing 2008). Particularmente el uso de nano-partículas provoca un mayor aumento del módulo de Young que el uso de micropartículas, incluso con una menor cantidad de las mismas (Zhang, Zhang, Friedrich & Eger 2006). Los nanocompuestos de matriz orgánica y carga inorgánica han mostrado ser capaces de proporcionar aumentos simultáneos en propiedades tales como la estabilidad térmica, el retardo a la llama, la temperatura de transición vítrea y la estabilidad dimensional, además de la disminución de la constante dieléctrica (Zhang et al. 2006, Liu, Wei, Hsu & Ho 1999, Preghenella, Pegoretti & Migliaresi 2005, Yousefi, Lafleur & Gauvin 1997). Debido a la mejora de sus propiedades mecánicas, el uso de materiales compuestos de matriz epoxídica es una de las opciones más utilizadas en la construcción de materiales ligeros para aplicaciones estructurales en la industria aeronáutica (aviones, transbordadores, satélites) y de los vehículos de transporte en general, como por ejemplo en automóviles y embarcaciones de altas prestaciones (Petrie 2006). Particularizando en este estudio, se ha buscado obtener mejoras en las propiedades de una determinada matriz epoxídica, siendo el primer paso para obtener nuevos materiales compuestos con aplicación, tanto en las industrias anteriormente mencionadas como en el sector de la construcción, fabricación de adhesivos o mismo en la industria eléctrica y electrónica, además de cumplir con una labor necesaria como es el reciclaje de un importante desecho industrial como es el humo de sílice.

En este trabajo, los compuestos de humo de sílice-resina epoxi fueron preparados y caracterizados mediante el empleo de técnicas termogravimétricas (TG), usuales en la evaluación de la estabilidad térmica de un material (Tarrio-Saavedra, López-Beceiro, Naya & Artiaga 2008). Para ello se ha escogido una resina epoxi no convencional, basada en el trimetilolpropano, especialmente indicada para la manufactura de materiales compuestos. Por otro lado, el humo de sílice empleado es un subproducto de la industria del silicio y ferrosilicio. Se forma en lo alto de los hornos de inducción de arco sumergido, siendo su forma de obtención diferente a los procesos convencionales para  $\text{SiO}_2$  sintético (Mohammad & Simon 2006). Por las características especiales del método de producción, el humo de sílice empleado tiene una pureza variable, dependiente de las condiciones de operación en los hornos. En todo caso, la proporción en peso de sílice nunca resulta menor del 95%. También es variable el tamaño de partícula, de hecho el humo de sílice consta de nano y micropartículas, teniendo en cuenta el criterio de Schadler (2003) (diámetro menor de 100 nm implica que es una nanopartícula). Esta especial distribución de tamaños hace esperar una posible combinación de efectos micro y nano.

Parte importante de este estudio es la determinación del grado de interacción del humo de sílice con la resina epoxi, relacionado con la interfase orgánica-inorgánica. Esta se define como aquella región que comienza en el punto de la carga que difiere del resto de la carga y termina en el punto de la matriz en el que

sus propiedades son las mismas que en el resto de la matriz (Schadler 2003). La existencia de esta interfase influye en propiedades tales como la estabilidad térmica o la temperatura de transición vítrea. De hecho, la variación de dichas propiedades puede tomarse como un índice de su existencia (Tarrío-Saavedra et al. 2008).

El resto del artículo se ha estructurado del siguiente modo. En la sección siguiente se presenta el modelo ANOVA funcional, seguidamente, en una sección dedicada a la parte experimental, se analizan los datos, curvas TG obtenidas en el estudio térmico de los nanomateriales, en otra sección se realiza un ajuste de los mismos en una base adecuada, con el fin de suavizarlos. A continuación se aplica el modelo ANOVA funcional a estas curvas termogravimétricas. Finalmente en una última sección se presentan los resultados, las conclusiones y se proponen nuevas líneas de investigación.

## 2. El modelo ANOVA funcional

Los datos funcionales, llamados también datos longitudinales, aparecen asociados a procesos monitorizados continuamente en el tiempo. Es decir, cuando se mide una variable en un conjunto discreto y finito de puntos ordenados; considerando que esa variable sigue una relación funcional continua. En este contexto, una variable aleatoria  $X$  se dice una variable funcional si toma valores en un espacio funcional  $F$  (normado o seminormado completo). Un caso particular se produce cuando la variable funcional  $X$  pertenece a un espacio de Hilbert, como es el caso de las funciones continuas en un intervalo (Ramsay & Silverman 1997). Este es el caso de las curvas de Análisis Termogravimétrico, conocidas como curvas TGA o TG, que serán las utilizadas en este trabajo, y que cumplen la propiedad de pertenecer al espacio de funciones medibles en un intervalo real, es decir,  $X \in L_2([0, T])$  (Naya, Cao & Artiaga 2003).

Como la variable de interés, masa de la muestra, se ha evaluado cada segundo, los datos pueden ser considerados funcionales. Dada la naturaleza funcional de los datos se propone la construcción de un ANOVA funcional (Cuevas, Febrero & Fraiman 2004). Una de las ventajas que presenta frente a un ANOVA clásico consiste en el aprovechamiento de toda la información de la curva en lugar de recurrir a indicadores puntuales sobre la misma.

Siguiendo la nomenclatura de Cuevas et al. (2004), cada dato funcional se puede representar como  $X_{ij}(t)$  con  $t \in [a, b]$ , extraídos de  $k$  procesos  $L_2$ ,  $X_i(t)$ , siendo  $i$  el subíndice que indica el nivel del factor y  $j$  el número de réplica ( $j = 1, 2, \dots, n_i$  e  $i = 1, 2, \dots, k$ ). Como, en este caso, la temperatura es directamente proporcional al tiempo (10 °C/min), puede considerarse a “ $t$ ” como los valores de la propia temperatura.

La media correspondiente a cada nivel o muestra independiente tiene la forma  $E(X_i(t)) = m_i(t)$ , mientras la covarianza entre dos instantes determinados dentro de una curva, en el caso más restrictivo de existencia de heterocedasticidad, se

identifica como  $K_i(s, t)$ :

$$K_i(s, t) = \sum_{j=1}^{n_i} \frac{(X_{ij}(s) - \bar{X}_{i\cdot}(s)) (X_{ij}(t) - \bar{X}_{i\cdot}(t))}{n_i - 1} \quad (1)$$

Por tanto, en este contexto funcional, el objetivo del contraste asociado al ANOVA funcional es probar que:

$$H_0 : m_1 = m_2 = \dots = m_k \quad (2)$$

El estadístico implementado por Cuevas et al. (2004) para contratar esta hipótesis es el siguiente:

$$V_n = \sum_{i < j} n_i \|\bar{X}_{i\cdot} - \bar{X}_{j\cdot}\|^2 \quad (3)$$

La utilización del estadístico (3) evita el requerimiento de la hipótesis de homocedasticidad en un ANOVA usual, siempre que se cumplan las propiedades siguientes:

- a)  $n_i, n \rightarrow \infty$  suponiendo  $p_i > 0$  para  $i = 1, \dots, k$ ,  $p_i = \frac{n_i}{n}$ .
- b) las observaciones  $X_{ij}(t)$  con  $j = 1, \dots, n_i$ , se corresponden con  $k$  muestras independientes de tamaño  $n_i$  procedentes de  $k$  procesos  $L_2$  con media cero y covarianza  $K_i(s, t)$ , se puede afirmar que la distribución de  $V_n$  bajo la hipótesis  $H_0$  coincide con la del estadístico:

$$V = \sum_{i < j} \|C_i(t) - C_{ij} Z_j(t)\|^2 \quad (4)$$

Donde  $C_{ij} = (p_i/p_j)^{1/2}$  y  $Z_1(t), \dots, Z_j(t)$  son procesos gaussianos independientes de media cero y covarianza  $K_i(s, t)$ .

El cálculo de la distribución del estadístico (4) bajo la hipótesis nula puede hacerse mediante la aplicación de un bootstrap paramétrico y del método de Monte Carlo, obteniéndose el valor del cuantil  $V_\alpha$ , que permite resolver el contraste, así si  $V_n > V_\alpha$  se rechazará la hipótesis  $H_0$  siendo el test significativo, es decir, las medias de los distintos niveles del factor serían diferentes.

### 3. Experimentación y obtención de datos

#### 3.1. Materiales

Los materiales empleados en este estudio están formados por una resina epoxi, que actúa como matriz del material compuesto. Este polímero está basado en el diglicidil eter de trimetilolpropano, Triepox GA, manufacturado por la empresa GAIRESA, SA. El agente de curado utilizado es una amina aromática, 1,3-bencenodimetanamina, suministrada por Aldrich. Triepox GA es una resina altamente tixotrópica que posee, además de una baja densidad, la capacidad de curar a temperatura ambiente en ausencia de plastificantes o aditivos.

El humo de sílice ha sido proporcionado por Ferroatlántica I+D, España. Se obtiene como un subproducto en la producción del silicio en hornos de arco eléctrico sumergido. Este proceso implica la reducción de cuarzo de alta pureza, a temperaturas superiores a 1800°C. El humo de sílice se forma cuando el gas SiO, resultante de la reducción de cuarzo, se mezcla con el oxígeno en la parte superior del horno, dando lugar a la producción de partículas esféricas de sílice. Es un polvo fino de color variable, desde casi negro a ligeramente blanquecino, en función de su contenido en carbono. Su tamaño medio de partícula es de 0.15 micras y el 41.9% de las partículas tienen un diámetro inferior a 0.2 micras. El área superficial se sitúa en torno a  $20 \text{ m}^2\text{g}^{-1}$ . En cuanto a la composición química, el humo de sílice consta de SiO<sub>2</sub> amorfo de pureza variable.

### 3.2. Obtención de las muestras

Las muestras han sido preparadas para contenidos del 0, 10, y 20% en peso de humo de sílice. Tanto la resina como el endurecedor se han mezclado en una proporción estequiométrica. Para la obtención de los compuestos correspondientes al 10 y al 20% en humo de sílice, las mezclas de sílice y resina se han agitado durante 15 minutos con el fin de obtener una distribución lo más uniforme posible. A continuación, se ha aplicado un tratamiento de ultrasonidos durante 5 minutos y a temperatura ambiente para dispersar a los aglomerados de sílice. La pasta así obtenida se ha vertido en un molde de silicona con cavidades de dimensiones  $0.8 \times 4 \times 30$  mm. En este emplazamiento, las muestras han sido curadas a temperatura ambiente durante 24 horas, aplicando posteriormente un postcurado a 90°C durante 2 horas.

Los experimentos TGA se han llevado a cabo por medio de una termo-balanza STA 1500, de Rheometric Scientific. Todas las muestras se han sometido a una rampa de calentamiento de 10°C/min, en un rango de temperaturas comprendido entre 20 y 600°C. Todos los experimentos se han realizado bajo atmósfera oxidante, manteniendo un flujo de aire de  $50 \text{ mL}\cdot\text{min}^{-1}$ .

## 4. Diseño de experimentos

Con el objeto de evaluar el efecto de la adición de humo de sílice en la degradación térmica del material resultante, se ha realizado un diseño de experimentos con datos funcionales. Se ha elegido un diseño completamente aleatorizado de una vía para probar las posibles diferencias en la respuesta al variar el nivel del factor. Se ha tomado como factor de estudio o variable explicativa el contenido en sílice de cada muestra, con tres niveles (0%, 10% y 20% en peso de humo de sílice). Se han efectuado entre 5 y 7 experimentos o réplicas por cada nivel con el fin de caracterizar la influencia de la cantidad de humo de sílice en la estabilidad térmica de los materiales compuestos resultantes: 7 correspondientes a la resina epoxi sin carga, otros 7 para un 10% en peso de humo de sílice y 5 para un 20% (ver figura 1). El número de réplicas elegidas se ha marcado principalmente con el fin de llegar a un compromiso aceptable entre la representación adecuada de la variabilidad dentro

de cada nivel y el tiempo de experimentación total requerido. Además, de entre los 19 experimentos efectuados en total, 4 han sido realizados variando ligeramente las condiciones experimentales. Mediante un posterior análisis de profundidad de datos funcionales se pretende identificar las curvas TG obtenidas mediante condiciones experimentales ligeramente diferentes. En el caso de poder identificarlas correctamente -probándose la utilidad del cálculo de profundidad de datos- se procede a eliminarlas para no introducir mayor variabilidad al modelo, con lo que quedarían 5 curvas por nivel (correspondiéndose con un diseño balanceado).

Para poder comparar convenientemente todos los datos, se ha expresado la masa de cada muestra en tanto por ciento de la cantidad inicial. Así, todas las curvas comienzan con un valor del 100 % en el eje de ordenadas.

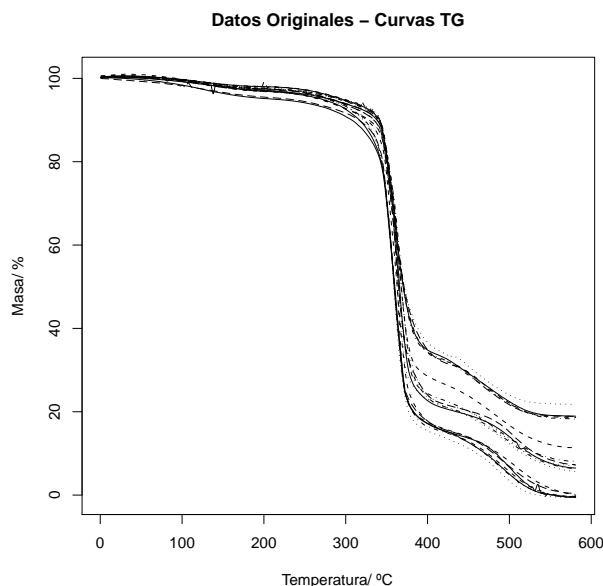


FIGURA 1: Datos experimentales: curvas TG de partida.

Cada uno de los experimentos se corresponde con un dato funcional donde se representa la masa de la muestra en función de la temperatura a la que está sometida. Como ya se ha indicado, cada muestra es calentada a  $10^{\circ}\text{C}/\text{min}$  en un rango de temperaturas que va de 20 a  $600^{\circ}\text{C}$ . Se observa que al final de cada ensayo, a  $600^{\circ}\text{C}$ , se degrada por completo la fase orgánica (resina epoxi), quedando únicamente la masa añadida de humo de sílice, mucho más resistente térmicamente. Es importante señalar que la masa de la muestra se representa en %, es decir, se asigna a la masa inicial el 100 % representándose todos los valores siguientes con respecto a esta. Cada curva consta de un número variable de puntos entorno a 3480, uno por segundo, dependiendo de la temperatura ambiente a la que se encuentre la máquina de ensayo.



La variable respuesta o dependiente es una variable funcional donde cada dato es una curva que representa la masa del material en función de la temperatura a la que está sometido. Para su obtención se ha programado un aumento constante en la temperatura de  $10^{\circ}\text{C}/\text{min}$ . Todas las curvas son decrecientes, pues el material se degrada, es decir, pierde masa, según aumenta la temperatura a la que está sometido. En sí, desde el punto de vista de un estudio de fiabilidad de los materiales, cada curva representa el camino particular de degradación de cada muestra ensayada. La finalidad de esta experimentación es dar respuesta a las siguientes preguntas: ¿son las medias de las curvas TG (camino de degradación de cada muestra) diferentes para distintos niveles del factor cantidad de sílice? Si la respuesta se responde afirmativamente, las diferencias en cuanto a estabilidad térmica observadas en la figura 4 tendrían una significación estadística.

#### 4.1. Suavización de los datos mediante bases splines

La representación de un dato funcional en una base ortonormal proporciona ventajas desde el punto de vista teórico y práctico sirviendo de puente entre la inevitable discretización del dato funcional para su tratamiento y su verdadera forma funcional. En este caso se representa cada dato funcional discretizado según una base finita, de modo que se obtiene una forma explícita para la función (Ramsay & Silverman 1997, Ferraty & Vieu 2006, Cuevas, Febrero & Fraiman 2006):

$$X(t) \approx \sum_{k=1}^K c_k \phi_k(t) \quad (5)$$

Donde la base  $\{\phi_k\}$ , con  $k \in \mathbb{N}$ , es un conjunto de funciones conocidas e independientes tales que cualquier función puede ser aproximada, mediante la combinación lineal de  $K$  de ellas (elementos de la base) y siendo  $c_k$  constantes.

Eligiendo una base apropiada se obtiene una forma más suave del dato funcional. Además, presenta la ventaja de poder ser derivable. De entre todos los tipos de base disponibles, se ha elegido ensayar una bases de b-splines y b-splines penalizados. Esto es debido a las características de los datos de partida. Ambas bases proporcionan la estructura flexible requerida. Las bases de Fourier, sin embargo, no serían tan apropiadas, ya que los datos no presentan una aparente trayectoria periódica. Además, dada la ya de por sí suavidad de los datos de partida, no parece tampoco adecuado ensayar bases wavelets. El ajuste según una base b-splines tiene la expresión:

$$S(t) \approx \sum_{k=1}^{m+L-1} c_k B_k(t, \tau) \quad (6)$$

Donde  $m$  corresponde al orden del polinomio,  $\tau$  son los valores de los  $L$  puntos de corte de los subintervalos (nodos) y  $k$  es el número de elementos de la base. Mientras que la suma de residuos del ajuste según una base b-spline penalizada responden a la expresión:

$$PRSS = \|Y - X\beta\|^2 + \lambda \int (f''(x))^2 dx \quad (7)$$

Donde  $Y$  son los datos a ajustar,  $X\beta$  es el ajuste según la base b-spline,  $f''(x)$  es la derivada segunda de dicho ajuste y  $\lambda$  es el parámetro de suavización (que penaliza la derivada segunda, es decir, restringe la varianza interna que pueda tener el ajuste). Más información al respecto se puede encontrar en los trabajos de Ferraty & Vieu (2006), además de los ya mencionados de Ramsay & Silverman(1997, 2005).

Para elegir el número óptimo de elementos de la base se ha implementado el criterio de Validación Cruzada Generalizada (GCV, por sus siglas en inglés). Se ha ensayado un número de elementos de la base que oscila entre 4 y 480. Se ha elegido el número de elementos que se corresponde con el GCV mínimo en cada dato funcional. Acto seguido se ha elegido el número de elementos de la base mínimos de entre todos los mínimos 19 GCV (uno por dato funcional). De esta forma se disminuye el riesgo de sobresuavizado. El resultado de la minimización de la expresión anterior puede verse en la figura 2 y en la tabla 1.

TABLA 1: Número de elementos de las bases óptimos según el criterio GCV.

Profundidad	GCV óptimo	N° de elementos	GCV (80)
b-splines	$2.0e - 07$	375	$3.1e - 03$
b-splines penal ( $\lambda = 0.5$ )	$3.1e - 04$	182	$6.4e - 03$

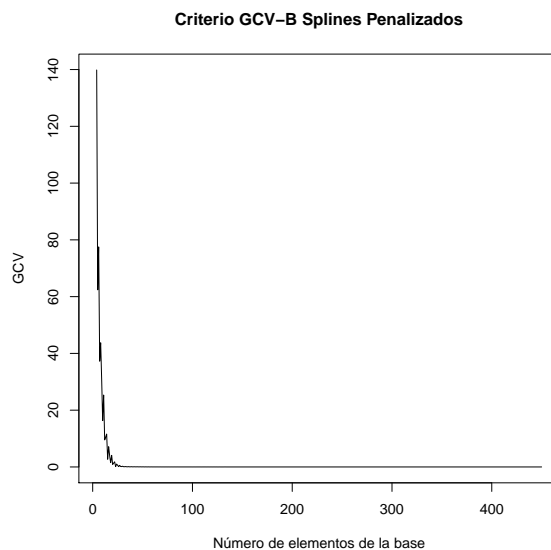


FIGURA 2: Para un dato funcional determinado, GCV frente al número de bases para una base b-spline penalizada.

En cuanto a los b-splines, el GCV óptimo resulta demasiado pequeño y el número de bases demasiado grande, con lo cual se corre el riesgo de interpolar los datos. En cambio se obtiene un GCV aceptable para una base de 80 elementos. Se consigue un ajuste más suave sin apartarse de la trayectoria de los datos originales.

Empleando b-splines penalizados se obtiene un valor para GCV mayor, con un número de elementos en la base también menor. Pero se observa que para 80 elementos se obtiene un GCV aceptable. La prueba está en la figura 2: para un número de elementos de la base de 80 se consigue seguir fielmente la trayectoria de los datos, suavizando además las discontinuidades experimentales. Asimismo, en la figura 2 se observa que el GCV cae abruptamente hasta estancarse para un valor determinado del número de elementos de la base. Este hecho apoya la decisión de optar por una base más pequeña, correspondiente a un número de elementos de la base más cercano al comienzo de ese estancamiento.

Si se disminuyera todavía más el número de elementos de la base, se incurriría en un error inaceptable. Los ajustes se alejan de los datos allí donde es crítico que los reproduzcan fielmente: en los cambios de pendiente de los escalones.

Se ha elegido una base b-spline penalizada de orden cuatro, con 80 elementos, como base de ajuste. El ajuste es muy bueno y, muy ligeramente, más suave que en una base de b-splines. Además, el número reducido de datos no origina un coste computacional muy grande.

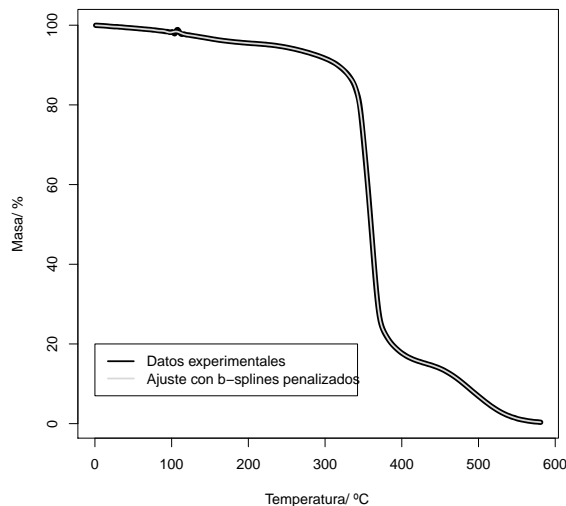


FIGURA 3: Curva TGA de resina epoxi y ajuste con una base b-spline penalizada con 80 elementos.

Ya de por sí, los datos presentan una forma muy suave, sin grandes discontinuidades. Se observa que los dos ajustes de la figura 3 reproducen fielmente la

trayectoria de los datos reales, obviando las pocas discontinuidades debidas al error experimental.

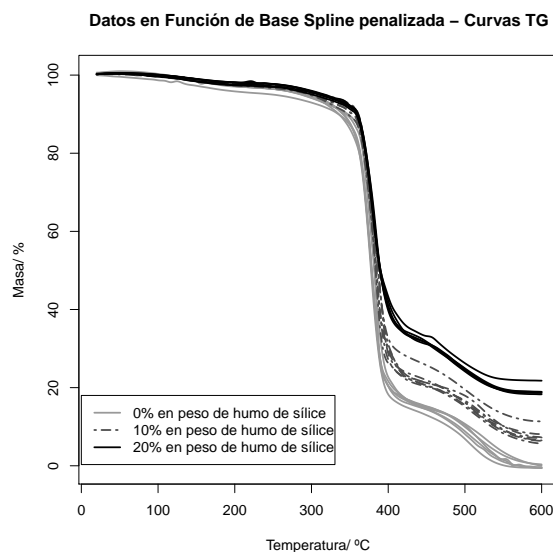


FIGURA 4: Dato experimental (curva TGA de resina epoxi) y ajuste con una base b-spline penalizada con 80 elementos.

## 4.2. Estudio de la profundidad de los datos

El concepto de profundidad, aplicado al análisis de datos funcionales, indica qué tan central es una curva o dato funcional particular en relación con un conjunto de curvas pertenecientes a una población. Siguiendo este criterio, pueden ordenarse los diferentes datos funcionales pertenecientes a una muestra de una población dada: se identifican como más profundos aquellas curvas más cercanas al centro (dato más profundo definido como mediana), mientras que se identifica como datos atípicos aquellas curvas que se generan por un proceso estocástico diferente al resto (Cuevas et al. 2006). En el presente estudio se ha calculado la profundidad asociada a cada dato funcional por tres métodos referenciados cuando se tratan datos de dimensión infinita (FDA Cuevas, Febrero & Fraiman (2007)). El resultado obtenido nos da una idea completa de la profundidad de cada curva TG con respecto al total.

El primer método utilizado es la profundidad de Fraiman & Muniz (2001), que está basado en la idea de profundidad funcional integrada:

$$I_i = \int_0^1 D_i(t) dt \quad (8)$$

donde  $D_i(t) = 1 - |\frac{1}{2} - F_{n,t}(x_i(t))|$  es la profundidad univariante correspondiente a la curva  $x_i(t)$  con  $t \in [0, 1]$  y  $F_{n,t}$  la distribución empírica de la muestra

$x_1(t), \dots, x_n(t)$ . Los datos funcionales  $x_i(t)$  se ordenan según el valor  $I_i$  correspondiente.

El método de la profundidad moda es una alternativa al uso de la profundidad de Fraiman y Muniz en el contexto del FDA. Según este procedimiento, se define la población de profundidades moda relativas a un dato  $z$  mediante la expresión:

$$f_h(z) = E(K_h(\|z - X\|)) \quad (9)$$

para el caso de dimensión infinita y

$$\hat{f}_h(z) = \frac{1}{n} \sum_{i=1}^n K_h(\|z - X_i\|) \quad (10)$$

para el caso de una muestra compuesta de  $n$  curvas donde  $X$  es el elemento aleatorio que describe la población,  $\|\cdot\|$  es, por lo general, la norma  $L_2$  y  $K_h$  un kernel reescalado que presenta la forma  $K_h = \frac{1}{h} K(\frac{\cdot}{h})$ , siendo  $K(t)$  una función kernel (el kernel gaussiano es el más utilizado) y  $h$  un parámetro de ajuste. La profundidad asignada a cada dato funcional  $z$  coincide con la moda de  $X$ , que es el producto de la maximización, en  $z$ , de la función  $f_h(z)$ .

Por último, a continuación se muestra el método de proyecciones aleatorias que consiste en, a partir de una muestra  $X_1, X_2, \dots, X_n$ , tomar una dirección aleatoria  $a$  independiente de  $X_i$  y proyectar los datos a lo largo de esa dirección. De este modo, la profundidad de cada dato funcional  $X_i$  es definida como la profundidad univariante correspondiente a su proyección unidimensional. En el caso particular de datos funcionales, suponiendo que  $X_i$  pertenece a un espacio de Hilbert, las proyecciones de cada dato se obtienen mediante el producto interior  $\langle a, S \rangle = \int_0^1 a(t)X(t) dt$ . Para más información, consultar Cuevas et al. (2007).

Cada nivel del factor aparece, en la figura 4, relativamente bien diferenciado de los restantes. Las diferencias se observan sobre todo a altas temperaturas. La resina se va degradando paulatinamente de forma que al final sólo va quedando el humo de sílice añadido. La resina sin humo de sílice se degrada completamente (en tono gris claro y trazo continuo) mientras que todavía queda un entorno a un 10 % y un 20 % de la masa para los grupos que tienen un 10 y un 20 % en carga de humo de sílice.

A simple vista se pueden observar tres escalones en las curvas. El primero, no siempre perceptible, se corresponde a la pérdida de sustancias volátiles y humedad (se localiza a temperaturas alrededor de 100 °C).

Es singularmente importante el segundo escalón de las gráficas, ya que indica en realidad el primer proceso de degradación. Es aquí donde la resina epoxi empieza a perder masa, y lo hace de forma abrupta. Se observa que cuando se pasa de 0 % a 10 % en humo de sílice, la masa remanente al comienzo de este escalón es ligeramente mayor a una temperatura dada (por ejemplo, en torno a 320 °C). Esto quiere decir que se ha conseguido aumentar aparentemente la estabilidad térmica del material compuesto resultante. En cambio, cuando se pasa de 10 % a 20 % de humo de sílice, este aumento es mucho más ligero.

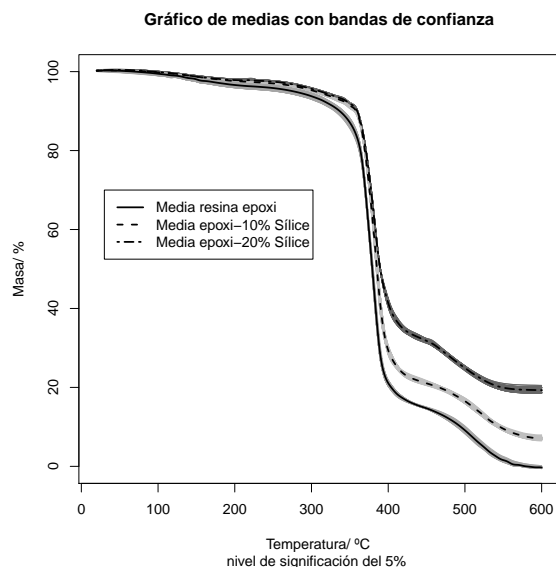


FIGURA 5: Datos funcionales suavizados con una base b-spline penalizada de 80 elementos.

El tercer escalón se corresponde con la desaparición del residuo carbonoso resultante de la reacción anterior. Se observa que aquí las diferencias son muy evidentes, debido principalmente a que el humo de sílice no se degrada a esas temperaturas: queda patente, al comparar las muestras, las diferencias que existen de contenido en sílice.

Para mayor sencillez de cálculo del estadístico de contraste, se pretende realizar un diseño balanceado. Por consiguiente, habría que desestimar dos datos funcionales correspondientes al nivel del 0% y otros dos del 10%. De hecho, existen algunas diferencias de experimentación en algunos datos, con lo cual sería interesante detectarlos como atípicos o estudiar simplemente su profundidad.

## 5. Resultados

### 5.1. Cálculo de medidas de profundidad

Se han calculado las profundidades de Fraiman & Muniz (2001), llamada profundidad mediana, la profundidad moda (el punto más profundo sería el equivalente a la moda de los datos) (Cuevas et al. 2006) y la profundidad de proyecciones aleatorias (Cuevas et al. 2007) para los grupos de 0% y de 10% por separado. Es en estos grupos donde interesa reducir el número de datos. Los resultados aparecen en las tablas 2 y 3. En estas tablas se presentan la profundidad mediana (F-M), la profundidad moda y la de proyecciones aleatorias (P-A) para las 14 muestras analizadas (M1, M2, ..., M14).

Todos los criterios coinciden en indicar los datos menos profundos tanto para el 0 % como para el grupo de 10 % de humo de sílice: para el 0 % es la muestra 7 la menos profunda por unanimidad, mientras que para el 10 % es la muestra 12 con diferencia (obsérvense los números en negrita dentro de las tablas).

TABLA 2: Profundidades para las 7 muestras de resina epoxi sin humo de sílice según los 3 criterios.

Prof.	M1	M2	M3	M4	M5	M6	M7
F-M	0.766	0.757	0.746	0.766	<b>0.718</b>	0.799	<b>0.661</b>
Moda	0.645	<b>0.345</b>	0.480	1.000	0.725	0.910	<b>0.000</b>
P-A	0.258	0.234	<b>0.200</b>	0.287	0.247	0.287	<b>0.202</b>

TABLA 3: Profundidades para las 7 muestras de 10 % en peso de humo de sílice según los 3 criterios.

Prof.	M8	M9	M10	M11	M12	M13	M14
F-M	<b>0.751</b>	0.721	0.814	0.755	<b>0.577</b>	0.764	0.833
Moda	0.705	0.751	0.745	0.942	<b>0.000</b>	<b>0.565</b>	1.000
P-A	0.232	0.271	0.260	0.280	<b>0.190</b>	<b>0.198</b>	0.284

Es necesario todavía eliminar otro punto dentro de estos dos grupos pero la unanimidad en los tres métodos ya no existe. Así, en el caso de 0 % se elige como posible atípico una solución de compromiso: la muestra 3, que coincide con el segundo dato menos profundo para las proyecciones aleatorias y con el tercero menos profundo según FM y la profundidad moda. Para el caso de 10 %, se ha elegido el segundo dato menos profundo según la profundidad de proyecciones aleatorias y según la profundidad moda, la muestra 13.

Es de destacar que en el caso de 0 % en humo de sílice, se ha identificado como menos profundo aquel cuyas condiciones experimentales eran más especiales: se ensayó muy poca masa de material con respecto a las demás muestras y, además, partida en dos trozos. La identificación de la muestra 3 puede deberse a que, en este caso, no se efectuó una calibración adicional en peso.

En cuanto a la de 10 %, se ha señalado como dato menos profundo aquel cuya masa en humo de sílice se aparta más de 10 % especificado. La dispersión de la carga sobre la matriz puede ser no del todo homogénea debido a la formación de aglomerados de partículas. Por lo tanto, dependiendo del trozo de muestra elegido puede haber más o menos cantidad de sílice. En definitiva, los puntos menos profundos han resultado ser los correspondientes a las condiciones más extremas de ensayo.

Una vez estudiada la profundidad de los datos, cabría la posibilidad que aquellos menos profundos pudieran clasificarse como atípicos. Para localizarlos se utiliza el siguiente procedimiento (Febrero, Galeano & González-Manteiga 2008):

- Se obtienen las medidas de profundidad por uno de los tres métodos.

- Si las profundidades  $D(x_{ik}) \leq C$  con  $k = 1, \dots, K$ , siendo  $C$  una profundidad prefijada, los datos funcionales  $x_{ik}$  se toman como datos atípicos.
- Se ha de elegir  $C$  de forma tal que el número de observaciones no atípicas clasificadas como atípicas sea aproximadamente igual al 1 %.

Para calcular  $C$  se realizan  $B$  remuestras bootstrap (suavizado) a partir de la muestra inicial (una vez descartadas las  $\alpha$  menos profundas):  $y_{ib}$ , donde  $i = 1, \dots, n$  y  $b = 1, \dots, B$ . Acto seguido se obtiene el cuantil  $C_b$  correspondiente al 1 % de la distribución de las profundidades  $D(y_{ib})$ . El valor  $C$  es la mediana de los valores de  $C_b$ .

Se ha realizado, por tanto, un bootstrap suavizado basado en el recorte (Cuevas et al. 2006) con  $B = 200$  y  $\alpha = 0.05$ . Se ha empleado la profundidad basada en Fraiman y Muniz dando como resultado la no existencia de atípicos. Hay que indicar que en este resultado pesa el hecho de disponer sólo de 5 datos funcionales por nivel.

Finalmente, se ha optado por prescindir de aquellos datos elegidos como menos profundos, así para el 0 % en humo de sílice prescindiremos de las muestra 3 y 7; mientras que para el 10 % en humo de sílice, se eliminarán las muestras 12 y 13. Con los 15 datos funcionales restantes, 5 por nivel, se calculan la media, mediana y varianza de los datos funcionales (ver figura 4, figura 5 y figura 6).

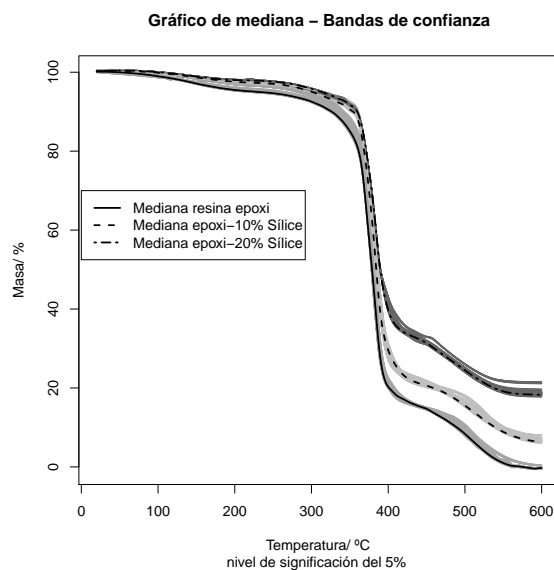


FIGURA 6: Medias para los distintos grupos con bandas de confianza bootstrap.

Mediante bootstrap simple (en el caso de la media) y bootstrap suavizado de parámetro  $h = 0.07$  (en los demás casos), se construyen bandas de confianza para cada uno de los estadísticos. Se ha empleado bootstrap suavizado en aquellos casos



donde el bootstrap simple no ofrece una idea clara de cómo puede ser el intervalo de confianza, siendo necesaria la adición de una componente aleatoria adicional.

Al realizar un bootstrap suavizado se está perturbando la muestra original sumándole un proceso gaussiano multiplicado por un factor de suavizado  $h$  (Cuevas et al. 2006):  $x_i^* = x_j(t) + hz(t)$  donde  $z(t)$  es un proceso gaussiano de media cero y matriz de covarianzas  $\Sigma$ . Así se combate en cierta forma la existencia de tan pocos datos para elaborar las bandas.

Se calculan la media y las medianas (Fraiman & Muniz 2001), tanto en la media como en la mediana se observa que las curvas correspondientes al 20 % en humo de sílice se sitúan por encima de las curvas correspondientes 10 %, que a su vez se sitúan por encima de las que no tienen humo de sílice. Esto es un indicador del aumento de la estabilidad térmica. Se observa que es mayor la diferencia entre los estadísticos de 0 % y de 10 % que entre 10 y 20 %, sobre todo en el escalón principal.

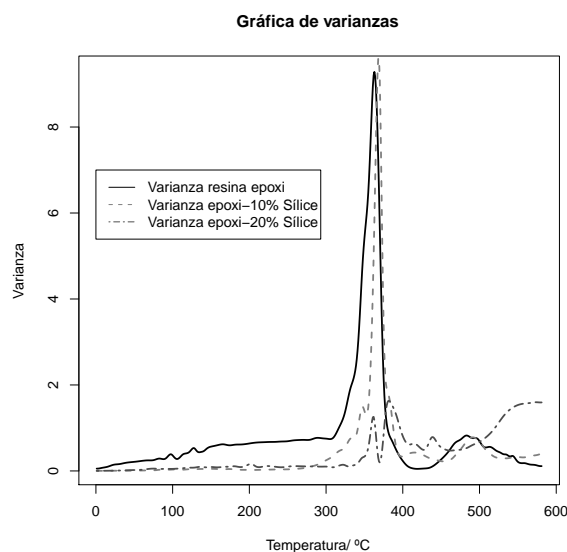


FIGURA 7: Varianza para los distintos grupos.

Se observa una mayor variabilidad general en los datos correspondientes a la resina epoxi sola. La variabilidad disminuye ligeramente para 10 % y, finalmente, es mucho menor para 20 % (esto se puede observar en media y mediana, además de en la varianza de la figura 6). Dicha tendencia puede deberse a la heterogeneidad de las propias muestras o al efecto aprendizaje del operador. De hecho, las primeras muestras que se realizaron corresponden a 0 % y las últimas a 20 % (aquellas que tienen menor variabilidad). Otra posible causa a tener en cuenta es el ensayo de muestras con diferentes contenidos de humedad, para el caso de 0 %.

La gráfica de las varianzas (figura 6) es en particular interesante porque resalta que allí donde difieren las sucesivas muestras es, principalmente, en los dos esca-

nes principales asociados a los dos primordiales procesos de degradación. Además, se observa que, contrariamente a lo que se había enunciado antes, la varianza para una temperatura cercana a los 600°C aumenta en la secuencia: resina epoxi→10% humo de sílice→20% humo de sílice. Esto se debe a que se forman aglomerados de partículas con lo que la distribución no es del todo uniforme. Así, para un 10% teórico, unas muestras tendrán un 8% en sílice, otras un 9% y otras un 11% (figura 5). La varianza es un indicador de la heterogeneidad de la muestra.

## 5.2. Aplicación del ANOVA funcional

Para construir el contraste, se han seguido los pasos marcados por Cuevas et al. (2004), teniendo en cuenta que la hipótesis nula a contrastar de igualdad de medias entre las curvas (7), donde  $m_i$  es la media de los datos funcionales dentro de cada uno de los tres niveles estudiados del factor “cantidad de humo de sílice”.

En primer lugar se construye el estadístico  $V_n$ , donde los tres niveles se corresponden con la adición del 0%, del 10% y del 20% en humo de sílice mientras que la respuesta está representada por la media de los datos funcionales para cada nivel. Con el fin de aproximar por Monte Carlo la distribución del estadístico  $V_n$ , se procede del modo siguiente:

- a) Se calcula la matriz de varianzas covarianzas de los datos muestrales. Si se supone que la estructura de covarianzas es diferente según el nivel del factor, nos encontramos en el caso heterocedástico; más restrictivo que el homocedástico. Suponiendo heterocedasticidad, se calcula una matriz de covarianzas diferente por nivel, siendo su estimador el siguiente:

$$\widehat{K}_i(s, t) = \frac{\sum_{j=1}^{n_i} (X_{ij}(s) - \overline{X}_{i\cdot}(s)) (X_{ij}(t) - \overline{X}_{i\cdot}(t))}{n_i - 1}$$

Siendo  $s$  y  $t$  dos instantes dados dentro de cada dato funcional.

- b) Bajo la hipótesis nula, el estadístico  $V$  sigue una distribución dependiente de  $k$  procesos gaussianos (Cuevas et al. 2004) difícil de conseguir. Pero dicha distribución se puede aproximar por Monte Carlo. Para ello se lleva a cabo el siguiente procedimiento bootstrap paramétrico:
  - b.1) partiendo de la muestra original, se calculan las matrices  $\widehat{K}_i(s_p, t_q)_{1 \leq p, q \leq n}$  y se generan  $B = 2000$  remuestras bootstrap por nivel, generadas según una distribución normal de media cero y matriz de covarianzas  $\widehat{K}_i(s_p, t_q)_{1 \leq p, q \leq n}$ .
  - b.2) se consiguen 2000 valores  $Z_{il}^* = (Z_{il}^*(t_1), \dots, Z_{il}^*(t_m))$  por nivel, con  $l = 1, \dots, 2000$  e  $i = 1, 2, 3$ . Estos valores aproximan las trayectorias continuas de  $Z_i(t)$  por versiones escalonadas evaluadas en una rejilla  $a \leq t_1, \dots, t_m \leq b$ .
  - b.3) se construyen 2000 réplicas de la forma  $\widehat{V}_l = \sum_{i < j} \left\| Z_{il}^*(t) - C_{ij} Z_{jl}^*(t_1) \right\|^2$ , que aproximan la distribución de  $V$  bajo  $H_0$ .

Mediante el proceso anterior se puede calcular  $V_\alpha$  tal que  $P(V > V_\alpha) = \alpha$  bajo la hipótesis nula. Si  $V_n > V_\alpha$ , el contraste resulta significativo.

El resultado de la implementación de este proceso al caso de las curvas TG sin reescalar ha proporcionado los siguientes valores  $V_n = 570315.1$  y  $V_\alpha = V_{0.95} = 1.402.292$ .

Por lo tanto, como  $V_n \gg V_\alpha$  (siendo el  $P$ -valor =  $5e - 6 < 0.05$ ), el test resulta significativo para un nivel de significación del 0.05, al menos una de las medias funcionales por factor es diferente a las demás. Desde el punto de vista ingenil podría afirmarse que la adición de humo de sílice provoca cambios en las medias funcionales de las curvas TG. Se podría concluir que aumenta significativamente la estabilidad térmica del material compuesto al aumentar la cantidad de sílice.

## 6. Conclusiones y líneas futuras

Como una primera conclusión de este estudio, podría decirse que la aplicación de técnicas de estadística funcional resulta determinante para el estudio de este tipo de datos. Además, el concepto de profundidad resulta muy adecuado para identificar ligeros cambios en las condiciones experimentales, posibilitando eliminar muestras anómalas.

Por otra parte el contraste ANOVA funcional empleado ha resultado significativo, lo cual avala el resultado del análisis realizado mediante el estudio de la profundidad de los datos. Es decir, el aumento en la estabilidad térmica del material compuesto, observado en el análisis descriptivo o, lo que en este caso es lo mismo, las diferencias provocadas en la respuesta funcional al aumentar la cantidad de humo de sílice, tienen significación estadística.

La variabilidad observada en el último tramo de las curvas, creciente conforme se aumenta la cantidad de humo de sílice, ha resultado ser un buen indicador del grado de dispersión de esta carga inorgánica dentro de la matriz polimérica.

Como línea de futura investigación sería interesante realizar un análisis similar con los datos de las curvas TGA realizando una sustracción matemática de la cantidad de humo de sílice añadida y el posterior reescalado con objeto de caracterizar el camino de degradación de la resina epoxi pura, dentro de cada muestra de material compuesto.

También resultaría de gran interés la realización de un estudio similar con las derivadas de estas curvas TGA, las denominadas DTG, ya que permitirá analizar si la velocidad de degradación térmica es diferente en función del humo de sílice añadido. Además, podrían plantearse estudios de interacción entre los factores estudiados.

## Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Educación y Ciencia MTM2008-00166. Los autores quieren expresar su especial agradecimiento a Manuel Febrero Bande por sus interesantes aportaciones para la aplicación de los programas en R utilizados (fda.usc), así como a Aldana González Montoro, por sus valiosos comentarios. Además, también agradecen a los árbitros que revisaron el trabajo sus interesantes sugerencias.

[Recibido: abril de 2010 — Aceptado: febrero de 2011]

## Referencias

- Cuevas, A., Febrero, M. & Fraiman, R. (2004), 'An anova test for functional data', *Computational Statistics and Data Analysis* **47**, 111–122.
- Cuevas, A., Febrero, M. & Fraiman, R. (2006), 'On the use of the bootstrap for estimating functions with functional data', *Computational Statistics and Data Analysis* **51**, 1063–1074.
- Cuevas, A., Febrero, M. & Fraiman, R. (2007), 'Robust estimation and classification for functional data via projection-based depth notions', *Computational Statistics* **22**(3), 481–496.
- Febrero, M., Galeano, P. & González-Manteiga, W. (2008), 'Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels', *Environmetrics* **19**, 331–345.
- Ferraty, F. & Romain, Y. (2010), *The Oxford Handbook of Functional Data Analysis*, Oxford University Press, Oxford, England.
- Ferraty, F. & Vieu, P. (2006), *Nonparametric Functional Data Analysis: Theory and Practice*, Series in Statistics, Springer, Berlin, Alemania.
- Fraiman, R. & Muniz, G. (2001), 'Trimmed means for functional data', *Test* **10**(2), 419–440.
- Harsch, M., Karger-Kocsis, J. & Holst, M. (2007), 'Influence of fillers and additives on the cure kinetics of an epoxy/anhydride resin', *European Polymer Journal* **43**, 1168–1178.
- Hsiue, G. H., Chen, J. K. & Liu, Y. L. (2000), 'Synthesis and characterization of nanocomposite of polyimide-silica hybrid film nanoaqueous sol-gel process', *Journal of Applied Polymer Science* **76**, 1609–1618.
- Lee, A. & Lichtenhan, J. D. (1999), 'Thermal and viscoelastic property of epoxy-clay and hybrid inorganic-organic epoxy nanocomposites', *Journal of Applied Polymer Science* **73**, 1993–2001.

- Liu, Y. L., Wei, W. L., Hsu, K. Y. & Ho, W. H. (1999), 'Thermal stability of epoxy-silica hybrid materials by thermogravimetric analysis', *Thermochimica Acta* **412**, 139–147.
- Mehta, S., Mirabella, F. M., Bafna, A. & Rufener, K. (2004), 'Thermoplastic olefin/clay nano-composites: morphology and mechanical properties', *Journal of Applied Polymer Science* **92**, 928–936.
- Mohammad, A. & Simon, G. P. (2006), Rubber-clay nanocomposites, in M. Yiu-Wing & Y. Zhong-Zhen, eds, 'Polymer Nanocomposites', Woodhead Publishing Limited.
- Naya, S., Cao, R. & Artiaga, R. (2003), 'Local polynomial estimation of TGA derivatives using logistic regression for pilot bandwidth selection', *Thermochimica Acta* **6**, 319–322.
- Petrie, E. M. (2006), *Epoxy Adhesive Formulations*, McGraw-Hill.
- Preghenella, M., Pegoretti, A. & Migliaresi, C. (2005), 'Thermo-mechanical characterization of fumed silica-epoxy nanocomposites', *Polymer* **46**, 12065–12072.
- Ramsay, J. O. & Silverman, B. W. (1997), *Functional Data Analysis*, Springer.
- Ramsay, J. O. & Silverman, B. W. (2005), *Applied Functional Data Analysis*, Springer.
- Schadler, L. S. (2003), Polymer-based and polymer-filled nanocomposites, in P. M. Ajayan, L. S. Schadler & P. V. Braun, eds, 'Nanocomposite Science and Technology', Wiley-VCH, Weinheim, pp. 77–135.
- Shao-Yun, F., Xi-Qiao, F., Bernd, L. & Yiu-Wing, M. (2008), 'Effects of particle size, particle/matrix interface adhesion and particle loading on mechanical properties of particulate-polymer composites', *Composites: Part B* **39**, 933–961.
- Tarrio-Saavedra, J., López-Beceiro, J., Naya, S. & Artiaga, R. (2008), 'Effect of silica content on thermal stability of fumed silica/epoxy composites', *Polymer Degradation and Stability* **93**, 2133–2137.
- Yousefi, A., Lafleur, P. G. & Gauvin, R. (1997), 'Kinetic studies of thermoset cure reactions: a review', *Polymer Composites* **18**, 157–168.
- Zhang, H., Zhang, Z., Friedrich, K. & Eger, C. (2006), 'Property improvements of in situ epoxy nanocomposites with reduced interparticle distance at high nanosilica content', *Acta Materials* **54**, 1833–1842.

## Comparación entre riesgos competitivos vía el estimador cópula-gráfico

### Comparison between Competing Risks via the Copula-Graphic Estimator

SERGIO YÁÑEZ<sup>1,a</sup>, HUGO BRANGO<sup>2,b</sup>, MARIO C. JARAMILLO<sup>1,c</sup>,  
CARLOS M. LOPERA<sup>1,d</sup>

<sup>1</sup>ESCUELA DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE COLOMBIA,  
MEDELLÍN, COLOMBIA

<sup>2</sup>DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA, UNIVERSIDAD DE CÓRDOBA, MONTERÍA,  
COLOMBIA

---

#### Resumen

En riesgos competitivos, el problema de identificabilidad asociado a la dependencia entre los modos de falla, se puede resolver utilizando el estimador cópula-gráfico que asume la forma de la cópula conocida. En este trabajo se compara, mediante simulación, este estimador para un modelo de riesgos competitivos dependientes con el estimador tradicional que asume independencia. Se muestra que ignorar la dependencia puede causar subestimación significativa de la función de confiabilidad. Se ilustra la temática con un ejemplo y se visualiza a futuro la posible utilidad del estimador cópula-gráfico en confiabilidad industrial para predecir en experimentos acelerados.

**Palabras clave:** confiabilidad, cópula, identificabilidad, riesgos competitivos.

#### Abstract

The identifiability problem in competing risks associated to the dependence measure between the failure modes, can be solved by means of the copula-graphic estimator which assumes the copula function known. This paper compares, through a simulation study, the copula-graphic estimator for a dependent competing risks model with the traditional one under the assumption of independence. It is shown that a significant subestimation of the reliability function can be made when the dependence is not taken into account. An example illustrates this point and as a future work it is seen the possibility of using copula-graphic in industrial reliability to predict in accelerated life tests.

**Key words:** Competing risks, Copula, Identifiability, Reliability.

---

<sup>a</sup>Profesor asociado. E-mail: syanez@unal.edu.co

<sup>b</sup>Profesor asistente. E-mail: habrango@hotmail.com

<sup>c</sup>Profesor asociado. E-mail: mcjarami@unal.edu.co

<sup>d</sup>Profesor asistente. E-mail: cmlopera@unal.edu.co

## 1. Introducción

Las causas o las diferentes maneras en que puede fallar un sistema o componente, reciben el nombre de modos de falla. Existen muchos sistemas, subsistemas y componentes que tienen más de un modo de falla; en ciertas aplicaciones y para algunos propósitos es importante distinguir entre esos diferentes modos de falla, con el objeto de mejorar la confiabilidad (Meeker & Escobar 1998). El tiempo de falla de un sistema con varios modos de falla puede ser modelado considerando un sistema en serie o un modelo de riesgos competitivos (“competing risks”), donde cada modo de falla es una componente del sistema en serie, y cada componente tiene un tiempo de vida aleatorio; por lo tanto, el sistema fallará cuando el modo de falla con el tiempo de vida más corto falle de manera que censura así a los otros modos de falla.

Muchas son las aplicaciones de riesgos competitivos en confiabilidad. Basu & Klein (1982) obtuvieron algunos resultados en la teoría de riesgos competitivos; Nelson (1990) presenta datos para el tiempo de falla en horas de calentadores industriales, los cuales tienen dos modos de falla; Bedford & Lindqvist (2004) trataron el problema de la identificabilidad en sistemas reparables cuando se presentan los riesgos competitivos; Bedford (2005) muestra como modelar la confiabilidad en presencia de riesgos competitivos; Manotas, Yáñez, Lopera & Jaramillo (2008), estudiaron el efecto en la estimación de la confiabilidad, cuando se asume el supuesto de independencia entre los tiempos de falla que compiten, que realmente son dependientes; Meeker, Escobar & Hong (2009) usaron pruebas de vida aceleradas para predecir la distribución del tiempo de falla de un nuevo producto con dos modos de falla.

En el caso no-paramétrico, la distribución conjunta en un problema con múltiples modos de falla no puede ser completamente identificada en la situación usual cuando sólo se conocen los tiempos mínimos de falla (Tsiatis 1975). Incluso en el caso paramétrico, los datos pueden contener poca información sobre el coeficiente de asociación entre las variables y es necesario hacer algunos supuestos al respecto (Meeker et al. 2009). Para el caso no-paramétrico el estimador cópula-gráfico de Zheng & Klein (1995) resuelve el problema bajo el supuesto del conocimiento de la forma funcional de la cópula que da cuenta de la estructura de asociación entre los riesgos competitivos. Es de anotar que bajo independencia el problema de identificabilidad no existe.

Yáñez, Manotas & Lopera (2006) y Manotas et al. (2008), estudiaron, para el caso Weibull y lognormal, el efecto en la estimación de la sobrevivencia cuando se asume el supuesto de independencia entre tiempos de falla que realmente son dependientes. Encontraron que bajo el supuesto de independencia, que es la metodología tradicional utilizada en muchos campos de la ciencia (Pintilie 2006), no genera errores significativos en la estimación de la función de confiabilidad del tiempo mínimo. Puesto que la literatura reciente reporta que el supuesto de independencia en riesgos competitivos generalmente da una visión pesimista del comportamiento del sistema, ellos comentan la importancia de estudiar en más detalle la temática.

En este trabajo se estima la función de confiabilidad bajo dependencia mediante el estimador cópula-gráfico propuesto en Zheng & Klein (1995). Se hace un estudio de simulación para comparar este estimador que permite incorporar la dependencia con el que considera independencia entre los modos de falla.

En la sección 2 se presentan los conceptos básicos de riesgos competitivos y se detalla el ejemplo que se usará como ilustración. El estimador cópula-gráfico se discute en la sección 3, con una ilustración que determina la cópula a usar en este artículo. En la sección 4 se presenta el esquema de simulación utilizado. El análisis de resultados se hace en la sección 5. Finalmente, en la sección 6 se presentan las conclusiones y posible trabajo futuro, donde se muestra que ignorar la dependencia puede causar una subestimación representativa de la función de confiabilidad.

## 2. Riesgos competitivos

En este artículo se trabaja con dos modos de falla. El tiempo de falla de un sistema con dos modos de falla puede ser modelado como un sistema en serie o un modelo de riesgos competitivos, como se ilustra en la figura 1. Cada unidad tiene un tiempo potencial de falla. El tiempo de falla observado es el mínimo de esos tiempos potenciales individuales.



FIGURA 1: Sistema con dos componentes en serie.

Así por ejemplo, para un sistema con dos modos de falla, sean  $X$  y  $Y$  los respectivos tiempos potenciales, entonces lo que se observa en la práctica son parejas de la forma  $(T, \delta)$ , donde  $T = \min(X, Y)$  y  $\delta$  es una variable discreta que toma valores de 0 cuando se presenta una censura a derecha, 1 cuando falla debido al primer modo de falla y 2 cuando falla debido al segundo modo de falla.

Para estudiar este modelo de riesgos competitivos, es necesario presentar algunas funciones básicas. La función de subdistribución o función de incidencia acumulada (FIA), para el modo de falla  $i$ ,  $i = 1, 2$ , está dada por:

$$G_i(t) = P(T \leq t, \delta = i) \quad (1)$$

La función de distribución total es la probabilidad que un evento de cualquier tipo ocurra en, o antes del tiempo  $t$ , es decir es igual a la suma de las FIA para los dos tipos de eventos, así:

$$F(t) = P(T \leq t) = \sum_{i=1}^2 P(T \leq t, \delta = i) = \sum_{i=1}^2 G_i(t) \quad (2)$$



Nótese que cuando no hay riesgos competitivos la función de distribución total se encuentra en el intervalo  $[0, 1]$ . En contraste, en presencia de riesgos competitivos la FIA sólo puede tomar valores hasta  $P(\delta = i)$  debido a que

$$\lim_{t \rightarrow \infty} G_i(t) = P(\delta = i) \quad (3)$$

Por lo tanto, la función  $G_i(t)$  no es una función de distribución propia.

La función de subsobrevivencia es la probabilidad que el modo de falla  $i$  no ocurra antes del tiempo  $t$

$$\bar{G}_i(t) = P(T > t, \delta = i) \quad (4)$$

La función de subdensidad para el modo de falla  $i$  y  $T$  continuo, se define como

$$g_i(t) = \frac{dG_i(t)}{dt} = -\frac{d\bar{G}_i(t)}{dt} \quad (5)$$

## 2.1. El problema de identificabilidad

La aproximación tradicional para especificar el modelo de riesgos competitivos es vía tiempos de falla latentes. En este contexto surge el denominado problema de identificabilidad. Para el caso bivariado tenemos dos tiempos de falla potenciales  $X, Y$  asociados a dos modos de falla, de manera que si ocurre el primero, el segundo no puede ser observado. En este sentido, se dice que los tiempos son latentes. Se supone que la pareja  $(X, Y)$  tiene una distribución multivariada de la forma  $F(x, y) = P(X \leq x, Y \leq y)$ . Así  $T = \min(X, Y)$  determina el tiempo de falla del sistema completo. De manera que una vez que el sistema haya fallado, el otro tiempo de falla no puede ser observado. Las funciones de distribución marginales para los dos modos de falla son  $F_1(t) = P(X \leq t)$  y  $F_2(t) = P(Y \leq t)$ , respectivamente.

El problema de identificabilidad (Tsiatis 1975) establece que para el caso de modelos de tiempos de falla latentes, dada una función de distribución conjunta con dependencia arbitraria entre las componentes, existe siempre una función de distribución conjunta en la cual las variables son independientes, y la cual reproduce las subdensidades  $g_i(t)$  de manera precisa. Así, uno no puede conocer a partir solamente de los datos observados  $(T, \delta)$ , cuál de los dos modelos es correcto, puesto que ambos se ajustan muy bien a los datos. Crowder (2001) da detalles técnicos sobre este problema que él denomina “El impase de Cox-Tsiatis”. En resumen, podemos decir que en la práctica se pueden encontrar las subdistribuciones marginales, asociados con datos de riesgos competitivos, y ese conjunto es consistente con un número infinito de distribuciones conjuntas de tiempos de falla potenciales. Esto quiere decir que es imposible identificar las distribuciones marginales a partir solamente de los datos de riesgos competitivos.

Supongamos que se tienen datos de la forma  $(T, \delta)$  y es necesario identificar el modelo conjunto  $F(x, y)$ . Algunas posibles soluciones son: i) Asuma independencia, lo cual es frecuentemente no realista (ver subsección 2.2.1); ii) Asuma una

estructura de dependencia conocida usando una cópula específica para la distribución conjunta de los modos de falla; Zheng & Klein (1995) proponen el llamado estimador cópula-gráfico, el cual es un estimador máximo verosímil y trabaja muy bien independiente de la cópula escogida (ver sección 3); iii) Especifique un modelo paramétrico. En este último caso, a pesar del fuerte supuesto, los datos de confiabilidad pueden contener poca información acerca de la asociación entre riesgos competitivos debido, a veces, a la alta censura y a la cantidad limitada de datos. Meeker et al. (2009) muestran un caso de este tipo donde el problema de identificabilidad prevalece.

En este artículo se van a comparar las soluciones i) y ii). También, como trabajo futuro se propone en la sección 6 incorporar los resultados de este artículo a la solución iii).

## 2.2. Metodología tradicional: un ejemplo en ingeniería

Se ilustra con un ejemplo de Meeker & Escobar (1998), la metodología tradicional en el análisis de modelos de riesgos competitivos, que supone independencia entre los dos modos de falla.

Se tienen tiempos de falla de un cierto tipo de equipos, los cuales se deben principalmente a dos modos o causas de falla, una asociada a fallas por descargas por acumulación de daños producidos por altos picos de voltaje durante tormentas eléctricas (se denota  $S$ ), y la otra falla debida al desgaste normal de los equipos (se denota  $W$ ). La tabla 1 presenta los tiempos y los modos de falla de 30 unidades que fueron instaladas en un ambiente de servicio estándar. Los tiempos de falla asociados al modo de falla  $S$  ocurren de forma temprana en la vida de los equipos. El modo de falla  $W$  empieza a aparecer después de los 100 kilociclos de uso.

TABLA 1: Tiempos de falla y modos de falla para los datos del ejemplo.

Kilociclos	Modo de falla	Kilociclos	Modo de falla	Kilociclos	Modo de falla
275	W	106	S	88	S
13	S	300	–	247	S
147	W	300	–	28	S
23	S	212	W	143	S
181	W	300	–	300	–
30	S	300	–	23	S
65	S	300	–	300	–
10	S	2	S	80	S
300	–	261	S	245	W
173	S	293	W	266	W

En la figura 2 se muestran los estimadores de Kaplan Meier, para cada modo de falla, considerando el otro modo de falla como censura. Se debe aclarar que cuando hay dependencia entre los modos de falla, los estimadores de Kaplan Meier no son estimadores consistentes de las distribuciones marginales.

En la figura 3 se muestra el estimador de Kaplan Meier, para el tiempo mínimo de los dos modos de falla.

Este ejemplo se usa más adelante en la subsección 5.4, bajo el supuesto de dependencia entre los dos modos de falla.

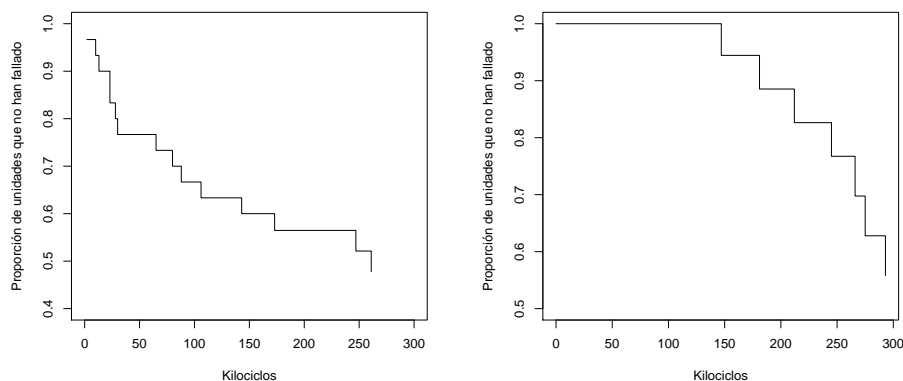


FIGURA 2: Estimación de la función de supervivencia de cada modo de falla.

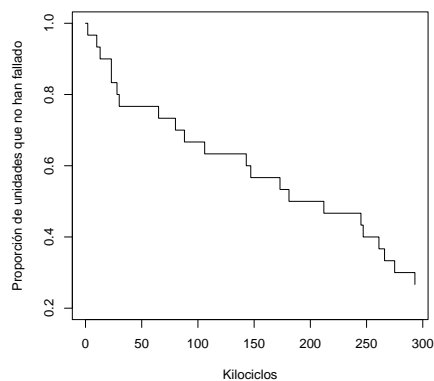


FIGURA 3: Estimación de la función de supervivencia del tiempo mínimo ignorando los modos de falla.

### 2.2.1. Estimador bajo independencia

Una estructura en serie con  $s$  componentes funciona, si y sólo si todas las componentes funcionan. Para un sistema en serie con dos componentes independientes en serie, la función de distribución es

$$\begin{aligned}
 F(t) &= P(T \leq t) = 1 - P(T > t) \\
 &= 1 - P(X > t, Y > t) \\
 &= 1 - P(X > t)P(Y > t) \\
 &= 1 - [1 - F_1(t)][1 - F_2(t)] = F_1(t) + F_2(t) - F_1(t)F_2(t)
 \end{aligned} \tag{6}$$

donde  $(X, Y)$  son los tiempos de vida asociados a las dos componentes y  $F_i, i = 1, 2$  son, respectivamente, las funciones de distribución para las componentes, y  $T = \min(X, Y)$  es el tiempo de falla observado del sistema.

La función de confiabilidad  $S(t) = 1 - F(t)$  para el caso de independencia es

$$\begin{aligned} S(t) &= P(T > t) = P(X > t, Y > t) \\ &= S_1(t)S_2(t) \end{aligned} \quad (7)$$

donde  $S_i(t), i = 1, 2$  son las funciones de confiabilidad para las dos componentes.

Para un sistema con dos componentes en serie y tiempos de falla dependientes,

$$\begin{aligned} F(t) &= P(T \leq t) = 1 - P(T > t) \\ &= 1 - P(X > t, Y > t) \end{aligned} \quad (8)$$

En este caso, la evaluación debe hacerse con respecto a la distribución conjunta de  $X$  y  $Y$ , que incluye los parámetros de dependencia.

La función de confiabilidad para este caso es

$$S(t) = P(T > t) = P(X > t, Y > t) \quad (9)$$

Para dos modos de falla en riesgos competitivos, se sabe que (Barlow & Proschan 1975)

$$\begin{aligned} S(t) &= P(X > t, Y > t) \geq P(X > t)P(Y > t) \\ &= S^*(t) = S_1(t)S_2(t) \end{aligned} \quad (10)$$

donde  $S^*(t)$  es la función de confiabilidad asumiendo independencia. Por lo tanto  $S^*(t)$  es pesimista en relación con  $S(t)$ , esto es para un mismo tiempo  $t$ , la probabilidad de sobrevivencia bajo el supuesto de independencia está siempre por debajo de la probabilidad conjunta que incluye la dependencia.

El estimador de  $S^*(t)$  se denotará  $\widehat{S}^*(t)$ , que es un estimador de la confiabilidad del tiempo mínimo de falla de un sistema con dos modos de falla que compiten. Éste se obtiene como el producto de las funciones de confiabilidad estimadas bajo una distribución específica, donde para cada función marginal se estiman los parámetros, considerando los tiempos del otro modo de falla como tiempos de censura, así:

$$\widehat{S}^*(t) = \widehat{S}_1(t)\widehat{S}_2(t) \quad (11)$$

donde  $\widehat{S}_1(t)$  y  $\widehat{S}_2(t)$  se suponen son las funciones de confiabilidad marginales estimadas para cada modo de falla.

### 3. Estimador cópula-gráfico

Zheng & Klein (1995) proponen un estimador no paramétrico para las distribuciones marginales, al cual denominan estimador cópula-gráfico, como una solución al problema de identificabilidad, descrito en la subsección 2.1. Para ello utilizan

cóputas como una función no paramétrica que captura la dependencia entre dos variables aleatorias. Bajo el supuesto que la cópula entre  $X$  y  $Y$  con su parámetro de asociación conocido y con los datos de riesgos competitivos disponibles, se resuelve el problema de identificabilidad.

Siguiendo a Zheng & Klein (1995), suponga que se tiene, en el contexto de riesgos competitivos, una muestra de tamaño  $n$  de tiempos de vida  $T_j = \min(X_j, Y_j)$ . Si  $F_1(t)$  y  $F_2(t)$  son las distribuciones marginales de  $X$  y  $Y$ , respectivamente, entonces, para cualquier cópula, tenemos

$$\mu_c(A_t) = P(X > t, Y > t) \quad (12)$$

donde

$$\mu_c(A_t) = \int_{I^2} \mathbf{1}_{E_t} \{F_1^{-1}(x), F_2^{-1}(y)\} dC(x, y)$$

con  $E_t = \{(x, y) : x > t, y > t\}$  y  $A_t = \{(x, y) : F_1(t) < x \leq 1, F_2(t) < y \leq 1\}$

$$\mu_c(B_t) = P(X \leq t, X < Y) \quad (13)$$

donde

$$\mu_c(B_t) = \int_{I^2} \mathbf{1}_{D_t} \{F_1^{-1}(x), F_2^{-1}(y)\} dC(x, y)$$

con  $D_t = \{(x, y) : x < y, 0 < x < t\}$  y  $B_t = \{(x, y) : 0 < x \leq F_1(t), F_2 F_1^{-1}(x) < y \leq 1\}$ .

Estas dos relaciones determinan de manera única a  $F_1(t)$  y  $F_2(t)$ . Por tanto, se pueden encontrar estimadores  $\hat{F}_1(t)$  y  $\hat{F}_2(t)$  para  $F_1(t)$  y  $F_2(t)$  respectivamente, que preservan estas relaciones sobre una rejilla seleccionada de  $m$  puntos  $0 < t_1 < t_2 < \dots < t_m < \max\{T_j, j = 1, \dots, n\}$ .

Para construir el estimador cópula-gráfico, sean

$$\hat{A}_t = \{(x, y) : \hat{F}_1(t) < x \leq 1, \hat{F}_2(t) < y \leq 1\}$$

$$\hat{B}_t = \{(x, y) : 0 < x \leq \hat{F}_1(t), \hat{F}_2 \hat{F}_1^{-1}(x) < y \leq 1\}$$

y

$$\text{est. } P(X > t, Y > t) = \frac{1}{n} \sum \mathbf{1}(T_j > t)$$

$$\text{est. } P(X \leq t, X < Y) = \frac{1}{n} \sum \mathbf{1}(T_j \leq t, \delta = 1)$$

las estimaciones empíricas de  $P(X > t, Y > t)$  y  $P(X \leq t, X < Y)$  respectivamente.

$\hat{F}_1(t_i)$  y  $\hat{F}_2(t_i)$  son la soluciones simultáneas de (14) y (15), siendo  $\hat{F}_1$  y  $\hat{F}_2$  líneas rectas en cada intervalo  $(t_i, t_{i+1})$ .

$$\mu_c(\hat{A}_{t_i}) - \text{est. } P(X > t_i, Y > t_i) = 0 \quad (14)$$

$$\mu_c(\hat{B}_{t_i}) - \text{est. } P(X \leq t_i, X < Y) = 0 \quad (15)$$

El estimador cópula-gráfico es de máxima verosimilitud y además bajo condiciones de independencia, coincide con el estimador de Kaplan Meier (Zheng & Klein 1995).

El algoritmo cópula-gráfico para la construcción de  $\widehat{F}_1$  y  $\widehat{F}_2$ , basado en un algoritmo de bisección para encontrar raíces, se puede ver en detalle en Zheng & Klein (1995).

### 3.1. Ilustración

Se estimará  $S(t)$  asumiendo la cópula Gama, la cual se muestra a continuación

$$C(F_1(x), F_2(y)) = F_1(x) + F_2(y) - 1 + \left\{ \left( \frac{1}{1 - F_1(x)} \right)^{\alpha-1} + \left( \frac{1}{1 - F_2(y)} \right)^{\alpha-1} - 1 \right\}^{-\frac{1}{\alpha-1}}, \quad \alpha > 1$$

donde  $F_1(t)$  y  $F_2(t)$  son las funciones de distribución marginales para los tiempos de falla  $X$  y  $Y$ , respectivamente,  $\alpha$  es el parámetro de dependencia de la cópula que mide el grado de asociación entre los tiempos de falla  $X$  y  $Y$ . Este parámetro se relaciona con el coeficiente de correlación  $\tau$  de Kendall de la siguiente forma:  $\tau = \frac{1}{\alpha-1}$ . Para estimar  $F_1(t)$  y  $F_2(t)$ , se aplica el algoritmo cópula-gráfico.

A partir de la ecuación de la cópula, se puede encontrar la función de confiabilidad,

$$S(t) = \left\{ \left( \frac{1}{1 - F_1(t)} \right)^{\alpha-1} + \left( \frac{1}{1 - F_2(t)} \right)^{\alpha-1} - 1 \right\}^{-\frac{1}{\alpha-1}} \quad (16)$$

Remplazando los estimadores  $\widehat{F}_1(t)$  y  $\widehat{F}_2(t)$  obtenidos del algoritmo cópula-gráfico se obtiene  $\widehat{S}(t)_{\text{cópula}}$ . Este último estimador es el que se utilizará en el estudio comparativo de la siguiente sección, puesto que la cópula Gama permite variables con rango  $(0, \infty)$  como es el caso de la Weibull que es de amplio uso en confiabilidad. Además, Zheng & Klein (1995) muestran que este estimador es razonablemente robusto a la escogencia de la cópula.

La programación del algoritmo cópula-gráfico se realizó usando el paquete estadístico R (R Development Core Team 2010). El código utilizado se suministra bajo pedido a los autores.

## 4. Estudio comparativo entre $\widehat{S}^*(t)$ y $\widehat{S}(t)_{\text{cópula}}$

Yáñez et al. (2006) y Manotas et al. (2008) estudiaron, para el caso Weibull y lognormal, el efecto en la estimación de la confiabilidad  $S(t)$ , cuando se asume el supuesto de independencia entre los tiempos de falla que compiten, que realmente son dependientes. Ellos encontraron que el estimador  $\widehat{S}^*(t)$  bajo el supuesto de independencia, que es la metodología tradicional, no genera errores significativos en la estimación de la función de confiabilidad del tiempo mínimo. Puesto que la

literatura reciente reporta que el supuesto de independencia en riesgos competitivos generalmente da una visión pesimista del comportamiento del sistema, ellos comentan la importancia de estudiar en más detalle las propiedades del estimador  $\widehat{S}^*(t)$ .

En este trabajo se estima  $S(t)$  bajo dependencia mediante el estimador cópula-gráfico, denotado por  $\widehat{S}(t)_{\text{cópula}}$  propuesto en Zheng & Klein (1995). Se hace un estudio de simulación para comparar este estimador con el estimador de la confiabilidad  $\widehat{S}^*(t)$ .

#### 4.1. Esquema de simulación

Puesto que el estimador cópula-gráfico estima las funciones de distribución marginales, haremos el estudio comparativo entre las funciones de distribución, lo cual es equivalente a trabajar con las funciones de sobrevivencia.

En el estudio de simulación se utiliza la distribución Weibull bivariada para dos variables  $X$  y  $Y$ , la cual se puede visualizar en varios contextos, tales como los tiempos hasta la primera y segunda falla de un equipo reparable, los tiempos de falla “breakdown” de generadores duales en una planta de energía, o los tiempos de sobrevivencia en un sistema de dos órganos tal como los pulmones o riñones, en el cuerpo humano (Lu & Bhattacharyya 1990).

Para simular los tiempos de falla Weibull bivariados se usó un algoritmo cópula (Frees & Valdez 1998, Jaramillo, Lopera, Manotas & Yáñez 2008). La función de confiabilidad conjunta de la Weibull bivariada utilizada es (Lu & Bhattacharyya 1990)

$$S(x, y) = \exp \left\{ - \left[ \left( \frac{x}{\theta_1} \right)^{\frac{\beta_1}{1-\lambda}} + \left( \frac{y}{\theta_2} \right)^{\frac{\beta_2}{1-\lambda}} \right]^{1-\lambda} \right\} \quad (17)$$

donde  $x > 0$  y  $y > 0$  son los tiempos de falla Weibull con funciones de distribución  $F_1(x)$  y  $F_2(y)$ ,  $\beta_1 > 0$  y  $\beta_2 > 0$  son los parámetros de forma,  $\theta_1 > 0$  y  $\theta_2 > 0$  son los parámetros de escala asociados a  $X$  y  $Y$  respectivamente, y  $0 \leq \lambda < 1$  es el parámetro de dependencia entre  $X$  y  $Y$ . Cuando el parámetro de dependencia  $\lambda$  es 0, entonces hay independencia entre  $X$  y  $Y$ . A medida que  $\lambda$  aumenta, la dependencia entre  $X$  y  $Y$  aumenta.

Se consideran 24 escenarios con tiempos de falla Weibull, cada uno de los cuales corresponde a la llamada función de confiabilidad del tiempo mínimo de falla del sistema  $S(t)$ , donde  $T = \min(X, Y)$ . Los parámetros de escala de la distribución Weibull bivariada se fijan en  $\theta_1 = \theta_2 = 1$  ya que el parámetro de dependencia  $\lambda$  no depende de los parámetros de escala (Lu & Bhattacharyya 1990). Los escenarios presentan situaciones donde compiten distribuciones Weibull con tasas de falla decrecientes y crecientes. Cuando el parámetro  $\beta < 1$ , la tasa de falla es decreciente, y cuando  $\beta > 1$ , la tasa de falla es creciente. Los escenarios toman como base las distintas combinaciones posibles de dicha tasa.

Los casos de tiempos Weibull a estudiar son:

1. Tiempos de falla con parámetros de forma  $\beta_i$  diferentes. Modo de falla con tasa de falla decreciente versus Modo de falla con tasa de falla creciente:  $\beta_1 = 0.5$  y  $\beta_2 = 2$ .
2. Tiempos de falla con parámetros de forma  $\beta_i$  diferentes. Modo de falla con tasa de falla decreciente versus Modo de falla con tasa de falla constante:  $\beta_1 = 0.5$  y  $\beta_2 = 1$ .
3. Tiempos de falla con parámetros de forma  $\beta_i$  iguales. Modo de falla con tasa de falla decreciente versus Modo de falla con tasa de falla decreciente:  $\beta_1 = 0.5$  y  $\beta_2 = 0.5$ .
4. Tiempos de falla con parámetros de forma  $\beta_i$  diferentes. Modo de falla con tasa de falla creciente versus Modo de falla con tasa de falla constante:  $\beta_1 = 2$  y  $\beta_2 = 1$ .

En cada uno de estos casos se toman valores del parámetro de dependencia:  $\lambda = 0.0, 0.2, 0.4, 0.6, 0.8, 0.9$ . De manera que en total se tienen 24 escenarios distintos para la distribución Weibull bivariada.

Para efectos de comparación se obtienen los intervalos de confianza empíricos basados en simulación montecarlo para  $S(t)$ , usando los estimadores  $\widehat{S}^*(t)$  y  $\widehat{S}(t)_{\text{cópula}}$ . El procedimiento se resume como sigue:

- Para cada uno de los escenarios descritos, se generaron 1000 muestras de tamaño  $n = 50$ , de tiempos bivariados Weibull. Se elige este tamaño muestral, ya que en general en experimentos de confiabilidad se manejan bases de datos no muy grandes, estos estudios son costosos y los datos de falla son difíciles de obtener.
- Luego, en cada escenario se obtienen 1000 estimaciones de  $S(t_p)$  usando ambos estimadores  $\widehat{S}^*(t_p)$  y  $\widehat{S}(t_p)_{\text{cópula}}$  en cada percentil  $t_p$  con  $p = 0.05, 0.25, 0.50, 0.75, 0.95$ . De esta forma se obtiene la distribución empírica de las estimaciones de  $S(t_p)$  en cada uno de los percentiles considerados.
- Finalmente, de estas distribuciones se obtienen los límites puntuales de confianza empíricos aproximados del 95 %.

## 5. Análisis de resultados

En aplicaciones de confiabilidad generalmente se tienen pocas fallas y muchos datos censurados. Bajo estas características, las pruebas formales de bondad de ajuste no son útiles (Meeker et al. 2009). En lugar de ello se realizaron pruebas de hipótesis empíricas en cada percentil, utilizando como región de aceptación los intervalos de confianza empíricos, descritos en la sección anterior.



### 5.1. Intervalos de confianza para $S(t)$

Note que en la generación de la Weibull bivariada, el parámetro  $\lambda$  es una medida de dependencia entre los tiempos de falla.

La figura 4 muestra los límites de confianza aproximados para los casos Weibull en que los parámetros de forma son  $\beta_1 = 0.5$ ;  $\beta_2 = 2.0$ , y parámetros de dependencia  $\lambda = 0.6$  y  $0.9$ .

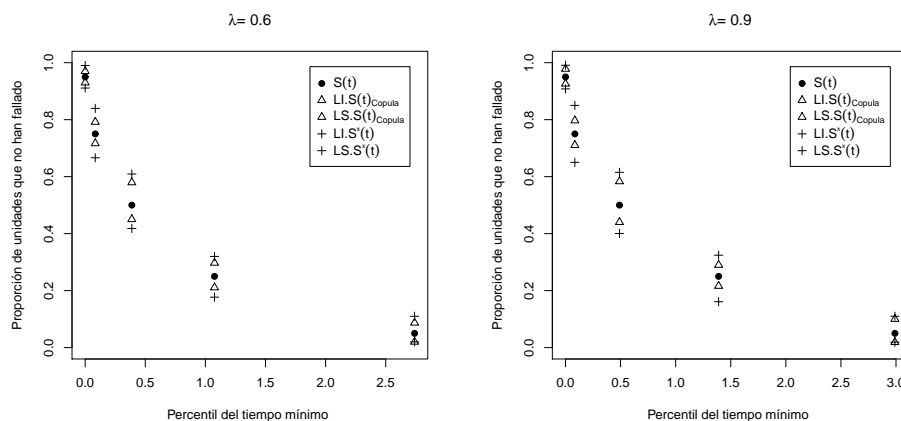


FIGURA 4: Intervalos de confianza empíricos aproximados del 95 % para  $S(t_p)$  en el caso Weibull  $\beta_1 = 0.5$  y  $\beta_2 = 2$  en los percentiles  $t_p$  con  $p = 0.05, 0.25, 0.50, 0.75$  y  $0.95$  para  $\lambda = 0.6$  y  $0.9$ .

En estas figuras, en cada percentil  $t_p$  se tiene el valor verdadero de  $S(t_p)$ , los límites inferior y superior de confianza aproximados para  $S(t_p)$  calculados con  $S^*(t_p)$  (denotados como  $LI.S^*(t_p)$  y  $LS.S^*(t_p)$ , respectivamente), y calculados con  $S(t)_{cópula}$  (denotados  $LI.S(t)_{cópula}$  y  $LS.S(t)_{cópula}$ , respectivamente). Note que las escalas en el eje de las abscisas son distintas.

Observe que los intervalos para  $S(t)$  basados en  $\hat{S}^*(t_p)$  son más amplios que aquellos basados en  $\hat{S}(t_p)_{cópula}$ , indicando que el estimador  $\hat{S}(t)_{cópula}$  tiene mayor precisión a la hora de estimar  $S(t)$ . La amplitud de estos intervalos se hace más evidente a medida que aumenta la dependencia  $\lambda$ . Al ser tan amplios los límites de confianza, la estimación de  $S(t)$  con  $\hat{S}^*(t)$  puede tomar diferentes comportamientos, sobreestimando o subestimando la confiabilidad verdadera.

### 5.2. Estimación marginal

Se compara la función de distribución marginal para el tiempo  $X$ ,  $S_1(t)$ , con las funciones de distribución marginales estimadas,  $\hat{S}_1(t)_{cópula}$  y  $\hat{S}_1(t)$ .

La figura 5 muestra los intervalos de confianza puntuales empíricos de la distribución marginal  $S_1(t)$ , basados en el método de simulación montecarlo, para tiempos de falla bivariados Weibull con parámetros de forma  $\beta_1 = 0.5$ ;  $\beta_2 = 2$  y

dependencia  $\lambda = 0.6$  estimados mediante el algoritmo cópula-gráfico (a la izquierda) y mediante la distribución “marginal” estimada  $S_1(t)$ , que asume independencia entre los tiempos de falla  $X$  y  $Y$  (a la derecha).

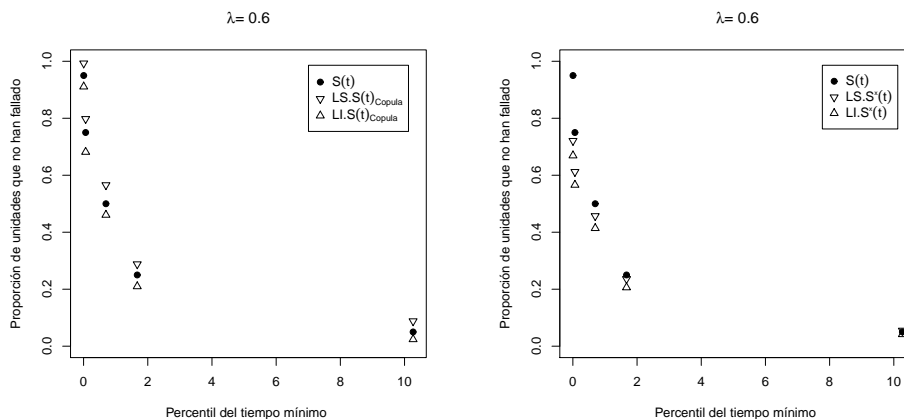


FIGURA 5: Intervalos de confianza empíricos aproximados del 95 % para la distribución marginal  $S_1(t)$  en el caso Weibull  $\beta_1 = 0.5$ ;  $\beta_2 = 2.0$  y  $\lambda = 0.6$  en los percentiles  $t_p$  con  $p = 0.05, 0.25, 0.50, 0.75$  y  $0.95$ .

Se observa como el estimador que emplea el algoritmo cópula-gráfico estima la marginal ya que  $S_1(t_p)$  cae dentro de los límites del intervalo empírico, mientras que el otro estimador realmente no estima marginales puesto que los límites del intervalo construido no contiene la distribución marginal  $S_1(t)$  como era de esperarse, ya que el algoritmo cópula-gráfico resuelve el problema de identificabilidad al asumir una cópula y la dependencia.

### 5.3. Predicción

En problemas de ingeniería donde se estudia la confiabilidad de algún sistema, es de interés evaluar el efecto sobre la sobrevivencia debido a un cambio en una marginal. En esta sección se realizan cambios en uno de los parámetros de forma de la distribución Weibull y el mismo cambio se le hace al estimador de ese parámetro en  $\hat{S}^*(t)$ , y se estima el cambio porcentual con relación a la confiabilidad verdadera en los percentiles  $p = 0.05, 0.25, 0.50, 0.75, 0.95$ .

Meeker et al. (2009) señalan: “Los ingenieros de diseño tienen herramientas que les permiten obtener predicciones razonables acerca de los efectos de un simple cambio geométrico o cambios de tamaño que una pieza de un producto mecánico tendrá sobre la confiabilidad”, y esto se puede traducir en cambios en los parámetros de las marginales.

Se realizó un estudio de simulación para mostrar el efecto de un cambio en el parámetro  $\beta_2$  de la distribución Weibull para predecir la nueva confiabilidad de acuerdo a dicho cambio. El esquema de simulación se diseñó como una réplica

ca de lo realizado por Meeker et al. (2009). La función de confiabilidad conjunta de la Weibull  $S(t)$  depende de los parámetros  $(\beta_1, \beta_2, \lambda)$  (con  $\theta_1 = \theta_2 = 1$ , de acuerdo a lo fijado en la subsección 4.1) y se realiza un cambio en  $\beta_2$  multiplicándolo por 0.3, obteniendo una función de confiabilidad nueva que denotamos por  $S(t)_{\text{nuevo}} = S(t; \beta_1, 0.3\beta_2, \lambda)$ . Ahora bien, el estimador bajo independencia nuevo (i.e. con el cambio en  $0.3\beta_2$ ) se obtiene de  $\widehat{S}^*(t)_{\text{nuevo}} = \widehat{S}_1(t)\widehat{S}_2(t)_{\text{nuevo}}(t; \widehat{\theta}_2, 0.3\widehat{\beta}_2)$ , donde  $\widehat{S}_2(t)_{\text{nuevo}}$  se obtiene al reemplazar el estimador  $\widehat{\beta}_2$  por  $0.3\widehat{\beta}_2$ , donde  $\widehat{\beta}_2$  es obtenido de los datos de la distribución marginal 2 y el estimador  $\widehat{\theta}_2$  se deja igual; como se describió en la subsección 2.2.1. Para comparar se obtuvieron los cambios porcentuales ( $CP_p$ ) en los percentiles  $t_p$ , así

$$CP_p = \frac{|S(t_p) - \widehat{S}^*(t_p)|}{S(t_p)} \times 100 \tag{18}$$

que corresponde al cambio relativo entre el valor verdadero  $S(t)$  y el valor estimado  $\widehat{S}^*(t)$  antes de hacer cambios en el parámetro  $\beta_2$ , lo que se interpreta como el sesgo relativo del estimador  $\widehat{S}^*(t)$ . El cambio porcentual nuevo  $CP_p(\text{nuevo})$  en los percentiles  $t_p$  se obtiene, así

$$CP_p(\text{nuevo}) = \frac{|S(t_p)_{\text{nuevo}} - \widehat{S}^*(t_p)_{\text{nuevo}}|}{S(t_p)_{\text{nuevo}}} \times 100. \tag{19}$$

En la tabla 2 se muestran los cambios porcentuales  $CP_p$  y  $CP_p(\text{nuevo})$ .

TABLA 2: Cambios porcentuales  $CP_p$  y  $CP_p(\text{nuevo})$  para  $S(t)$  y  $S(t)_{\text{nuevo}}$  en tiempos de falla Weibull con parámetros  $\beta_1 = 0.5$ ;  $\beta_2 = 2$  y  $\theta_1 = \theta_2 = 1$ .

$\lambda$	$p = 0.05$		$p = 0.25$		$p = 0.5$	
	$CP_p(\text{nuevo})$	$CP_p$	$CP_p(\text{nuevo})$	$CP_p$	$CP_p(\text{nuevo})$	$CP_p$
0.0	1.14279	1.2891	1.12186	1.1097	1.0822	1.0714
0.2	71.8997	49.3511	16.3558	15.2779	7.4906	3.7173
0.4	147.1947	79.6577	34.1554	31.4991	15.9938	5.8585
0.6	226.0480	97.6448	51.7345	49.1091	20.1146	7.5282
0.8	348.2726	99.9365	78.7846	74.7928	22.5856	11.5464
0.9	379.5557	99.9736	83.7510	81.8147	31.7949	12.1519
0.98	389.0409	99.9731	93.4373	82.6682	35.2286	15.7612

$\lambda$	$p = 0.75$		$p = 0.95$	
	$CP_p(\text{nuevo})$	$CP_p$	$CP_p(\text{nuevo})$	$CP_p$
0.0	0.4668	0.3255	1.0234	0.9419
0.2	0.9193	0.8953	2.3615	2.1102
0.4	1.3844	1.2676	5.3259	2.8302
0.6	1.8549	1.5949	3.6267	3.5698
0.8	1.9075	1.8547	6.5551	3.7108
0.9	3.0218	2.9878	8.3761	4.0974
0.98	3.8366	3.7381	11.8975	6.0977

Analizando por columna para cada percentil, se observa que a medida que aumenta la dependencia  $\lambda$  entre los tiempos de falla Weibull, el cambio porcentual

$CP_p$  es siempre menor que el cambio porcentual  $CP_p(\text{nuevo})$ . En el percentil  $t_{0.05}$  la diferencia entre los dos llega a ser hasta de casi cuatro veces. De lo cual se puede decir que en la práctica el estimador  $\widehat{S}^*(t)_{\text{nuevo}}$ , que es el estimador de la confiabilidad bajo independencia, no logra captar correctamente el cambio en el parámetro  $\beta_2$ , luego para efectos de predicción en este contexto es muy mal estimador. En esta situación lo que ocurre es que  $\widehat{S}_2(t)_{\text{nuevo}}$  se calcula cambiando  $\widehat{\beta}_2$  obtenido de los datos simulados para la marginal del modo de falla 2, por  $0.3\widehat{\beta}_2$ , pues al no tener datos nuevos, se tienen una estimación inapropiada de  $S_2(t)_{\text{nuevo}}$ . Por ello afirmamos que  $\widehat{S}^*(t)_{\text{nuevo}}$  no es útil para predecir cuando hay cambios en los parámetros.

Obsérvese además, que el estimador  $\widehat{S}(t)_{\text{cópula}}$  no sirve para fines de predicción, ya que este estimador es no paramétrico y al hacer cambios en los parámetros de la marginales no hay forma de incluir dicho cambio en el estimador. Pero se puede pensar que con las estimaciones de las distribuciones marginales obtenidas con el estimador cópula-gráfico se puede hacer un modelamiento paramétrico, como paso previo al proceso de predicción de la confiabilidad bajo cambios en los parámetros, lo cual mejoraría el análisis exploratorio, ya que se utilizan estimaciones adecuadas de las distribuciones marginales en lugar de las estimaciones basadas en asumir independencia. Estos ajustes paramétricos permitirían hacer predicciones en experimentos acelerados como los descritos en Meeker et al. (2009).

### 5.4. Ejemplo utilizando los datos de la subsección 2.2.

En la figura 6, las distribuciones marginales bajo dependencia son diferentes a las obtenidas bajo independencia.

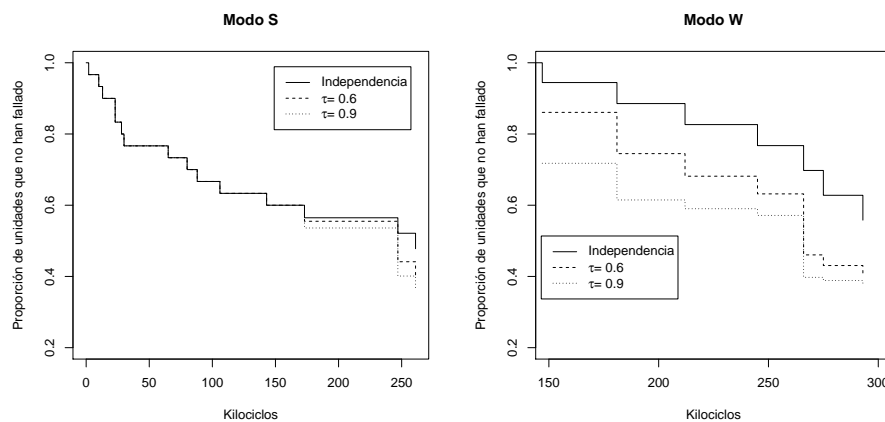


FIGURA 6: Estimación de las funciones de distribución marginales con independencia y con  $\tau = 0.6$  y  $0.9$ .

De ahí, como se muestra en la figura 7 las funciones de distribución acumuladas del sistema (tiempo mínimo) con independencia generan una sobrestimación de las

funciones de distribución, lo cual tiene graves implicaciones en la determinación, por ejemplo, del esquema de mantenimiento preventivo de los equipos, con las consecuencias económicas que tales equivocaciones generarían.

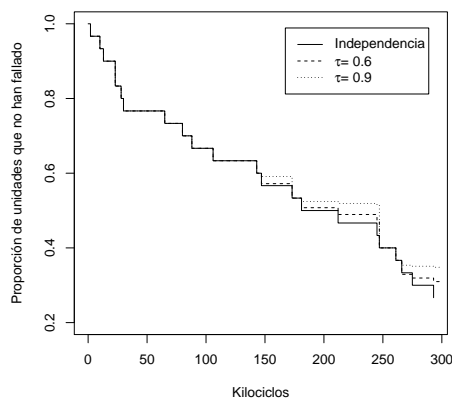


FIGURA 7: Estimación de la función de distribución del tiempo mínimo considerando independencia y con  $\tau = 0.6$  y  $0.9$ .

## 6. Conclusiones y trabajo futuro

El estimador  $\hat{S}(t)_{\text{cópula}}$ , es mucho mejor estimador que  $\hat{S}^*(t)$ . Por ello recomendamos para trabajos de confiabilidad en ingeniería, utilizar el estimador cópula-gráfico, cuando no se tenga certeza sobre el supuesto de independencia, ya que tal supuesto puede generar sobrecostos y mala planeación en diseño, por ejemplo, en mantenimiento de equipos.

Para efectos de predicción en problemas de ingeniería, que implican cambios en los parámetros de las marginales para mejorar la confiabilidad conjunta, ninguno de los dos estimadores funciona. Este sería un tópico de importancia para trabajo futuro tratando de adecuar el estimador cópula-gráfico para tales efectos, y comparar con resultados como los reportados por Meeker et al. (2009). Se piensa que con las distribuciones marginales estimadas se pueden hacer ajustes paramétricos que permitan incluir la dependencia en la estimación de la predicción de la vida útil de un equipo sujeto a cambios en su diseño original, en el contexto de experimentos acelerados.

También, de acuerdo a lo sugerido por uno de los árbitros, se planteará un estudio de simulación haciendo uso de la distribución lognormal bivariada generada por cópulas, siguiendo los lineamientos de Johnson (1987) y Nelsen (2006).

## Agradecimientos

Los autores agradecen a la Dirección de Investigaciones de la Universidad Nacional de Colombia, sede Medellín (DIME) por la financiación. Además, agradecen a los árbitros y al editor por sus valiosos comentarios que enriquecieron el texto del manuscrito.

Este trabajo es parte del proyecto “Riesgos competitivos en la modelación de la confiabilidad de un sistema con dos modos de falla”, código Quipu: 20101007773.

[Recibido: agosto de 2010 — Aceptado: febrero de 2011]

## Referencias

- Barlow, R. & Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, Inc, New York.
- Basu, A. & Klein, P. (1982), ‘Some recent results in competing risks theory’, *Proceedings on Survival Analysis. IMS monograph series* **2**, 216–229.
- Bedford, T. (2005), Competing risk modelling in reliability, in ‘Modern Statistical and Mathematical Methods in Reliability’, World Scientific Books, United Kingdom.
- Bedford, T. & Lindqvist, B. (2004), ‘The identifiability problem for repairable systems subject to competing risks’, *Advances in Applied Probability* **36**(3), 774–790.
- Crowder, M. (2001), *Classical Competing Risks*, first edn, Chapman & Hall, London.
- Frees, E. & Valdez, E. (1998), ‘Understanding relationships using copulas’, *North American Actuarial Journal* **2**(1), 1–25.
- Jaramillo, M. C., Lopera, C. M., Manotas, E. C. & Yáñez, S. (2008), ‘Generación de Tiempos de Falla Dependientes Weibull Bivariados Usando Cópulas’, *Revista Colombiana de Estadística* **31**(2), 169–181.
- Johnson, M. (1987), *Multivariate Statistical Simulation*, John Wiley & Sons, New York.
- Lu, J. & Bhattacharyya, G. (1990), ‘Some new constructions of bivariate weibull models’, *Annals of the Institute of Statistical Mathematics* **42**(3), 543–559.
- Manotas, E. C., Yáñez, S., Lopera, C. M. & Jaramillo, M. C. (2008), ‘Estudio del Efecto de la Dependencia en la Estimación de la Confiabilidad de un Sistema con Dos Modos de Falla Concurrentes’, *DYNA* **75**(154), 5–21.
- Meeker, W. & Escobar, L. (1998), *Statistical Methods for Reliability Data*, John Wiley and Sons, Inc, New York.

- Meeker, W., Escobar, L. A. & Hong, Y. (2009), 'Using accelerated life tests results to predict product field reliability', *Technometrics* **51**(2), 146–161.
- Nelsen, R. B. (2006), *An Introduction to Copulas*, second edn, Springer, New York.
- Nelson, W. (1990), *Accelerated Testing: Statistical Models Test Plans and Data Analyses*, John Wiley & Sons, New York.
- Pintilie, M. (2006), *Competing Risks: A Practical Perspective*, John Wiley & Sons, United Kingdom.
- R Development Core Team (2010), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.  
\*<http://www.R-project.org/>, accedido en 01/08/2010
- Tsiatis, A. (1975), 'A nonidentifiability aspect of the problem of competing risks', *Proceedings of the National Academy of Sciences* **72**, 20–22.
- Yáñez, S., Manotas, E. C. & Lopera, C. M. (2006), 'Estudio de la Dependencia en la Estimación de la Confiabilidad de un Sistema con Dos Modos de Falla', *Memorias del XV Simposio de Estadística*.
- Zheng, M. & Klein, J. P. (1995), 'Estimates of marginal survival for dependent competing risks based on an assumed copula', *Biometrika* **82**(1), 12–38.

## Aplicación del análisis de datos recurrentes sobre interruptores FL245 en Interconexión Eléctrica S.A.

### Application of the Recurrent Data Analysis on FL245 Switches at Interconexión Eléctrica S.A.

CARLOS M. LOPERA<sup>1,a</sup>, EVA CRISTINA MANOTAS<sup>2,b</sup>

<sup>1</sup>ESCUELA DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE COLOMBIA, MEDELLÍN, COLOMBIA

<sup>2</sup>ESCUELA DE INGENIERÍA DE LA ORGANIZACIÓN, FACULTAD DE MINAS, UNIVERSIDAD NACIONAL DE COLOMBIA, MEDELLÍN, COLOMBIA

---

#### Resumen

Los datos recurrentes surgen cuando una unidad (o un grupo de tales unidades) es monitoreada a través del tiempo y un evento particular (o grupo de eventos) ocurre en varios puntos del periodo de observación, por ejemplo, los tiempos de episodios recurrentes de una enfermedad en pacientes o los tiempos de reparación de un producto manufacturado. Muchos sistemas, subsistemas y componentes (que genéricamente son denominadas “unidades”) tienen asociadas más de una causa o modo de falla. En algunas aplicaciones, y para ciertos propósitos, es importante distinguir entre las causas o modos de falla. Para mejorar la confiabilidad, es esencial identificar la causa de falla hasta el nivel de componente, y en muchas aplicaciones, hasta la causa física real de una falla. En este trabajo, se presenta una aplicación del análisis de datos recurrentes realizado sobre interruptores tipo FL245 (unidades reparables en Interconexión Eléctrica S.A., ISA), que incluye el uso de métodos estadísticos no paramétricos y paramétricos considerando varios modos de falla.

**Palabras clave:** confiabilidad, cópula, identificabilidad, riesgos competitivos.

#### Abstract

Recurrent data arise when a unit (or group of such units) is monitored over time and a particular event (or group of events) occurs at several points

---

<sup>a</sup>Profesor Asistente. E-mail: cmlopera@unal.edu.co

<sup>b</sup>Profesor Asistente. E-mail: ecmanota@unal.edu.co



of the observation period, for example, times of recurrent episodes of a disease in patients or times of repair of a manufactured product. Many systems, subsystems, and components (which we generically refer to as “units”) have more than one cause of failure. In some applications and for some purposes it is important to distinguish between the different failure causes (sometimes referred to as “failure modes”). For purposes of improving reliability, it is essential to identify the cause of failure up to the component level and, in many applications, up to the actual physical cause of failure. This paper presents an application of the recurrent data analysis performed on type FL245 switches (repairable units of Interconexión Eléctrica S.A., ISA), that includes the use of non-parametric and parametric statistical methods, considering several failure modes.

**Key words:** Competing risks, Copula, Identifiability, Reliability.

## 1. Introducción

Los datos de tiempos de eventos son importantes en muchos campos de aplicación. Algunos ejemplos incluyen tiempos de falla de productos no reparables, el tiempo de supervivencia de pacientes médicos, tiempo en que las cuentas de un banco se mantienen abiertas, tiempo en que los suscriptores pagan por un servicio de cable, etc. La metodología para estos datos se denomina análisis de datos de confiabilidad, análisis de datos de vida, análisis de supervivencia, o análisis de tiempos a eventos. La literatura en su mayoría se dedica a presentar modelos y métodos de análisis de datos para las unidades simples que sólo experimentan un evento, muerte o falla (Nelson 2003).

Los datos recurrentes surgen en diversas áreas del conocimiento, particularmente surgen cuando un sujeto o unidad (o grupo de ellos) es monitoreado a través del tiempo y un evento particular (o grupo de eventos) ocurre en varios puntos del período de observación (Escobar, Villa & Yañez 2003). Algunas de las aplicaciones de datos recurrentes comprenden ensayos clínicos, estudios de confiabilidad e investigaciones sociológicas (Andersen, Borgan, Gill & Keiding (1993); Cook, Lawless & Nadeau (1996); Lin, Wei, Yang & Ying (2000); Wang, Qin & Chiang (2001)). Ejemplos en confiabilidad de productos incluyen: aplicaciones en la industria automotriz (Kalbfleisch, Lawless & Robinson (1991); Nelson (1998), Nelson (2000); Hu, Lawless & Suzuki (1998); Lawless (1995a); Lawless (1995b)); en informática/electrónica (Tobias & Trindade (1995); Trindade & Haugh (1980); Vallarino (1988); Nelson (1988)); en energía eléctrica (Nelson (1990); Ross (1989); Kvam, Singh & Whitaker (2002)), entre otros.

Interconexión Eléctrica S.A. (ISA) es una empresa colombiana, cuya actividad principal es el transporte de energía eléctrica. Los interruptores de potencia son equipos muy importantes para garantizar el cumplimiento de tal actividad. El funcionamiento inadecuado de estos equipos podría llevar al sistema eléctrico a estar fuera de servicio, lo que genera graves sanciones económicas para las empresas de este sector. Por tanto, es muy importante estudiar la confiabilidad de los interruptores, a fin de apoyar la toma de decisiones en relación con programas

de mantenimiento preventivo. Este artículo describe una aplicación del análisis de datos recurrentes, no paramétrico y paramétrico, realizado sobre los interruptores FL245 en ISA (Manotas & Lopera 2004). El análisis considera varios modos de falla de los interruptores y se realizó con el módulo SPLIDA del software S+ (Meeker & Escobar 2003).

La sección 2 describe el análisis no paramétrico para este tipo de eventos. El análisis paramétrico es presentado en la sección 3. La sección 4 introduce los datos de interruptores FL245 (equipos reparables en Interconexión Eléctrica S.A.) y presenta los resultados de la aplicación del análisis de datos recurrentes a estos datos. Finalmente, algunas conclusiones y recomendaciones acerca del trabajo son presentadas en la sección 5.

## 2. Análisis no paramétrico de datos recurrentes

En esta sección se describen algunos métodos gráficos no paramétricos, los cuales son útiles para cuantificar empíricamente la confiabilidad total del sistema.

### 2.1. Función promedio acumulada

Es una función que representa el número esperado de recurrencias acumuladas hasta un tiempo  $t$ . Usualmente esta función se denota por  $M(t)$  y es dada por:

$$M(t) = E[N(t)] = E[N(0, t)] \quad (1)$$

donde,  $N(0, t)$  representa el número acumulado de fallas de un equipo durante el periodo  $(0, t]$ .

#### 2.1.1. Tasa de recurrencia

Se denota  $m(t)$  y se define como la derivada de la función promedio acumulada (MCF, por sus siglas en inglés) poblacional respecto al tiempo, por ello se conoce como la función de intensidad. La expresión para la tasa de recurrencia es:

$$m(t) \equiv \frac{dM(t)}{dt} \quad (2)$$

donde  $M(t)$  es la MCF poblacional. Así la tasa de recurrencia representa el número promedio de recurrencias por unidad de tiempo por unidad poblacional. Para datos discretos se asume que tal derivada existe. La figura 1 ilustra tres tipos comunes de comportamientos de la tasa de recurrencia en la MCF poblacional, que pueden observarse de forma aproximada en la MCF estimada.

De la figura 1 se observa que de acuerdo con la MCF poblacional, la tasa de recurrencia en el tiempo  $t$  (que es precisamente la tangente a la curva en ese preciso instante del tiempo) puede comportarse de tres formas típicas: a) La tasa de recurrencia disminuye conforme aumenta el tiempo, b) La tasa de recurrencia es constante en el tiempo, y c) La tasa de recurrencia aumenta con el tiempo.

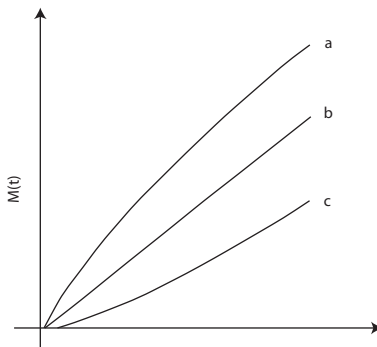


FIGURA 1: Caracterización de la tasa de recurrencia.

El análisis anterior permite definir bajo ciertos supuestos, que se discutirán más adelante en la sección 3.1.2., si el análisis paramétrico se puede hacer utilizando un proceso Poisson homogéneo (HPP) o un proceso Poisson no homogéneo (NHPP). En el sentido que si la tasa de recurrencia es constante, se aplicaría un HPP, pero si ésta no lo es, entonces el análisis adecuado es un NHPP.

## 2.2. Estimación puntual e intervalos de confianza para la MCF

Suponga que se ha observado una colección de  $n \geq 1$  equipos que conforman una muestra de una población mayor de equipos. Entonces el método que se describe a continuación se puede usar para estimar la MCF poblacional.

Denote a  $N_i(t)$  el número acumulado de recurrencias por equipo antes del tiempo  $t$ , y sea  $t_{ij}, j = 1, \dots, r_i$ , los tiempos de recurrencias para el equipo  $i$ . Nelson (1988) proporcionó un estimador insesgado, que permite diferentes longitudes de los periodos de observación entre los equipos. La estimación de Nelson para la MCF poblacional puede ser calculada mediante el siguiente algoritmo (Meeker & Escobar 1998):

1. Ordene los tiempos de recurrencia únicos  $t_{ij}$  entre todos los  $n$  equipos. Denote  $r$  el número de tiempos únicos. Estos tiempos únicos ordenados son denotados por  $t_1 < \dots < t_r$ .
2. Calcule  $d_i(t_k)$ , el número total de recurrencias para el equipo  $i$  en el tiempo  $t_k$ .
3. Sea  $\delta_i(t_k) = 1$  si el equipo  $i$  está aún siendo observado en el tiempo  $t_k$  y  $\delta_i(t_k) = 0$  en otro caso.
4. Calcule

$$\widehat{M}(t_j) = \sum_{k=1}^j \left[ \frac{\sum_{i=1}^n \delta_i(t_k) d_i(t_k)}{\sum_{i=1}^n \delta_i(t_k)} \right] = \sum_{k=1}^j \frac{d_{\cdot}(t_k)}{\delta_{\cdot}(t_k)} = \sum_{k=1}^j \bar{d}(t_k) \quad (3)$$

para  $j = 1, \dots, r$ , donde  $d.(t_k) = \sum_{i=1}^n \delta_i(t_k)d_i(t_k)$ ,  $\delta.(t_k) = \sum_{i=1}^n \delta_i(t_k)$  y  $\bar{d}(t_k) = d.(t_k)/\delta.(t_k)$ .

Observe que  $d.(t_k)$  es el número total de recurrencias de equipos en el tiempo  $t_k$ ,  $\delta.(t_k)$  es el tamaño del conjunto a riesgo en el tiempo  $t_k$ , y  $\bar{d}(t_k)$  es el número promedio de recurrencias por equipo en el tiempo  $t_k$ , de manera que el estimador de la MCF es el número promedio acumulado de recurrencias por equipo en cada intervalo de tiempo.

Cuando se consideran tiempos de censuras en el periodo de observación de cada equipo, se puede utilizar una versión del algoritmo anterior donde los tiempos de censura modifican el conjunto de equipos a riesgo. Para detalles del algoritmo modificado ver Manotas & Lopera (2004).

Existen varios métodos para la obtención de los intervalos de confianza para la estimación no paramétrica de la MCF, entre ellos están: los intervalos de confianza Naive, descritos en Nelson (2003); los intervalos de confianza Poisson, descritos por Manotas & Lopera (2004); así como los propuestos por Nelson (1995) y por Lawless & Nadeau (1995). La elección del método empleado está sujeta a la estructura de los datos, la facilidad de implementación del método para la obtención de los límites de confianza, el calcular siempre límites de confianza positivos, etc. Para más información, ver Nelson (2003) y Meeker & Escobar (1998).

A continuación se describe la construcción de los intervalos de confianza propuestos por Lawless & Nadeau (1995).

Éste método considera que una estimación de la varianza real (basada en estimadores de momentos) de  $\widehat{M}(t_j)$  es:

$$\widehat{V}[\widehat{M}(t_j)] = \sum_{i=1}^n \left\{ \sum_{k=1}^j \frac{\delta_i(t_k)}{\delta.(t_k)} [d_i(t_k) - \bar{d}(t_k)] \right\}^2 \tag{4}$$

Esta expresión para la varianza estimada del estimador de la MCF siempre resulta en valores mayores o iguales que cero.

Luego, intervalos puntuales del  $(1 - \alpha)100\%$  de confianza usando la aproximación normal basados en que  $Z_{\widehat{M}(t)} = [\widehat{M}(t) - M(t)]/\widehat{se}_{\widehat{M}(t)} \sim N(0, 1)$  es:

$$\widehat{M}(t) \pm z_{(1-\alpha/2)} \widehat{se}_{\widehat{M}(t)} \tag{5}$$

donde  $\widehat{se}_{\widehat{M}(t)} = \sqrt{\widehat{V}[\widehat{M}(t)]}$ .

Un intervalo alternativo basado en  $Z_{\log[\widehat{M}(t)]} \sim N(0, 1)$  es:

$$[\widehat{M}(t)/w, \widehat{M}(t) \times w] \tag{6}$$

donde  $w = \exp[z_{(1-\alpha/2)} \widehat{se}_{\widehat{M}(t)}/\widehat{M}(t)]$ . Los intervalos construidos usando esta última expresión siempre tendrán límites positivos y en algunos procesos positivos, se pueden esperar probabilidades de cobertura más cercanas al nivel de confianza nominal.

### 2.3. Estrategia de análisis

El análisis de los datos recurrentes se puede realizar considerando que todas las fallas son del mismo tipo (análisis combinado). En este caso, el análisis no paramétrico incluye la estimación no paramétrica de la MCF, su gráfico y la estimación de los intervalos de confianza no paramétricos para la MCF, como fue descrito en la subsección 2.2.

El análisis gráfico de la MCF estimada y de los intervalos de confianza respectivos da pie para observar el comportamiento de la tasa de recurrencia para las unidades bajo estudio, lo que puede señalar el camino hacia un modelamiento paramétrico de la tasa de recurrencia, que permita establecer predicciones del número de recurrencia en periodos de tiempo futuros.

Sin embargo, el análisis combinado puede llevar a una importante pérdida de información, en especial cuando en realidad se está en presencia de diferentes causas o modos de falla. Esto requiere una definición operacional clara de cada modo de falla. La elección de las causas o modos de falla depende de la aplicación, y diferentes elecciones pueden ser usadas para diferentes propósitos con la misma base de datos. El análisis no paramétrico considerando varios modos de falla es el mismo que el análisis combinado, solo que se hace por separado para cada uno de los modos de falla considerados.

## 3. Análisis paramétrico de datos recurrentes

Los modelos más utilizados para datos de recurrencia son los procesos Poisson (homogéneos y no homogéneos) y los procesos de renovación.

### 3.1. Procesos Poisson

El proceso Poisson es un modelo paramétrico simple que es comúnmente utilizado para el análisis de ciertas clases de datos de recurrencia. Un proceso Poisson se presenta cuando se satisfacen las siguientes condiciones:

- El número de recurrencias acumuladas en el tiempo cero es cero (notado como  $N(0) = 0$ ).
- El número de recurrencias en intervalos de tiempos disjuntos, son estadísticamente independientes. Un proceso con esta propiedad tiene incrementos independientes.
- La tasa de recurrencia  $m(t)$  es positiva y

$$M(a, b) = E[N(a, b)] = \int_a^b m(t)dt < \infty, \text{ cuando } 0 \leq a < b < \infty$$

por lo tanto para un proceso Poisson,  $N(a, b)$ : número de recurrencias en el intervalo  $(a, b]$ , tiene distribución Poisson con distribución de probabilidad:

$$\Pr [N(a, b) = d] = \frac{[M(a, b)]^d}{d!} \exp [-M(a, b)], d = 0, 1, 2, \dots$$

### 3.1.1. Proceso Poisson homogéneo

Un proceso Poisson homogéneo (HPP, por sus siglas en inglés) es un proceso Poisson con tasa de recurrencia constante, es decir,  $m(t) = 1/\theta$ . En este caso:  $N(a, b)$  tiene distribución Poisson con  $M(a, b) = (b - a)/\theta$ .

El número esperado de recurrencias en  $(a, b]$  es  $M(a, b)$ . Equivalentemente, el número esperado de recurrencias por unidad de tiempo en  $(a, b]$  es constante e igual a  $1/\theta$ . Esta propiedad se conoce como incrementos estacionarios.

Los tiempos entre recurrencias,  $t_j = T_j - T_{j-1}$ , son independientes e idénticamente distribuidos (iid), con distribución exponencial.

### 3.1.2. Proceso Poisson no homogéneo

Un proceso Poisson no homogéneo (NHPP, por sus siglas en inglés) es un modelo Poisson con tasa de recurrencia no constante  $m(t)$ . En este caso los tiempos entre recurrencias no son independientes ni idénticamente distribuidos. El número de recurrencias por unidad de tiempo sobre el intervalo  $(a, b]$  es:

$$\frac{M(a, b)}{b - a} = \frac{1}{b - a} \int_a^b m(t) dt$$

Un modelo NHPP es frecuentemente especificado en términos de la tasa de recurrencia  $m(t)$ . Para especificar un modelo NHPP se usa  $m(t) = m(t, \theta)$ , como una función de un vector de parámetros desconocidos  $\theta$ .

Por ejemplo, el modelo de tasa de recurrencia potencial es

$$m(t; \beta, \eta) = \frac{\beta}{\eta} \left( \frac{t}{\eta} \right)^{\beta-1}, \beta > 0, \eta > 0$$

La correspondiente MCF sobre el intervalo  $(0, t]$  es  $M(t; \beta, \eta) = (t/\eta)^\beta$ . Cuando  $\beta = 1$ , este modelo se reduce al modelo HPP.

El modelo log-lineal para la tasa de recurrencia es,

$$m(t; \gamma_0, \gamma_1) = \exp(\gamma_0 + \gamma_1 t)$$

La correspondiente expresión para la MCF poblacional en el intervalo  $(0, t]$  es  $M(t; \gamma_0, \gamma_1) = [\exp(\gamma_0)] [\exp(\gamma_1 t) - 1] / \gamma_1$ . Cuando  $\gamma_1 = 0$ ,  $m(t; \gamma_0, 0) = \exp(\gamma_0)$ , el cual es un HPP.

El método usual de estimación para los procesos Poisson no homogéneos es el de máxima verosimilitud, que está implementado en la mayoría de paquetes estadísticos, en particular en el módulo SPLIDA del paquete S+. Detalles acerca del proceso de estimación pueden ser consultados en Meeker & Escobar (1998).

### 3.2. Procesos de renovación

Una secuencia de recurrencias de un sistema y edades del sistema  $T_1, T_2, \dots$  es un proceso de renovación si los tiempos entre recurrencias  $t_j = T_j - T_{j-1}$ , para  $j = 1, 2, \dots$  (con  $T_0 = 0$ ) son iid. La MCF para un proceso de renovación es también conocida como una función de renovación. Observe que un HPP es un proceso de renovación (y los tiempos entre recurrencias tienen distribución exponencial) pero el NHPP no es un proceso de renovación.

Antes de usar un modelo de renovación, es importante chequear los supuestos del modelo tales como la tendencia y la no independencia de los tiempos entre recurrencias. Si las pruebas de tendencia y no independencia sugieren un proceso de renovación, entonces uno podría usar métodos para describir la distribución de los tiempos entre recurrencias. Detalles acerca de la evaluación de supuestos para determinar si hay un proceso de renovación se pueden encontrar en Manotas & Lopera (2004).

## 4. Aplicación en interruptores FL245 en ISA

### 4.1. Datos

Para presentar una aplicación del análisis de datos recurrentes se consideran datos de la confiabilidad de interruptores tipo FL245 en Interconexión Eléctrica S.A. (ISA), la cual es una empresa colombiana, cuya actividad principal es el transporte de energía eléctrica. Los interruptores de potencia (ilustrados en la figura 2) son usados para interrumpir el flujo de corriente y desconectar algún elemento de la subestación, para lo cual se pueden interrumpir tanto corrientes de carga normales como debidas a fallas eléctricas.



FIGURA 2: Interruptores de potencia tipo FL245.

Los datos consisten de tiempos transcurridos desde el tiempo de puesta en servicio hasta eventos de mantenimiento correctivos en este tipo de interruptores recolectados entre los años 1985 y 2003. Se hizo un seguimiento a un total de 40 interruptores (Manotas & Lopera 2004). La estructura de los datos se muestra para un equipo en la tabla 1.

TABLA 1: Estructura de los datos de recurrencias para uno de los interruptores FL245

EQUIPO	TIEMPO	ESTATUS	FALLA	EQUIPO	TIEMPO	ESTATUS	FALLA
<i>i</i>	0	Start		<i>i</i>	1913	Fail	2
<i>i</i>	141	Fail	1	<i>i</i>	1970	Fail	3
<i>i</i>	709	Fail	3	<i>i</i>	1983	Fail	2
<i>i</i>	716	Fail	2	<i>i</i>	2017	Fail	1
<i>i</i>	743	Fail	2	<i>i</i>	2017	Fail	2
<i>i</i>	820	Fail	2	<i>i</i>	2059	Fail	2
<i>i</i>	1115	Fail	2	<i>i</i>	2627	Fail	1
<i>i</i>	1251	Fail	2	<i>i</i>	2627	Fail	3
<i>i</i>	1282	Fail	2	<i>i</i>	3200	Fail	1
<i>i</i>	1334	Fail	2	<i>i</i>	3211	Fail	1
<i>i</i>	1384	Fail	2	<i>i</i>	3246	Fail	1
<i>i</i>	1433	Fail	2	<i>i</i>	3314	Fail	1
<i>i</i>	1480	Fail	2	<i>i</i>	3519	Fail	1
<i>i</i>	1549	Fail	2	<i>i</i>	3519	Fail	3
<i>i</i>	1605	Fail	2	<i>i</i>	3594	Fail	1
<i>i</i>	1654	Fail	2	<i>i</i>	3601	Fail	1
<i>i</i>	1706	Fail	3	<i>i</i>	3708	Fail	1
<i>i</i>	1758	Fail	2	<i>i</i>	3708	Fail	3
<i>i</i>	1810	Fail	3	<i>i</i>	4383	Fail	1
<i>i</i>	1814	Fail	2	<i>i</i>	5369	Fail	2
<i>i</i>	1863	Fail	2	<i>i</i>	5470	End	

Los eventos de mantenimiento obedecen a fallas cuyas causas se clasifican principalmente en fugas de aceite en sistema hidráulico, fugas de gas SF6 y deterioro en los instrumentos de calibración, los cuales fueron codificados como: 1, baja de presión de aceite; 2, baja de presión de gas SF6; y 3, deterioro en los instrumentos de calibración.

A continuación se realiza un análisis de datos recurrentes para los eventos de mantenimiento correctivos con base en los métodos descritos en secciones 2 y 3.

#### 4.2. Análisis no paramétrico para los datos de interruptores

En primer lugar se realizará un análisis no paramétrico ignorando los modos de falla, esto es, todas las fallas se asumen de un mismo tipo. La figura 3 muestra la MCF y los intervalos de confianza respectivos, estimados como fue descrito en la subsección 2.2, usando el módulo SPLIDA del paquete S+.

Para el caso de los interruptores FL245, se tiene una tasa de recurrencia que aumenta con el tiempo, de manera que parece que el análisis paramétrico adecuado es el NHPP. A continuación se presentan los análisis considerando varios modos de falla para los interruptores FL245.



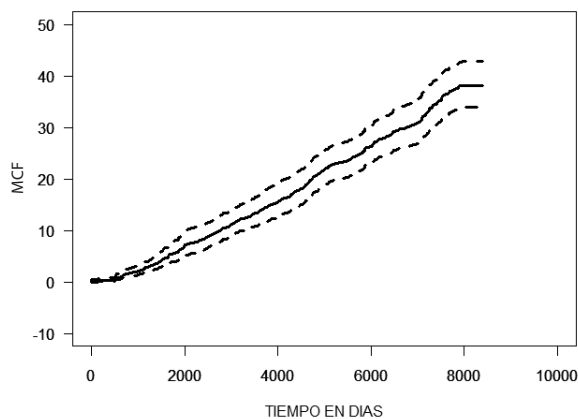


FIGURA 3: MCF estimada en SPLIDA, ignorando los modos de falla, para los interruptores FL245 con intervalos de confianza del 95 %.

#### 4.2.1. Análisis no paramétrico para varios modos de falla

En el estudio que involucra a los interruptores FL245 se ha decidido realizar la partición de eventos de acuerdo con el modo de falla y se han escogido tres modos de falla: baja presión de aceite (1), baja presión de SF6 (2) y deterioro en los instrumentos de calibración (3).

La metodología aplicada para obtener la MCF estimada para cada modo de falla es la misma que se utilizó en los datos de recurrencia con todos los modos de falla. A continuación se muestran los resultados obtenidos del análisis de los Interruptores FL245 por cada modo de falla.

##### Baja presión de aceite

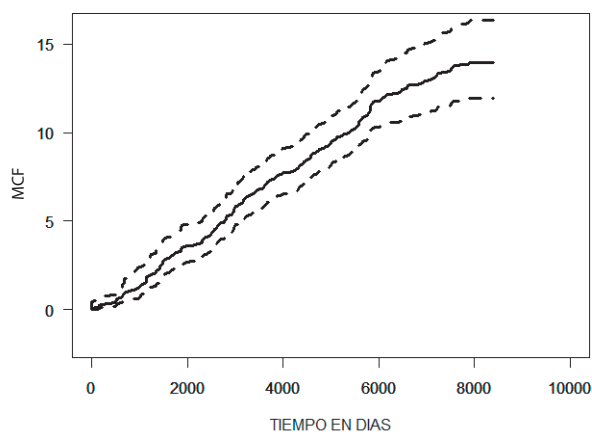


FIGURA 4: MCF estimada en SPLIDA para los interruptores FL245 por baja presión de aceite con intervalos de confianza del 95 %.

El gráfico de la MCF presenta una tendencia creciente para la tasa de recurrencia. En poco más de los 8000 días de funcionamiento de los equipos se han acumulado en promedio 14 fallas asociadas a baja presión de aceite.

**Baja presión de gas SF6**

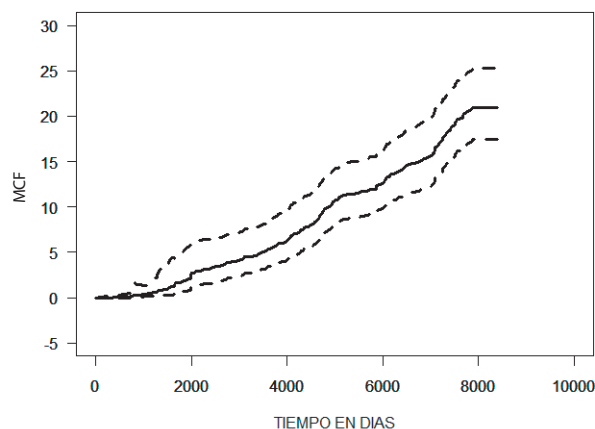


FIGURA 5: MCF estimada en SPLIDA para los interruptores FL245 por baja presión de gas SF6 con intervalos de confianza del 95 %.

La MCF acumula más fallas en promedio que los otros modos de fallas, y en poco más de los 8000 días de funcionamiento de los equipos se han acumulado en promedio 21 fallas asociadas a baja presión de SF6.

**Deterioro de los instrumentos de calibración**

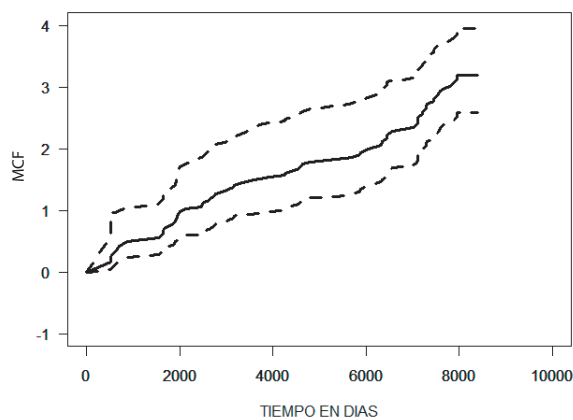


FIGURA 6: MCF estimada en SPLIDA para los interruptores FL245 por deterioro en los instrumentos de calibración con intervalos de confianza del 95 %.

La MCF acumula en promedio un poco más de 3 fallas por equipo en más de 8000 días de funcionamiento.

### 4.3. Comparación de las MCF estimadas para interruptores FL245 por cada modo de falla

La figura 7 contiene las cuatro MCF estimadas para los interruptores FL245. Las MCF corresponden a la obtenida al combinar todos los modos de falla, y a las obtenidas para cada modo de falla de manera independiente. Observe que el modo de falla que tiene mayor influencia sobre la MCF combinada, es el de baja presión de SF6. Esta curva presenta una  $m(t)$  muy baja en los primeros 2000 días de funcionamiento de los equipos, pero posteriormente presenta una tendencia creciente y llega a acumular más fallas por equipo que los otros dos modos de falla. Se observa, además que esta curva es la más parecida a la MCF con todos los modos de falla combinados.

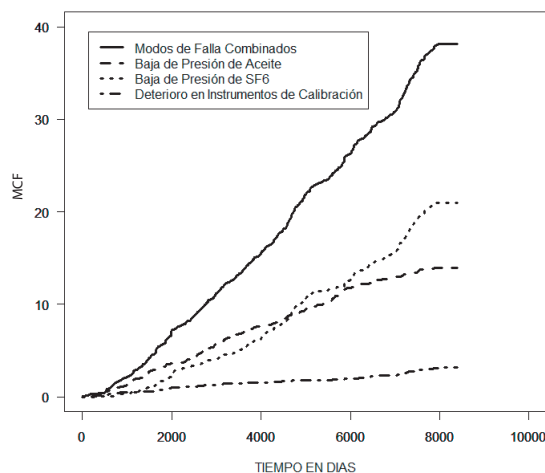


FIGURA 7: MCF estimadas para los interruptores FL245 para cada modo de falla.

La MCF asociada a baja presión de aceite tiene una  $m(t)$  más alta que los otros modos de falla en los primeros días de funcionamiento de los equipos; posteriormente la  $m(t)$  tiene una tendencia semiestable y es más baja que la  $m(t)$  asociada a baja presión de gas SF6.

Finalmente la MCF asociada al modo de falla de deterioro en los instrumentos de calibración presenta una MCF con tendencia casi constante, con una tasa de recurrencia muy cercana a cero. Este modo de falla es el que menos peso tiene en la MCF combinada.

### 4.4. Análisis paramétrico para los datos de interruptores

En esta sección se presenta un análisis paramétrico realizado a los datos de eventos de mantenimiento en los interruptores FL245, de manera que se comienza haciendo la validación de supuestos y con base en ello se establece cuál de los modelos paramétricos para la tasa de recurrencia se debe utilizar, y a partir de ese

punto se puede establecer el procedimiento para realizar predicciones del número de eventos futuros en un intervalo de tiempo de interés.

**4.4.1. Pruebas sobre la tendencia (HPP vs. NHPP)**

Existen varias pruebas para testear un modelo HPP vs NHPP. La prueba “Military Handbook” permite probar HPP vs. NHPP con modelo de regla de la potencia, y la prueba de “Laplace” testea HPP vs. NHPP con modelo log-lineal.

Otra forma de testear el modelo HPP contra el NHPP es a través de las estimaciones obtenidas para  $\beta$  y  $\gamma_1$  en los ajustes NHPP con modelos de regla de la potencia y log-lineal. De acuerdo con la teoría si  $\beta = 1$  ó  $\gamma_1 = 0$ , entonces se tiene un modelo HPP.

**4.4.2. Estimaciones de los parámetros del NHPP con modelo de regla de la potencia**

La tabla 2, presenta las estimaciones de los parámetros utilizando el módulo SPLIDA del paquete S+.

```

FL245 data
NHPP Power Rule model
Response units: Tiempo
Maximum likelihood estimation results
Appears to have converged; relative function convergence
Log likelihood at maximum point: -5732

```

	MLE	Std.Err.	95% Lower	95% Upper
eta	463.563	56.84271	352.153	574.972
beta	1.276	0.05191	1.174	1.378

TABLA 2: Parámetros estimados por máxima verosimilitud asumiendo un modelo NHPP de regla de la potencia para los datos de interruptores FL245.

Observe que el parámetro  $\beta$  es mayor que 1 (el intervalo de confianza al 95 % para  $\beta$  es: [1.174, 1.378]), por lo tanto se descarta HPP en la prueba de HPP vs. NHPP con el modelo de regla de la potencia.

**4.4.3. Estimaciones de los parámetros del NHPP con modelo log-lineal**

Utilizando SPLIDA se obtienen los resultados para el modelo log-lineal mostrado en la tabla 3.

Note que la estimación de  $\gamma_1$  es mayor que 0 (el intervalo de confianza del 95 % para  $\gamma_1$  es: [0.00004239, 0.0001049]); por tanto, se descarta entonces HPP en la prueba de HPP vs. NHPP con modelo log-lineal.

TABLA 3: Parámetros estimados por máxima verosimilitud asumiendo un modelo NHPP log-lineal para los datos de interruptores FL245

FL245 data  
 NHPP Log Linear model  
 Response units: Tiempo  
 Maximum likelihood estimation results  
 Appears to have converged; relative function convergence  
 Log likelihood at maximum point: -5738

	MLE	Std.Err.	95% Lower	95% Upper
gamma0	-5.64493304	0.08060249	-5.80291101	-5.4869551
gamma1	0.00007365	0.00001595	0.00004239	0.0001049

#### 4.4.4. Elección del modelo paramétrico NHPP para los datos de recurrencia de los interruptores FL245

La elección del modelo paramétrico NHPP que mejor ajusta a los datos de recurrencia, es una decisión que se toma de acuerdo con las estimaciones obtenidas y a los gráficos de las curvas ajustadas.

De las tablas 2 y 3, se observa que el valor de la función de log-verosimilitud para el modelo de regla de la potencia ( $-5732$ ) es mayor que el correspondiente al modelo log-lineal ( $-5738$ ), de manera que el modelo de regla de la potencia sugiere un mejor ajuste. A continuación se presentan las gráficas del ajuste de estos modelos.

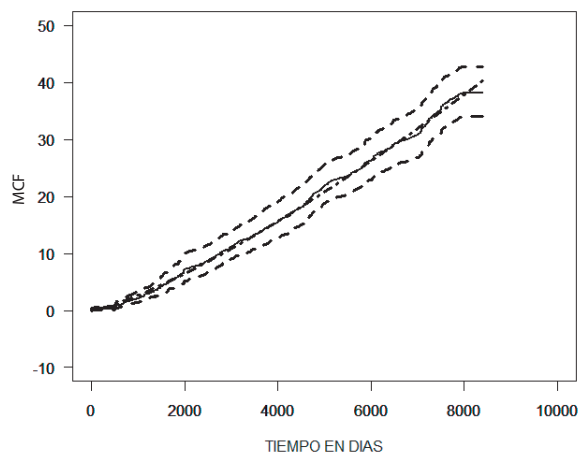


FIGURA 8: Gráfico de la MCF no paramétrica y la MCF ajustada con el modelo NHPP de regla de la potencia.

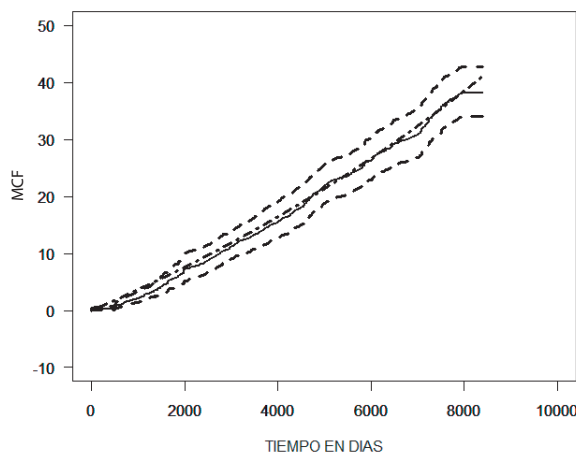


FIGURA 9: Gráfico de la MCF no paramétrica y la MCF ajustada con el modelo NHPP log-lineal.

De las figuras 8 y 9 se aprecia, que el mejor ajuste lo hace el modelo NHPP de regla de la potencia, ratificando lo observado en los ajustes por máxima verosimilitud en las tablas 2 y 3.

Eligiendo este modelo paramétrico se tendría la siguiente tasa de recurrencia:

$$m(t; \hat{\theta}) = \left( \frac{1.276}{463.563} \right) \left( \frac{t}{463.563} \right)^{1.276-1} = 0.00275 \left( \frac{t}{463.563} \right)^{0.276}$$

#### 4.4.5. Estimación de la predicción del número futuro de recurrencias

El número esperado de recurrencias en un intervalo  $[a, b]$  es  $\int_a^b m(t, \theta) dt$ . La correspondiente predicción puntual de máxima verosimilitud es  $\int_a^b m(t, \hat{\theta}) dt$ .

Una predicción puntual para el número de recurrencias futuras usando el modelo NHPP de regla de la potencia es:

$$\int_a^b m(t; \hat{\theta}) dt = \left( \frac{1}{\hat{\eta}} \right)^{\hat{\beta}} (b^{\hat{\beta}} - a^{\hat{\beta}})$$

De manera análoga, una predicción puntual para el número de recurrencias futuras usando el modelo NHPP log-lineal es:

$$\int_a^b m(t; \hat{\theta}) dt = \frac{\exp(\hat{\gamma}_0)}{\hat{\gamma}_1} [\exp(\hat{\gamma}_1 b) - \exp(\hat{\gamma}_1 a)]$$

Para el caso de los interruptores FL245, se tiene que la predicción del número de fallas en un intervalo futuro, está dada por el modelo NHPP de regla de la

potencia, es decir:

$$\int_a^b m(t; \hat{\theta}) dt = \left( \frac{1}{463.563} \right)^{1.276} (b^{1.276} - a^{1.276})$$

Supóngase que se desea estimar el número de fallas futuras para el año siguiente al final del seguimiento, es decir entre  $a = 8400$  días y  $b = 8765$  días, por lo tanto el número de recurrencias esperadas en este intervalo es:

$$\int_{8400}^{8765} m(t; \hat{\theta}) dt = \left( \frac{1}{463.563} \right)^{1.276} (8765^{1.276} - 8400^{1.276}) = 2.25$$

Es decir, para ese año se esperaban tener por cada interruptor FL245 entre 2 y 3 recurrencias.

## 5. Conclusiones y recomendaciones

La aplicación de métodos estadísticos para el análisis de datos recurrentes realizada sobre los interruptores FL245, puede servir como guía para el desarrollo de otras aplicaciones sobre este tipo de datos.

De los resultados obtenidos en la aplicación en Interconexión Eléctrica S.A., tanto en el análisis no paramétrico como en el paramétrico, se concluye lo siguiente:

- La tasa de recurrencia de los equipos FL245 es creciente; por lo tanto, indica que éstos han experimentado desgaste. El modelo paramétrico obtenido (NHPP con regla de la potencia), describe a los interruptores FL245 como equipos cuyos eventos de reparación solo estaban enfocados hacia que el equipo funcionara, llevándolo, a la condición de tan malo como viejo, es decir, que sigan funcionando de la misma manera en que venían trabajando, sin mejorar sus condiciones de operación. Este mismo modelo estima que para el año siguiente al final del estudio se tendrían entre 2 y 3 eventos de falla por equipo.
- De otra parte, el estudio para los modos de falla señaló a la baja presión de gas SF<sub>6</sub>, como el modo de falla de mayor influencia en el comportamiento de la MCF combinada de los interruptores, esto debido a que la mayoría de los eventos de reparación están asociados a este modo de falla, y su MCF es muy parecida a la MCF combinada. La MCF de baja presión de aceite presenta un comportamiento más lineal, indicando posiblemente un proceso de renovación o HPP para la modelación paramétrica de este modo de falla. Finalmente, para el modo de falla deterioro en los instrumentos de calibración, se encontró que éste era el de menor influencia en la MCF combinada y que su tasa de recurrencia era muy cercana a cero.
- Lo anterior sugiere tomar decisiones que busquen la mejora de los procedimientos de mantenimiento para el modo de falla baja presión de gas SF<sub>6</sub>,

que permitan reducir la tasa de recurrencia asociada a éste y, por lo tanto, mejorar el comportamiento de la MCF combinada y disminuir los costos de mantenimiento asociados a los interruptores FL245.

Toda la metodología aplicada en este estudio a eventos de mantenimiento puede ser replicada para los costos de tales eventos, de manera que se pueden realizar estimaciones acerca de los costos promedio por equipo en un intervalo de tiempo futuro, y esto ayudaría en la planeación de presupuestos para el mantenimiento de la familia de equipos en estudio.

## Agradecimientos

Los autores agradecen de manera especial a la Dirección de Gestión de Mantenimiento de Interconexión Eléctrica S.A., por facilitar los datos para realizar este trabajo. Además, agradecen a los árbitros y al editor por sus valiosos comentarios que enriquecieron el texto del manuscrito.

[Recibido: septiembre de 2010 — Aceptado: febrero de 2011]

## Referencias

- Andersen, K., Borgan, O., Gill, R. D. & Keiding, N. (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.
- Cook, R. J., Lawless, J. F. & Nadeau, C. (1996), 'Robust tests for treatment comparisons based on recurrent event responses', *Biometrics* **52**, 557–571.
- Escobar, L., Villa, E. & Yañez, S. (2003), 'Confiabilidad: historia, estado del arte y desafíos futuros', *Dyna* **70**(140), 5–21.
- Hu, X. J., Lawless, J. F. & Suzuki, K. (1998), 'Nonparametric estimation of a lifetime distribution when censoring times are missing', *Technometrics* **40**, 3–13.
- Kalbfleisch, J. D., Lawless, J. F. & Robinson, J. A. (1991), 'Methods for the analysis and prediction of warranty claims', *Technometrics* **33**, 273–285.
- Kvam, P. H., Singh, H. & Whitaker, L. R. (2002), 'Estimating distributions with increasing failure rate in an imperfect repair model', *Lifetime Data Analysis* **8**, 53–67.
- Lawless, J. F. (1995a), 'The analysis of recurrent events for multiple subjects', *Journal of Applied Statistics* **44**, 487–498.
- Lawless, J. F. (1995b), 'Adjustments for reporting delays and the prediction of occurred but not reported events', *Canadian Journal of Statistics* **22**, 15–31.



- Lawless, J. F. & Nadeau, C. (1995), 'Some simple robust methods for the analysis of recurrent events', *Technometrics* **37**, 158–168.
- Lin, D. Y., Wei, L. J., Yang, I. & Ying, Z. (2000), 'Semiparametric regression for the mean and rate functions of recurrent events', *Journal of the Royal Statistical Society: Series B* **62**, 711–730.
- Manotas, E. & Lopera, C. (2004), 'Metodología para el análisis de datos de recurrencia aplicada a los interruptores FL245', *Interconexión Eléctrica S.A. E.S.P.* pp. 1–145. Documento DGM-2004-I-S-ET-023-V1.
- Meeker, W. & Escobar, L. (1998), *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.
- Meeker, W. & Escobar, L. (2003), 'SPLIDA: SPLUS Life Data Analysis'.  
\*Disponible en [www.public.iastate.edu/~splida](http://www.public.iastate.edu/~splida), accedido en 01/09/2010
- Nelson, W. (1988), 'Graphical analysis of system repair data', *Journal of Quality Technology* **20**, 24–35.
- Nelson, W. (1990), 'Hazard plotting of left truncated data', *Journal of Quality Technology* **22**, 230–238.
- Nelson, W. (1995), 'Confidence limits for recurrence data: Applied to cost or number of repairs', *Technometrics* **37**, 147–157.
- Nelson, W. (1998), 'An application of graphical analysis of repair data', *Quality and Reliability Engineering International* **14**, 49–52.
- Nelson, W. (2000), 'Graphical comparison of sets of repair data', *Quality and Reliability Engineering International* **16**, 235–241.
- Nelson, W. (2003), *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*, ASA-SIAM, Philadelphia, ASA, Alexandria.
- Ross, D. P. (1989), 'Pilot Study of Commercial Water-Loop Heat Pump Compressor Life', *Report under EPRI contract RP 2480-06*, Electric Power Research Institute, Palo Alto, CA .
- Tobias, P. A. & Trindade, D. C. (1995), *Applied Reliability*, 2nd edn, CRC Press, Boca Raton, FL.
- Trindade, D. C. & Haugh, L. D. (1980), 'Estimation of the reliability of computer components from field renewal data', *Microelectronics Reliability* **20**, 205–218.
- Vallarino, C. R. (1988), Confidence Bands for a Mean Value Function Estimated from a Sample of Right-Censored Poisson Processes, Technical report 02.1471, IBM General Products Division, San Jose, California.
- Wang, M. C., Qin, J. & Chiang, C. T. (2001), 'Analyzing recurrent events data with informative censoring', *Journal of the American Statistical Association* **96**, 1057–1065.

## Análisis bivariado de confiabilidad basado en cópulas

### Reliability Bivariate Analysis Based on Copulas

ENRIQUE RAÚL VILLA-DIHARCE<sup>1,a</sup>, PEDRO ENRIQUE MONJARDIN<sup>1,b</sup>

<sup>1</sup>CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A. C., GUANAJUATO, MÉXICO

---

#### Resumen

En la industria se requieren estudios de confiabilidad para asegurar el nivel de calidad de los productos que se fabrican. Algunas veces tales productos tienen varios modos de falla que deben considerarse en el análisis de confiabilidad. Este análisis se complica cuando se incluye el patrón de dependencia de los modos de falla. En este artículo se muestra el análisis de confiabilidad de componentes que presentan dos modos de falla dependientes, expresando el patrón de dependencia de ambos modos de falla por medio de una función de cópula. Esta representación es adecuada, debido a que las distribuciones de los tiempos de falla son diferentes, Lognormal y Weibull. Se presenta un ejemplo donde se ilustra el análisis de confiabilidad de componentes con dos modos de falla.

**Palabras clave:** cópula, distribución bivariada, estimación en dos etapas, función de confiabilidad, función de verosimilitud.

#### Abstract

In industry, reliability studies are done to assure the quality level of products manufactured. Sometimes these products have several failure modes that must be considered in reliability analysis. This analysis is complicated when the dependence pattern is included. In this article a reliability analysis of components that have two dependent failure modes is proposed. The dependence pattern is expressed through a copula function, this representation is appropriate, because the marginal distributions of failure times are from different families, Lognormal and Weibull. We provide an example which illustrates the reliability analysis of components with two failure modes.

**Key words:** Bivariate distribution, Copula, Likelihood function, Reliability function, Two-stage estimation.

---

<sup>a</sup>Profesor e investigador. E-mail: villadi@cimat.mx

<sup>b</sup>Estudiante de doctorado. E-mail: pemo@cimat.mx

## 1. Introducción

En la industria existe la preocupación de mejorar los procesos de producción y aumentar la calidad de sus productos, lo que se busca es disminuir costos e incrementar la confiabilidad de los equipos o componentes que producen. Para el segundo objetivo se realizan estudios de confiabilidad para verificar si el producto satisface los requerimientos del cliente. En este trabajo comentamos una forma de analizar la confiabilidad de un producto que tiene dos modos de falla. Usualmente los equipos están formados por varios componentes, y cuando falla uno de los componentes, falla el equipo (o el desempeño del mismo disminuye), a esto se le puede llamar un sistema con varios modos de falla, de hecho un componente también puede tener varios modos de falla. Como los componentes de un sistema o producto operan bajo las mismas condiciones, se espera que sus tiempos a la falla sean dependientes. Para el análisis de este tipo de información se pueden utilizar técnicas de riesgos en competencia. En estas técnicas, los diferentes modos de falla están en competencia, esto es, solo se observa el primero que ocurre, ya que una vez que la unidad falla por alguno de los modos de falla, deja de ser observada. El tiempo a la falla de una unidad que tiene varios modos de falla, será el mínimo de los tiempos de falla potenciales asociados a cada modo de falla. En algunos estudios de confiabilidad, sobre todo en los que se realizan en laboratorio, es posible corregir los modos de falla que se van presentando, esto permite continuar con el estudio hasta que se observen todos los modos de falla. En este caso, los datos son multivariados y se pueden usar modelos multivariados para su análisis, los datos que analizamos en este trabajo son de este tipo.

En este artículo se presenta un estudio de confiabilidad de encendedores de horno de estufa a gas, que tienen dos modos de falla, que se describen posteriormente en la sección 4. En este estudio, se observan ambos modos de falla y se registran sus tiempos representados por las variables aleatorias  $T_1$  y  $T_2$ , respectivamente. Como la falla de un componente no censura la ocurrencia de una falla en el otro componente, los datos en este estudio son bivariados.

El objetivo en este estudio consiste en conocer diferentes puntos de interés como son: el patrón de dependencia de los dos tiempos a la falla, la confiabilidad en dos tiempos de referencia importantes, uno y diez años, así como los cuantiles para diferentes probabilidades.

Las distribuciones identificadas para ambas variables aleatorias, son de diferentes familias: Weibull y Lognormal. Esto nos lleva a usar el enfoque de cópulas para modelar la dependencia entre ambas variables. Las cópulas separan la estructura de dependencia de los efectos de las marginales, por esta razón, es posible considerar marginales de familias de diferentes distribuciones, como ocurre en el presente trabajo.

En la sección 2, se discute la relación entre función de supervivencia y cópulas, también se presentan algunas de las cópulas arquimedeanas comunes. Se presenta además una técnica para identificar una cópula que modele el patrón de dependencia de los datos. En la sección 3, se describe la manera de hacer inferencia, cuando tenemos funciones de supervivencia basadas en cópulas. En la sección 4, se

discuten los detalles del análisis de los datos de tiempos de falla bivariados, para los dos modos de falla de los encendedores de horno. En la sección 5 se presentan las conclusiones del presente trabajo.

## 2. Función de confiabilidad bivariada y cópulas

El objetivo de un estudio de confiabilidad es estimar cantidades de interés como cuantiles para diferentes probabilidades y niveles de confiabilidad en diferentes tiempos. Con frecuencia, en la industria la confiabilidad bivariada se estima asumiendo independencia entre los tiempos a la falla  $T_1$  y  $T_2$  de los modos de falla. Por supuesto que esto no es del todo correcto, ya que al trabajar bajo las mismas condiciones, los modos de falla generalmente son dependientes. Así, es necesario considerar modelos que tomen en cuenta la estructura de dependencia que pueden tener los datos. Además de esto, la censura puede estar presente. De manera más precisa, si hay censura por la derecha y si  $C_1$  y  $C_2$  son las variables de censura para  $T_1$  y  $T_2$ , respectivamente, los datos que se observan son parejas  $(X, Y)$  e indicadores de censura  $(\delta_1, \delta_2)$ , con  $X = \min(T_1, C_1)$ ,  $Y = \min(T_2, C_2)$ ,  $\delta_1 = \mathbb{1}_{[X=T_1]}$  y  $\delta_2 = \mathbb{1}_{[Y=T_2]}$ , donde  $\mathbb{1}_A$  es la función indicadora del evento  $A$ .

Si las distribuciones marginales de  $T_1$  y  $T_2$  son continuas, podemos unir las con una función para formar una función de confiabilidad bivariada como sigue (*e. g.* Wang & Wells 2000).

$$S(t_1, t_2) = C_\theta[S_1(t_1), S_2(t_2)] \quad (1)$$

donde  $S_j(\cdot)$ ,  $j = 1, 2$ , son las funciones de confiabilidad marginales de  $T_1$  y  $T_2$ ,  $C_\theta(\cdot, \cdot)$  es una función que une a las funciones de confiabilidad marginales, llamada función cópula, y  $\theta \in \mathbb{R}$  es un parámetro de asociación desconocido. Algunos ejemplos de estas funciones de confiabilidad bivariada se encuentran en Clayton (1978), Frank (1979), Joe (1993) y Hougaard (1986). Note que la función de cópula  $C_\theta(\cdot, \cdot)$  es una función de supervivencia en  $[0, 1] \times [0, 1]$ . En la literatura, existen una gran cantidad de trabajos relacionados con las propiedades de las funciones cópulas y el modelado de variables aleatorias usando éstas (*e.g.* Roger (2006), Joe (1997), Escarela & Hernández (2009), Jaworski, Durante, Hädel & Rychlik (2010)). Una característica especial de las funciones cópula es que la estructura de dependencia es separada de los efectos de las marginales, así que la relación de dependencia se puede estudiar sin especificar las distribuciones marginales.

La función de densidad conjunta correspondiente a la función de confiabilidad dada en la ecuación (1) está dada por

$$f(t_1, t_2) = f_1(t_1)f_2(t_2) \times c_\theta[S_1(t_1), S_2(t_2)]$$

donde  $f_j(z) = -\frac{d}{dz}S_j(z)$ ,  $j = 1, 2$ , son las densidades de los tiempos a la falla y  $c_\theta(u, v) = \frac{\partial^2}{\partial u \partial v}C_\theta(u, v)$ .

El parámetro  $\theta$  se puede ver como un parámetro de asociación global, y está relacionado con la  $\tau$  de Kendall a través de la ecuación (ver Escarela & Hernández 2009)

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u, v) dC_\theta(u, v) - 1 \equiv \Lambda(\theta) \tag{2}$$

Dado un nivel de asociación medido por  $\tau$  o  $\theta$ ,  $C_\theta(u, v)$  determina el grado de dependencia local en  $(u, v) \in [0, 1] \times [0, 1]$ . Cuando la asociación se aproxima a los niveles extremos, todos los modelos toman la misma forma: bajo independencia ( $\tau = 0$ ),  $C_\theta(u, v) = uv \equiv C^\perp(u, v)$  y bajo la máxima dependencia positiva ( $\tau = 1$ ),  $C_\theta(u, v) = \min(u, v) \equiv C^+(u, v)$ .

Existe una clase de cópulas que engloba a la mayoría de las cópulas bivariadas más comunes, su característica principal es que es indexada por una función univariada, por lo que sus propiedades matemáticas y computacionales son fáciles, esta es la llamada clase de cópulas arquimedeanas. Una función cópula es arquimedea si puede expresarse como

$$C_\theta(u, v) = \varphi_\theta^{-1} [\varphi_\theta(u) + \varphi_\theta(v)], \quad (u, v) \in [0, 1] \times [0, 1]$$

donde  $\varphi_\theta(\cdot)$  es una función convexa definida en  $[0, 1]$  tal que  $\varphi_\theta(1) = 0$  y  $\varphi_\theta^{-1}(\cdot)$  denota la pseudoinversa de  $\varphi_\theta(\cdot)$ . A  $\varphi_\theta$  se le llama la función generadora de la cópula.

Algunas de las familias de cópulas más comunes que se pueden expresar en una representación de cópulas arquimedeanas se muestran en la tabla 1.

TABLA 1: Ejemplos de cópulas arquimedeanas.

Familia	$C_\theta(u, v)$	$\varphi_\theta(t)$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta} - 1)$
Frank	$-\frac{1}{\theta} \ln \left( 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right)$	$-\ln \left( \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right)$
Gumbel	$\exp \left\{ - [(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta} \right\}$	$(-\ln t)^\theta$
Joe	$1 - [(1 - u)^\theta + (1 - v)^\theta - (1 - u)^\theta(1 - v)^\theta]$	$-\ln[1 - (1 - t)^\theta]$

Una función de confiabilidad bivariada modelada con una función cópula arquimedea toma la forma

$$\mathbf{S}(t_1, t_2) = \varphi_\theta^{-1} \{ \varphi_\theta[S_1(t_1)] + \varphi_\theta[S_2(t_2)] \}$$

Para poder usar el modelo dado en (1) es necesario elegir una cópula. Genest & Riverst (1993) proponen un método no-paramétrico que sin tomar en cuenta las marginales, estima una función que identifica a una cópula arquimedea. Ellos prueban, cuando no hay datos censurados, que la función generadora  $\varphi_\theta(\cdot)$  de una cópula arquimedea se puede identificar usando la función  $K(z) = \Pr[\mathbf{S}(T_1, T_2) \leq z]$  a través de la ecuación diferencial

$$\lambda(z) = z - K(z) = \frac{\varphi_\theta(z)}{\varphi'_\theta(z)}$$

donde  $\varphi'_\theta(z) = \frac{d}{dz}\varphi_\theta(z)$ . Este método se usa en Arcila, Jaramillo & Lopera (2009) para analizar datos de seguros.

Esta propuesta no considera la censura en los datos, lo cual se tiene con gran frecuencia en estudios de confiabilidad. Wang & Wells (2000) proponen una técnica similar a la de Genest & Rivest (1993), pero para el caso en que hay datos censurados por la derecha, el método es el siguiente: Sean  $x_{(1)} \leq \dots \leq x_{(n)}$  y  $y_{(1)} \leq \dots \leq y_{(n)}$  las observaciones ordenadas de los datos  $\{(X_i, Y_i) : i = 1, \dots, n\}$ .

1. Calcular el estimador  $\tilde{K}(z)$  de  $K(z) = 1 - \Pr[\mathbf{S}(T_1, T_2) > z]$  como

$$\tilde{K}(z) = 1 - \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_{[\tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) > z]} \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)})$$

donde  $\tilde{\mathbf{S}}$  es un estimador no-paramétrico de  $\mathbf{S}$  y

$$\begin{aligned} \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)}) &= \tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) - \tilde{\mathbf{S}}(x_{(i-1)}, y_{(j)}) \\ &\quad - \tilde{\mathbf{S}}(x_{(i)}, y_{(j-1)}) + \tilde{\mathbf{S}}(x_{(i-1)}, y_{(j-1)}) \end{aligned}$$

es la masa estimada en  $(x_{(i)}, y_{(j)})$ . Existen varias propuestas de estimadores no paramétricos de  $\mathbf{S}(t_1, t_2)$  cuando los datos son censurados, por ejemplo Dabrowska (1988), Lin & Ying (1993), Prentice & Cai (1992), Wang & Wells (1997) y Batún-Cutz (2005). Luego calcular  $\tilde{\lambda}(z) = z - \tilde{K}(z)$  y graficar  $\tilde{\lambda}(z_i)$  para  $z_i = \tilde{\mathbf{S}}(x_i, y_i)$ ,  $i = 1, \dots, n$ .

2. Sean  $C_{\theta_l}(\cdot, \cdot)$ ,  $l = 1, \dots, m$  las  $m$  funciones cópulas bajo consideración. Obtener un valor  $\theta_l$ , digamos  $\tilde{\theta}_l$ , tal que  $\hat{\tau}_0 = \Lambda(\tilde{\theta}_l)$  donde  $\hat{\tau}_0$  es un estimador de  $\tau$ . Si  $\tau = \Lambda(\theta)$  es una función uno a uno,  $\tilde{\theta}_l = \Lambda^{-1}(\hat{\tau}_0)$ . Podemos usar el estimador de  $\tau$  propuesto por Wang y Wells,

$$\hat{\tau}_0 = 4 \sum_{i=1}^n \sum_{j=1}^n \tilde{\mathbf{S}}(x_{(i)}, y_{(j)}) \tilde{\mathbf{S}}(\Delta x_{(i)}, \Delta y_{(j)}) - 1 \quad (3)$$

Si  $x_{(n)}$  y  $y_{(n)}$  son no censuradas,  $\hat{\tau}_0$  es un estimador consistente de  $\tau$ . Para cada  $\tilde{\theta}_l$ , obtener el estimador paramétrico  $\lambda_{\tilde{\theta}_l}(z)$ .

3. Comparar la gráfica de  $(z_i, \tilde{\lambda}(z_i))$  con las gráficas teóricas  $\lambda_{\tilde{\theta}_l}(z)$  para los modelos bajo consideración. Elegir el modelo que mejor ajuste los puntos  $(z_i, \tilde{\lambda}(z_i))$ , por ejemplo el que tenga menor suma de cuadrados

$$\sum_{i=1}^n \left[ \tilde{\lambda}(z_i) - \lambda_{\tilde{\theta}_l}(z_i) \right]^2$$

### 3. Inferencia

Lo primero que hay que hacer para ajustar un modelo basado en cópulas es elegir la función cópula que usaremos en el modelo (1); para esto se puede usar

el método propuesto por Wang y Wells, descrito anteriormente. El siguiente paso es determinar las distribuciones marginales apropiadas, una buena referencia para el ajuste de modelos paramétricos univariadas es el libro de Meeker & Escobar (1998).

Para estimar los parámetros  $\boldsymbol{\theta}$  de la confiabilidad bivariada  $\mathbf{S}$  usaremos el método de máxima verosimilitud. Con el modelo basado en cópulas

$$\mathbf{S}(t_1, t_2; \boldsymbol{\theta}) = C_{\theta} \left( S_1(t_1; \boldsymbol{\theta}^{(1)}), S_2(t_2; \boldsymbol{\theta}^{(2)}) \right)$$

donde  $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \theta)$ , siendo,  $\boldsymbol{\theta}^{(1)}$  y  $\boldsymbol{\theta}^{(2)}$  los parámetros correspondientes a las confiabilidades marginales, y  $\theta$  es el parámetro de la función cópula.

Sea  $\mathbf{d} = \{(x_i, \delta_{i1}, y_i, \delta_{i2}) : i = 1, \dots, n\}$  el conjunto de datos. Para especificar la función de verosimilitud es necesario distinguir entre las observaciones que son tiempos de falla y las que son censuradas.

Si ambas observaciones  $(x_i, y_i)$  son tiempos de falla ( $\delta_{i1} = 1, \delta_{i2} = 1$ ), la contribución de esta pareja a la función de verosimilitud es

$$c_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right) f_1 \left( x_i; \boldsymbol{\theta}^{(1)} \right) f_2 \left( y_i; \boldsymbol{\theta}^{(2)} \right)$$

donde  $f_j(z; \boldsymbol{\theta}^{(j)}) = -\frac{\partial}{\partial z} S_j(z; \boldsymbol{\theta}^{(j)})$ ,  $j = 1, 2$ , son las densidades de los tiempos de falla y  $c_{\theta}(u, v) = \frac{\partial^2}{\partial u \partial v} C_{\theta}(u, v)$ .

Observar que  $x_i$  es tiempo de falla y  $y_i$  es censurada ( $\delta_{i1} = 1, \delta_{i2} = 0$ ), es equivalente a observar el evento  $\{x_i < T_1 \leq x_i + h, T_2 > y_i\}$ , para  $h$  pequeña. Pero

$$\begin{aligned} \Pr(x_i < T_1 \leq x_i + h, T_2 > y_i) &= \mathbf{S}(x_i, y_i; \boldsymbol{\theta}) - \mathbf{S}(x_i + h, y_i; \boldsymbol{\theta}) \\ &\approx h \cdot \left[ -\frac{\partial}{\partial x_i} \mathbf{S}(x_i, y_i; \boldsymbol{\theta}) \right] \end{aligned}$$

y la contribución a la verosimilitud por esta pareja es

$$-\frac{\partial}{\partial x_i} \mathbf{S}(x_i, y_i; \boldsymbol{\theta}) = \frac{\partial}{\partial S_1(x_i; \boldsymbol{\theta}^{(1)})} C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right) f_1(x_i; \boldsymbol{\theta}^{(1)})$$

De la misma forma, si  $x_i$  es una censura y  $y_i$  es tiempo de falla, la contribución de esta pareja a la verosimilitud es

$$-\frac{\partial}{\partial y} \mathbf{S}(x_i, y_i; \boldsymbol{\theta}) = \frac{\partial}{\partial S_2(y_i; \boldsymbol{\theta}^{(2)})} C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right) f_2(y_i; \boldsymbol{\theta}^{(2)})$$

Por último, si  $x_i$  y  $y_i$  son censuras, su contribución a la verosimilitud es

$$\Pr(T_1 > x_i, T_2 > y_i) = \mathbf{S}(t_1, t_2; \boldsymbol{\theta}) = C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)$$

Por lo tanto, la función de verosimilitud es

$$\begin{aligned}
 L(\boldsymbol{\theta}; \mathbf{d}) &= \prod_{i=1}^n \left\{ c_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right) f_1(x_i; \boldsymbol{\theta}^{(1)}) f_2(y_i; \boldsymbol{\theta}^{(2)}) \right\}^{\delta_{i1} \delta_{i2}} \\
 &\times \left\{ \frac{\partial C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)}{\partial S_1(x_i; \boldsymbol{\theta}^{(1)})} f_1(x_i; \boldsymbol{\theta}^{(1)}) \right\}^{\delta_{i1} (1 - \delta_{i2})} \\
 &\times \left\{ \frac{\partial C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)}{\partial S_2(x_i; \boldsymbol{\theta}^{(2)})} f_2(y_i; \boldsymbol{\theta}^{(2)}) \right\}^{(1 - \delta_{i1}) \delta_{i2}} \\
 &\times C_{\theta} \left( S_1(x_i; \boldsymbol{\theta}^{(1)}), S_2(y_i; \boldsymbol{\theta}^{(2)}) \right)^{(1 - \delta_{i1})(1 - \delta_{i2})}
 \end{aligned} \tag{4}$$

El estimador máximo verosimil para  $\boldsymbol{\theta}$  es

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta}; \mathbf{d})$$

con  $\Theta$  el espacio paramétrico. Equivalentemente

$$\widehat{\boldsymbol{\theta}}_{\text{MV}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta}; \mathbf{d}) \tag{5}$$

donde  $\ell(\boldsymbol{\theta}; \mathbf{d}) = \ln L(\boldsymbol{\theta}; \mathbf{d})$  es la función log-verosimilitud. Si la función log-verosimilitud alcanza su máximo en un punto interior de  $\Theta$ , el estimador máximo verosimil es la solución de la ecuación

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}; \mathbf{d}) = \mathbf{0}.$$

Sea  $\boldsymbol{\theta}_0$  el vector de valores reales. Bajo ciertas condiciones de regularidad,

$$n^{1/2}(\widehat{\boldsymbol{\theta}}_{\text{MV}} - \boldsymbol{\theta}_0) \underset{\text{as.}}{\sim} N(\mathbf{0}, \mathcal{J}^{-1}(\boldsymbol{\theta}_0))$$

donde  $\mathcal{J}(\boldsymbol{\theta}_0)$  es la matriz información.

Con este método (la estimación máxima verosimilitud completa), los parámetros  $\boldsymbol{\theta}$  se estiman simultáneamente. Sin embargo, el enfoque con cópulas sugiere hacer la estimación en dos pasos:

1. primero estimar  $\boldsymbol{\theta}^{(1)}$  y  $\boldsymbol{\theta}^{(2)}$  separadamente maximizando las verosimilitudes

$$\begin{aligned}
 L_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)}) &= \prod_{i=1}^n f_1(x_i; \boldsymbol{\theta}^{(1)})^{\delta_{i1}} S_1(x_i; \boldsymbol{\theta}^{(1)})^{(1 - \delta_{i1})} \\
 L_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)}) &= \prod_{i=1}^n f_2(y_i; \boldsymbol{\theta}^{(2)})^{\delta_{i2}} S_2(x_i; \boldsymbol{\theta}^{(2)})^{(1 - \delta_{i2})}
 \end{aligned}$$

de los datos univariados. Equivalentemente, se puede maximizar  $\ell_j(\boldsymbol{\theta}^{(j)}; \mathbf{d}^{(j)}) = \ln L_j(\boldsymbol{\theta}^{(j)}; \mathbf{d}^{(j)})$ ,  $j = 1, 2$ .



2. luego, estimar  $\theta$  maximizando  $L(\theta; \hat{\boldsymbol{\theta}}^{(1)}, \hat{\boldsymbol{\theta}}^{(2)}; \mathbf{d})$  dadas las estimaciones  $\hat{\boldsymbol{\theta}}^{(1)}$  y  $\hat{\boldsymbol{\theta}}^{(2)}$

Este método es llamado inferencia por funciones marginales (IFM) por Joe & Xu (1996). Los estimadores  $\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}$  y  $\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}$  están dados por

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)} &= \arg \text{máx } \ell_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)}) \\ \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)} &= \arg \text{máx } \ell_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)})\end{aligned}\tag{6}$$

y se sigue que el estimador para el parámetro de la cópula es

$$\hat{\theta}_{\text{IFM}} = \arg \text{máx } \ell(\theta; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})\tag{7}$$

Este método es computacionalmente más simple que la estimación simultánea de todos los parámetros. Una optimización numérica con muchos parámetros es más costosa en tiempo de computo comparada con varias optimizaciones con pocos parámetros (Joe 1997, pp. 300).

Quizás el parámetro de mayor interés es el parámetro  $\theta$ , ya que refleja la dependencia que hay en los datos. La verosimilitud relativa del parámetro  $\theta$  se define como

$$R(\theta) = \frac{L(\theta; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})}{L(\hat{\theta}_{\text{IFM}}; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}, \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}, \mathbf{d})}$$

y asintóticamente  $r^*(\theta) = -2 \ln[R(\theta)]$  tiene una distribución ji-cuadrada con un grado de libertad, cuando se evalúa en el valor real de  $\theta$ . Al hacer una prueba de razón de verosimilitudes, se rechaza la hipótesis nula  $\theta = \theta_0$ , al nivel de significancia  $\alpha$ , si

$$r^*(\theta) = -2 \ln[R(\theta)] > \chi_{(1-\alpha, 1)}^2$$

con  $\chi_{(1-\alpha, 1)}^2$  el cuantil  $1-\alpha$  de la distribución ji-cuadrada con un grado de libertad. También es posible, usando esta aproximación asintótica, construir intervalos de confianza aproximada del  $100(1-\alpha)\%$ .

La principal ventaja del método IFM es la simplificación de los cálculos numéricos. Sin embargo los estimadores IFM tienen otras propiedades deseables. Al igual que los estimadores de máxima verosimilitud, los estimadores IFM también son asintóticamente eficientes, y se tiene que

$$n^{1/2}(\hat{\boldsymbol{\theta}}_{\text{IFM}} - \boldsymbol{\theta}_0) \underset{\text{as.}}{\sim} N(\mathbf{0}, \mathfrak{B}(\boldsymbol{\theta}_0))$$

donde  $\mathfrak{B}(\boldsymbol{\theta}_0)$  es llamada inversa de la matriz de información de Godambe. Si se define

$$\mathbf{g}(\boldsymbol{\theta}; \mathbf{d}) = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{\theta}^{(1)}} \ell_1(\boldsymbol{\theta}^{(1)}; \mathbf{d}^{(1)})^T \\ \frac{\partial}{\partial \boldsymbol{\theta}^{(2)}} \ell_2(\boldsymbol{\theta}^{(2)}; \mathbf{d}^{(2)})^T \\ \frac{\partial}{\partial \theta} \ell(\theta; \boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \mathbf{d}) \end{bmatrix}$$

entonces

$$\mathfrak{B}(\boldsymbol{\theta}_0) = \mathbf{D}^{-1} \mathbf{M} (\mathbf{D}^{-1})^T$$

donde  $\mathbf{D} = \text{E} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$  y  $\mathbf{M} = \text{E} \left[ \mathbf{g}(\boldsymbol{\theta}; \mathbf{d}) \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$ .

Para estimar la matriz de covarianzas asintóticas  $n^{-1} \mathfrak{B}(\boldsymbol{\theta}_0)$  de  $\hat{\boldsymbol{\theta}}_{\text{IFM}}$  es necesario calcular  $\text{E} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta}; \mathbf{d})^T \right]$ , lo cual puede ser tedioso y complicado. Para evitar esto se pueden usar métodos de remuestreo como jackknife y bootstrap, en este trabajo decidimos usar jackknife, ya que, de acuerdo con Shao & Tu (1995), el procedimiento bootstrap no da mejores estimadores de la varianza que el procedimiento jackknife y además este último es computacionalmente más sencillo. Sea  $\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}$  el estimador de  $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta})$  con la  $i$ -ésima observación  $d_i$  excluida de los datos,  $i = 1, \dots, n$ . Si  $\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}$  y  $\hat{\boldsymbol{\theta}}_{\text{IFM}}$  son vectores renglones, el estimador jackknife de  $n^{-1} \mathfrak{B}(\boldsymbol{\theta}_0)$  es (ver Joe 1997)

$$\sum_{i=1}^n \left( \hat{\boldsymbol{\theta}}_{(-i)\text{IFM}} - \hat{\boldsymbol{\theta}}_{\text{IFM}} \right)^T \left( \hat{\boldsymbol{\theta}}_{(-i)\text{IFM}} - \hat{\boldsymbol{\theta}}_{\text{IFM}} \right)$$

El método jackknife también se puede usar para calcular los errores estándar de funciones de los parámetros. El método delta requiere las derivadas de las funciones respecto a los parámetros, el método jackknife evita la necesidad de estas. Sea  $h(\boldsymbol{\theta})$  la cantidad de interés (de valores reales), el estimador del error estándar de  $h(\hat{\boldsymbol{\theta}}_{\text{IFM}})$  es

$$\left\{ \sum_{i=1}^n \left[ h(\hat{\boldsymbol{\theta}}_{(-i)\text{IFM}}) - h(\hat{\boldsymbol{\theta}}_{\text{IFM}}) \right]^2 \right\}^{1/2} \quad (8)$$

Ya que la unidad falla cuando ocurre por primera vez alguno de los modos de falla, la función de confiabilidad de la unidad está dada por

$$\begin{aligned} S(t) &= \Pr(T_1 > t, T_2 > t) \\ &= \mathbf{S}(t, t) = C_{\theta}(S_1(t), S_2(t)) \end{aligned}$$

si se usa el modelo (1). Entonces, un estimador de la función de confiabilidad de la unidad es

$$\hat{S}(t) = C_{\hat{\boldsymbol{\theta}}_{\text{IFM}}} \left[ S_1(t; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(1)}), S_2(t; \hat{\boldsymbol{\theta}}_{\text{IFM}}^{(2)}) \right]$$

Usando el método jackknife, específicamente la ecuación (8), podemos encontrar intervalos de confianza para la función de confiabilidad estimada de la unidad  $\hat{S}(t)$ .

Dado  $p \in [0, 1]$ , si resolvemos la ecuación  $p = \hat{S}(t_p)$  para  $t_p$  se obtiene una estimación del tiempo en el que la unidad tiene una confiabilidad del  $100p\%$ . Ya que la función de confiabilidad es una función monótona, un intervalo de confianza aproximado del  $100(1 - \alpha)\%$  para  $t_p$  es el conjunto  $\{t : A(p) \leq \hat{S}(t) \leq B(p)\}$ , donde  $(A(p), B(p))$  es un el intervalo de confianza aproximado del  $100(1 - \alpha)\%$  para  $\hat{S}(t_p)$ .

## 4. Aplicación a un estudio de encendedores

Se considera un estudio de confiabilidad de encendedores de hornos de estufas a gas, que tienen dos modos de falla. El motivo del estudio fue conocer la confiabilidad en dos tiempos de referencia, un año y diez años. Estos tiempos son de interés, debido a que las unidades donde se instalan los encendedores, tienen un tiempo de garantía de un año y un tiempo de vida de diez años. El fabricante del electrodoméstico donde se instalan los encendedores, planea cambiar los encendedores con el fin de reducir el número de reclamos de los usuarios, por fallas de los electrodomésticos debidas a la falla del encendedor. El fabricante planea introducir un nuevo encendedor que satisfaga las confiabilidades objetivo a uno y diez años, de 0.992 y 0.974, respectivamente. Un proveedor asegura que su producto cumple con estas exigencias, así que se llevó a cabo un estudio de confiabilidad de los encendedores que oferta el proveedor.

Los datos constan de 27 parejas de tiempos a la falla  $(x_i, y_i)$ , donde  $x_i$  corresponde a una falla debida a que el encendedor no alcanza el amperaje necesario para que se active el control electrónico del horno y  $y_i$  corresponde a la falla debida a que el encendedor no alcanza en el tiempo reglamentario (20 segundos) la temperatura mínima de operación. Los encendedores que se estudian, son resistencias de grafito en las que se aplica un voltaje determinado y su temperatura va creciendo hasta alcanzar un nivel estable determinado. Consideramos aquí dos modos de falla. Un modo de falla se presenta cuando la corriente que circula por el encendedor en operación es menor que el nivel mínimo de la corriente necesaria para activar la tarjeta de control electrónico. El otro modo de falla se presenta cuando la temperatura del encendedor a los 20 segundos de operación es menor a la temperatura mínima de operación, de 1050 °C.

Con el envejecimiento de los encendedores, ocurre que al aplicar un nivel determinado de voltaje, el crecimiento de su temperatura es cada vez más lento, y la corriente que circula por él, es cada vez menor. En este conjunto de datos tenemos varias censuras por la derecha, debido a que el proceso de medición de la corriente o la temperatura se interrumpió. Para acelerar las fallas en este estudio, no se llevó a cabo un estudio de vida acelerada, sino que se incrementó la tasa de uso de los componentes. Después de hacer un estudio de campo, se determinó el patrón de operación y el número promedio de horas de uso del encendedor al año. Posteriormente, en el laboratorio las unidades se mantuvieron en operación, con una frecuencia de uso mayor a la real y los tiempos a la falla registrados se expresaron en años de operación real. La gráfica de los datos se muestra en la figura 1. En ella se observa que hay una dependencia moderada entre ambos tiempos de falla.

Lo primero que hacemos es seleccionar la función cópula que se usará en el modelo (1). Recurrimos al método propuesto por Wang & Wells (2000) descrito en la sección 2 para elegir una de las funciones cópula mostradas en la tabla 1, para lo cual es necesario tener una estimación no-paramétrica de la función de confiabilidad bivariada  $\mathcal{S}(\cdot, \cdot)$ . El estimador de Dabrowska (1988) es el más común, aun cuando en algunas ocasiones el resultado obtenido no es una confiabilidad propia. Esto es lo que ocurre para los datos de encendedores, por lo que optamos por usar el estimador de  $\mathcal{S}(\cdot, \cdot)$  propuesto por Batún-Cutz (2005) en su tesis doctoral.

Este estimador de la confiabilidad bivariada es una modificación del estimador de Dabrowska (1988) que corrige este problema.

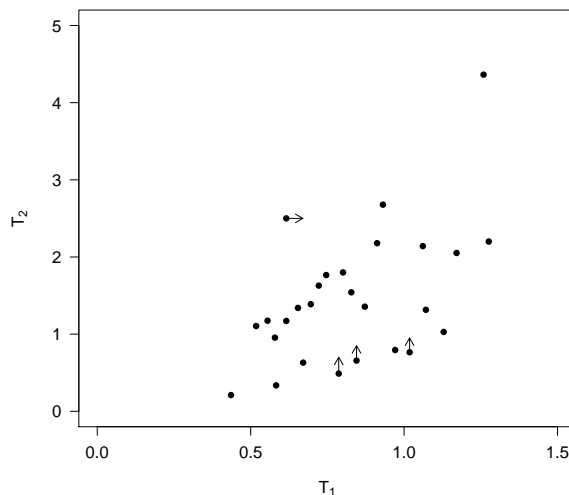


FIGURA 1: Tiempos de falla observados para los encendedores. La flecha indica el tiempo censurado.

Con el estimador  $\tilde{S}(t_1, t_2)$  de Batún-Cutz (2005) calculamos  $\tilde{\lambda}(z_i)$  para  $z_i = \tilde{S}(x_i, y_i)$ ,  $i = 1 \dots, 27$ . El estimador de la  $\tau$  de Kendall dado por la ecuación (3) es  $\tilde{\tau} = 0.3939$ , con éste obtenemos los estimadores  $\tilde{\theta}$ 's de los parámetros de las cópulas en la tabla 1 como se describe en el punto 2 del método para elegir una función cópula. Luego construimos los estimadores  $\lambda_{\tilde{\theta}}(z)$  de las funciones  $\lambda(z)$ . La gráfica de las curvas  $\lambda_{\tilde{\theta}}(z)$  y de los puntos  $(z_i, \tilde{\lambda}(z_i))$  se muestran en la Figura 2 y en la tabla 2 se dan las distancias de los puntos a las curvas. Como los puntos  $(z_i, \tilde{\lambda}(z_i))$  están más cerca de la curva  $\lambda_{\tilde{\theta}}(z)$  que corresponde a la función cópula de Gumbel, elegimos esta función cópula para hacer la inferencia.

TABLA 2: Suma de las diferencias al cuadrado,  $\sum_{i=1}^n [\tilde{\lambda}(z_i) - \lambda_{\tilde{\theta}}(z_i)]^2$

Cópula	Distancia $\times 10^3$
Frank	34.5867
Joe	28.0921
Clayton	96.7035
Gumbel	25.8142

Cuando se analizan los modos de falla por separado, se encuentra que las distribuciones más adecuadas para los tiempos de falla  $T_1$  y  $T_2$  son la Lognormal y la Weibull, respectivamente. Luego, las funciones de confiabilidad marginales para

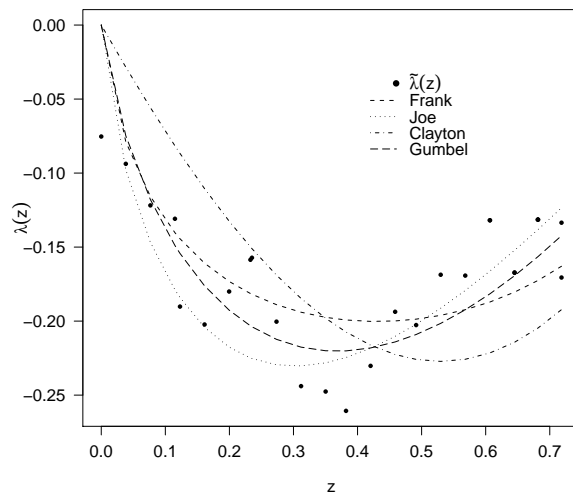


FIGURA 2: Gráfica de  $\tilde{\lambda}(z)$  y  $\lambda_{\hat{\theta}}(z)$  para las cópulas consideradas.

los tiempos de falla serán

$$\begin{aligned} S_1(t_1; \mu, \sigma) &= 1 - \Phi\left(\frac{\ln t_1 - \mu}{\sigma}\right) \\ S_2(t_2; \beta, \eta) &= \exp\left[-\left(\frac{t_2}{\eta}\right)^\beta\right] \end{aligned} \quad (9)$$

donde  $\Phi(\cdot)$  es la función de distribución normal estándar. En la notación de la sección anterior,  $\boldsymbol{\theta}^{(1)} = (\mu, \sigma)$  y  $\boldsymbol{\theta}^{(2)} = (\beta, \eta)$  son los parámetros de las marginales.

La calidad del ajuste de las distribuciones marginales puede examinarse con una comparación gráfica de las confiabilidades ajustadas y de sus versiones empíricas como se muestra en la figura 3. Debido a la censura usamos el estimador de Kaplan-Meier para las confiabilidades empíricas. Note que la confiabilidad del primer modo de falla es menor, es decir, el primer modo de falla es el que determina el comportamiento de la unidad.

Entonces la función de confiabilidad bivariada de la unidad toma la forma

$$\mathcal{S}(t_1, t_2; \boldsymbol{\theta}) = \exp\left\{-\left[\left(-\ln S_1(t_1; \boldsymbol{\theta}^{(1)})\right)^\theta + \left(-\ln S_2(t_2; \boldsymbol{\theta}^{(2)})\right)^\theta\right]^{1/\theta}\right\}$$

donde  $S_j(t_j; \boldsymbol{\theta}^{(j)})$ ,  $j = 1, 2$  están dadas por las ecuaciones (9).

Los valores de los estimadores IFM de los parámetros  $\boldsymbol{\theta} = (\mu, \sigma, \beta, \eta, \theta)$  se dan en la tabla 3, y la gráfica de contornos de la confiabilidad bivariada estimada se muestra en la figura 4. De la gráfica vemos que hay indicios de una dependencia entre los tiempos a la falla.

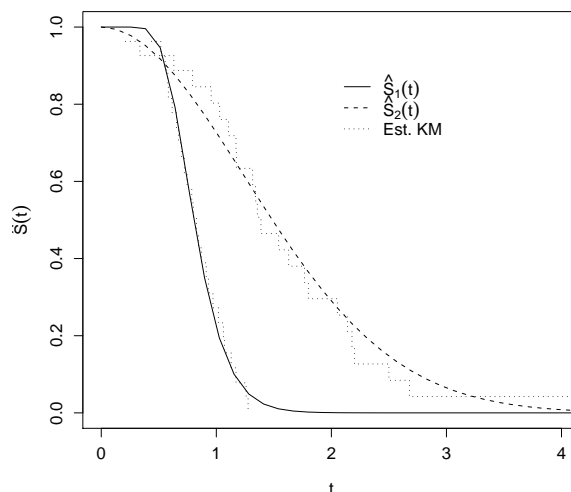


FIGURA 3: Funciones de confiabilidad univariadas estimadas de los modos de falla. La línea punteada corresponde a los estimadores Kaplan-Meier.

TABLA 3: Estimadores IFM, errores estándar e intervalos de confianza aproximados del 95 % para los datos de encendedores.

Parámetro	Estimación	Error	
		Estándar	Intervalo de confianza 95 %
$\mu$	-0.2160	0.0566	(-0.327, -0.105)
$\sigma$	0.2806	0.0334	(0.2152, 0.3461)
$\beta$	1.9561	0.5591	(0.8603, 3.0519)
$\eta$	1.7932	0.2126	(1.3765, 2.2100)
$\theta$	1.4732	0.2862	(0.9123, 2.0341)

Usamos la técnica jackknife para calcular la matriz de covarianzas asintótica de Godambe de las estimaciones IFM de los parámetros, la matriz de correlaciones correspondiente se da en la tabla 4. También es posible calcular intervalos de confianza aproximados del 95 % para las estimaciones de los parámetros, que se muestran en la tabla 3.

Como ya comentamos, el método IFM facilita los cálculos numéricos para estimar los parámetros. Esto es cierto sobre todo si el número de parámetros es grande. En el ejemplo aquí tratado sólo es necesario estimar cinco parámetros, y es posible estimarlos conjuntamente sin mucho esfuerzo. Los estimadores máximo verosímiles dados al resolver la ecuación (5) se dan en la tabla 5, observe que básicamente no hay diferencias en las estimaciones obtenidas por los dos métodos.

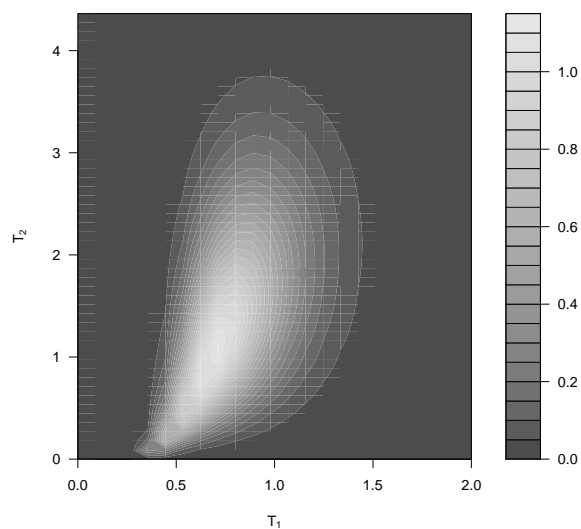


FIGURA 4: Gráfica de contornos de la densidad conjunta estimada de los tiempos a la falla de los encendedores.

TABLA 4: Matriz de correlaciones de las estimaciones IFM de los parámetros

	$\mu$	$\sigma$	$\beta$	$\eta$	$\theta$
$\mu$	1.0000	-0.1682	-0.0451	0.6445	-0.2240
$\sigma$	-0.1682	1.0000	-0.4964	-0.0648	0.6323
$\beta$	-0.0451	-0.4964	1.0000	-0.3953	-0.7322
$\eta$	0.6445	-0.0648	-0.3953	1.0000	0.1373
$\theta$	-0.2240	0.6323	-0.7322	0.1373	1.0000

TABLA 5: Estimadores máximo verosímiles, errores estándar e intervalos de verosimilitud-confianza aproximados del 95 % para los datos de encendedores.

Parámetro	Estimación	Intervalo del 95 % de confianza
$\mu$	-0.2078	(-0.3165, -0.0939)
$\sigma$	0.2802	(0.2216, 0.3720)
$\beta$	1.9613	(1.4248, 2.5539)
$\eta$	1.8033	(1.4351, 2.2473)
$\theta$	1.4687	(1.0429, 2.3307)

En la cópula de Gumbel tenemos el caso de independenciam cuando  $\theta = 1$ , por lo que podemos probar la hipótesis de modos de falla independientes, probando la

hipótesis  $\theta = 1$ . El estadístico de la prueba de la razón de verosimilitudes  $r^*(\theta)$  toma el valor  $r^*(1) = 9.6808$  que es mayor que  $\chi^2_{(0.95,1)} = 3.84$ , por lo que debemos rechazar la hipótesis nula  $\theta = 1$ , ya que tenemos una fuerte evidencia en contra de la independencia de los modos de falla.

Por la ecuación (2), la relación entre la  $\tau$  de Kendall y el parámetro de la función cópula de Gumbel es  $\tau = 1 - \theta^{-1}$ . Entonces, es posible calcular una estimación de  $\tau$  como  $\hat{\tau} = 1 - (\hat{\theta}_{\text{IFM}})^{-1}$ . Para los datos de ejemplo obtenemos  $\hat{\tau} = 0.3212$ , lo que indica que existe una dependencia moderada entre los modos de falla, como ya lo vimos en la prueba de hipótesis anterior.

En la figura 5 se muestran las estimaciones de la función de confiabilidad  $\hat{S}(t) = \mathcal{S}(t, t; \hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \theta)$  considerando el modelo con la cópula de Gumbel. Si se asume que los modos de falla son independientes, la función de confiabilidad de los encendedores toma la forma  $\hat{S}^\perp(t) = S_1(t; \theta_{\text{IFM}}^{(1)})S_2(t; \theta_{\text{IFM}}^{(2)})$ . Su gráfica también se muestra en la figura 5. Note que  $\hat{S}^\perp(t) \leq \hat{S}(t)$ , lo cual indica que al asumir independencia se subestima la función de confiabilidad del encendedor.

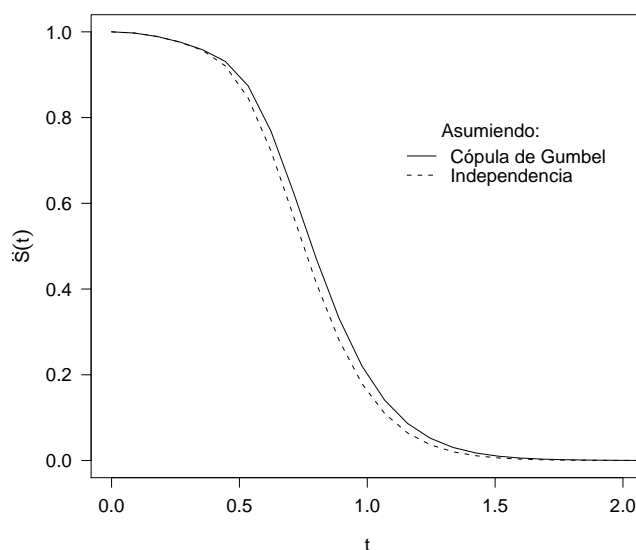


FIGURA 5: Confiabilidades estimadas de encendedores, se muestra la función de confiabilidad estimada con la cópula de Gumbel y la confiabilidad asumiendo independencia.

Con el fin de verificar las afirmaciones del proveedor, usamos la función de confiabilidad estimada  $\hat{S}(t)$  para calcular estimaciones de la confiabilidad del encendedor en los tiempos de uno y diez años. Las estimaciones de la confiabilidad para estos tiempos, además de un mes y dos años, se dan en la tabla 6. Vemos que las confiabilidades estimadas están muy por debajo de las confiabilidades ob-



jetivos. Al estimar el intervalo de confianza de la confiabilidad a los 10 años con el método jackknife, encontramos que las estimaciones de la confiabilidad a los 10 años prácticamente no cambian, al excluir cualquiera de las observaciones. En consecuencia la desviación estándar calculada con este método es cero. Así, el intervalo de confianza aproximado del 95 % para la confiabilidad a 10 años consta sólo del punto cero.

TABLA 6: Confiabilidades estimadas para tiempos dados.

$t_p$	Confiabilidad	Desviación	Intervalo
	Estimada	Estándar	Confianza 95 %
Un mes	0.9975	0.0027	(0.9923, 1.0000)
Un año	0.1992	0.0580	(0.0856, 0.3129)
Dos años	0.0004	0.0006	(0.0000, 0.0154)
10 años	0.0000	–	–

Para ver los tiempos en que se consiguen las confiabilidades objetivo que plantea el cliente, resolvemos la ecuación  $c = \widehat{S}(t_c)$  donde  $c$  es una confiabilidad dada. Así mismo, podemos dar un intervalo de confianza aproximado del 95 % para el tiempo  $t_c$  como se describe al final de la sección anterior. Los resultados para estas, y algunas otras, se muestran en la tabla 7. En ella observamos que los tiempos a los que se consiguen las confiabilidades objetivos son muy bajos. Por ejemplo, la confiabilidad de 0.992 se alcanza en  $t_c = 0.1522$ ; esto significa que se esperan ocho fallas, de cada 1000 unidades, a los 55 días, aproximadamente. Más grave aún es lo que indica el último renglón en la tabla, el 20 % de las unidades fallarán antes de 7.3 meses.

TABLA 7: Cuantiles obtenidos para confiabilidades dadas.

Confiabilidad	Intervalo	
dada	$t_c$	Confianza 95 %
0.992	0.1522	(0.0000, 0.1818)
0.974	0.2794	(0.0000, 0.2931)
0.950	0.3887	(0.0000, 0.3576)
0.900	0.5014	(0.1408, 0.4012)
0.800	0.6005	(0.3159, 0.6811)

El análisis de confiabilidad de estos encendedores, muestra que éstos no cumplen con los niveles requeridos por el fabricante de electrodomésticos que planea hacer el cambio. Los resultados del estudio se enviaron al proveedor con la intención de que modificara el diseño del encendedor, para que su confiabilidad satisfaga los requerimientos del cliente.

## 5. Conclusiones

Los fabricantes siempre han estado preocupados por incrementar la confiabilidad de sus productos. Para lo cual se ven en la necesidad de hacer estudios de confiabilidad de los artículos que producen, así como de exigir estudios que avalen la confiabilidad de aquellos que adquieren. En este trabajo se ha mostrado el uso de la función cópula para modelar la confiabilidad de un artículo con dos modos de falla. Este enfoque es muy útil en la industria ya que es muy flexible con respecto a la selección de las distribuciones marginales.

Hemos utilizado funciones de cópula para modelar la estructura de dependencia para dos distribuciones marginales provenientes de diferentes familias, lognormal y Weibull. Lo cual no es posible con los procedimientos tradicionales que suponen marginales de la misma familia. Además, utilizamos el procedimiento de Wang y Wells para seleccionar la función de cópula entre un conjunto de cópulas propuestas. Esta selección se basa en las discrepancias entre la cópula empírica y las cópulas propuestas.

Hemos estimado los parámetros con dos métodos, IFM y máxima verosimilitud, y hemos obtenido resultados similares. Esto nos lleva a proponer el uso del método IFM debido a que es computacionalmente más simple que el de máxima verosimilitud y tiene propiedades similares a este como lo han demostrado Joe & Xu (1996).

Comparamos gráficamente las confiabilidades estimadas suponiendo independencia y dependencia modelada con la cópula de Gumbel. En esta comparación observamos que sistemáticamente es menor la confiabilidad en el caso de independencia.

Hemos ilustrado el método propuesto con un conjunto de datos reales, provenientes de la industria de electrodomésticos.

Como trabajo futuro consideramos hacer una comparación más exhaustiva entre los modelos que asumen dependencia e independencia. Esta comparación es interesante debido a que el análisis bajo el modelo de independencia es más simple que cuando suponemos un determinado patrón de dependencia, y es conveniente conocer cuanto se pierde al usar el modelo de independencia.

## Agradecimientos

Agradecemos al Dr. Luis A. Escobar R. (Louisiana State University) sus comentarios y sugerencias, ya que fueron de gran utilidad para mejorar nuestra presentación; así mismo, queremos expresar nuestra gratitud al Editor invitado y a los árbitros por su lectura cuidadosa y recomendaciones que contribuyeron a mejorar el trabajo.

[Recibido: agosto de 2010 — Aceptado: febrero de 2011]

## Referencias

- Arcila, L. D., Jaramillo, M. C. & Lopera, C. M. (2009), ‘Selección de un modelo cópula para el ajuste de datos bivariados dependientes’, *Dyna* **76**(158), 253–263.
- Batún-Cutz, J. L. (2005), Algunos problemas en inferencia no paramétrica: estimación de distribuciones ordenadas por picudez y supervivencia bivariada con censura, Tesis de doctorado, Centro de Investigación en Matemáticas, A. C., Guanajuato, México.
- Clayton, D. G. (1978), ‘A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence’, *Biometrika* **65**, 141–151.
- Dabrowska, D. M. (1988), ‘Kaplan-Meier estimates on the plane’, *The Annals of Statistics* **16**, 1475–1489.
- Escarela, G. & Hernández, A. (2009), ‘Modelado de parejas aleatorias usando cópulas’, *Revista Colombiana de Estadística* **32**(1), 33–58.
- Frank, M. J. (1979), ‘On the simultaneous associativity of  $f(x, y)$  and  $1 - x + y - f(x, y)$ ’, *Aequationes Mathematicae* **19**, 194–226.
- Genest, C. & Rivest, L. P. (1993), ‘Statistical inference procedures for bivariate Archimedean copulas’, *Journal of the American Statistical Association* **88**, 1034–1043.
- Hougaard, P. (1986), ‘A class of multivariate failure time distributions’, *Biometrika* **73**, 671–678.
- Jaworski, P., Durante, F., Hädel, W. & Rychlik, T. (2010), *Copula Theory and its Applications*, Vol. 1, Springer, Germany.
- Joe, H. (1993), ‘Parametric families of multivariate distributions with given margins’, *Journal of Multivariate Analysis* **46**, 573–581.
- Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall.
- Joe, H. & Xu, J. J. (1996), The estimation method of inference functions for margins for multivariate models, Technical Report 166, Department of Statistics, University of British Columbia.
- Lin, D. Y. & Ying, Z. (1993), ‘A simple nonparametric estimator of the bivariate survival function under univariate censoring’, *Biometrika* **80**, 573–581.
- Meeker, W. L. & Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, John Wiley & Sons.
- Prentice, R. L. & Cai, J. (1992), ‘Covariance and survivor function estimation using censored multivariate failure time data’, *Biometrika* **79**, 495–512.

- Roger, B. N. (2006), *An Introduction to Copulas*, second edn, Springer.
- Shao, J. & Tu, D. (1995), *The Jackknife and Bootstrap*, Springer-Verlag.
- Wang, W. & Wells, M. T. (1997), 'Nonparametric estimation of the bivariate survival function under simplified censoring conditions', *Biometrika* **84**, 863–880.
- Wang, W. & Wells, M. T. (2000), 'Model selection and semiparametric inference for bivariate failure-time data', *Journal of the American Statistical Association, Theory and Methods* **95**(449), 62–72.



## Process Capability Analysis Plot for a Product with Bilateral Specifications

Análisis de capacidad gráfico para un producto con especificaciones  
bilaterales

RUBÉN DARÍO GUEVARA<sup>a</sup>, JOSÉ ALBERTO VARGAS<sup>b</sup>

DEPARTAMENTO DE ESTADÍSTICA, FACULTAD DE CIENCIAS, UNIVERSIDAD NACIONAL DE  
COLOMBIA, BOGOTÁ, COLOMBIA

---

### Abstract

A multi-process performance analysis plot based on process capability indices has been developed to analyse the manufacturing performance for multiple processes, which conveys critical information regarding the departure of the process mean from the target value and process variability, and provides a guideline for capability improvement. Existing performance analysis plots usually assume normal distributions and symmetric tolerances. In this paper, we present a graphical method to evaluate the process capability of a product composed of independent multiple processes with nominal-the-best specifications and symmetric or asymmetric tolerances. The proposed plot aims to get information about the location and spread of the studied characteristics which can follow any continuous distribution. An example showing the effect of processes with asymmetric tolerances on the capability of the product is considered.

**Key words:** Process capability indices, Process improvement, Statistical process control.

### Resumen

El análisis gráfico de la capacidad de un proceso multivariado suministra información crítica acerca de la distancia que existe entre la media y su valor objetivo, para cada uno de los procesos; así como la variabilidad de cada uno de ellos. Esta información provee luces para mejorar la capacidad de todos los procesos que hacen forma del producto. En los gráficos existentes para analizar la capacidad de un producto, se asume que cada uno de los procesos que conforman el producto son independientes, tienen distribución normal y presentan tolerancia simétrica. En este artículo, proponemos

---

<sup>a</sup>Professor. E-mail: rdguezarag@bt.unal.edu.co

<sup>b</sup>Professor. E-mail: javargasn@unal.edu.co

un método gráfico para evaluar la capacidad de un producto conformado por procesos independientes, bajo cualquier distribución continua para tolerancias simétricas o asimétricas. El gráfico propuesto suministra información acerca de la localización y dispersión de cada una de las características que conforman el producto. Se presenta un ejemplo que ilustra el efecto de la asimetría sobre la capacidad de los procesos.

**Palabras clave:** control estadístico de procesos, índices de capacidad de procesos, mejoramiento de procesos.

## 1. Introduction

In the quality improvement processes, to establish if the products meet some preset specifications is important. Process capability indices (PCIs) have been widely used to measure whether the product quality meets the preset specifications.

In process capability studies there are two characteristics, –the process location, usually the process mean, in relation to its target value and the process dispersion–. Some process capability indices combine these two characteristics. The larger the value of a process capability index the more capable the process. A low value of a process capability index indicates that a process is non-capable, but it does not give us information if this situation is due to the variability, the deviation from target or both. See Deleryd & Vännman (1999). This drawback is overcome by graphical displays, which allow getting information about the location and spread of the studied characteristics.

Vännman (1997b) introduced a process capability plot, called the  $(\delta, \gamma)$ -plot. The  $(\delta, \gamma)$ -plot was defined for a normally distributed process output, bilateral specifications and symmetric tolerances. This is, the target value  $T$  is located at the mid-point of the specification interval. Chen, Huang & Li (2001) developed a chart to analyse the capability of a whole product, called PCAC. The PCAC evaluates simultaneously independent characteristics, with bilateral or unilateral specifications, symmetric or asymmetric specifications, but each process must have normal distribution. Huang & Chen (2003) developed a process capability chart for a multi-process product with bilateral specifications, under normal distributions. Chen, Hsu & Wu (2006) proposed a multi-process capability chart to evaluate non-normal processes capability, with symmetric tolerances. Chen, Huang & Hung (2008) took the PCAC of Chen et al. (2001) and proposed the  $PCAC/C_{pm}$  chart. It can evaluate the capability of the product employing the univariate indices  $C_{pu}$ ,  $C_{pl}$  and  $C_{pm}$ .  $PCAC/C_{pm}$  is adequate for processes with normal distributions and symmetric tolerances.

In this paper we propose the  $(\delta, \gamma, \varepsilon)$ -plot to evaluate the capability of a product with  $n$  independent characteristics with bilateral specifications. The proposed plot can be applied to normal or non-normal distributions and symmetric or asymmetric tolerances.

The remainder of this paper is organized as follows: in Section 2 some capability indices are presented. Then in Section 3 a process capability index for a multi-

process product with bilateral specifications is proposed. Section 4 describes the  $(\delta, \gamma, \varepsilon)$ -plot. The practical advantages of the  $(\delta, \gamma, \varepsilon)$ -plot are shown through an example in Section 5. Finally, some conclusions are given.

## 2. Process Capability Indices

Vännman (1995) defined a family of capability indices, depending on two non-negative parameters,  $u$  and  $v$ , as

$$C_p(u, v) = \frac{d - u |\mu - m|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}$$

where  $\mu$  is the process mean,  $\sigma$  the process standard deviation,  $T$  is the target value,  $d = \frac{USL - LSL}{2}$  is half of the length of the specification interval, where  $USL$  and  $LSL$  are the upper and lower specification limits respectively, and  $m = \frac{USL + LSL}{2}$  is the mid-point of the specification interval. By setting  $u$  and  $v$  equal to 0 or 1, we obtain the four basic indices, this is,  $C_p(0, 0) = C_p$ ,  $C_p(1, 0) = C_{pk}$ ,  $C_p(0, 1) = C_{pm}$  and  $C_p(1, 1) = C_{pmk}$ . The indices  $C_p(u, v)$  are appropriate for normal processes and symmetric tolerances, see Pearn, Wu & Wang (2005), but they have been shown to be inappropriate otherwise.

Pearn & Chen (1997) considered the generalization of  $C_p(u, v)$ , called  $C_{Np}(u, v)$ , which can be applied to processes with arbitrary distributions.  $C_{Np}(u, v)$  is defined as

$$C_{Np}(u, v) = \frac{d - u |M - m|}{3\sqrt{\left(\frac{P_{99.865} - P_{0.135}}{6}\right)^2 + v(M - T)^2}}$$

where  $M$  is the process median,  $P_\alpha$  is the  $(100\alpha)$ th percentile of the distribution and  $\frac{P_{99.865} - P_{0.135}}{6}$  is the process spread. If the process follows the normal distribution, then  $C_{Np}(u, v)$  reduces to  $C_p(u, v)$ .

Both,  $C_p(u, v)$  and  $C_{Np}(u, v)$  are appropriate for processes with symmetric tolerances. To handle asymmetric tolerances Vännman (1997a) proposed the index  $C_{pa}(u, v)$ , defined as

$$C_{pa}(u, v) = \frac{d - |\mu - m| - u |\mu - T|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}$$

Pearn et al. (2005) proposed the index  $C''_{Np}(u, v)$ , which is expressed as

$$C''_{Np}(u, v) = \frac{d^* - uA^*}{3\sqrt{\left(\frac{P_{99.865} - P_{0.135}}{6}\right)^2 + vA^2}} \quad (1)$$

where  $A = \max\left\{\frac{d(M-T)}{d_u}, \frac{d(T-M)}{d_l}\right\}$ ,  $A^* = \max\left\{\frac{d^*(M-T)}{d_u}, \frac{d^*(T-M)}{d_l}\right\}$ ,  $d_u = USL - T$ ,  $d_l = T - LSL$ ,  $d^* = \min(d_u, d_l)$ . This index can be applied under non-normal distributions with symmetric or asymmetric tolerances. By letting  $u = 0$  or 1 and  $v = 0$  or 1, we denote  $C''_{Np}(0, 0) = C''_{Np}$ ,  $C''_{Np}(1, 0) = C''_{Npk}$ ,  $C''_{Np}(0, 1) = C''_{Npm}$ , and  $C''_{Np}(1, 1) = C''_{Npmk}$ . If  $T = m$ , then  $C''_{Np}(u, v)$  reduces to  $C_{Np}(u, v)$ .



### 3. Process Capability Index for an Entire Product

Since a product usually holds many quality characteristics, Chen et al. (2001) defined an integrated process capability index,  $C_T$ , for independent characteristics, each of them following a normal distribution. Chen et al. (2006) modified the index  $C_T$ , to allow non-normal distributions. They assume that the product has  $t_n$  characteristics with bilateral specifications evaluated by  $N_{pn}$ ,  $t_u + t_l$  characteristics with unilateral specifications, where  $t_u$  is the number of smaller-the-better characteristics evaluated by  $N_{pu}$  and  $t_l$  is the number of larger-the-better characteristics evaluated by  $N_{pl}$ . The new index, denoted by  $N_T$  is defined as

$$N_T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\left[ \left( \prod_{i \in S} \prod_{j=1}^{n_i} [2\Phi(3N_{pij}) - 1] \right) + 1 \right]}{2} \right\}$$

where  $\Phi^{-1}$  is the inverse cumulative distribution function of the standard normal distribution,  $S = \{u, l, n\}$  denotes the type of process, smaller-the-better, larger-the-better and nominal-the-best respectively, and

$$N_{pij} = \begin{cases} \frac{USL - M}{(P_{99.865} - P_{0.135})/2}, & i = u, j = 1, 2, \dots, t_u \\ \frac{M - LSL}{(P_{99.865} - P_{0.135})/2}, & i = l, j = 1, 2, \dots, t_l \\ \frac{\min(USL - M, M - LSL)}{3\sqrt{\left[\frac{(P_{99.865} - P_{0.135})}{6}\right]^2 + (M - T)^2}}, & i = n, j = 1, 2, \dots, t_n \end{cases}$$

$N_T$  can be calculated for unilateral and bilateral specifications. However, for bilateral specifications, the processes must have symmetric tolerances.

We propose an integrated process capability index for a product with  $n$  normal or non-normal characteristics, bilateral specifications and symmetric or asymmetric tolerances. This index, denoted by  $C_{Np}^T$ , is defined as

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\prod_{j=1}^n [2\Phi(3C_{Npmk_j}^n) - 1] + 1}{2} \right\}$$

where  $C_{Npmk}^n = C_{Np}^n(1, 1)$ .

If the process capability for the entire product is greater or equal to a fixed value, let's say  $a$ , then

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\prod_{j=1}^n [2\Phi(3C_{Npmk_j}^n) - 1] + 1}{2} \right\} \geq a$$

Specifically, when the preset minimum values of process capabilities for individual characteristics are equal, we have

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\left( \left[ 2\Phi \left( 3C_{Npmk_j}^n \right) - 1 \right] \right)^n + 1}{2} \right\} \geq a$$

The critical value  $a_0$  ( $C_{Npmk_j}^n \geq a_0$ ) for individual process capability can be attained by solving the previous inequality when the integrated process yield exceeds  $a$  ( $C_{Np}^T \geq a$ ), where

$$a_0 = \frac{1}{3} \Phi^{-1} \left[ \frac{(2\Phi(3a) - 1)^{\frac{1}{n}} + 1}{2} \right] \quad (2)$$

#### 4. The $(\delta, \gamma, \varepsilon)$ -plot for a Multi-Process Product

The index  $C_{Np}^n(u, v)$  defined in (1) can be written as

$$C_{Np}^n(u, v) = \frac{1 - u|\delta|}{3\sqrt{\gamma^2 + v\varepsilon^2\delta^2}}$$

where

$$\delta = \begin{cases} \frac{M-T}{d_u} & M \geq T, \\ \frac{M-T}{d_l} & M \leq T \end{cases} \quad (3)$$

$$\gamma = \left( \frac{P_{99.865} - P_{0.135}}{6d^*} \right) \quad (4)$$

and

$$\varepsilon = \frac{d}{d^*} \quad (5)$$

$\delta$  measures the distance from the target value to the process median relative to  $d_u$  or  $d_l$ ,  $\gamma$  measures how large the process spread is relative to  $d^*$ , and  $\varepsilon$  represents the process type of tolerance.

The  $(\delta, \gamma, \varepsilon)$ -plot is a three dimensional contour plot which aims to visualize the capability of a product holding several quality characteristics. If the capability of the entire product is preset at  $a$ ,  $C_{Np}^T = a$ , the minimum process capability index,  $a_0$ , of each characteristic is calculated from equation (2).

The level curves are obtained solving the equation  $C_{Np}^n(u, v) = a_0$  with respect to  $\gamma$ . This is by, setting

$$C_{Np}^n(u, v) = \frac{1 - u|\delta|}{3\sqrt{\gamma^2 + v\varepsilon^2\delta^2}} = a_0$$

we obtain

$$\gamma = \sqrt{\frac{(1 - u|\delta|)^2}{9a_0^2} - v\varepsilon^2\delta^2} \quad (6)$$

where  $|\delta| \leq \frac{1}{u+3a_0\varepsilon\sqrt{v}}$ .

The plot provides information instantly about the location, spread and tolerance of a multiprocess product, where each characteristic has bilateral specifications and follows a normal or non-normal distribution.

The region inside the contour plot represents capable processes, while the zone located outside the level plots corresponds to non-capable processes. The  $(\delta, \gamma, \varepsilon)$ -plot depends on the values of  $u$ ,  $v$  and  $a_0$ . Figures 1a and 1b present the  $(\delta, \gamma, \varepsilon)$ -plots for values of  $C_{Np}^r(0, 1)$  and  $C_{Np}^r(1, 1)$  equal to 1.0 and 2.0, according to the equation (6). These figures show us that the larger  $\varepsilon$  becomes the smaller the capable region is. For symmetric tolerances, that is  $\varepsilon = 1$ , we obtain the largest capable region. This region begins to reduce as the target value  $T$  is located far away from the mid-point of the specification interval.

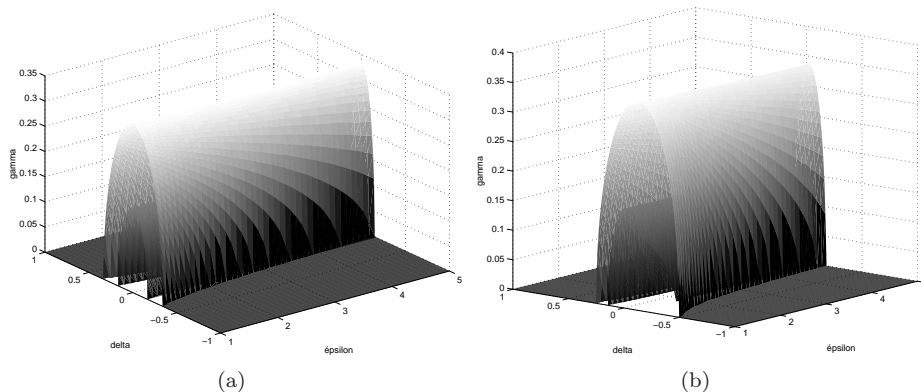


FIGURE 1: Example of a two Level curves for: (a)  $C_{Np}^r(0, 1) = 1.0, 2.0$  and (b)  $C_{Np}^r(1, 1) = 1.0, 2.0$ ; from top to bottom.

When each one of the characteristics measured in a product has symmetric tolerances, the  $(\delta, \gamma, \varepsilon)$ -plot can be visualized in two dimensions as it is shown in Figure 2. Figure 2 coincides with the graphics of Deleryd & Vännman (1999), for the case of Gaussian processes.

Any characteristic has a point  $(\delta, \gamma, \varepsilon)$  in a three-dimensional space. If the corresponding value of  $(\delta, \gamma, \varepsilon)$  is inside the capability region then the characteristic meets the preset specifications. If the value is outside, we analyse the  $\delta$ -axis to observe how far away the process median is relative to the target value, and the  $\gamma$ -axis to study the process spread. These analyses give us the key elements to be improved.

If the three-dimensional graphic is not easily visualized, we can plot two-dimensional graphics for each value of  $\varepsilon$ , so if we have  $r$  different values of  $\varepsilon$ , we construct  $r$   $(\delta, \gamma)$ -plots and analyse them. R code for implementing the proposed procedure is available from the first author upon request.

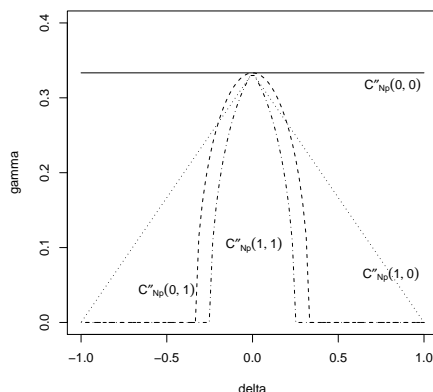


FIGURE 2:  $(\delta, \gamma, \varepsilon)$ -plot for  $C_{Np}^*(u, v) = 1$ , with  $\varepsilon=1, u=0,1$  and  $v=0,1$ .

### 5. Example

We consider the data set presented by Huang & Chen (2003) and compare their results with those obtained by our graphical method. The application consists of 10 key characteristics of a Sea island microfiber product. Table 1 gives the specifications of these characteristics. To emphasize the cases with asymmetric tolerances we have slightly modified the target value of the 8<sup>th</sup> characteristic from 3.0000 to 3.0010.

First, we set a value for the integrated process capability, say  $C_{Np}^T = 1.00$ . Then, the corresponding minimum individual process capability for each quality characteristic is  $a_0 = 1.214$ , applying equation (2). By using equations (3), (4) and (5), the  $\delta, \gamma$ , and  $\varepsilon$  coordinates for the 10 independent characteristics are calculated. See Table 1. The  $(\delta, \gamma, \varepsilon)$  points are located on the graph as shown in Figure 3.

TABLE 1: Process capabilities for Sea-island micro-fiber TS2-135N.

Characteristic	LSL	T	USL	$M$	$\sigma$	$\delta$	$\gamma$	$\varepsilon$	$C_{Np}^*(1,1)$	$C_{pmk}^*$
1. Denier	3.3000	3.5000	3.7000	3.4900	0.0500	-0.0500	0.2500	1.0	1.2421	1.2421
2. Tenacity	4.5000	5.0000	5.5000	4.8000	0.1500	-0.4000	0.3000	1.0	0.4000	0.4000
3. Elongation	47.5000	50.0000	52.5000	50.7500	0.2035	0.3000	0.0814	1.0	0.7506	0.7506
4. OPU	1.4925	1.5000	1.5075	1.4992	0.0015	-0.1067	0.2000	1.0	1.3137	1.3137
5. Crumple number	10.0000	11.0000	12.0000	11.1245	0.0341	0.1245	0.0341	1.0	2.2608	2.2608
6. Crimple	11.7600	12.0000	12.2400	12.0150	0.0500	0.0625	0.2083	1.0	1.4367	1.4367
7. Rate of crimple elasticity	40.5000	45.0000	49.5000	46.5500	0.3575	0.3444	0.0794	1.0	0.6182	0.6182
8. Water	2.9850	3.0010	3.0090	2.9980	0.0002	-0.1875	0.0250	1.5	0.9592	1.4318
9. Hot-air shrinkage	4.9000	5.0000	5.0500	4.9800	0.0090	-0.2000	0.1800	1.5	0.7622	0.9911
10. Length	47.9400	51.0000	54.0600	50.6250	0.3500	-0.1225	0.1144	1.0	1.7448	1.7448

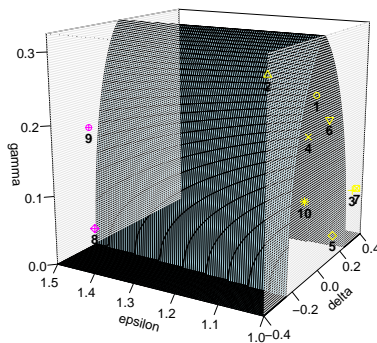


FIGURE 3:  $(\delta, \gamma, \epsilon)$ -plot for the quality status of Sea island micro-fiber TS2-135N.

Since there are two different values of  $\epsilon$ , 1 and 1.5, as it can be seen in Table 1, we can construct a  $(\delta, \gamma)$ -plot for each  $\epsilon$ . See Figures 4 and 5.

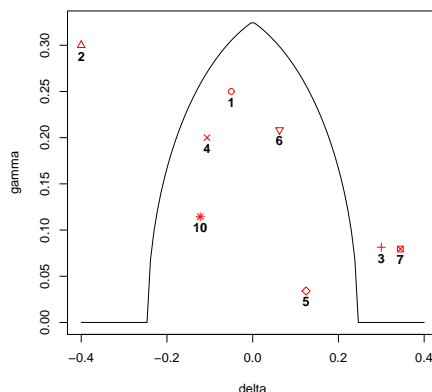


FIGURE 4:  $(\delta, \gamma)$ -plot for the quality status of Sea island micro-fiber TS2-135N when  $\epsilon=1$ .

From Figures 4 and 5 we observe that five characteristics (2, 3, 7, 8 and 9) are not within the contour line. Actions must be taken to improve the product quality. The process centers of characteristics 3, 7, 8 and 9 should be shifted toward the target to improve process capability. The spread of the distribution of characteristic 2 must be reduced and its mean moved closer to the target.

In column 9 of Table 1 we show the values of the index  $C''_{pmk}$  used by Huang & Chen (2003) and in column 8 we show the values of index  $C''_{Npmk}$  used in our proposal. Comparing columns 8 and 9 we see that both indices are equal when the tolerance is symmetric, but in presence of asymmetric tolerance the value of the index  $C''_{Npmk}$  is less than  $C''_{pmk}$ . This situation makes that the capability region used in our proposal for characteristics 8 and 9 be smaller than the capability

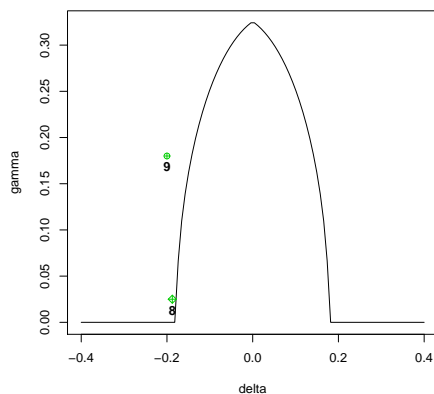


FIGURE 5:  $(\delta, \gamma)$ -plot for the quality status of Sea island micro-fiber TS2-135N when  $\varepsilon=1.5$ .

region used for the others characteristics. See Figures 4 and 5. Besides, we see that characteristic 8 is not capable ( $C_{Npmk}'' = 0.9592 < a_0$ ). In the methodology of Huang & Chen (2003), the capability region is the same for all characteristics and in particular the 8<sup>th</sup> characteristic is capable ( $C_{pmk}'' = 1.4318 > a_0$ ). When the tolerance is symmetric, the capability region is the same for the two methods, therefore there is not difference in the analysis for characteristics 2, 3 and 7.

## 6. Conclusions

Process capability plots have been proved to provide more information than process capability indices. In this paper, we have proposed a three-dimensional process capability graphic, called the  $(\delta, \gamma, \varepsilon)$ -plot. The  $(\delta, \gamma, \varepsilon)$ -plot allows simultaneous monitoring of several independent characteristics of a product with bilateral specifications and symmetric or asymmetric tolerances, which follow normal or non-normal distributions. The plot defines contour lines, which define the capability regions, and provides information about the location and spread of the characteristics. This information, provided to the practitioners, generates more efficient actions for improving the product quality. An issue, for a future research, is to extend this method to several correlated characteristics.

## Acknowledgements

The authors gratefully acknowledge the financial support from the National University of Colombia under research project 10493. The authors thank the reviewers and the editor for their valuable comments that improved this article.

[Recibido: agosto de 2010 — Aceptado: febrero de 2011]

## References

- Chen, K. S., Hsu, C. H. & Wu, C. C. (2006), 'Process capability analysis for a multi-process product', *International Journal of Advanced Manufacturing Technology* **27**(11), 1235–1241.
- Chen, K. S., Huang, M. L. & Hung, Y. H. (2008), 'Process capability analysis chart with the application of Cpm', *International Journal of Production Research* **46**(16), 4483–4499.
- Chen, K. S., Huang, M. L. & Li, R. K. (2001), 'Process capability analysis for an entire product', *International Journal of Production Research* **39**(17), 4077–4087.
- Deleryd, M. & Vännman, K. (1999), 'Process capability plots—a quality improvement tool', *Quality and Reliability Engineering International* **15**(3), 213–227.
- Huang, M. & Chen, K. (2003), 'Capability analysis for a multi-process product with bilateral specifications', *International Journal of Advanced Manufacturing Technology* **21**(10), 801–806.
- Pearn, W. L. & Chen, K. S. (1997), 'Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing', *Microelectronics Reliability* **37**(12), 1853–1858.
- Pearn, W. L., Wu, C. W. & Wang, K. H. (2005), 'Capability measure for asymmetric tolerance non-normal processes applied to speaker driver manufacturing', *International Journal of Advanced Manufacturing Technology* **25**(5), 506–515.
- Vännman, K. V. (1995), 'A unified approach to capability indices', *Statistica Sinica* **5**(2), 805–820.
- Vännman, K. V. (1997a), 'A general class of capability indices in the case of asymmetric tolerances', *Communications in statistics-theory and methods* **26**(8), 2049–2072.
- Vännman, K. V. (1997b), Interpreting process capability using graphical methods, Technical Report 7, Division of Quality Technology and Statistics, Lulea University of Technology.

## Appendix

### A1. Introducción

En los procesos de mejoramiento de calidad es importante determinar si los procesos se ajustan a unas especificaciones previamente establecidas. Los índices de capacidad de procesos, PCI por su sigla en inglés, miden este ajuste. Algunos de ellos combinan dos características básicas en los estudios de capacidad de procesos: la localización del proceso con relación a su valor objetivo y la dispersión del proceso. Un valor bajo del índice indica que el proceso no es capaz, pero desafortunadamente no expresa si esta información se debe a la variabilidad del proceso o que su localización está alejada del valor objetivo. Véase Deleryd & Vännman (1999). Esta deficiencia es superada mediante el empleo de métodos gráficos.

Vännman (1997b) introdujo un gráfico para representar la capacidad de un proceso, denominado el gráfico  $(\delta, \gamma)$ . Este gráfico requiere que el proceso tenga distribución normal, con especificaciones bilaterales y tolerancias simétricas (el valor objetivo se localiza en el punto medio del intervalo de especificación). Chen et al. (2001) desarrollaron un gráfico para analizar la capacidad de un producto, denominado PCAC. Este gráfico evalúa simultáneamente características independientes con especificaciones bilaterales o unilaterales, pero se requiere que la distribución de cada característica sea normal. Huang & Chen (2003) propusieron un gráfico de capacidad de procesos, bajo cualquier distribución pero sujeto a especificaciones bilaterales. Chen et al. (2006) plantearon un gráfico multiproceso para evaluar la capacidad de procesos no normales, pero con tolerancias simétricas. Chen et al. (2008) tomaron el PCAC planteado por Chen et al. (2001) y proponen el gráfico  $PCAC/C_{pm}$ , que puede evaluar la capacidad de un producto empleando los índices univariados básicos  $C_{pu}$ ,  $C_{pl}$  and  $C_{pm}$ . El  $PCAC/C_{pm}$  es adecuado para procesos con distribuciones normales y tolerancias simétricas.

En este artículo se propone el gráfico  $(\delta, \gamma, \varepsilon)$  para evaluar la capacidad de un producto conformado por  $n$  características independientes cuyas especificaciones son bilaterales. Este gráfico no requiere que cada proceso tenga distribución normal ni que las tolerancias sean simétricas.

El resto del artículo está organizado como sigue: en la sección A2 algunos índices de capacidad son presentados. Posteriormente, en la sección A3 un índice para la medir capacidad de un producto con especificaciones bilaterales es propuesto. En la sección A4 través de un ejemplo se presentan las ventajas del gráfico  $(\delta, \gamma, \varepsilon)$ . Este apéndice no incluye la descripción del gráfico  $(\delta, \gamma, \varepsilon)$  ni las conclusiones del artículo.



## A2. Índices de capacidad de proceso

Vännman (1995) definió una familia de índices de capacidad, dependiendo de dos parámetros no negativos,  $u$  and  $v$ , como:

$$C_p(u, v) = \frac{d - u |\mu - m|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}$$

donde  $\mu$  es la media del proceso,  $\sigma$  la desviación estándar del proceso,  $T$  el valor objetivo,  $d = \frac{USL - LSL}{2}$  la mitad de la longitud del intervalo de especificación,  $USL$  y  $LSL$  los límites de especificación superior e inferior, respectivamente; y  $m = \frac{USL + LSL}{2}$  es el punto medio del intervalo de especificación. Cuando  $u$  y  $v$  toman valores de 0 o 1 se obtienen los cuatro índices básicos:  $C_p(0, 0) = C_p$ ,  $C_p(1, 0) = C_{pk}$ ,  $C_p(0, 1) = C_{pm}$  y  $C_p(1, 1) = C_{pmk}$ . Los índices  $C_p(u, v)$  son apropiados para procesos con distribución normal y tolerancias simétricas. Véase Pearn et al. (2005).

Pearn & Chen (1997) consideran la generalización de  $C_p(u, v)$ , denominada  $C_{Np}(u, v)$ , la que puede ser aplicada a procesos con distribución arbitraria.  $C_{Np}(u, v)$  es definido como,

$$C_{Np}(u, v) = \frac{d - u |M - m|}{3\sqrt{\left(\frac{P_{99.865} - P_{0.135}}{6}\right)^2 + v(M - T)^2}}$$

donde  $M$  es la media del procesos,  $P_\alpha$  es el  $(100\alpha)$  percentil de la distribución y la fracción  $\frac{P_{99.865} - P_{0.135}}{6}$  constituye la dispersión del proceso. Si el proceso sigue una distribución normal, entonces  $C_{Np}(u, v)$  se reduce a  $C_p(u, v)$ . Ambos,  $C_p(u, v)$  y  $C_{Np}(u, v)$  son apropiados para procesos con tolerancia simétrica. Para trabajar con tolerancias asimétricas, Vännman (1997a) planteó el índice  $C_{pa}(u, v)$ , definido como:

$$C_{pa}(u, v) = \frac{d - |\mu - m| - u |\mu - T|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}$$

Pearn et al. (2005) proponen el índice  $C''_{Np}(u, v)$ , que se expresa como:

$$C''_{Np}(u, v) = \frac{d^* - uA^*}{3\sqrt{\left(\frac{P_{99.865} - P_{0.135}}{6}\right)^2 + vA^2}}$$

donde  $A = \max\left\{\frac{d(M-T)}{d_u}, \frac{d(T-M)}{d_l}\right\}$ ,  $A^* = \max\left\{\frac{d^*(M-T)}{d_u}, \frac{d^*(T-M)}{d_l}\right\}$ ,  $d_u = USL - T$ ,  $d_l = T - LSL$ ,  $d^* = \min(d_u, d_l)$ . Este índice puede ser aplicado a procesos con distribuciones no normal y con tolerancia simétrica o asimétrica. Cuando  $u$  o  $v$  toman valores de 0 o 1, este índice puede ser denotados como  $C''_{Np}(0, 0) = C''_{Np}$ ,  $C''_{Np}(1, 0) = C''_{Npk}$ ,  $C''_{Np}(0, 1) = C''_{Npm}$ , y  $C''_{Np}(1, 1) = C''_{Npmk}$ . Si  $T = m$ , then  $C''_{Np}(u, v)$  se reduce a  $C_{Np}(u, v)$ .

### A3. Índice de capacidad de procesos para todo un producto

Puesto que un producto usualmente contiene varias características, Chen et al. (2001) definen un índice de capacidad integrado,  $C_T$ , para características independientes, cada una de las cuales sigue una distribución normal. Chen et al. (2006) modifican el índice  $C_T$  para permitir procesos con distribuciones diferentes a la normal. Ellos asumieron que el producto tiene  $t_n$  características con especificaciones bilaterales evaluadas por el índice  $N_{pn}$ ,  $t_u + t_l$  características con especificaciones unilaterales, donde  $t_u$  es el número de especificaciones tales que cuanto más pequeña sea es mejor la característica evaluadas por el índice  $N_{pu}$ , y  $t_l$  es el número de características tales que cuanto mayor sea el valor que tomen mejor es la característica. Estas son evaluadas por el índice  $N_{pl}$ . El nuevo índice es denotado por  $N_T$  y es definido como

$$N_T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\left[ \left( \prod_{i \in S} \prod_{j=1}^{n_i} [2\Phi(3N_{pij}) - 1] \right) + 1 \right]}{2} \right\}$$

donde  $\Phi^{-1}$  es la función de distribución acumulada inversa de una distribución normal estándar,  $S = \{u, l, n\}$  denota el tipo de proceso, cuanto más pequeño mejor, cuanto más grande mejor y procesos con especificaciones bilaterales, respectivamente, y

$$N_{pij} = \begin{cases} \frac{USL - M}{(P_{99,865} - P_{0,135})/2}, & i = u, j = 1, 2, \dots, t_u \\ \frac{M - LSL}{(P_{99,865} - P_{0,135})/2}, & i = l, j = 1, 2, \dots, t_l \\ \frac{\min(USL - M, M - LSL)}{3\sqrt{\left[\frac{(P_{99,865} - P_{0,135})}{6}\right]^2 + (M - T)^2}}, & i = n, j = 1, 2, \dots, t_n \end{cases}$$

$N_T$  puede ser calculado para especificaciones unilaterales y bilaterales. Sin embargo, para especificaciones bilaterales, el proceso debe tener tolerancia simétrica.

Se propone un índice de capacidad integrado para un producto con características normales o no normales, especificaciones bilaterales y tolerancias simétricas o asimétricas. Este índice, denotado por  $C_{Np}^T$ , es definido como

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\prod_{j=1}^n [2\Phi(3C_{Npmk_j}^T) - 1] + 1}{2} \right\}$$

donde  $C_{Npmk}^T = C_{Np}^T(1, 1)$ .

Si el índice de capacidad para todo el producto es mayor o igual a valor fijo, digamos  $a$ , entonces

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\prod_{j=1}^n \left[ 2\Phi \left( 3C_{Npmk_j}'' \right) - 1 \right] + 1}{2} \right\} \geq a$$

Específicamente, cuando los valores mínimos preestablecidos para las capacidades de proceso para cada característica individual son iguales, se tiene:

$$C_{Np}^T = \frac{1}{3} \Phi^{-1} \left\{ \frac{\left( \left[ 2\Phi \left( 3C_{Npmk_j}'' \right) - 1 \right] \right)^n + 1}{2} \right\} \geq a$$

El valor crítico  $a_0$  ( $C_{Npmk_j}'' \geq a_0$ ) para la capacidad individual de procesos puede ser obtenido solucionando la desigualdad previa cuando la capacidad integrada del producto excede a ( $C_{Np}^T \geq a$ ), donde

$$a_0 = \frac{1}{3} \Phi^{-1} \left[ \frac{(2\Phi(3a) - 1)^{\frac{1}{n}} + 1}{2} \right]$$

## A4. Ejemplo

Se consideran los datos presentados por Huang & Chen (2003) y comparamos sus resultados con los obtenidos en la propuesta. La aplicación consta de 10 características fundamentales de una microfibra de Sea island. La tabla 1 presenta las especificaciones de estas características. Para enfatizar el caso con tolerancia asimétrica, se ha modificado ligeramente el valor objetivo de la característica 8 de 3.0000 a 3.0010.

Primero, se establece un valor para la capacidad integrada del proceso, digamos  $C_{Np}^T = 1,00$ . Entonces, la mínima capacidad de proceso individual para cada característica es  $a_0 = 1,214$ , aplicando la ecuación (2). Usando las ecuaciones (3), (4) y (5), se calculan las coordenadas para las 10 características independientes (tabla 1). Los puntos ( $\delta$ ,  $\gamma$ ,  $\varepsilon$ ) son ubicados en el gráfico como se muestra en la figura 3.

Dado que hay dos valores diferentes de  $\varepsilon$ , 1 y 1.5, como se puede observar en la tabla 1, se construye un gráfico ( $\delta$ ,  $\gamma$ ) para cada valor de  $\varepsilon$ , ver figuras 4 and 5.

A partir de las figuras 4 y 5 se observa que cinco características (2, 3, 7, 8 y 9) no están dentro de la línea de contorno; por consiguiente deben emprenderse acciones para mejorar la calidad del producto. Así, las medianas de las características 2, 3, 7, 8 y 9 deben desplazarse hacia el valor objetivo, y debe reducirse la dispersión de la característica 2.

En la columna 9 de la tabla 1 se muestran los valores del índice  $C_{pmk}''$  usado por Huang & Chen (2003); en la columna 8 se muestran los valores del índice  $C_{Npmk}''$

empleado en esta propuesta. Comparando las columnas 8 y 9 se observa que ambos índices son iguales cuando la tolerancia es simétrica, pero en presencia de tolerancia asimétrica el valor del índice  $C_{Npmk}^*$  es menor que el valor del índice  $C_{pmk}^*$ . Esta situación hace que la región de capacidad usada en nuestra propuesta para las características 8 y 9 sea más pequeña que la región de capacidad usada para las otras características (figuras 4 y 5). Además, se observa que la característica 8 no es capaz ( $C_{Npmk}^* = 0,9592 < a_0$ ). En la metodología de Huang & Chen (2003), la región de capacidad es igual para todas las características; en particular la característica 8 es capaz ( $C_{pmk}^* = 1,4318 > a_0$ ). Cuando la tolerancia es simétrica, la región de capacidad es igual para los dos métodos; por consiguiente, no hay diferencia en el análisis para las características 2, 3 y 7.



## Skeptical and Optimistic Robust Priors for Clinical Trials

Ensayos clínicos bajo un enfoque bayesiano robusto con previas escépticas y optimistas

JOHN COOK<sup>1,a</sup>, JAIRO FÚQUENE<sup>2,b</sup>, LUIS PERICCHI<sup>3,c</sup>

<sup>1</sup>DIVISION OF QUANTITATIVE SCIENCES, M.D. ANDERSON CANCER CENTER, UNIVERSITY OF TEXAS, HOUSTON

<sup>2</sup>INSTITUTE OF STATISTICS, SCHOOL OF BUSINESS ADMINISTRATION, UNIVERSITY OF PUERTO RICO, SAN JUAN, PUERTO RICO

<sup>3</sup>DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PUERTO RICO, SAN JUAN, PUERTO RICO

---

### Abstract

A useful technique from the subjective Bayesian viewpoint, is to ask the subject matter researchers and other parties involved, such as pharmaceutical companies and regulatory bodies, for reasonable optimistic and pessimistic priors regarding the effectiveness of a new treatment. Up to now, the proposed skeptical and optimistic priors have been limited to conjugate priors, though there is no need for this limitation. The same reasonably adversarial points of view can take with robust priors. Robust priors permit a much faster and efficient resolution of the disagreement between the conclusions based on skeptical and optimistic priors. As a consequence, robust Bayesian clinical trials tend to be shorter. Our proposal in this paper is to use Cauchy and intrinsic robust priors for both skeptical and optimistic priors leading to results more closely related with the sampling data when prior and data are in conflict. In other words, the use of robust priors removes the dogmatism implicit in conjugate priors. Dogmatism here has very precise meaning: Conjugate priors affect the posterior conclusions by a fixed rate, regardless if there is a conflict between prior and data. Robust priors are automatically discounted by Bayes Theorem in the presence of conflict.

**Key words:** Clinical trials, Bayesian robustness, Prior distribution.

---

<sup>a</sup>Professor. E-mail: [jdcook@mdanderson.org](mailto:jdcook@mdanderson.org)

<sup>b</sup>Instructor. E-mail: [jairo.a.fuquene@uprrp.edu](mailto:jairo.a.fuquene@uprrp.edu)

<sup>c</sup>Professor. E-mail: [luarpr@gmail.com](mailto:luarpr@gmail.com)

### Resumen

Ensayos clínicos bajo un enfoque de estadística Bayesiana están adquiriendo cada vez mayor importancia. Anteriormente se sugirió una idea que ha dado ventaja al uso de previas bayesianas: suponer dos densidades a priori con información muy distinta sobre la efectividad de un nuevo tratamiento: una previa escéptica que refleje pesimismo sobre la superioridad del tratamiento nuevo. Esta es la posición de los organismos reguladores que autorizan o no los nuevos medicamentos que deben ser vendidos en el mercado. También proponen una densidad previa optimista (o “entusiasta”) que refleja la posición de los investigadores y compañías farmacéuticas que proponen el nuevo tratamiento. Esta diversidad de opiniones es sumamente útil y queremos preservarla. En la propuesta original, sin embargo, se supone que la forma funcional de las densidades previas es normal, lo cual es más simple de analizar y asignar. Infortunadamente, la simplicidad exige un precio muy grande. Para que haya acuerdo entre los dos análisis se necesita muchísima información muestral. En este artículo se propone mantener las dos previas que representan puntos de vista adversos, pero con una forma funcional de colas gruesas, como la densidad Cauchy. Las previas robustas permiten una resolución del desacuerdo de forma más rápida y eficiente; por tanto, los ensayos clínicos tienden a ser más cortos. Se asume un concepto de dogmatismo muy preciso: las previas conjugadas afectan las conclusiones en una tasa fija, sin importar si hay conflicto entre la verosimilitud y la densidad previa. Además, las previas robustas de colas gruesas son automáticamente descontadas en presencia de conflicto, por el teorema de Bayes en favor de la información dada por los datos.

**Palabras clave:** distribución *a priori*, ensayos clínicos, robustez bayesiana.

## 1. Introduction

Clinical trials are contentious. Pharmaceutical companies are eager to show that their new drug, on which they may have been invested millions of dollars, is a substantial improvement over the current treatment. Government regulatory agencies take the opposite view and ask for substantial evidence that the new drug is not less effective than the current standard.

Bayesian statistics permits a useful technique for modeling adversarial positions by using *two* prior distributions on the parameters of interest. One, the optimistic prior, corresponds to positive expectations. The other, the pessimistic prior, corresponds to a more skeptical position.

To be more specific, let us consider the following one-sided hypothesis test. Suppose that we have two medical treatments with corresponding probabilities of events  $p_1$  and  $p_2$ . The events may be for instance “disease recurrence” or “deaths”. The Odds Ratio is defined as

$$OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$$

and to make the assumption of normality more realistic, the focus is on the parameter  $\theta = \log(OR)$ . Often the test of interest is the following:

$$H_0 : \theta \geq \theta_H, \text{ vs } H_1 : \theta < \theta_H \quad (1)$$

Here,  $H_1$  is the region of substantial improvement achieved by the new treatment over the standard one. Here,  $\theta_H < 0$ , corresponding to  $OR < 1$ , for example  $OR = 0.5$ , the point of 50% of improvement.

Part of the clever proposal by Spiegelhalter, Abrams & Myles (2004) is to assume two priors:  $\pi_S$  and  $\pi_O$ . The skeptical position (say the government regulatory body attitude) is to first center the prior on the point of no difference, i.e.  $\theta = 0$ , and secondly to give only a small probability, denoted by  $\xi$ , say  $\xi = 0.05$  or  $\xi = 0.025$  to the improvement (i.e  $H_1$ ), that is

$$P_S(H_1) = \int_{-\infty}^{\theta_H} \pi_S(\theta) d\theta = \xi \quad (2)$$

The optimistic position is centered on  $\theta_H$  and it gives only probability  $\xi$  that there is no benefit with the new treatment (the investigator position), so

$$P_O(\theta > 0) = \int_0^{\infty} \pi_O(\theta) d\theta = \xi \quad (3)$$

If under the skeptical prior, the posterior probability of  $H_1$  is larger than  $1 - \gamma$ , say  $1 - \gamma = 0.95$ , that is, if under the skeptical prior  $P(H_1 | \text{Data}, \pi_S) > 1 - \gamma$ , then to decide in favor of  $H_1$  is safe. If with the optimistic prior  $P(\theta > 0 | \text{Data}, \pi_O) > 1 - \gamma$ , to decide that there is not benefit with the new treatment is safe. If there is disagreement between the procedures then the trial is inconclusive, given the information so far collected.

This approach is intuitively satisfying. However, this framework may be overly cautious and lead to an enormous delay in the decisions. The implementation proposed in Spiegelhalter et al. (2004) uses conjugate priors, which lead to simple computations. However, we show in this article that we may preserve the useful framework of skeptical and optimistic priors without the dogmatism inherent in conjugate priors.

Pericchi & Smith (1992) showed some aspects of the robustness of the Student- $t$  prior for a normal location parameter and provided approximations to the posterior moments in the model Student- $t$ /normal. The Cauchy prior, as a Student- $t$  with one degree of freedom, can be used in this context as well. However, for normal log-odds there is a robust prior proposed by Berger (1985) that leads to a closed form posterior and moments, a sort of the “best of both worlds”. The “intrinsic prior” was obtained in Berger & Pericchi (1996) as the implicit prior to which the arithmetic intrinsic Bayes factor converges, and it turns out that it is a limiting case of Berger’s prior. In Fúquene, Cook & Pericchi (2009), robust priors are proposed for clinical trials. But the priors studied there are clinical priors based on previous trials. Here, we propose a class of robust priors, a novel proposal to the



best of our knowledge, that improves the class of Normal skeptical and optimistic priors originally proposed by Spiegelhalter et al. (2004). Note that the original proposal by Spiegelhalter et al. (2004) propose a class of priors of two elements: the skeptical and the optimistic normal priors. We propose a class of robust priors, again of two elements: the skeptical and the optimistic robust prior. Effectively, however, both these classes allow a substantial variation on prior assessments. We illustrate the improvement in terms of an example and also provide mathematical results that lead us to expect that the speed of convergence of the robust class is substantially higher than the normal class.

## 2. Illustration: Skeptical and Optimistic Robust Priors

A useful suggestion under the *subjective Bayesian viewpoint*, taken by Spiegelhalter, Freedman & Parmar (1994), is to ask the subject matter researchers, for reasonably optimistic and pessimistic priors (regarding the effectiveness of a new treatment). On the log-odds scale, a skeptical prior on the amount of improvement has mean zero (i.e., no difference between treatments,  $\theta = 0$ ) and a substantial probability that the new treatment is not better. The prior scale is assessed in reference to an optimistic hypothesis  $\theta_H$  (say  $\theta_H = -0.69$ , which corresponds to a 50% improvement, see below). Then a small (skeptical) probability  $\xi$  is assessed, for example  $\xi = 0.025$ : that the effect of the treatment is equal or better than  $\theta_H$ . For the Cauchy prior, the skeptical parameters are very easy to assess. The location  $\mu$  is zero and the scale  $\beta = \theta_H / \tan(\pi(\xi - 1/2))$ . In Fúquene et al. (2009), other robust heavy tailed priors are considered, apart from the Cauchy, and several mathematical results are presented. Two such priors are the Intrinsic Prior, calculated by Berger & Pericchi (1996), and the so called Berger's prior (Berger 1985). There is a close relationship between the Intrinsic and Berger's prior, and for the most part we present the results for the Intrinsic prior, for which the location ( $\mu$ ) is zero and the scale,  $\tau$ , is found by solving the following equation

$$\int_{-\infty}^{\theta_H} \frac{1}{2\sqrt{\pi}} \frac{1 - \exp\left\{-\frac{(\theta - \mu)^2}{\tau^2}\right\}}{(\theta - \mu)^2/\tau} d\theta = \xi \quad (4)$$

For the optimistic priors, the assessments are similar, except that the location is changed to  $\mu = \theta_H = -0.69$ , that is the location now placed at the point of 50% improvement. The scales are obtained by placing  $1 - \xi$  probability on  $H_1$ . For example, in equation (3) the right hand side is changed from  $\xi$  to  $1 - \xi$ .

**Illustrative example.** Suppose that a new drug is compared with a conventional treatment. With the conventional treatment, 26 out of 97 patients died (26.8%). Only 13 out of 193 patients died with the new drug (6.7%). We wish to test the hypothesis that the improvement with the new treatment is at least 50%

of reduction of the odds of death. Is there enough evidence to make a skeptical prior to give over 90% of posterior probability?

We now make the comparison of two analysis: The first combining the normal log odds coupled with a normal skeptical prior (with the current data). Alternatively, we assume heavy-tailed Cauchy and intrinsic priors, with the same location and the same probabilities of the region of interest as the normal prior. Thus we assume that the normal, Cauchy, and intrinsic skeptical priors have all mean  $\log(OR) = 0$  and with a 95% interval running from 50% reduction in odds of death ( $\log(OR) = -0.69$ ) to a 50% increase ( $\log(OR) = 0.69$ ).

In the normal prior case, we assume  $\pi_n(\theta) = N[0, \sigma^2/n_0]$ , and  $n_0$  is referred to as “the number of prior observations”. On a  $\log(OR)$  scale, this prior has a 95% interval from  $-0.69$  to  $0.69$  and so has a standard deviation  $0.69/1.96 = 0.35$  (for  $\xi = 0.025$ ) and hence the number of prior observations is  $n_0 = 4/0.35^2 = 32.3$ .

For Cauchy and intrinsic priors the scale is 0.05 and 0.06 respectively, so that the prior probability of the region on which there is a reduction of risk of 50% or more is 0.025. We use the normal approximation for binary data for the log-odds with the approximate standard error recommended in Spiegelhalter et al. (2004, pp. 26) for  $2 \times 2$  tables, following their suggestion of an standard error of the likelihood normal and normal/normal posterior model equal to  $\sigma = 2$ .

As we anticipated above, assume that the evidence arising from the study about 30-day mortality was 26/97 on control and 13/193 on new treatment. If the ratio of the odds of death following the new treatment to the odds of death on the conventional is  $OR < 1$  then the data favors the new treatment. We have that the estimated  $\log(OR)$  is  $\bar{y}_n = -1.6$  ( $OR = 0.2$  or 80% risk reduction) with estimated standard error 0.36 and  $n = 4/0.36^2 = 30.5$ , which in this case is approximately the same weight of the prior  $n_0 = 32.3$ , as calculated above.

We use the R (R Development Core Team 2009) package named ClinicalRobustPriors, available from the Comprehensive R Archive Network at <http://CRAN.R-project.org/package=ClinicalRobustPriors>, developed by one of the authors, which can be used to compute probabilities and figures for the prior, likelihood and posterior models.

The posterior mean in the normal/normal model is  $(n_0\mu + n\bar{X}_n)/(n_0 + n) = -0.77$  with standard deviation  $\sigma/\sqrt{n_0 + n} = 0.25$ ; the estimated odds ratio is  $e^{-0.77} = 0.46$  or 54% risk reduction.

In the Cauchy/normal and intrinsic/normal, models the posterior mean is  $-1.48$  ( $e^{-1.48} = 0.22$  or 78% risk reduction) with standard deviation 0.32. In the normal/normal an 95% credible interval on the  $\log(OR)$  scale is between  $-1.27$  and  $-0.28$ , that corresponds to odds ratios from 0.28 to 0.75, or a 95% probability that the true risk reduction lies between 25% and 72%. For the Cauchy/normal and intrinsic/normal posterior models, the 95% credible interval shows that the true risk reduction lies between 57% and 88%. The likelihood shows a risk reduction between 60% and 90%.

In Figure 1 we can see that the normal skeptical prior is more dogmatic than both the Cauchy and intrinsic skeptical priors. When there is a discrepancy be-

tween prior and data, the robust priors are discarded to some extent in favor of the likelihood, but the normal prior is not.

As a counterbalance to the skeptical priors, Spiegelhalter et al. (1994) suggest an “enthusiastic” (or “optimistic” as we call it here) prior centered on the alternative hypothesis and with a high chance that the true treatment benefit is over 50%. In this example, the alternative hypothesis is  $\theta_H = -0.69$  (50% risk reduction) and the high chance is  $1 - \xi = 0.975$ . The scale and prior sample size are the same as with the skeptical priors. Figure 2 display the results. The posterior mean in the normal/normal conjugate model is  $-1.13$  ( $e^{-1.13} = 0.32$  or 68% risk reduction), much closer than the  $-1.39$  ( $e^{-1.39} = 0.24$  or 76% risk reduction) of the Cauchy/normal or intrinsic/normal models. The scale for the posterior models are 0.25 and 0.34, respectively, with normal and Cauchy (or intrinsic) priors. For the Optimistic prior, there is no strong conflict with the likelihood and the inference with normal and robust priors is similar.

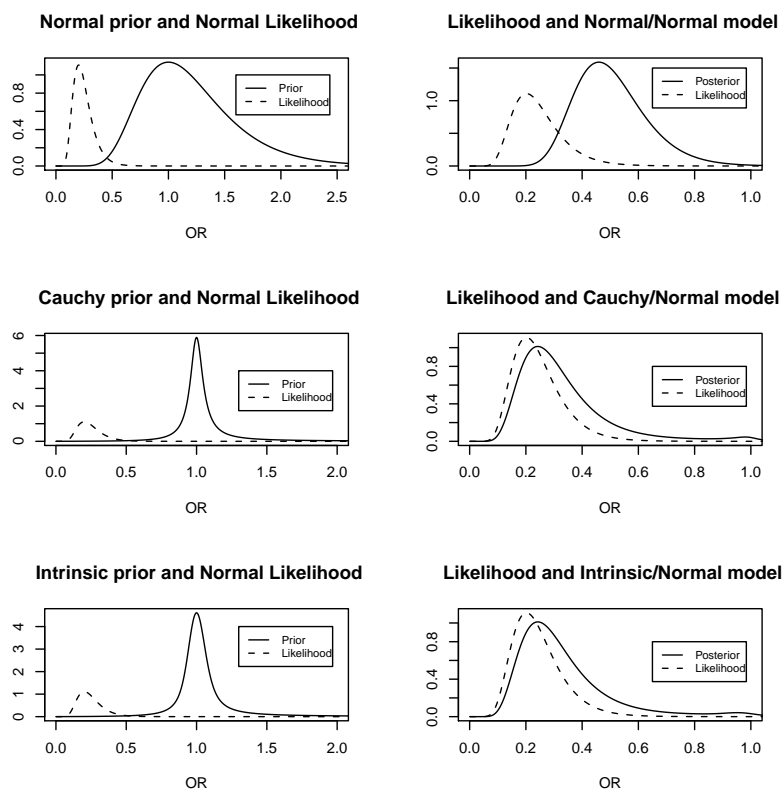


FIGURE 1: Skeptical priors, likelihoods and posterior models: normal/normal, cauchy/normal and intrinsic/normal.

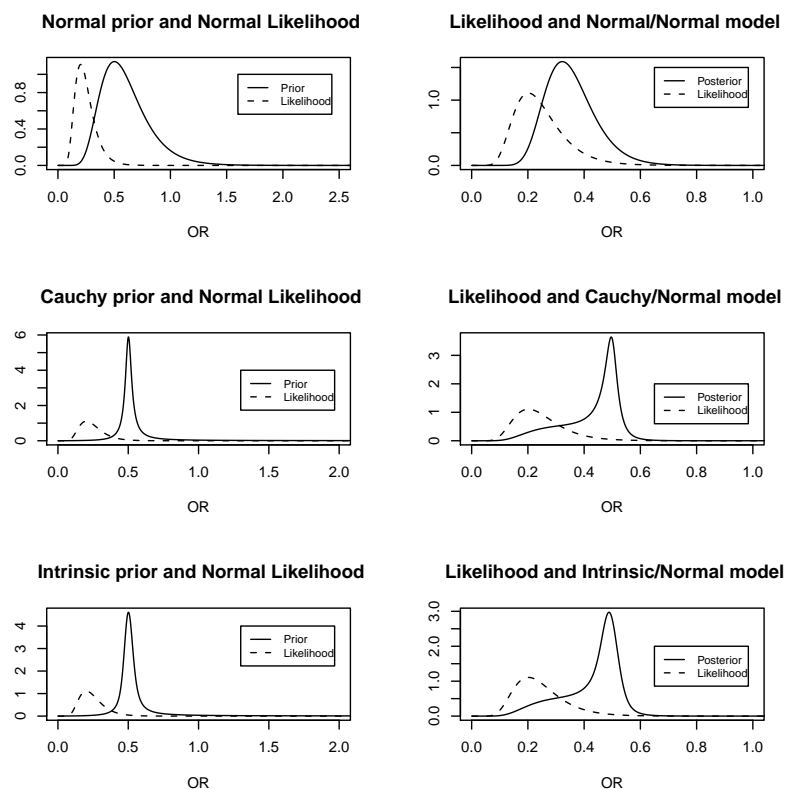


FIGURE 2: Optimistic priors, likelihoods and posterior models: normal/normal, cauchy/normal and intrinsic/normal.

**Conclusion for the example.** The quantity of interest is the set of a reduction of risk of at least of 50%, that is  $H_1$ . For the normal skeptical prior, the posterior probability can be calculated in the log-odds scale as  $\Phi((-0.69 + 0.77)/0.25) = \Phi(0.32) = 0.62$ . This contrast with both the Cauchy and intrinsic skeptical robust priors, for which this probability is 0.95. In other words, the hypothesis  $H_1$  is supported by the data, even assuming the Cauchy and intrinsic skeptical robust priors. This is not the case with the normal skeptical prior. Next consider the optimistic priors case. This probability with the normal prior is  $\Phi((0 + 1.13)/0.25) = \Phi(4.52) = 0.99$ . The same probability is obtained with both the Cauchy and the intrinsic. Thus, in the example, the skeptical and optimistic robust priors have converged in over 95% of probability of the region of substantial improvement. Unfortunately, using the normal priors, the procedure has not yet converged, and much more evidence is still required for a definitive conclusion, and this is due entirely to the particular functional form chosen for the prior, and not to its condition of skepticism. In our experience, the results obtained by using any of the considered robust priors, Cauchy or intrinsic or Berger's prior, yield essentially the same results.

In next section, we demonstrated the behavior observed in the illustration is a general feature of robust and normal priors, rather than specific phenomenon of this example.

### 3. Asymptotic Results

In this section we show that robust priors can “change their mind” more readily than conjugate priors by looking at asymptotic properties of posterior means under each type of prior. When both the skeptical and optimistic prior are robust, they may reach agreement more quickly than if the two priors were not robust.

Consider a single sample  $y$  from a normal( $\theta$ ,  $\sigma^2$ ) distribution where  $\theta$  has a conjugate normal( $0$ ,  $\tau^2$ ) prior. It is well-known that the posterior distribution on  $\theta$  has mean

$$\frac{\tau^2}{\tau^2 + \sigma^2}y.$$

If we use a Cauchy( $0$ ,  $1$ ) prior rather than a normal( $0$ ,  $\tau^2$ ) prior on  $\theta$  above, the posterior mean of  $\theta$  is

$$y - \mathcal{O}\left(\frac{1}{y}\right)$$

as  $y \rightarrow \infty$ . See the technical report Cook (2010) for the derivation of this asymptotic result.

Under the normal prior, the posterior mean of  $\theta$  under-estimates the data  $y$  by a constant ratio that depends on the relative scales of the sampling distribution and the prior distribution. Under the Cauchy prior, however, the posterior mean of  $\theta$  asymptotically approaches the value of  $y$ , independent of the scales of the sampling and prior distributions. This is illustrated by Figure 3.

Next we consider the case of  $n$  samples  $y_i$  with mean  $\bar{y}$  and consider  $n \rightarrow \infty$  the behavior of the posterior mean. Under a normal( $0$ ,  $\tau^2$ ) prior, the posterior mean is given by

$$\frac{1}{1 + \frac{\sigma^2}{n\tau^2}}\bar{y} = \left(1 - \frac{\sigma^2}{n\tau^2} + \mathcal{O}\left(\frac{1}{n^2}\right)\right)\bar{y}$$

Under a Cauchy prior, however, the posterior mean is

$$\bar{y} + \frac{(\bar{y}^2 - 3)\bar{y}}{(1 + \bar{y}^2)^2} \frac{\sigma^2}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

The rate at which the posterior mean converges to  $\bar{y}$  depends on  $\tau$  in the case of the normal prior and on  $\bar{y}$  in the case of the Cauchy prior. For any value of  $\tau$ , the convergence is faster under the Cauchy prior for sufficiently large values of  $\bar{y}$ .

We now obtain results for Berger’s prior that are quite similar to those above for the Cauchy prior. If we observe  $n$  samples from a normal( $\theta$ ,  $\sigma^2$ ) distribution

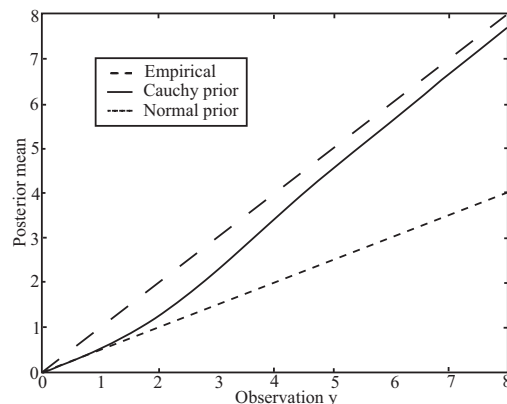


FIGURE 3: Posterior mean of  $\theta$  under robust and conjugate priors compared to empirical value.

and  $\theta$  has Berger's prior with location 0, the posterior mean of  $\theta$  is

$$\frac{2\sigma^2\bar{y}}{(\sigma^2 + n\beta^2) \left( \exp\left(\frac{n\bar{y}}{\sigma^2 + n\beta^2}\right) - 1 \right)} - \frac{2\sigma^2}{n\bar{y}}$$

where  $\bar{y}$  is the sample mean of the observations as shown in Fúquene et al. (2009). Therefore as a single sample  $y \rightarrow \infty$ , the posterior mean of  $\theta$  is

$$y - \frac{2\sigma^2}{y} + \mathcal{O}(\exp(-y^2))$$

Also, as  $n \rightarrow \infty$ , the posterior mean of  $\theta$  is

$$\bar{y} - \frac{\sigma^2}{n} \frac{2}{\bar{y}} + \mathcal{O}(\exp(-n))$$

These results show that, for Cauchy and Berger priors, the influence of the prior diminishes more quickly when the data are far from the location parameter of the prior. When skeptical and optimistic priors using either of these robust distributions, the priors will reach consensus more quickly than corresponding conjugate priors when the data are in conflict with one of the priors.

## 4. Conclusions

The idea of skeptical and optimistic priors is an important one, and useful to increase the impact of Bayesian statistics in the whole of medical statistics. As introduced by Spiegelhalter et al. (2004), this idea has two components. The first is that the prior probability of the region where the new treatment is substantially better than the current treatment is high for the optimistic prior and low for the skeptical prior. The second component is the specific shape of the priors

(effectively a class of two priors), that was assumed normal priors by Spiegelhalter et al. (2004). We show here, to the best of our knowledge for the first time, that the behavior of the procedure, even complying with the first component, crucially depends on the shape of the priors, particularly the tail size. We illustrate that with an example, similar as those in Spiegelhalter et al. (except that the equivalent sample sizes of the likelihood and priors are about the same in our example) and show it mathematically in Section 3. Furthermore, we put forward that the idea of skeptical and optimistic prior is actually improved by the use of heavy tailed priors, because the procedure is “less stubborn” and even though it places a handicap on both hypothesis, it is more willing to change its mind when there is clear information in favor of one of the hypothesis. As a consequence, the robust procedure put forward here is able to embrace better treatments or to discard bad ones in a more efficient manner: it is more likely that the optimistic and skeptical procedure reach a consensus under robust heavy tailed priors than under normal priors. Finally, the use of different heavy tailed priors presented here, the Cauchy, Berger’s or Intrinsic priors lead to essentially equivalent statistical behavior, and thus the choice among them is more a matter of convenience and taste. But the difference between normal and robust skeptical/optimistic priors is substantial, and the robust version of the idea seems to us much more acceptable for the players involved.

## Acknowledgements

We are grateful to two anonymous referees and to the special editor for their constructive suggestions that made our article much clearer and easier to read, and to the editor in chief for inviting us to send our work to the *Revista Colombiana de Estadística*. This work was supported by a grant from the pharmaceutical MERCK in Puerto Rico, Comprehensive Cancer Center of the University of Puerto Rico and by NIH Grant: P20-RR016470. Jairo Fúquene was supported by PII grant - School of Business Administration, UPR-RRP.

[Recibido: septiembre de 2010 — Aceptado: febrero de 2011]

## References

- Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, second edn, Springer-Verlag, New York, United States.
- Berger, J. O. & Pericchi, L. R. (1996), ‘The intrinsic Bayes factor for model selection and prediction’, *Journal of the American Statistical Association* **91**, 112–115.
- Cook, J. D. (2010), ‘Asymptotic results for normal-cauchy model’. UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series. Working Paper 61.  
\*<http://www.bepress.com/mdandersonbiostat/paper61>, Web 2 March 2011

- Fúquene, J., Cook, J. & Pericchi, L. R. (2009), 'A case for robust Bayesian priors with applications to clinical trials', *Bayesian Analysis* (4), 817–846.
- Pericchi, L. R. & Smith, A. F. M. (1992), 'Exact and approximate posterior moments for a normal location parameter', *Journal of the Royal Statistical Society, Series B* **54**, 793–804.
- Spiegelhalter, D. J., Abrams, K. R. & Myles, J. P. (2004), *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*, Wiley, London.
- Spiegelhalter, D. J., Freedman, L. S. & Parmar, M. K. B. (1994), 'Bayesian approaches to randomized trials (with discussion)', *Journal of the Royal Statistical Society, Series A* **157**, 357–416.

## Appendix

### A1. Resumen

Los ensayos clínicos son contenciosos y controversiales. Las compañías farmacéuticas están ansiosas de probar que el tratamiento nuevo, en el que han invertido fortunas, es una mejora sustancial sobre el tratamiento actual. Por el contrario, las agencias reguladoras gubernamentales adoptan la posición contraria y exigen, por evidencia inequívoca, que el nuevo tratamiento sea sustancialmente mejor que el actual.

La estadística bayesiana permite emplear una técnica muy útil para modelar posiciones encontradas, usando no una sino dos distribuciones *a priori* sobre los parámetros. Una es la previa optimista, que corresponde a la expectativa favorable. La otra, la previa pesimista, que corresponde a una posición escéptica.

Para ser mas específicos, consideremos el siguiente test unilateral. Supongamos que tenemos dos tratamientos médicos que corresponden a probabilidades de eventos  $p_1$  y  $p_2$ . Los eventos pueden ser, por ejemplo, "recurrencia del evento" o "muerte". La Tasa de Chances (Odds Ratio) está definida como

$$OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$$

y para hacer el supuesto de normalidad más realista consideramos los log odds:  $\theta = \log(OR)$ . Con frecuencia, el test de interés es

$$H_0 : \theta \geq \theta_H, \text{ vs } H_1 : \theta < \theta_H \quad (5)$$

Acá,  $H_1$  es la región de mejoría sustancial del nuevo tratamiento sobre el standard. Entonces,  $\theta_H < 0$ , corresponde a  $OR < 1$ ; por ejemplo,  $OR = 0,5$  es el punto de 50 % de mejoría.

La primera parte de la propuesta de Spiegelhalter et al. (2004) es suponer dos densidades *a priori*,  $\pi_S$  and  $\pi_O$  (escéptica y optimista respectivamente), que



cumplan con las siguientes especificaciones. La escéptica se centra en el punto de no diferencia, es decir, en  $\theta = 0$ , y le da una probabilidad pequeña, denotada por  $\xi$ , digamos  $\xi = 0,05$  o  $\xi = 0,025$ , a la hipótesis alternativa de mejoría sustancial, es decir a  $H_1$ , esto es:

$$P_S(H_1) = \int_{-\infty}^{\theta_H} \pi_S(\theta) d\theta = \xi \quad (6)$$

Por otro lado, la previa optimista esta centrada en  $\theta_H$  y le da probabilidad  $\xi$ , de que no hay beneficio con el tratamiento nuevo, y entonces

$$P_O(\theta > 0) = \int_0^{\infty} \pi_O(\theta) d\theta = \xi \quad (7)$$

Si bajo la previa escéptica,  $P(H_1 \mid \text{datos}, \pi_S) > 1 - \gamma$ , con  $1 - \gamma$  grande, digamos  $1 - \gamma = 0,95$ , entonces, habiendo convencido a un escéptico es seguro decidir en favor de  $H_1$ . Si por el contrario, con la previa optimista  $P_O(\theta > 0 \mid \text{datos}) > 1 - \gamma$ , es seguro concluir que no hay beneficio con el tratamiento nuevo. Si hay desacuerdo entre los procedimientos, el ensayo no es inconcluyente dada la información acumulada hasta el momento.

Este enfoque es intuitivamente muy satisfactorio. Sin embargo, tiene el defecto potencial de ser demasiado cauteloso y conservador, dilatando innecesariamente las decisiones. En Spiegelhalter et al. (2004) se proponen previas normales, las cuales simplifican las cuentas. Sin embargo, en este artículo mostramos que podemos preservar las ventajas del enfoque optimista vs. escéptico, pero eliminando el dogmatismo inherente en las previas conjugadas (normales en este caso).

**Algunos antecedentes.** Pericchi & Smith (1992) estudiaron algunos aspectos de la robustez de la previa Student- $t$  y un parámetro de localización normal, y proveyeron de aproximaciones los momentos de la densidad posterior bajo el modelo Student- $t$ /normal. La previa Cauchy, como Student- $t$  con un grado de libertad, puede ser utilizada en este contexto. También para normal “log-odds” (el parámetro  $\theta$  definido arriba) existe una previa definida por Berger (1985), que además de tener colas pesadas, da resultados que pueden ser calculados en forma cerrada, esto es, con expresiones matemáticas concretas en lugar de resultados numéricos. Otra previa de colas pesadas ha sido propuesta por Berger & Pericchi (1996); la llamada previa intrínseca. Todas estas previas de colas pesadas dan resultados prácticamente idénticos, hecho ilustrado por Fúquene et al. (2009). En este último artículo, se estudian previas robustas, basadas en resultados de ensayos clínicos previos: las así llamadas previas clínicas. Las previas estudiadas en el presente artículo modelizan no historia previa, sino posiciones adversas frente a un ensayo clínico nuevo; las previas escépticas vs optimistas. En este trabajo ilustramos que suponiendo estas previas, respetando los requerimientos probabilísticos de Spiegelhalter et al. (1994), pero cambiando su forma funcional a previas robustas, obtenemos ensayos clínicos más eficientes en los que se produce la convergencia de criterios en forma más rápida y eficiente.

## A2. Conclusiones

La idea de una previa escéptica vs. otra previa optimista es muy útil para incrementar la utilización de la estadística bayesiana en la estadística médica. Esta idea introducida por Spiegelhalter y colaboradores, tiene dos aspectos: el primero es que la región donde el tratamiento nuevo es sustancialmente mejor tiene alta probabilidad bajo la previa optimista y baja probabilidad bajo la previa escéptica. El segundo aspecto es la forma funcional de las dos previas. Eso fue supuesto por Spiegelhalter et al. (1994) como previas normales. Se muestra acá, lo inédito, que el comportamiento del procedimiento depende crucialmente de la forma funcional de las previas optimistas y escépticas, particularmente del tamaño de las colas. Lo ilustramos con un ejemplo en la sección 2 y lo demostramos matemáticamente en la sección 3. Más aún, se sostiene que el comportamiento con colas gruesas mejora y hace más eficiente el procedimiento, ya que las previas con colas gruesas están más dispuestas a cambiar de opinión que las previas de colas finas. Finalmente, el uso de la previa de Cauchy, de Berger o Intrínseca resultados análogos para efectos prácticos.

## Agradecimientos

Agradecemos a dos referees anónimos y al editor especial por sugerencias constructivas que hicieron mucho más claro nuestro artículo, y a la editora jefe por invitarnos a enviar nuestro trabajo a la *Revista Colombiana de Estadística*. Este trabajo ha sido parcialmente financiado por la farmacéutica MERCK, el Centro Comprehensivo de Cáncer de la Universidad de Puerto Rico y por NIH Grant: P20-RR016470. Jairo Fúquene fue financiado por PII - Facultad de Administración de Empresas, UPR-RRP.



## A Multivariate Analysis Approach to Forecasts Combination. Application to Foreign Exchange (FX) Markets

Una aproximación a la combinación de pronósticos basada en técnicas  
de análisis multivariante

CARLOS G. MATÉ<sup>a</sup>

INSTITUTO DE INVESTIGACIÓN TECNOLÓGICA (IIT), ESCUELA TÉCNICA SUPERIOR DE  
INGENIERÍA (ICAI), UNIVERSIDAD PONTIFICIA COMILLAS, MADRID, ESPAÑA

---

### Abstract

Forecasting is characterized by the availability of a lot of methods and the fact that technological and economic forecast horizons are increasingly more different from each other. Combining forecasts is an adequate methodology for handling the above scenario, which is conceptually suitable for the application of several methods of multivariate analysis. This paper reviews some main problems in combining forecasts efficiently from the multivariate analysis view. In particular, a methodology to produce combined forecasts with a large number of forecasts is proposed. The usefulness of such a methodology is assessed in exchange rates forecasting. Further research is suggested for finance as well as for other practical contexts such as energy markets.

**Key words:** Combining forecasting, Factor analysis, Forecasting methodology, Principal components analysis, Time series.

### Resumen

El cálculo de pronósticos se caracteriza por la disponibilidad de muchos métodos y porque los horizontes de los pronósticos o las predicciones (económicas, tecnológicas, etc.) son cada vez más diferentes. Combinar pronósticos es una metodología adecuada para manejar el escenario anterior, el cual es conceptualmente adecuado para la aplicación de varios métodos de análisis multivariante. Este artículo revisa algunos problemas principales al combinar pronósticos de manera eficiente, empleando el marco del análisis multivariante. En concreto, se propone una metodología para generar pronósticos combinados con un gran número de pronósticos y se analiza una aplicación al mercado de divisas. Se valora la utilidad de esta metodología en finanzas y varios contextos prácticos, abriéndose posibilidades futuras de investigación a otros contextos aplicados, como los mercados de energía.

**Palabras clave:** análisis de componentes principales, análisis factorial, combinar pronósticos, metodología para pronósticos, series temporales.

---

<sup>a</sup>Professor and Researcher. E-mail: cmate@upcomillas.es

## 1. Introduction

The 21<sup>st</sup> century is providing diversity and multiple options for every problem and situation in the real world. Obvious examples are the fast transference in know-how from the academic to the non-academic world, improved knowledge in the scientific community, easy and flexible software, or increased availability due to new technologies (see, as an example of the above facts, the film *The Social Network*).

Forecasting is now characterised by the availability of a large number of methods, an updated review of which is available from several sources: Armstrong (2001), Diebold (2001), Elliott, Granger & Timmerman (2006), Makridakis, Wheelwright & Hyndman (1998).

Furthermore, forecast horizons for companies and institutions are increasingly different and yet interrelated. Decision-makers now require related predictions for the immediate and short run, or for the medium or long run, or even for the three standard horizons: short, medium and long ones. For example, Sánchez-Ubeda & Berzosa (2007) provide an application to forecasting industrial end-use natural gas consumption in Spain, where the company Enagás requires forecasts in a medium-term horizon (1–3 years) with a very high resolution (days). Other works covering this problem are Kang (2003), Andrawis, Atiya & El-Shishiny (2011) and references cited there.

Combining forecasts is generally regarded as an appropriate methodology for handling the above scenario, having reached maturity and recognition over the last forty years (Bates & Granger 1969, Bunn 1989, Clemen 1989, de Menezes, Bunn & Taylor 2000, Armstrong 2001, Timmerman 2006). In economics, a recent application is devoted to forecasting US employment growth (Rapach & Strauss 2008). On the other hand, Kapetanios, Labhard & Price (2008) report about that forecasts combination generally leads to a reduction in forecast error.

Such a setting is conceptually suitable for the application of *multivariate analysis methods* (MAM in the following). In the last fifteen years, these techniques have benefited from big improvements with regard to easiness of use, the overcoming of limitations and the recognition of enlarged usefulness (Hair, Anderson, Tatham & Black 1998). Hence, an adaptation of MAM to new technologies such as databases, Internet and others is an emerging area.

Given this scenario, this article proposes the employment of several methods of classical multivariate analysis to obtain combined forecasts. A recent paper (Poncela, Rodríguez, Sánchez-Mangas & Senra 2011) proposes the use of dimension reduction techniques in forecasts combination. However, our approach is different because it considers several forecast horizons, analyzes the possibilities in forecasts combination of the key set of multivariate methods and suggests different fields of applications such as financial markets and energy markets.

The paper is organized as follows. Section 2 presents the foreign exchange market and how forecasting is key in this market. Section 3 is devoted to the analysis of one forecast combination framework and highlights the multivariate analysis methods approach. Section 4 reviews one of the problems in combining forecasts

efficiently and proposes a solution to this problem using principal component analysis. In addition, one application for EURUSD forecasting is presented. Section 5 outlines a new methodology for producing combined forecasts. In Section 6, the usefulness of such a methodology is considered when its practical implications in several contexts are analyzed. Finally, some conclusions from this research are provided.

## 2. Forecasting in the Foreign Exchange Market

According to Wikipedia, the **foreign exchange market (forex, FX, or currency market)** is a worldwide decentralized over-the-counter financial market for the trading of currencies. Financial centers around the world function as anchors of trading between a wide range of different types of around-the-clock buyers and sellers, with the exception of weekends. The foreign exchange market determines the relative values of different currencies. The primary purpose of foreign exchange is to assist international trade and investment, by allowing businesses to convert one currency to another currency. For example, it permits a US business to import British goods and pay Pound Sterling, even though its income is in US dollars. It also enables investors to invest in foreign exchange, and facilitates the carry trade, in which investors borrow low-yielding currencies and lend (invest in) high-yielding currencies, and which (it has been claimed) may lead to loss of competitiveness in some countries.

The foreign exchange market is highly competitive. In the second semester of 2010 media around the world devoted a lot of news to a possible new currency war. But this is not a new problem. For example, Posen (2004) analyzed the issue of avoiding a currency war.

With respect to the size and importance of the foreign exchange market, Eun & Sabherwal (2002) stated the following: “The market for foreign exchange is the largest financial market in the world. According to the Bank for International Settlements (2001) the average worldwide daily trading in trading foreign exchange markets is estimated to be US\$1.2 trillion”. More recently, King & Rime (2010) have analyzed the so called “\$4 trillion question” giving explanations to the FX growth since 2007.

Subsequent to Eun & Sabherwal (2002), and in response to the increasing importance of this market, there have been a number of studies on forecasting exchange rates. The substance of many of these studies is the evaluation of the forecasting performance of one or more exchange rate models. These evaluations suggest that the best predictor of future spot rate is the current spot rate, i.e., the random-walk or “no-change” model, the most commonly used benchmark. This conclusion has also been obtained by Kilian & Taylor (2003).

There are a lot of websites producing forecasts for different time horizons. Some of them include forecasts for 3 (3M), 6 month (6M) and 12 (12M) month periods. For example, <http://www-2.danskebank.com/danskemarketsresearch>

Some of these sites include forecasts for a week, a month and a quarter. For example, the following website includes these forecasts by 16 experts.

<http://www.fxstreet.com/technical/forex-forecasts/experts-forecast-currencies-poll/2010/12/17/03/>

Table 1 shows an example of the information given in the above website.

TABLE 1: Elaborated from the information given in [www.fxstreet.com](http://www.fxstreet.com)

EUR/USD	1 week	1 month	1 quarter	1 week	1 month	1 quarter
		Levels			Trend	
Expert 1	1.36	1.38	1.32	Bullish	Bullish	Bearish
Expert 2	1.315	1.291	1.274	Bullish	Bullish	Bearish
...	...	...	...	Sideways	Bullish	Sideways
...	...	...	...	...	...	...
Expert 16	1.32	1.30	1.26	Very Bullish	Bearish	Bearish

King & Rime (2010) explain why electronic trading, based on electronic execution methods, is transforming FX markets. In particular, high-frequency trading (HFT) is one algorithmic trading strategy that profits from incremental price movements with frequent, small trades executed in milliseconds for investment horizons of typically less than one day. They state that HFT emerged with the advent of the 21<sup>st</sup> century, becoming an important source of FX growth from 2004.

As we can see generating accurate forecasts for the next 24 hours, or the next 60 minutes or the next 60 seconds, is a real challenge in FX markets.

### 3. A Multivariate Methods Approach for Combining Forecasts

In what follows, it will be assumed that we are interested in forecasting a magnitude, denoted by  $Y$ , for a large number of forecast horizons, denoted by  $n$ , from the information about it in different periods of time  $1, 2, \dots$  until the present moment (denoted by  $t$ ). With this objective in mind, the forecaster has several forecasts obtained using a number of forecasting methods, denoted by  $p$ . In the case of the application provided in Sánchez-Ubeda & Berzosa (2007), currently used by Enagás,  $Y$  is the daily demand for industrial end-use natural gas consumption and  $n$  is  $365 \times 3 = 1095$ .

TABLE 2: Structure of the information in the problem.

	$M1$	$M2$	...	...	$Mp$
$H1$	$f_{11}(t)$	$f_{12}(t)$	...	...	$f_{1p}(t)$
$H2$	$f_{21}(t)$	$f_{22}(t)$	...	...	$f_{2p}(t)$
...	...	...	...	...	...
...	...	...	...	...	...
$Hn$	$f_{n1}(t)$	$f_{n2}(t)$	...	...	$f_{np}(t)$

Hence, we will assume that we have  $p$  unbiased forecasts of the magnitude  $Y$  obtained by  $p$  forecasting methods (or experts or a combination of methods and experts)  $(M1, M2, \dots, Mp)$  for  $n$  forecast horizons  $(H1, H2, \dots, Hn)$ . The notation  $f_{ij}(t)$  stands for the forecast obtained in the  $i$ th forecast horizon by the  $j$ th method or expert with  $i = 1, \dots, n$  and  $j = 1, \dots, p$ . All this information can be shown in a matrix structure (Table 2), denoted by  $\mathbf{F}(t)$ . Table 1 is a particular case of Table 2 transposing rows and columns.

In the above setting, the application of classical multivariate analysis methods (see, for example, Johnson & Wichern 2002, Rencher 2002, and Hair et al. 1998), where the variables are the forecast methods and the units are the forecast horizons, is appropriate.

From this information, MAM have been developed for the study of different problems in data analysis. In our case, the forecasts are obtained using several methods in different forecast horizons. One of the problems studied is the reduction of the dimensions of the problem (forecasting methods), the goal of this reduction being to facilitate the assimilation process of the relationships discovered by the analyst.

Another problem is the process of searching for dependence relationships which would facilitate the understanding of the behavior of some forecasting methods by taking into account the values that other methods show. Finally, an interesting problem is the analysis of possible interdependence relationships, which will permit the detection of subtle relationships between the forecast horizons.

An initial classification of MAM applied to the structure of Tables 1 and 2, which can be seen as a particular case of the general structure for MAM, is summarized in Table 3, which collects the proposed use of MAM in forecasting. It concerns the forecasting of horizons, forecast methods, and the tasks of analyzing dependence and interdependence between methods, or the problem of reducing the number of methods, or that of involving the grouping or classifying of forecast horizons.

In the following, the potential of MAM in the five problems in Table 3 will be briefly considered.

**Problem 1.** In our forecasting scenario the clearest task would be the reduction of the  $p$  unbiased forecasts to a smaller number of forecasts (ideally one or two) for every forecast horizon considered. With this goal in mind, methods such as Principal Component Analysis (PCA) or Factor Analysis (FA), linear or non-linear, or even Cluster Analysis applied to forecast methods, could deserve our attention. In a general context, the clustering approach in prediction has a long tradition (see, for example, Kusiak & Li 2010 for a recent application in wind power forecasting).

**Problem 2.** Granger & Ramanathan (1984) suggest that conventional forecast combining methods could be considered within the regression framework. Hence, the alternative methods for handling this problem in Table 2 must be explored. For example, MANOVA could be used to analyze if there are significant differences between a group of forecasts obtained using the same methodology (ARIMA, Neural Nets, Expert Opinions, ...) in different groups of time horizons such as short,



medium or long ones. Depending on the conclusion reached, the decision of the forecaster as to the best way to combine would be different.

CCA could be used to determine the kind of relationship (linear or non-linear) between groups of forecasts to decide on the possibility of reducing the number of forecasts to be considered in the problem.

**Problem 3.** Depending on the different forecasting methods used, the usefulness of knowing the profile of the forecast horizons with respect to accuracy is beyond doubt. As many authors have said, there are no methods or methodologies that outperform the rest for every problem or every time horizon. It seems that DA and LR can help in future research into these questions. In particular, approaches in FOREX pose new challenges in International Finance.

**Problem 4.** Identifying groups of similar forecast horizons with respect to a series of forecasting methods or methodologies would make the creation of different types of combination for every cluster of forecast horizons possible.

**Problem 5.** This is one of the most complex problems in MAM and the same occurs in Table 3. Here, Partial Least Squares could help in solving the problem, commented on in the above paragraph, of deciding which methods should be included in clustering forecast horizons.

#### 4. A Solution from Multivariate Analysis Methods to the Problem of Efficient Combined Forecasts

According to several authors Bunn (1988), Clemen & Winkler (1986), Winkler & Makridakis (1983), in addition to the notation introduced in Section 2, if we have  $p$  unbiased forecasts of the same variable  $Y$  for the same time horizon (let this be  $i$ ), denoted by  $f_{i1}(t), \dots, f_{ip}(t)$ , then the composite forecast, denoted by  $c_i(t)$ , based upon the  $p \times 1$  vector of the linear weights,  $\mathbf{w}_i(t)$ , will adopt the expression

$$c_i(t) = \mathbf{w}'_i(t) \mathbf{f}_i(t) \quad (1)$$

and will be optimum in the sense of having a minimum forecast error variance if  $\mathbf{w}_i(t)$  is determined according to

$$\mathbf{w}_i(t) = \frac{\mathbf{S}_i^{-1}(t) \mathbf{e}_i(t)}{\mathbf{e}'_i(t) \mathbf{S}_i^{-1}(t) \mathbf{e}_i(t)} \quad (2)$$

where  $\mathbf{e}_i(t) = (1, 1, \dots, 1)'$  is an  $\mathbf{p} \times 1$  unit vector and  $\mathbf{S}_i(t)$  is an  $\mathbf{p} \times \mathbf{p}$  covariance matrix of forecast errors between the  $\mathbf{p}$  forecasts for the  $i$ th forecast horizon.

According to a previous study (Clemen & Winkler 1986), if a normal or Gaussian is assumed for the forecasting errors, then the combined forecast  $c_i(t)$  is a weighted average of the individual forecasts  $f_{i1}(t), \dots, f_{ip}(t)$  with the vector of weights

$$\mathbf{w}'_i(t) = \frac{\mathbf{e}'_i(t) \left( \widehat{\Sigma}_i(t) \right)^{-1}}{\mathbf{e}'_i(t) \left( \widehat{\Sigma}_i(t) \right)^{-1} \mathbf{e}_i(t)} \quad (3)$$

TABLE 3: Brief summary of the proposed use of MAM in forecasting.

Problem	MAM	Comments
1. To reduce the original information to a smaller set of new forecasting methods (factors)	- Factor Analysis (FA) - Principal Component Analysis (PCA)	Factors usually are determined by principal component analysis or by a common factor analysis.
2. To determine the existence of dependence relationships between two groups of forecasting methods	- Regression - Analysis of variance (ANOVA) - Canonical Correlation Analysis (CCA) - Conjoint Analysis (CoA)	The differences between methods are due to the number and kind of forecasting methods that are included in every group of forecasting methods.
3. To assign forecast horizons to groups characterized by a series of forecasting methods	- Discriminant Analysis (DA) - Logistic Regression (LR)	The first method is more general than the second one, however the latter raises more flexible hypotheses.
4. To make or to identify groups of forecast horizons similar between them	- Cluster Analysis (CIA) - Multidimensional Scaling (MS) - Correspondence Analysis (CA) - CHAID Analysis	The first method provides groups of forecast horizons with maximum internal homogeneity and maximum heterogeneity between them. The second one facilitates the discovery of the number of dimensions where the forecasts horizons (units) are similar. While, the third one allows accommodating different type of data and non-linear relationships. The last one proposes groups according to a dependent variable or forecasting method, making easy the classification of new forecast horizons in the corresponding groups.
5. Searching for crossed relationships of multiple dependence and the representation of non-observed concepts in such relationships	- Structural Equation Modelling (SEM) - Partial Least Squares (PLS)	SEM makes it possible to analyse several relationships at the same time. It is an extension of several MAM like FA or Regression. PLS facilitates the discovery of the number of dimensions where the forecast horizons are similar.

depending on  $\widehat{\Sigma}_i(t)$ . Thus, the estimate  $\widehat{\Sigma}_i(t)$  which can be determined from past data on estimation errors, prior information or some combination thereof plays an important role in the determination of the combined forecast.

#### 4.1. Combined Forecasts with Principal Component Analysis

Principal Component Analysis (PCA), initially introduced by Pearson in 1901 and independently by Hotelling in 1933, is one of the oldest MAM. This tech-

nique, applied to our problem, will involve describing or explaining the structure of variation in a multivariate data set (with evidence of interdependence between the variables, see Table 2 in Section 3) in terms of a set of uncorrelated variables where every one is a specific linear combination of the original forecasting methods. That is, a linear combination of forecasts can be obtained as an alternative to usual ways of proceeding.

In the following, we denote by  $\mathbf{f}(t)$  the vector of the  $p$  forecast methods with the available information until  $t$ , in such a way that if the  $j$ th method or column of Figure 1 is designed by  $f_j(t)$ , it follows that

$$\mathbf{f}(t) = (f_1(t), f_2(t), \dots, f_p(t))$$

Furthermore, the population covariance matrix associated with the  $\mathbf{F}(t)$  matrix in Table 2 will be denoted by  $\mathbf{S}(t)$ .

In this situation, the first component extracted will be a linear combination of the  $p$  forecast methods, designed by  $C1(t)$ , such that

$$C1(t) = \mathbf{a}'_1 \cdot \mathbf{f}(t) \quad (4)$$

whose sample variance is the biggest for all vector of coefficients  $\mathbf{a}_1$ . Consequently, the sample variance can be increased without limit and a restriction to these coefficients must be considered. That the sum of squares will be one, i.e.,

$$\mathbf{a}'_1 \cdot \mathbf{a}_1 = 1 \quad (5)$$

Similarly, the second component denoted by  $C2(t)$  will be

$$C2(t) = \mathbf{a}'_2 \cdot \mathbf{f}(t)$$

with maximum variance restricted to the two conditions  $\mathbf{a}'_2 \cdot \mathbf{a}_2 = 1$  and  $\mathbf{a}_2 \cdot \mathbf{a}_1 = 0$ .

Extending the above reasoning to the  $i$ th component, it will be verified that it is a linear combination of the  $p$  original variables, denoted by  $C^j(t)$ , such that

$$C^j(t) = \mathbf{a}'_j \cdot \mathbf{f}(t) \quad (6)$$

with maximum variance restricted to the following conditions

$$\mathbf{a}'_j \cdot \mathbf{a}_j = 1 \quad \text{and} \quad \mathbf{a}'_j \cdot \mathbf{a}_i = 0 \quad (i < j) \quad (7)$$

Thus, the problem of finding the principal components (PC) in a matrix like that of Table 2 consists in obtaining the vectors of coefficients  $\mathbf{a}_j$ , for  $j = 1, 2, \dots$ ; with a limit in  $p$ . We can show (Johnson & Wichern 2002) that this problem is equivalent to that of obtaining the eigenvectors of the correlation matrix between the forecast methods. The associated eigenvalues will be  $\lambda_1 > \dots > \lambda_p$  and the total variance of the  $p$  PC will be equal to the variance of the original

forecasts. Moreover, the trace (sum of the diagonal elements) of the covariance matrix,  $tr(\mathbf{S}(t))$ , will verify

$$\sum_{i=1}^p \lambda_i = tr(\mathbf{S}(t)) \quad (8)$$

Hence, the  $j$ th PC accounts for a proportion of variance

$$\frac{\lambda_i}{tr(\mathbf{S}(t))} \quad (9)$$

of the total variance in the original forecast methods.

The usefulness of PCA lies in that the components, the different combination of the forecast methods in our setting, employ the variance in a decreasing proportion. Thus, the important issue for obtaining an appropriate reduction of the forecasts consists of answering the following question: How many components should be extracted?

There are a series of formal techniques (see, for example, Basilevsky 2004), but we will restrict ourselves to informal rules used in practical contexts Gorsuch (1983). They are the following.

1. The inclusion of a sufficient number of components to explain some large relative percentage of the total variance. The quotes between 70 and 90 percent are usually suggested. However, lower values arise when  $p$  or  $n$  increase.
2. The exclusion of those components whose eigenvalues are less than the average.
3. Plotting the eigenvalues  $\lambda_i$  (in decreasing order) versus the index  $i$ , called *scree plot*, to select the number of components needed to extract the value of  $i$  corresponding to an elbow in the curve. That is, the point where big eigenvalues soften and small eigenvalues begin.

The above development is based on terms of eigenvectors and eigenvalues of the covariance matrix  $\mathbf{S}(t)$ . However, these components are usually extracted from the correlation matrix (that associated to  $\mathbf{F}(t)$  in our setting), which will be denoted by  $\mathbf{R}(t)$ .

The reasons for the last statement are not difficult to understand if we imagine a multivariate data set where the variables represent very different approaches to the problem. In such a case, the structure of the PC obtained from the covariance matrix will depend on the choice, essentially arbitrary, of the units of measurement. Moreover, if there are great differences between the variances of the forecasting methods, those whose variances are biggest will tend to dominate the first CP.

Extracting PC by the eigenvalues of  $\mathbf{R}(t)$  will exclude the above difficulty. This procedure is equivalent to calculating the PC of the original methods after standardizing for having unit variance.

However, in the rest of the situations, the PC must be extracted from the covariance matrix, because this is closer to the spirit and substance of the PCA, especially if the components will be employed in further computations, which is very common when forecasting with several forecasts.

With respect to the PC extracted from  $\mathbf{R}(t)$  and  $\mathbf{S}(t)$ , the following can be stated.

1. Eigenvectors, eigenvalues and coefficients from the PC extracted from  $\mathbf{S}(t)$  will differ from the PC extracted from  $\mathbf{R}(t)$ .
2. The percentage of variance that components of  $\mathbf{R}(t)$  account for will differ from the percentage that account for the extracted components from  $\mathbf{S}(t)$ .
3. The extracted components from a matrix  $\mathbf{R}(t)$  are not unique or exclusive to that matrix.

Finally, several measuring to indicate the adequacy of the data matrix  $\mathbf{F}(t)$  to PCA have been developed. One of the most simple is the Kaiser-Meyer-Olkin or KMO test. According to Hair et al. (1998), this test assesses the adequacy of using PCA on data. Values greater than 0.7 suggest that it is appropriate to apply PCA to the considered matrix. Moreover, measures to elucidate the question of whether a forecasting method must be included in the problem of the reduction of the information are available. The most commonly used is MSA (measure of sample adequacy) and recommendations about this amount have been given Rencher (2002).

## 4.2. One Basic Application to FX markets

Several cases have been run to evaluate the above approach. To show a particular case, the daily EURUSD close rate from January 1999 to December 2010 has been considered. This series will be denoted by  $Y(t)$ . To analyze this series, the series of differences should be obtained. That is,  $\Delta Y(t) = Y(t) - Y(t - 1)$ .

Plots of these series are shown in Figures 1 and 2.

A simple and visual analysis of Figures 1 and 2 allows us to conclude that the EURUSD series has the following features:

1. Soft linear trend, with several long periods of decreasing trend and other periods of increasing trend. In consequence, several turning points can be detected.
2. Due to sharp decreasing or increasing, presence or absence of structural breaks should be considered.
3. The differentiated series is not a white noise. Therefore, the EURUSD series is not a random walk.

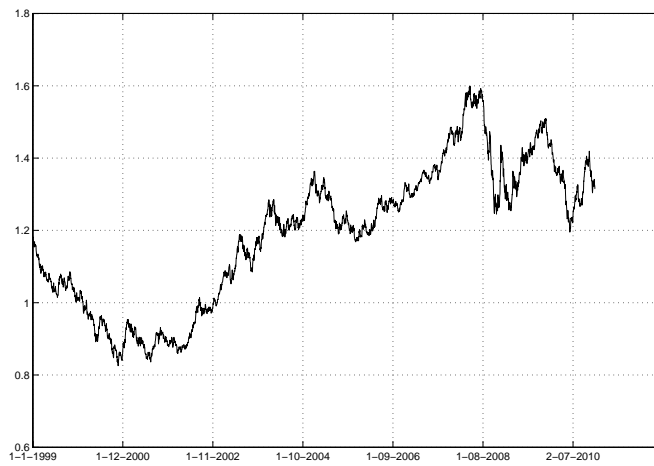


FIGURE 1: Plot of daily EURUSD.

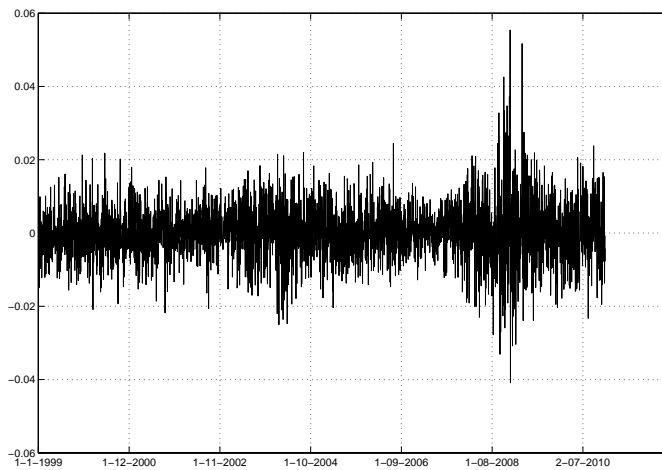


FIGURE 2: Plot of the series EURUSD differenced.

These facts lead to conclude that the EURUSD series is predictable, but not for all years (e.g., the first half of 2007 in the differenced series is close to a white noise). In addition, several methods should be considered and information from the past which is not longer valid could be removed.

Figure 3 shows the boxplots of both datasets,  $Y(t)$  and  $\Delta Y(t)$ .

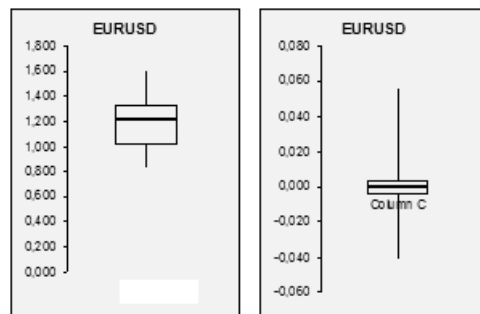


FIGURE 3: Boxplots of series 1 and series 1 differentiated.

In the remainder of this article, the results will be shown using the STATGRAPHICS CENTURION software. There are, however, a variety of software packages and we do not wish to give the impression that we endorse one over any other.

The forecasting methods employed are very basic, but usual in FOREX, to give more relevance to the application proposed itself instead of using sophisticated or complicated methods. They came from the exponential smoothing approach, polynomial trend and the ARIMA methodology, using the Forecast method procedure in STATGRAPHICS in an automatic or a manual way. Forecasts were obtained using six methods, see Table 4 for details, and a matrix like that in Table 2 was available. ARIMA models are usually considered in FOREX markets (see, e.g., Dunis & Williams 2005).

TABLE 4: Accuracy measures with the six methods employed.

Model	#	RMSE	MAE	MAPE	AIC
Random walk	0	0.00767929	0.0056745	0.479297	-9.73846
ARIMA (3,0,2) with const.	1	0.00765786	0.0056721	0.478967	-9.74405
ARIMA (1,0,0)	2	0.00767834	0.00567289	0.47914	-9.7387
Holt, alpha = 0.9966 and beta = 0.0075	3	0.00769148	<b>0.00566384</b>	<b>0.478415</b>	-9.734
ARIMA (6,0,5) with const.	4	<b>0.00765561</b>	0.00567007	0.478764	<b>-9.74463</b>
ARIMA (4,1,6)	5	0.007668	0.00567807	0.479561	-9.7414
Quadratic trend $-6.29 + 0.00058 t$ $-1.0399E-8 t^2$	6	0.104693	0.0849887	7.52837	-4.51345

As a result of extracting the first PC of the five first methods considered, and using the covariance matrix associated to the forecasting matrix, Figure 4 is obtained. Using only a PC, we will retain practically all the information in the five forecasting methods considered.

The main conclusion obtained from Table 5 is that combining by means of the proposed approach in this paper is the same as combining by means of the simple

average. That is, the following important property can be stated: combining with the simple average theoretically will produce the maximum variance between linear combinations of the considered forecasting methods. Timmerman (2006) concludes that the equal weights average is a benchmark that has proved surprisingly difficult to beat. An interesting research issue is to connect both results from a theoretical point of view.

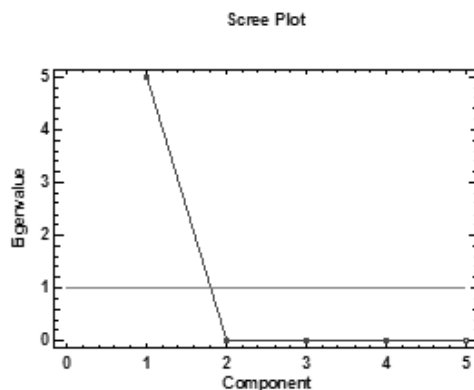


FIGURE 4: Scree plot for 5 forecasting methods with the EURUSD series.

TABLE 5: Matrix of weights for the combination of the five forecasting methods with the EURUSD series.

	<i>Component</i>
	1
FORECASTS_M1	0.447214
FORECASTS_M2	0.447214
FORECASTS_M3	0.447213
FORECASTS_M4	0.447214
FORECASTS_M5	0.447214

Another use of PCA is in the detection of forecasting methods with null or low influence on the structure of variability in the forecasting matrix, where PCA will permit us to discard those methods in the combination. For example, Table 6 shows that situation with the sixth method used to obtain the PC.

## 5. A New Methodology to Produce Combined Forecasts for Several Forecast Horizons

As has been mentioned in the introduction to this paper, forecasters of the 21<sup>st</sup> century will have at their disposal a large number of forecasts that they may want to combine. See, for example, the ‘suite of statistical forecasting models’ (the ‘Suite’) of The Bank of England in Kapetanios et al. (2008).



TABLE 6: Matrix of weights for the combination of the six forecasting methods with the EURUSD series.

	<i>Component</i>
	1
FORECASTS_M1	0.415608
FORECASTS_M2	0.415609
FORECASTS_M3	0.4156
FORECASTS_M4	0.415608
FORECASTS_M5	0.415608
FORECASTS_M6	0.369264

In 1989, Granger reviewed some alternatives to original methods of forecasting, stating that the latter were too complicated for using with a large number of forecasts available for combination.

In this section, a methodology for the production of combined forecasts is proposed, for a context in which a set of forecasting methods are available for several forecast horizons. In the sequel, it will be assumed that we are interested in forecasting a magnitude  $Y$  for a large number of forecast horizons, denoted by  $n$ , using the information available up to the present moment (designed by  $t$ ). With this objective in mind, the forecaster has obtained several forecasts using a number of forecasting methods, denoted by  $p$ . All this information is shown in the matrix structure of Table 2, introduced in Section 3.

In some MAM it is customary to consider that  $p \ll n$ , but situations where  $p \gg n$  are frequent and important analyses can be conducted. In forecasting, every situation for all possible values of  $p$  and  $n$  must be considered.

**Case  $p \geq n$**  (there are at least as many forecasting methods as forecast horizons)

In this situation, experience in the use of MAM indicates that it would be appropriate to use procedures for the reduction of the dimension of the problem as PCA or FA do. One possibility would be to cluster the forecasting methods and then to explore a way of constructing combined forecasts from the centroids of the clusters.

**Case  $p < n$**  (there are fewer forecasting methods than forecast horizons)

Here, we will assume that the ratio is less than five to one (no more than five forecast horizons for every forecasting method). Hence, the use of techniques for reducing the dimension of the problem as PCA or FA would not be appropriate. However, these techniques could be used in the way that is analysed in the following case.

**Case  $p \ll n$**  (there are far more forecast horizons than forecasting methods)

In this situation, the practice of MAM suggests the use of techniques (such as PCA or FA) for reducing the dimension of the problem. The literature on these methods is much too vast to adequately cite here; but there are excellent general

developments (Johnson & Wichern 2002, Rencher 2002, Hair et al. 1998) and a thorough coverage of these techniques (Gorsuch 1983, Basilevsky 2004).

In our context, factor analysis would be concerned with the identification of an underlying structure within a set of forecasting methods. Hence, we will try to simplify the complicated interrelations among a large number of forecasting methods, in a search of certain unobservable common factors which are not related to each other. In consequence, the goal of this technique is the reduction of forecasting methods and the ascertaining of the relevant characteristics of the original set of methods by the construction of a smaller number of new common factors or latent variables with a minimum loss of information.

The adaptation of the general model of factor analysis (Johnson & Wichern 2002, Jobson 1992, Basilevsky 2004) to the forecasting context developed in Section 3 will involve handling three sets of variables: a set of  $p$  forecasting methods  $FM_1, \dots, FM_p$  with mean vector  $\mu$  and covariance matrix  $\Sigma$ ; a set of  $m$  unobserved variables or combined forecasts called common factors  $C_1, \dots, C_m$  where  $m \leq p$ ; and a set of  $p$  unique but unobserved factors  $e_1, \dots, e_p$ . The model will be

$$\begin{aligned}
 FM_1 - \mu_1 &= l_{11}C_1 + l_{12}C_2 + \dots + l_{1m}C_m + e_1 \\
 FM_2 - \mu_2 &= l_{21}C_1 + l_{22}C_2 + \dots + l_{2m}C_m + e_2 \\
 &\quad \circ \\
 &\quad \circ \\
 FM_p - \mu_p &= l_{p1}C_1 + l_{p2}C_2 + \dots + l_{pm}C_m + e_p
 \end{aligned}$$

or, in matrix notation,

$$\mathbf{FM} - \mu = \begin{matrix} \mathbf{L} & \mathbf{C} \\ \text{---} & \text{---} \\ (p \times 1) & (\mathbf{p} \times \mathbf{m}) \text{ } (m \times 1) \end{matrix} + \begin{matrix} \mathbf{e} \\ \text{---} \\ (\mathbf{p} \times 1) \end{matrix}$$

The coefficient  $l_{ij}$  accompanying  $C_j$  in the linear combination describing  $FM_i - \mu_i$  is called the loading of the  $i$ th forecasting method on the  $j$ th common factor, so the matrix  $\mathbf{L}$  is the matrix of factor loadings or the factor pattern matrix.

In this model,

$$\begin{aligned}
 E(\mathbf{C}) &= \mathbf{0}, \text{Cov}(\mathbf{C}) = \mathbf{I} \\
 E(\mathbf{e}) &= \mathbf{0}, \text{Cov}(\mathbf{e}) = \Psi \\
 E(\mathbf{e} \cdot \mathbf{C}) &= \mathbf{0}
 \end{aligned}$$

Then,  $\Sigma = \mathbf{L}\mathbf{L}' + \Psi$ . Hence,

$$\begin{aligned}
 Var(FM_i) &= \sum_{j=1}^m l_{ij}^2 + \Psi_i; \quad i = 1, \dots, p \\
 Cov(FM_i, FM_k) &= \sum_{j=1}^m l_{ij}l_{kj}
 \end{aligned}$$

Hence, the variance of the  $i$ th forecasting method can be decomposed into two parts: the first part  $\sum_{j=1}^m l_{ij}^2$  denotes the portion of  $Var(FM_i)$  related to the common factors and is called the  $i$ th communality; and the second term  $\Psi_i$  is the unique variance, which is not related to the common factors.

### The FA Methodology applied to the Forecasting Data Matrix

From the observations on the above  $p$  forecasting methods obtained for  $n$  forecast horizons, which are denoted by  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ; and with the application of the linear model discussed previously, the FA method is designed to find a small number of combined forecasts (factors) that represent the data adequately. Noting the sample covariance matrix by  $\mathbf{S}$ , we have an estimation of the unknown population covariance matrix  $\Sigma$ . If the off-diagonal elements of  $\mathbf{S}$  are small or those of the sample correlation matrix  $\mathbf{R}$  essentially zero, the forecasting methods are not related and factor analysis will not be appropriate. In another case, a factor model can be established and the first problem is to estimate the matrices  $\mathbf{L}$  and  $\Psi$ .

Methods giving solutions to this problem are principal component analysis (PCA), principal factors (PF), maximum likelihood factoring, alpha factoring, unweighted least square factoring, image factoring and so on. The most commonly used are PCA and PF. PCA is commonly employed as a preliminary extraction technique, followed by one or more of the other procedures, perhaps varying the number of factors, the communality estimates, and the rotational methods in each run Tabachnik & Fidell (1996). The basic difference between PCA and PF is that the former focuses on variance and the latter on covariance (communality).

The goal of PCA in the proposed forecasting context would be to extract maximum variance from the forecasts matrix  $\mathbf{F}(t)$  with a few orthogonal combined forecasts, while the aim of PF would be to reproduce the correlation matrix  $\mathbf{R}(t)$  between forecasting methods with a few orthogonal factors that will act as combined forecasts. Thus, the proposed methodology consists of the following steps.

1. The reduction of the dimensions of the forecasting matrix  $\mathbf{F}(t)$  data to a few factors or PC (very probably dependents).
2. The repetition of '1' in order to extract one or two factors, called in the context of FA *higher-order factors*, and the taking of these factors of the combined forecasts as the final combination of forecasts. That is, new combinations are created based on the different combined forecasts.

There is a practical case of extracting higher-order factors, in a different setting to that of this paper (Maté & Calderón 2000).

## 6. Some Practical Contexts and Application Issues

In this section, three practical forecasting contexts will be considered. The usefulness of the proposed methodology for these contexts and for some patterns of time series will be analyzed.

### Energy Markets

Forecasting is a very important task for people involved in the management of energy markets, even when there is a stable economic setting. Energy producers urgently need automated and accurate forecasts so that they can anticipate demand, prices and other market variables.

All over the world, the entire sector is currently undergoing an important restructuring process towards deregulation and competition. The emergence and development of this new context implies that the management of energy producing companies can no longer be based on administrative and centralized procedures: their profit now depends directly on their own managerial decisions. Decision-making processes become more complex and risky due to the higher uncertainty which results from the new operational framework.

It has been pointed out in previous research that a careful calculation of expected demand and supply can lead to contracts that enhance the profitability of companies. Hence, more research for very short run forecasts will be required (Ramanathan, Engle, Granger, Vahid-Araghi & Brace 1997). In this article, the development of a number of models to produce very short run forecasts of hourly system loads was considered. The required forecast horizons vary from 16 to 40 hours from Monday to Thursday, and from 16 to 88 hours on Friday. However, the availability of forecasts for all week, 168 hours, is normal in energy companies. For example, Hahn, Meyer-Nieberg & Pickl (2009), is a recent survey of mathematical methods and ideas which have been used for electricity load forecasting.

The methodology proposed in our paper can be applied to price forecasts in energy companies. According to Zhao, Dong, Xu & Wong (2008), electricity price forecasting is a difficult yet essential task for market participants in a deregulated electricity market. Pindyck (1999) analyzes the influence of price forecasts on investment decisions and the choice of products to produce.

### Stock Markets

Forecasting is a crucial task for brokers and people involved in the management of stocks. The planning of strategies might be improved with the availability of forecasts for returns, prices and volatility, in those periods where the day is divided to provide information about the evolution of a stock or groups of stocks. For example, if a quarter of an hour is taken as a unit of time division, forecasts for 32 horizons are required in an 8-hour trading day. Obviously, other divisions of time, like 5 minutes, could be used.

If, as has been suggested (Makridakis & Winkler 1983), forecasts for the 32 time horizons obtained by using four or five methods are available for the broker, the use of the proposed methodology might provide more accurate forecasts.

This approach opens a new alternative in this context because the standard practice in stock markets is to generate one-step-ahead forecasts. For example, with respect to the daily stock market data, the use of Artificial Neural Networks to combine time series forecasts of market volatility from USA, Canada, Japan and UK has been discussed (Donaldson & Kamstra 1996).

### **Sales**

Issues concerning the design of forecasting methodologies have been looked at by other researchers (Winklhofer, Diamantopoulos & Witt 1996). One of them was the time horizon. Some conclusions were that the majority of firms prepared sales forecasts on a yearly basis, that the most popular short-term sales forecasting horizon was one month and that firms operating in highly competitive markets put more emphasis on short term than on long-term forecasts.

The above suggests that the proposed approach in this paper will not be useful for many companies that employ sales forecasts. However, the increasingly competitive context in which some firms develop their activity, in a situation with some parallelism with that of the energy or stock markets, can increase the need to consider the desirability of using more forecasting methods and a large number of forecast horizons. If adopted, these changes in approach could be handled with the methodology proposed in this paper.

### **Seasonal Series**

Series with a seasonal component reproduce the behaviour of the magnitude every certain period of time. So we can assume that the forecast horizons act as units of study similar to the problems where multivariate analysis is usually applied.

### **Trending Series**

Series with a trend component do not behave with the same pattern through several forecast horizons. So the forecast horizons perhaps do not show the same structure as the units in the usual multivariate analysis problems.

### **Cyclical Series**

Series with a cyclical component must come back to a previous structure in the behaviour of the magnitude at regular periods of time. So, as with seasonal series, we can assume that forecast horizons act as units of study similar to the problems where MAM are usually applied.

### **Automation**

One of the key points for a successful implementation of the proposed methodology is its automation. For this reason, the issues of versatility, accuracy, speed and intelligibility must be strongly considered, for as long as a limited amount of resources (software, knowledge, data and so on) are available.

### **Forecasting Methods to Be Included**

Several classic statistical forecasting models (exponential smoothing, ARIMA, GARCH, and so on) are widely used as part of forecasting systems, giving satisfactory results when applied in different fields. They provide many possibilities of modeling for one or several variables since they allow different behaviors and horizons be taken into account. Likewise, they have very intuitive use features and are very user-friendly, because the parameters they involve represent the underlying process in a simple way.

In addition, many of the variables included in the practical applications behave well in many cases or moments of their history, justifying the view that linear or

classic forecasting models may be seen suitable for such contexts. However, the possibility of using artificial neural networks and other non-linear methods should be considered.

There are several references for a further study of general forecasting methods (Armstrong 2001, Diebold 2001, Elliott et al. 2006, Makridakis et al. 1998).

#### **Integration with Rule-Based Forecasting and Expert Systems**

A forecasting technique which integrates judgement and statistical extrapolation is the so-called rule-based forecasting technique. Rule-based forecasting automates some tasks associated with maintaining a complex body of knowledge, given in an accessible and modifiable form, to provide more accurate forecasts. There are several examples of rule-based forecast systems (Collopy & Armstrong 1992, Armstrong, Adya & Collopy 2001), and we think the proposed approach could provide a new way of looking at them.

The potential and benefits of using expert systems in forecasting tasks, both to formulate the model and to select the forecasting model, has been demonstrated (Arinze 1994, Armstrong et al. 2001, Collopy, Adya & Armstrong 2001). The use of intelligent decision-support systems based on neural network technology for model selection and forecasting has been explored (Venkatachalam & Sohl 1999).

The framework developed in this article might be incorporated in the future into expert systems as well as into intelligent forecasting systems.

## **7. Conclusions**

Competitive settings, real-time econometrics and deregulated management frameworks will require an increasing number of forecast horizons, particularly in the globalised FX market.

Economists, engineers, forecasters and managers will have at their disposal a set of forecasts for the magnitudes of interest obtained by using several methods.

The aim of this paper has been to provide a starting point for the handling of the above scenario by showing the possibility of using multivariate analysis (especially principal components analysis and factor analysis) in combining forecasts. A conceptual framework has been developed and suggestions for using several multivariate analysis methods to solve the problem of how to effectively combine forecasts have been made. In addition, a new methodology for the construction of combined forecasts has been proposed.

In particular, it has been shown that the employment of PCA in the considered scenario of FX markets is very promising for:

1. Reinforcing the properties of easy combinations like that of simple average.
2. Selecting the methods to include in the combination.

Moreover, some practical contexts have been considered and the usefulness of the proposed methodology in those contexts and for some patterns of time series commented on.

It seems that multivariate analysis may be very useful in combining forecasts. However, much more research is required to evaluate the exciting possibilities offered by these methods in forecasting, in general, and in combining economic, financial or technological forecasts, in particular.

## Acknowledgements

The author acknowledges the Editor of this special issue and two referees all the suggestions that have proposed, which have helped to improve the first version.

[Recibido: septiembre de 2010 — Aceptado: abril de 2011]

## References

- Andrawis, R. R., Atiya, A. F. & El-Shishiny, H. (2011), ‘Combination of long term and short term forecasts, with application to tourism demand forecasting’, *International Journal of Forecasting*. Article in Press.
- Arinze, B. (1994), ‘Selecting appropriate forecasting models using rule induction’, *Omega* **22**(6), 647–658.
- Armstrong, J. S. (2001), Combined forecasts, *in* ‘Principles of Forecasting: A Handbook for Researchers and Practitioners’, Kluwer Academic Publishers, Norwell, MA.
- Armstrong, J. S., Adya, M. & Collopy, F. (2001), Rule-based forecasting: using domain knowledge in time series extrapolation, *in* ‘Principles of Forecasting: A Handbook for Researchers and Practitioners’, Kluwer Academic Publishers, Norwell, MA.
- Bank for International Settlements (2001), ‘Central bank survey of foreign exchange and derivatives market activity in April 2001’, *Press release, 31/2001E, Bank for International Settlements*.
- Basilevsky, A. (2004), *Statistical Factor Analysis and Related Methods. Theory and Applications*, John Wiley & Sons, New York.
- Bates, J. & Granger, C. (1969), ‘The combination of forecasts’, *Operations Research Quarterly* **20**(6), 451–468.
- Bunn, D. W. (1988), ‘Combining forecasts’, *European Journal of Operational Research* **33**(3), 223–229.
- Bunn, D. W. (1989), ‘Forecasting with more than one model’, *Journal of Forecasting* **8**(3), 161–166.
- Clemen, R. T. (1989), ‘Combining forecasts: A review and annotated bibliography’, *International Journal of Forecasting* **5**(4), 559–583.

- Clemen, R. T. & Winkler, R. L. (1986), 'Combining economic forecasts', *Journal of Business and Economic Statistics* **4**, 39–46.
- Collopy, F., Adya, M. & Armstrong, J. S. (2001), *Expert Systems for Forecasting. In Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer Academic Publishers, Norwell, MA.
- Collopy, F. & Armstrong, J. S. (1992), 'Rule-based forecasting: Development and validation of an expert system approach to combining time series extrapolations', *Management Science* **38**(10), 1392–1414.
- de Menezes, L. M., Bunn, D. W. & Taylor, J. W. (2000), 'Review of guidelines for the use of combined forecasts', *European Journal of Operational Research* **120**(1), 190–204.
- Diebold, F. X. (2001), *Elements of Forecasting*, South-Western, Cincinnati, Ohio.
- Donaldson, R. G. & Kamstra, M. (1996), 'Forecast combining with neural networks', *Journal of Forecasting* **15**(1), 49–61.
- Dunis, C. & Williams, M. (2005), 'Modelling and trading the EUR/USD exchange rate: Do neural network models perform better?', *Derivatives Use, Trading and Regulation* **8**(3), 211–239.
- Elliott, G., Granger, C. W. J. & Timmerman, A. (2006), *Handbook of Economic Forecasting*, Elsevier, Amsterdam.
- Eun, C. S. & Sabherwal, S. (2002), 'Forecasting exchange rates: Do banks know better?', *Global Finance Journal* **13**(2), 195–215.
- Gorsuch, R. L. (1983), *Factor Analysis*, second edn, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Granger, C. W. J. & Ramanathan, R. (1984), 'Improved methods of combining forecasts', *Journal of Forecasting* **3**(2), 197–204.
- Hahn, H., Meyer-Nieberg, S. & Pickl, S. (2009), 'Electric load forecasting methods: Tools for decision making', *European Journal of Operational Research* **199**(3), 902–907.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L. & Black, W. C. (1998), *Multivariate Data Analysis with Readings*, 5th edn, Prentice Hall: Englewood Cliffs, New Jersey.
- Jobson, J. D. (1992), *Applied Multivariate Data Analysis. Vol. II: Categorical and Multivariate Methods*, Springer-Verlag, New York.
- Johnson, R. A. & Wichern, D. W. (2002), *Applied Multivariate Statistical Analysis*, 5th edn, Prentice Hall: Englewood Cliffs, New Jersey.
- Kang, I. B. (2003), 'Multi-period forecasting using different models for different horizons: an application to U.S. economic time series data', *International Journal of Forecasting* **19**(3), 387–400.



- Kapetanios, G., Labhard, V. & Price, S. (2008), 'Forecast combination and the bank of England's suite of statistical forecasting models', *Economic Modelling* **25**(4), 772–792.
- Kilian, L. & Taylor, M. P. (2003), 'Why is it difficult to beat the random walk forecast of exchange rates?', *Journal of International Economics* **60**(1), 85–107.
- King, M. R. & Rime, D. (2010), 'The \$4 trillion question: what explains FX growth since the 2007 survey?', *BIS Quarterly Review* (4), 27–42.
- Kusiak, A. & Li, W. (2010), 'Short-term prediction of wind power with a clustering approach', *Renewable Energy* **35**(10), 2362–2369.
- Makridakis, S., Wheelwright, S. C. & Hyndman, R. (1998), *Forecasting Methods and Applications*, John Wiley & Sons, New York.
- Makridakis, S. & Winkler, R. L. (1983), 'Averages of forecasts: Some empirical results', *Management Science* **29**, 987–996.
- Maté, C. & Calderón, R. (2000), 'Exploring the characteristics of rotating electric machines with factor analysis', *Journal of Applied Statistics* **27**(8), 991 – 1006.
- Pindyck, R. S. (1999), 'Long-run evolution of energy prices', *Energy Journal* **20**(2), 1–27.
- Poncela, P., Rodríguez, J., Sánchez-Mangas, R. & Senra, E. (2011), 'Forecast combination through dimension reduction techniques', *International Journal of Forecasting* **27**(2), 224–237.
- Posen, A. S. (2004), 'Avoiding a currency war', *The International Economy* pp. 10–15.
- Ramanathan, R., Engle, R., Granger, C. W. J., Vahid-Araghi, F. & Brace, C. (1997), 'Short-run forecasts of electricity loads and peaks', *International Journal of Forecasting* **13**(2), 161–174.
- Rapach, D. E. & Strauss, J. K. (2008), 'Forecasting U.S. employment growth using forecast combining methods', *Journal of Forecasting* **27**(1), 75–93.
- Rencher, A. C. (2002), *Methods of Multivariate Analysis*, second edn, John Wiley & Sons, New York.
- Sánchez-Ubeda, E. F. & Berzosa, A. (2007), 'Modeling and forecasting industrial end-use natural gas consumption', *Energy Economics* **29**(4), 710–742.
- Tabachnik, B. G. & Fidell, L. (1996), *Using Multivariate Statistics*, 3rd edn, HarperCollins Publishers, New York.
- Timmerman, A. (2006), Forecast combinations, in G. Elliott, C. W. J. Granger & A. Timmerman, eds, 'Handbook of Economic Forecasting', Vol. 1, Elsevier, Amsterdam.

- Venkatachalam, A. R. & Sohl, J. E. (1999), 'An intelligent model selection and forecasting system', *Journal of Forecasting* **18**(3), 167–180.
- Winkler, R. L. & Makridakis, S. (1983), 'The combination of forecasts', *Journal of the Royal Statistical Society, Series A* **146**, 150–157.
- Winklhofer, H., Diamantopoulos, A. & Witt, S. (1996), 'Forecasting practice: A review of the empirical literature and an agenda for future research', *International Journal of Forecasting* **12**(2), 193–221.
- Zhao, J. H., Dong, Z. Y., Xu, Z. & Wong, K. P. (2008), 'A statistical approach for interval forecasting of the electricity price', *IEEE Transactions on Power Systems* **23**(2), 267–276.

## Appendix

### A1. Introducción

El siglo XXI está caracterizándose por ofrecer múltiples opciones para cada problema y situación del mundo real. Muchos hechos corroboran la afirmación anterior; por ejemplo una rápida transferencia de conocimientos del mundo académico al no académico, una comunidad científica con un conocimiento cada vez más amplio, programas de ordenador sencillos y flexibles y un incremento en cuanto a disponibilidad de nuevas tecnologías (véase, como ejemplo de lo anterior, el caso de la película *La red social*).

En nuestros días la predicción se caracteriza por la disponibilidad de un gran número de métodos. Se pueden consultar varias fuentes para actualizar el conocimiento de los mencionados métodos (por ejemplo, Armstrong 2001, Diebold 2001, Elliott et al. 2006, Makridakis et al. 1998).

Además, los horizontes de predicción para las empresas y las instituciones son cada vez más diferentes, aunque se encuentran relacionados. Quienes toman las decisiones necesitan predicciones relacionadas con el plazo inmediato, así como para el corto, el mediano y el largo plazo, o incluso para los tres horizontes estándar: corto, mediano y largo. Por ejemplo, Sánchez-Ubeda & Berzosa (2007) proponen una aplicación para pronosticar el consumo de gas natural de los usuarios finales en España, donde la empresa ENAGAS necesita predicciones para un horizonte de mediano plazo (1 a 3 años) con una resolución muy elevada (días). Otros trabajos que abordan dicho problema son Kang (2003), Andrawis et al. (2011) y las referencias citadas en ambos.

La combinación de pronósticos se considera una metodología apropiada para manejar el escenario anterior, habiendo alcanzado madurez y reconocimiento en los últimos 40 años (Bates & Granger 1969, Bunn 1988, Clemen 1989, de Menezes et al. 2000, Armstrong 2001, Timmerman 2006). En economía, una aplicación reciente se dedica a pronosticar el crecimiento del empleo en Estados Unidos (Rapach &

Strauss 2008). Además, Kapetanios et al. (2008) constatan que la combinación de pronósticos por lo general conduce a una reducción en el error de predicción.

Los escenarios anteriores son conceptualmente adecuados para la aplicación de los *métodos de análisis multivariante* (en lo sucesivo MAM). En los últimos quince años estas técnicas han experimentado de grandes mejoras en cuanto a la facilidad de uso, la superación de ciertas limitaciones y el reconocimiento de una utilidad mayor (Hair et al. 1998). En consecuencia, una adaptación de los MAM a las nuevas tecnologías como bases de datos, Internet, etc., es un área nueva.

Dado este escenario, este artículo propone el empleo de varios métodos de análisis multivariante clásico para obtener pronósticos combinados. Un artículo reciente (Poncela et al. 2011) propone el uso de técnicas de reducción de la dimensión en la combinación de pronósticos. Nuestro enfoque, sin embargo, es diferente porque considera distintos horizontes de predicción, analiza las posibilidades en la combinación de pronósticos del conjunto clave de métodos multivariantes y sugiere diferentes campos de aplicación, como los mercados financieros o los mercados de energía.

Esta versión reducida en español se organiza de la siguiente forma. La sección A2 se dedica al análisis de un marco de combinación de pronósticos y resalta la aproximación de los métodos del análisis multivariante. La sección A3 presenta la utilidad de la metodología propuesta en diferentes contextos prácticos. Por último, se muestran algunas conclusiones de la investigación.

## A2. Combinación de pronósticos bajo el marco del análisis multivariante

En lo que sigue se supondrá que estamos interesados en pronosticar una magnitud,  $Y$ , para un número grande de horizontes de pronósticos, denotado por  $n$ , a partir de la información sobre la magnitud en diferentes periodos  $1, 2, \dots$  hasta el momento (denotado por  $t$ ). Con este objetivo en mente, el profesional de la predicción dispone de varios pronósticos obtenidos mediante un número de métodos de predicción, notado por  $p$ .

Por tanto, supondremos que tenemos  $p$  pronósticos centrados o no sesgados de la magnitud  $Y$  obtenidos mediante  $p$  métodos (o expertos o una combinación de métodos y expertos) ( $M1, M2, \dots, Mp$ ) para  $n$  horizontes de pronóstico ( $H1, H2, \dots, Hn$ ). La notación  $f_{ij}(t)$  se refiere al pronóstico dado para el horizonte  $i$ -ésimo por el método o experto  $j$ -ésimo, con  $i = 1, \dots, n$  y  $j = 1, \dots, p$ . Esta información se puede mostrar en una estructura de matriz (tabla 2), notada por

$$F(t)$$

En el escenario anterior, resultará apropiada la aplicación de los métodos de análisis multivariante (Johnson & Wichern 2002, Rencher 2002, Hair et al. 1998), considerando variables los métodos de predicción y unidades los horizontes de pronóstico.

A partir de esta información, los MAM se han desarrollado para el estudio de diferentes problemas en el análisis de datos. En nuestro caso, los pronósticos se obtienen empleando varios métodos para diferentes horizontes de pronóstico. Uno de los problemas estudiado es la reducción de las dimensiones del problema (métodos de pronóstico), cuyo objetivo es facilitar el proceso de asimilación de las relaciones descubiertas por el analista.

Otro problema es el proceso de búsqueda de relaciones de dependencia que simplificaría la comprensión del comportamiento de algunos métodos de predicción, teniendo en cuenta los valores que muestren otros métodos. Por último, un problema interesante es el análisis de las posibles relaciones de interdependencia que permitirá detectar relaciones sutiles entre los horizontes de pronóstico.

Una clasificación inicial de los MAM, aplicada la estructura de la tabla 2 se resume en la tabla 3, que recoge el uso propuesto de los MAM en la predicción. Nos referimos a la predicción para diferentes horizontes, los métodos de pronóstico y las tareas de analizar dependencia e interdependencia entre métodos, o al problema de reducir el número de métodos, o al de agrupar o clasificar horizontes de pronóstico.

A continuación se considera brevemente el potencial de los MAM en los cinco problemas de la tabla 3.

**Problema 1.** En nuestro escenario de predicción, la tarea más evidente sería la reducción de los  $p$  pronósticos nosegados a un número más pequeño de pronósticos (idealmente uno, o a lo sumo 2), para cada horizonte de pronóstico considerado. Con este objetivo en mente, métodos como el Análisis de Componentes Principales (PCA) o el Análisis Factorial (FA), lineal o no lineal, o incluso el Análisis de Conglomerados aplicado a los métodos de pronóstico, llamarían nuestra atención. En un contexto general, la aproximación de conglomerados en predicción tiene una larga tradición. Se puede consultar Kusiak & Li (2010) para una aplicación reciente en la predicción de la producción de energía eólica.

**Problema 2.** Granger & Ramanathan (1984) han sugerido que los métodos convencionales de combinar pronósticos podrían considerarse en el marco de la regresión. Por tanto, deberían explorarse los métodos alternativos para manejar este problema en la Tabla 1.

Por ejemplo, MANOVA se podría emplear para analizar si hay diferencias significativas entre un grupo de pronósticos obtenidos utilizando la misma metodología (ARIMA, redes neuronales, opiniones expertas, etc.) en diferentes grupos de horizontes de tiempo como corto, mediano o largo plazo. En función de la conclusión alcanzada, sería diferente la decisión del profesional de la predicción sobre la mejor forma de combinar.

CCA se podría utilizar para determinar el tipo de relación (lineal o no lineal) entre grupos de pronósticos a fin de decidir sobre la posibilidad de reducir el número de pronósticos que van a considerarse en el problema.

**Problema 3.** En función de los métodos de predicción utilizados, la utilidad de conocer el perfil de los horizontes de pronóstico con respecto a la exactitud está fuera de duda. Como han manifestado muchos autores, no hay ningún método ni metodología que supere o rebase al resto en cada problema considerado o en cada

horizonte de tiempo. Por consiguiente, parece que el DA y la LR podrían ayudar en investigaciones futuras sobre estas cuestiones. En concreto, las aproximaciones en FOREX plantean nuevos retos en finanzas internacionales.

**Problema 4.** Identificar grupos de horizontes de pronóstico similares con respecto a una serie de métodos o metodologías de pronósticos haría posible la configuración de diferentes tipos de combinación para cada conglomerado de horizontes de pronósticos.

**Problema 5.** Este es uno de los problemas más complejos en los MAM. Eso sucede en la tabla 3. Aquí, los Mínimos Cuadrados Parciales (PLS) podrían ser una metodología útil al resolver el problema comentado de decidir qué métodos se deberían incluir al agrupar en conglomerados los horizontes de pronóstico.

### A3. Algunos contextos prácticos y algunas cuestiones de aplicación

En esta sección se consideran tres contextos prácticos de predicción. Se analiza la utilidad de la metodología propuesta en dichos contextos, así como en algunos patrones de series temporales.

#### Mercados de energía

La predicción es una tarea muy importante para las personas que trabajan gestionando los mercados de energía, incluso cuando hay un escenario económico estable. Los productores de energía necesitan con urgencia pronósticos exactos que se encuentren automatizados, de manera que puedan anticipar la demanda, los precios y otras variables del mercado.

En la mayor parte de las zonas económicas más relevantes del mundo, el sector de la energía está acometiendo un importante proceso de reestructuración cuya meta es la competencia y la liberalización de cuestiones de regulación que impiden la libre competencia de los productores y proveedores de la energía. Todo esto implica que la gestión de las empresas productoras de energía ya no puede estar basada en procedimientos administrativos y centralizados; su beneficio ahora depende directamente de las decisiones gerenciales que se tomen. Los procesos de la toma de decisiones son cada vez más complejos y arriesgados, debido a la mayor incertidumbre que resulta del nuevo marco de operaciones en los mercados de energía del siglo XXI.

En investigaciones previas se ha resaltado que un cálculo cuidadoso de la demanda esperada y de la oferta en los mercados de energía puede conducir a contratos que incrementen los beneficios de las empresas. Por tanto, se va a requerir más investigación para pronósticos a muy corto plazo (Ramanathan et al. 1997). En este artículo, se ha considerado el desarrollo de una serie de modelos para producir pronósticos a muy corto plazo en sistemas de demanda horaria. La disponibilidad de pronósticos para toda la semana, 168 horas, es el estándar en las empresas de energía. Por ejemplo, Hahn et al. (2009) recogen información de una encuesta so-

bre métodos matemáticos e ideas que se han empleado para predecir la demanda de electricidad.

La metodología propuesta en este artículo se puede aplicar al cálculo de pronósticos de precios en las empresas de energía. Según Zhao et al. (2008), obtener estos pronósticos de precios constituye una tarea difícil pero esencial para los participantes en un mercado de electricidad desregulado. Pindyck (1999) analiza la influencia de los pronósticos de precios sobre las decisiones de inversión y la elección de los productos que se van a producir.

#### **Mercados de acciones**

El cálculo de pronósticos resulta una tarea crucial para los agentes de bolsa y las personas involucradas en la gestión de las acciones. Se podría mejorar la planificación de estrategias mediante la disponibilidad de pronósticos para retornos, precios y volatilidad, en algunos periodos del día. Por ejemplo, si un cuarto de hora se toma como unidad de división del tiempo, se van a necesitar pronósticos para 32 horizontes en un día de 8 horas. Obviamente, se utilizan con mucha frecuencia otras divisiones de tiempo; por ejemplo, 5 minutos. Por tanto, si como se ha sugerido (Makridakis & Winkler 1983) se obtuvieran pronósticos para los 32 horizontes de tiempo utilizando 4 o 5 métodos, el uso de la metodología propuesta podría proporcionar pronósticos más exactos a los agentes de bolsa.

#### **Ventas**

En algunas investigaciones (Winklhofer et al. 1996) se ha estudiado la cuestión del horizonte de tiempo en el ámbito de los pronósticos en ventas. Algunas conclusiones fueron que la mayoría de las empresas preparaba pronósticos de ventas sobre una base anual, que el horizonte más popular en los pronósticos de ventas a corto plazo era un mes y que las empresas que operan en mercados muy competitivos ponen más énfasis en el corto plazo que en el largo.

Esto sugiere que la aproximación propuesta en este artículo no será útil para muchas empresas que emplean pronósticos de ventas. Sin embargo, el contexto crecientemente competitivo donde algunas empresas desarrollan su actividad, en una situación con cierto paralelismo con la de los mercados de energía y la de los mercados de acciones, puede incrementar la necesidad de considerar interesante emplear más métodos de predicción y un gran número de horizontes de pronóstico. Una vez que estos cambios fueran adoptados se podrían manejar con la aproximación propuesta.

#### **Series estacionales**

Series con una componente estacional reproducen el comportamiento de la magnitud cada cierto periodo. En consecuencia, se puede suponer que los horizontes de pronóstico actúan como unidades de estudio, de manera similar a los problemas donde el análisis multivariante se aplica habitualmente.

#### **Series con tendencia**

Las series que muestran una componente de tendencia no se comportan con el mismo patrón a través de varios horizontes de pronóstico. Por consiguiente, en esta situación los horizontes de pronóstico quizás no muestran la misma estructura que las unidades en los problemas usuales del análisis multivariante.

### **Series cíclicas**

Las series que presentan una componente de ciclo con sus correspondientes puntos de giro deben retornar a una estructura previa en el comportamiento de la magnitud. Por tanto, como ocurre con las series estacionales, se puede suponer que los horizontes de pronóstico actúan como unidades de estudio, de manera similar a los problemas donde los MAM se aplican habitualmente.

### **Automatización**

Uno de los puntos clave para una implantación con éxito de la metodología propuesta es la automatización de la misma. Por esta razón, las cuestiones de versatilidad, exactitud, velocidad e inteligibilidad se deben considerar de manera rotunda, porque solo habrá disponible una cantidad limitada de recursos (software, conocimiento, datos, etc.).

### **Métodos de cálculo de pronósticos a incluir**

Varias metodologías de pronóstico clásicas (alisado exponencial, ARIMA, GARCH, etc.) se emplean ampliamente como parte de los sistemas de predicción, dando resultados satisfactorios cuando se aplican en diferentes campos. Proporcionan muchas posibilidades de modelado para una o varias variables, puesto que permiten considerar horizontes y comportamientos diferentes. Además, presentan rasgos muy intuitivos y son muy favorables, ya que los parámetros que manejan representan al proceso de manera simple.

Además, muchas de las variables incluidas en las aplicaciones prácticas se comportan bien en muchos casos o momentos de su historia, justificando la opinión de que los modelos lineales o clásicos de cálculo de pronósticos se pueden ver adecuados para tales contextos. Sin embargo, debería considerarse la posibilidad de emplear Redes Neuronales Artificiales y otros métodos no lineales.

Hay varias referencias para un estudio detallado de los métodos generales de cálculo de pronósticos (Makridakis et al. 1998, Diebold 2001, Elliott et al. 2006).

### **Integración con predicción basada en reglas y sistemas expertos**

Una técnica de pronóstico que integra opinión experta y extrapolación estadística se denomina técnica de pronóstico basada en reglas. La predicción basada en reglas automatiza algunas tareas asociadas al mantenimiento de un cuerpo complejo de conocimiento, dado en una forma accesible y modificable, a fin de proporcionar pronósticos más exactos. Existen varios ejemplos de sistemas de elaboración de pronósticos basados en reglas (Collopy & Armstrong 1992, Armstrong et al. 2001). El enfoque propuesto podría proporcionar una nueva perspectiva a los citados sistemas.

En distintos trabajos (Arinze 1994, Armstrong et al. 2001, Collopy et al. 2001) se ha mostrado el potencial y los beneficios de emplear sistemas expertos en las tareas de predicción, tanto para formular el modelo como para seleccionarlo. También, se ha explorado el uso de sistemas inteligentes de apoyo a la decisión basados en redes neuronales para la selección del modelo y el cálculo de pronósticos (Venkatachalam & Sohl 1999).

El marco desarrollado en este artículo se podría incorporar en el futuro tanto en sistemas expertos como en sistemas inteligentes de cálculo de pronósticos.



## A4. Conclusiones

Entornos competitivos, econometría en tiempo real y marcos de gestión liberalizados van a requerir un número creciente de horizontes de pronósticos; particularmente en el mercado globalizado de divisas.

Economistas, ingenieros, gerentes y profesionales de la predicción tendrán a su disposición, para las magnitudes de interés, un conjunto de pronósticos obtenidos mediante el empleo de varios métodos.

El objetivo de este artículo es proporcionar un punto inicial para el manejo del escenario anterior, mostrando la posibilidad de utilizar análisis multivariante (especialmente análisis de componentes principales y análisis factorial) al combinar pronósticos. Se ha presentado un marco conceptual y se han dado sugerencias para utilizar métodos de análisis multivariante para resolver el problema de combinar pronósticos de forma eficaz. Además, se ha propuesto una nueva metodología para la construcción de pronósticos combinados.

En concreto, se ha mostrado que el empleo de PCA en el escenario considerado de los mercados de divisas es muy prometedor para lo siguiente.

1. Reafirmar ciertas propiedades que tienen las combinaciones sencillas de pronósticos, como el promedio simple.
2. Seleccionar los métodos que se van a incluirse en la combinación.

Asimismo, se han considerado algunos contextos prácticos y la utilidad de la metodología propuesta en estos para algunos patrones de series temporales.

Parece que el análisis multivariante puede ser muy útil a la hora de combinar pronósticos. Sin embargo, se requiere mucha más investigación a fin de evaluar las excitantes posibilidades que ofrecen estos métodos en la predicción, en general, y al combinar pronósticos económicos, financieros o tecnológicos, en particular.

## Agradecimientos

El autor agradece muy sinceramente al Editor de este número especial y a los dos árbitros que han revisado el manuscrito todas las sugerencias que han propuesto, las cuales han contribuido a mejorar notablemente la primera versión del artículo.





## Funciones de producción, análisis de economías a escala y eficiencia técnica en el eje cafetero colombiano: una aproximación con frontera estocástica

### Estimation of the Production Functional Form, Returns to Scale and Technical Efficiency in Colombian Coffee Zone by Means Stochastic Frontier

JORGE ANDRÉS PERDOMO<sup>a</sup>, DARRELL LEE HUETH<sup>b</sup>

DEPARTAMENTO DE ECONOMÍA AGRÍCOLA Y RECURSOS NATURALES, UNIVERSIDAD DE MARYLAND, MARYLAND, ESTADOS UNIDOS  
FACULTAD DE ECONOMÍA, UNIVERSIDAD DE LOS ANDES, BOGOTÁ, COLOMBIA

---

#### Resumen

El objetivo principal de este estudio consiste en estimar distintas formas (convencionales y flexibles) funcionales de producción cafetera en Colombia, mediante frontera estocástica. Con los resultados, calcular y analizar las economías a escala por unidad cafetera (productores pequeños, medianos, grandes y sector general cafetero), así como obtener y evaluar la eficiencia técnica de los caficultores asentados en Caldas, Quindío y Risaralda.

**Palabras clave:** elasticidad, estimación máximo-verosimil, forma funcional, función de producción, regresión estocástica.

#### Abstract

This article is an empirical study about flexible and conventional functional forms of coffee production, minflex Laurent Translog function econometrically has been established in Colombia coffee zone for the farm size (smallholders, medium and large farms, general sector), using a stochastic frontier model through standard maximum likelihood method. Likewise, their returns to scale and technical efficiency were derived.

**Key words:** Elasticity, Functional form, Maximum likelihood estimation, Production function, Stochastic regression.

---

<sup>a</sup>Profesor. E-mail: jor-perd@uniandes.edu.co

<sup>b</sup>Profesor emérito. E-mail: dhueth@arec.umd.edu.

## 1. Introducción

Colombia es el tercer productor cafetero y el principal cultivador de café arábica lavado mundial (Banco Mundial 2002). Desde 1870 empezó a desarrollar su producción comercial; actualmente, representa el 2 % del producto interno bruto nacional (PIB). Según la Federación Nacional de Cafeteros de Colombia (Fedecafé), existen aproximadamente 560.000 fincas dedicadas a cultivar café, en las cuales se encuentran pequeños<sup>1</sup> (minifundistas), medianos<sup>2</sup> (campesinos) y grandes productores<sup>3</sup> (empresariales).

Por esta razón, el sector es importante en la economía, la cultura y el ámbito social colombiano. En estas circunstancias y para determinar adecuadamente las economías a escala por unidad de producción (pequeños, medianos, grandes y sector general), establecer la eficiencia técnica (ET) y diseñar con los resultados recomendaciones que puedan mejorar la productividad en la actividad, el objetivo principal del estudio consiste en estimar, mediante frontera estocástica (FE), la forma funcional de producción cafetera en los departamentos de Caldas, Quindío y Risaralda (Colombia).

Esto ocurre porque en el eje cafetero colombiano falta información estadística microeconómica sobre la actividad lo que limita la elaboración de un diagnóstico y análisis microeconómico confiable del sector y de los caficultores colombianos. Por consiguiente, disminuye las probabilidades de éxito para cualquier estrategia dirigida a ellos. Así, profundizar el tema permitirá diseñar políticas y programas institucionales acertados y pertinentes (García & Ramírez 2002). En la zona existen hipótesis sobre fragmentación de tierras cafeteras y diversificación de cultivos (Guhl 2004).

Estos factores posiblemente están disminuyendo la eficiencia, la productividad y la competitividad del grano en Colombia, ocasionando altos costos de producción, comparado con Vietnam y Brasil. Por consiguiente, para el diseño de políticas apropiadas al sector debe estimarse una función de producción cafetera correcta desde la perspectiva microeconómica y econométrica. En este sentido, conviene especificar una forma funcional adecuada que no afecte significativamente los resultados y los análisis de las economías a escala y eficiencia técnica, obtenidas mediante el método paramétrico de frontera estocástica (Konstantinos, Kien & Vangelis 2003).

Para alcanzar estadística y econométricamente lo expuesto, el documento se encuentra dividido de la siguiente manera: la sección 2 comprende el estado del arte mediante una revisión literaria nacional e internacional en el tema. La sección 3 presenta el marco teórico y metodológico con los aspectos más destacados sobre formas funcionales de producción y frontera estocástica. La sección 4 contiene re-

---

<sup>1</sup>El 64 % de los caficultores son minifundistas con menos de media hectárea sembrada en café (Ramírez, Silva, Valenzuela, Villegas & Villegas 2002).

<sup>2</sup>El 31 % de los productores corresponden a unidades empresariales cafeteras campesinas con un promedio de 2,2 hectáreas sembradas en café (Ramírez et al. 2002).

<sup>3</sup>Los cafeteros empresariales son el 5 % del total de unidades productivas, con fincas que fluctúan entre 7 y 35 hectáreas cultivadas en café (Ramírez et al. 2002).

sultados empíricos. En la 5 están expuestas las conclusiones y sugerencias derivadas del trabajo.

## 2. Literatura relacionada

La literatura sobre eficiencia productiva o técnica en la actividad cafetera colombiana está surgiendo (Perdomo 2006, Perdomo & Hueth 2010). Igualmente, la técnica de frontera estocástica (FE) no ha sido aplicada en el sector con el fin de estimar formas funcionales de producción, economías a escala y eficiencia técnica, para diseñar políticas que ayuden a mejorar los rendimientos en el cultivo del grano. No obstante, Perdomo, Hueth & Mendieta (2007) y Perdomo & Mendieta (2007) obtuvieron la eficiencia técnica y asignativa mediante el método no paramétrico DEA (análisis envolvente de datos) en la zona cafetera de Colombia. Para nuestro conocimiento, hay pocos estudios nacionales al respecto.

Sin embargo, internacionalmente existe una amplia investigación empírica sobre la eficiencia económica de agricultores para países desarrollados y en desarrollo utilizando FE (Kumbhakar 1993, Battese & Broca 1991, Reinhard, Lovell & G. 1999, Thanda & Matthias von 1999, Donnell & Griffiths 2006, Lohr & Park 2006). Aunque en producción cafetera, la evidencia empírica indica que para medir eficiencia los autores se han centrado especialmente en especificar y estimar funciones translogarítmica y Cobb-Douglas (Coelli & Fleming 2003, Saravia 2007, Wollni 2007, Cárdenas, Vedenov & Houston 2008).

Por consiguiente, solo dos estudios llevados a cabo en las fincas de olivo griegas (Konstantinos et al. 2003) y las granjas de Estados Unidos (Ornelas, Shumway & Ozuna 1994) consideraron los efectos sobre las medidas de economías a escala y la eficiencia obtenidas desde la técnica paramétrica de frontera estocástica por elegir formas funcionales de producción inapropiadas. Este análisis no ha sido realizado nacional e internacionalmente en producción cafetera. Como consecuencia, estos trabajos están estrechamente relacionados con nuestro documento. Finalmente, mediante análisis envolvente de datos (DEA), Joachim, Kalilou, Ibrahim & Gwendoline (2003), Rios & Shively (2005) y Mosheim (2002) estimaron la eficiencia técnica en fincas productoras de café en Costa de Marfil (África), Vietnam (Asia) y Costa Rica, respectivamente.

## 3. Funciones de producción y frontera estocástica

Esta sección comprende el marco teórico microeconómico y metodológico sobre frontera estocástica para analizar la producción por unidad productiva y sector cafetero colombiano. Así, la relación entre la cantidad de insumos (hectáreas cultivadas  $-x_{1i}-$ , mano de obra utilizada  $-x_{2i}-$ , maquinaria requerida  $-x_{3i}-$  y fertilizantes aplicados  $-x_{4i}-$  para la finca  $i$ ) y la producción de café ( $q_i$ , cantidad de café en arro-

bas para el año 2003 para la finca  $i$ )<sup>4</sup> bajo formas funcionales convencionales<sup>5</sup> y flexibles (Diewert 1974) puede apreciarse en la tabla 1<sup>6</sup>.

TABLA 1: Funciones de producción convencionales y flexibles.

Funciones de producción convencionales
Proporciones fijas - Leontief
$(\sigma = 0, \rho \rightarrow -\infty) : q_i = \min(\beta_1 x_{1i}, \beta_2 x_{2i}) \Rightarrow \beta_1, \beta_2 > 0, ESC(\sigma)$
Cobb-Douglas (CD)
$(\sigma = 1, \rho = 0) : q_i = f(x_{1i}, x_{2i}) = \beta_0 x_{1i}^{\beta_1} x_{2i}^{\beta_2} \Rightarrow \beta_0, \beta_1, \beta_2 > 0$
ESC (Elasticidad de Sustitución Constante, $\sigma$ )
$q_i = f(x_{1i}, x_{2i}) = [\beta_1 x_{1i}^\rho + \beta_2 x_{2i}^\rho]^{\frac{1}{\rho}} \Rightarrow \rho \leq 1, \rho \neq 0, \rho \rightarrow -\infty, \sigma > 0$
Cuadrática
$q_i = f(x_{1i}, x_{2i}) = \beta_1 x_{1i} x_{2i} + \beta_2 x_{1i}^2 + \beta_3 x_{2i}^2$
Funciones de producción flexibles
Leontief generalizada (Diewert)
$q_i = \beta_0 + 2\beta_1 \sqrt{x_{1i}} + 2\beta_2 \sqrt{x_{2i}} + 2\beta_3 \sqrt{x_{1i} x_{2i}}$
Translogarítmica
$\ln q_i = \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + \beta_3 \frac{1}{2} \ln x_{1i}^2 + \beta_4 \frac{1}{2} \ln x_{2i}^2 + \beta_5 \ln x_{1i} \ln x_{2i}$
Cuadrática generalizada
$q_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 \frac{1}{2} x_{1i} x_{2i}$
Cuadrática, raíz cuadrada (CRC)
$q_i = [\beta_0 + 2\beta_1 x_{1i} + 2\beta_2 x_{2i} + \beta_3 x_{1i} x_{2i}]^{0.5}$
Cuadrática generalizada Box-Cox
$\frac{q_i^{2\theta} - 1}{2\theta} = \beta_0 + \beta_1 \frac{x_{1i}^\lambda - 1}{\lambda} + \beta_2 \frac{x_{2i}^\lambda - 1}{\lambda} + \beta_3 \frac{1}{2} \frac{x_{1i}^\lambda - 1}{\lambda} \frac{x_{2i}^\lambda - 1}{\lambda}$
ESC, multifactores
$q_i = [\beta_1 x_{1i}^\rho + \beta_2 x_{2i}^\rho + \beta_3 x_{3i}^\rho + \beta_4 x_{4i}^\rho]^{\frac{1}{\rho}}$

$\sigma$  se refiere a la elasticidad de sustitución constante (esc) y  $\rho$  al parámetro de sustitución; así cada forma funcional depende de los valores tomados por estos coeficientes.

$$\sigma = \frac{1}{1+\rho}, \text{ si } \begin{cases} -1 < \rho < 0 \rightarrow \sigma > 1 \\ \rho = 0 \rightarrow \sigma = 1 \\ 0 < \rho < \infty \rightarrow \sigma < 1 \end{cases}$$

Para más detalles, véase Chiang (1984, p. 426) y Nicholson (2002, p. 280-284).

Fuente: autores a partir de Nicholson (2002), Diewert (1974), Konstantinos et al. (2003)

Perdomo, (2010, 2011), Mendieta & Perdomo (2008, p. 88).

Donde  $\beta_0, \beta_1, \beta_2, \beta_3$  y  $\beta_4$  describen los respectivos parámetros en cada modelo; sin embargo, su valor e interpretación difiere en cada función. Por ejemplo,  $\beta_0, \beta_1, \beta_2, \beta_3$  y  $\beta_4$  toma distintos valores e interpretaciones entre las formas Cobb-Douglas, cuadrática generalizada o cuadrática raíz cuadrada, lo que significa elasticidades en la primera, mientras que en las generalizadas carece de definición por ser funciones no lineales. Igualmente, es importante resaltar que los valores y

<sup>4</sup>El subíndice  $i$  define el corte transversal y el tamaño de la muestra, desde la finca 1 hasta la 999 ( $i = 1, 2, \dots, 999$ ), para los insumos ( $x_{1i}, x_{2i}, x_{3i}$  y  $x_{4i}$ ) y producción ( $q_i$ ) especificados.

<sup>5</sup>Generalmente simplifica el modelo mediante el análisis de uno o dos insumos ( $x_{1i}$  y  $x_{2i}$ ).

<sup>6</sup>Asumiendo dos insumos de producción ( $x_{1i}$  y  $x_{2i}$ ) para simplificar y facilitar su comprensión, teniendo en cuenta que la evidencia empírica expuesta en la sección 4 implicó los cuatro principales insumos empleados en el cultivo de café (hectáreas cultivadas -  $x_{1i}$  -, mano de obra utilizada -  $x_{2i}$  -, maquinaria requerida -  $x_{3i}$  - y cantidad de fertilizantes aplicados -  $x_{4i}$  -). Igualmente, en la tabla 1,  $\ln$  describe el logaritmo natural de la respectiva variable.

signos de los coeficientes estimados de las funciones no lineales (convencionales y flexibles) carecen de interpretación directa.

Tampoco puede inferirse algún tipo de conclusión sobre la intensidad y el relevancia entre un insumo y el café producido si los signos parciales de los parámetros estimados resultan negativos en los términos cuadráticos e interacciones y estadísticamente no significativos. Ante esto, para las funciones no lineales (convencionales y flexibles) deben calcularse los efectos marginales, elasticidades y economías a escala con su respectiva desviación estándar<sup>7</sup>, y así determinar la importancia y relación (directa o inversa, parcial y conjunta) entre los insumos y la producción cafetera (Greene 2002, p. 103-104). Prosiguiendo la descripción de la tabla 1,  $\theta$  y  $\lambda$  son los coeficientes de transformación no lineales para la función cuadrática generalizada Box-Cox y  $\rho$  representa el parámetro de sustitución, cuyo valor determina la elasticidad de sustitución constante (ESC o  $\sigma$ ) en su presentación convencional y flexible (Chiang 1984).

Por consiguiente, con los datos microeconómicos (insumo-producción) disponibles sobre los caficultores en el eje cafetero colombiano y mediante el método de frontera estocástica pueden estimarse las formas funcionales exhibidas en la tabla 1, y con sus resultados determinar las economías a escala y la eficiencia técnica para los pequeños, medianos y grandes productores del sector de cafeteros.

### 3.1. Frontera estocástica

La frontera estocástica (FE) consiste en ajustar las formas funcionales de producción descritas en la tabla 1 utilizando técnicas econométricas mediante máxima verosimilitud. Es una aproximación paramétrica propuesta por Aigner, Lovell & Schmidt (1977) y Meesen & Van den Broeck (1977). En este sentido y de manera simplificada, la eficiencia técnica resulta gráficamente (figura 1) encontrando la función distancia entre el valor observado ( $x - q$ , punto B) y óptimo derivado del modelo econométrico ( $x - q^*$ , punto A), para la cantidad insumo-producción cafetera. Esto indica que, con la misma intensidad del factor ( $x$ ), es posible lograr una mayor producción ( $q^*$ ), o disminuyendo su monto ( $x^*$ ), se obtiene la cantidad real de café ( $q$ ).

Asimismo, la medida de eficiencia técnica para cada caficultor por unidad productiva en Colombia puede obtenerse mediante frontera estocástica, especificando previamente una forma funcional de producción adecuada como las presentadas en la tabla 1, Debido a que los resultados de las economías a escala y la eficiencia técnica, derivados del método FE, son sensibles a la elección y la estimación funcional (Konstantinos et al. 2003, Rosales, Perdomo, Morales & Urrego 2010).

<sup>7</sup>Como ejemplo para  $x_{1i}$  y  $x_{2i}$ , los efectos marginales y las elasticidades están dados por  $\frac{\partial q_i}{\partial x_{1i}}$  y  $\frac{\partial q_i}{\partial x_{2i}}$  y  $\frac{\partial q_i}{\partial x_{1i}} \frac{x_{1i}}{q_i}$  y  $\frac{\partial q_i}{\partial x_{2i}} \frac{x_{2i}}{q_i}$ , respectivamente; la economía a escala es concebida de sumar las elasticidades parciales; por consiguiente, si es mayor que 1 exhibe rendimientos crecientes, menor que 1, decrecientes; e igual a 1 constantes a escala. Una conocidos vez los resultados, la importancia estadística de las elasticidades puede inferirse con el valor de la desviación estándar asociada a cada coeficiente implicado en su cálculo y multiplicado, respectivamente, por los valores promedios ( $\bar{x}_{1i}$  y  $\bar{x}_{2i}$ ) de las variables involucradas en el mismo. véase más detalles en Greene (2002, p. 104), y en el apéndice de este documento.

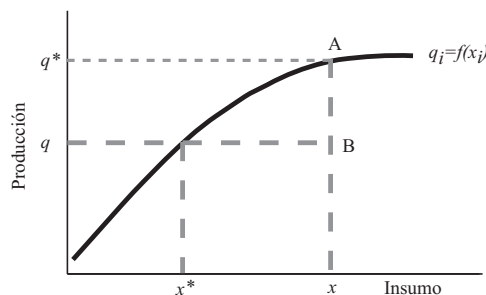


FIGURA 1: Función distancia.

Fuente. Reinhard et al. (1999, p. 51)

$$q_i = f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \beta) + \eta_i, \text{ donde } \eta_i = v_i - u_i \quad (1)$$

Así, siguiendo el esquema de Aigner et al. (1977), la frontera estocástica es representada en la ecuación (1) para cualquier función ( $f$ ) de producción cafetera convencional o flexible en la tabla 1. Igualmente,  $q_i$  indica la cantidad de café conseguido por la finca  $i$  ( $i = 1, 2, \dots, 999$ );  $x_{1i}$ ,  $x_{2i}$ ,  $x_{3i}$ , y  $x_{4i}$  son los principales insumos aplicados en la actividad, como fueron descritos inicialmente en esta sección.  $\beta$  es un vector de parámetros que debe estimarse empleando máxima verosimilitud que difiere en tamaño, resultados e interpretación dependiendo de la función especificada y estimada en la tabla 1;  $\eta_i$  hace referencia al componente estocástico compuesto por los elementos independientes  $v_i$  y  $u_i$ . Este procedimiento permite calcular el nivel máximo producido de café ( $q_i^*$ ), como lo representa la figura 1.  $v_i$  almacena las variaciones aleatorias<sup>8</sup> de la producción cafetera ( $q_i$ ), es simétrica e independientemente distribuida (iid) con media cero y varianza constante [ $v_i \sim N(0, \sigma_v^2)$ ] tomando valores positivos y negativos hacia el infinito ( $-\infty < v_i < \infty$ ). Mientras  $u_i$  acumula la ineficiencia técnica<sup>9</sup> observada en la caficultura, ( $q_i$ ) es un término asimétrico iid [ $u_i \geq 0 \sim N(0, \sigma_u^2)$ ] mayor que 0 e independiente de  $v_i$ .

Entonces, dadas estas características de  $\eta_i$ , los estimadores ( $\hat{\beta}$ ) de la frontera estocástica deben obtenerse mediante máxima verosimilitud, con el fin de conseguir parámetros eficientes, nosesgados y consistentes (Aigner et al. 1977). Por consiguiente, el logaritmo de la función de verosimilitud ( $\ln f$ ) es:

$$\ln(\sigma_s^2, \hat{\beta}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma_s^2) + \sum_{i=1}^n \ln[1 - \varphi(z_i)] - \frac{n}{2\sigma_s^2} \sum_{i=1}^n \left[ q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta}) \right]^2 \quad (2)$$

<sup>8</sup>Se refiere a eventos independientes de la producción cafetera, que no pueden controlar directamente los productores (cambio de políticas institucionales en los mercados, volatilidad de los precios de café, incertidumbre, riesgo, clima desfavorable, desastres naturales, entre otros).

<sup>9</sup>Dada por los insumos o variables de producción que pueden controlar los cafeteros, como la cantidad e intensidad de factores utilizados para producir.

Donde  $n$  denota el número de observaciones o fincas cafeteras encuestadas (999),  $\sigma_s^2$  la varianza del modelo<sup>10</sup> y  $\varphi(z_i)$ <sup>11</sup> la distribución normal estándar acumulada, en la cual gamma ( $\gamma$ )<sup>12</sup> representa el parámetro de eficiencia proveniente de las fuentes ( $v_i$  y  $u_i$ ) del error ( $\eta_i$ ) en la ecuación (1). En este sentido, cuando el efecto aleatorio predomina ( $\sigma_u^2 \rightarrow 0$  y  $\gamma = 0$ ), significa eficiencia o ausencia de ineficiencia técnica. Esto indica que los caficultores emplean adecuadamente la cantidad de insumos maximizando su producción, ubicados en cualquier sitio (punto A) sobre la frontera de la figura 1.

Sin embargo, simultáneamente puede existir ineficiencia por eventos inesperados en el cultivo de café porque la varianza aleatoria posiblemente tiende a infinito ( $\sigma_v^2 \rightarrow \infty$ ) y es predominante, haciendo que  $\eta_i$  converja a una distribución normal. En el caso contrario, cuando la variación del componente asimétrico ( $u_i$ ) tiende a infinito ( $\sigma_u^2 \rightarrow \infty$  y  $\gamma \geq 1$ ), la ineficiencia técnica es la principal fuente de variación en el modelo.

En otras palabras, la cantidad de café producido por gran parte de las fincas está ubicada debajo de la frontera en la figura 1 (punto B). Esto describe un comportamiento inefectivo de los caficultores en el manejo de sus factores productivos, porque con esta cantidad de insumos pueden alcanzar mayor producción o, reduciendo su intensidad, mantener la cuantía observada de café cultivado. No obstante, la evidencia de eficiencia técnica se establece a través del estadístico de razón de verosimilitud (RV)<sup>13</sup> probando la hipótesis nula sobre el parámetro de eficiencia ( $H_0 : \gamma = 0$ ). Por consiguiente, si es rechazada  $H_0$  la respectiva unidad productora de café opera con ineficiencia técnica.

$$ET_i = \frac{q_i}{q_i^*} = \frac{e^{f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta}) + v_i - u_i}}{e^{f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta}) + v_i}} = e^{-u_i} \quad (3)$$

La frontera estocástica permite encontrar el nivel de eficiencia técnica ( $ET_i$ ), como se representa en la ecuación (3) (Battese & Coelli 1988) para cada finca cafetera  $i$  ( $i = 1, 2, \dots, 999$ ), la cual se concibe como la relación entre el producto conseguido ( $q_i$ ) y el máximo a alcanzarse ( $q_i^*$ ) con eficiencia técnica ( $u_i = 0$ ). En otras palabras, representa la proporción entre la producción real respecto a la potencial si los caficultores utilizaran eficientemente la cantidad de insumos requeridos en la actividad; por tanto,  $ET_i$  contiene valores entre 0 y 1 ( $0 \leq ET_i \leq 1$ ). Así, cuando  $ET_i$  tiende o equivale a uno ( $ET_i \rightarrow 1$ ), el caficultor exhibe eficiencia técnica, mientras que si su valor es menor que 1 o cercano a 0 ( $ET_i \rightarrow 0$ ), es considerado ineficiente. De esta manera, puede establecerse el productor cafetero más eficiente entre los pequeños, medianos y grandes productores del sector cafetero colombiano.

<sup>10</sup>  $\sigma_s^2 = \sigma_u^2 + \sigma_v^2$ .

<sup>11</sup>  $z_i = \frac{(q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta}))}{\sigma_s} \sqrt{\frac{\gamma}{1-\gamma}}$ .

<sup>12</sup>  $\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ .

<sup>13</sup>  $RV = -2(\ln f_R - \ln f_{NR})$ ;  $RV \sim \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$ , donde  $\ln f$  es el logaritmo de la función de verosimilitud. El subíndice  $R$  hace referencia al modelo restringido (con  $\gamma = 0$ ),  $NR$  al no restringido (con  $\gamma \neq 0$ ) y  $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$  a la distribución asintótica que resulta de combinar las distribuciones Ji-Cuadrado con cero y un grado de libertad, respectivamente (Coelli 1995, p. 292).



Una de las grandes dificultades para estimar la frontera estocástica es seleccionar una función de producción apropiada como cualquiera de las expresadas en la tabla 1. Porque la forma funcional afecta considerablemente los resultados para las elasticidades, economías a escala, parámetro de eficiencia ( $\gamma$ )<sup>14</sup> y  $ET_i$  (Konstantinos et al. 2003). Ante esto, el menor valor del criterio de Akaike<sup>15</sup> (AIC), elegido entre las distintas funciones de producción cafeteras estimadas, determinará la forma funcional adecuada con la cual se establecieron los resultados y las conclusiones del estudio.

## 4. Datos y análisis empírico

Los datos para este estudio fueron concebidos a partir de información primaria de la encuesta cafetera aplicada en 999<sup>16</sup> fincas de los departamentos de Caldas, Quindío y Risaralda por la Facultad de Economía-Centro de Investigaciones Sobre el Desarrollo Económico (CEDE) de la Universidad de los Andes<sup>17</sup> (entre marzo y abril de 2004<sup>18</sup>). El formulario contiene las principales características<sup>19</sup> sobre los diferentes caficultores<sup>20</sup> entrevistados, entre las cuales están las variables relacionadas con producción de café tomadas para este trabajo. Por consiguiente, la variable dependiente es la producción anual de café ( $q_i$ ) medida en arrobas (por 25 libras), la cantidad de insumos empleados en la actividad e incluidos como variables explicativas: tierra (hectáreas dedicadas al cultivo del grano), trabajo total (incluye trabajadores permanentes, temporales, casuales, núcleo familiar y contratos por labores<sup>21</sup>, medidos en cantidad de gente empleada), fertilizantes (incluye

<sup>14</sup>Dado que el parámetro de eficiencia es muy sensible a la forma funcional especificada, lleva a resultados espurios y a cometer errores tipo I y II en las pruebas de hipótesis, sobrevalorando o subestimando la medida de eficiencia técnica ( $ET_i$ ).

<sup>15</sup> $AIC = -2\frac{\ln f}{n} + 2\frac{k}{n}$ , donde  $\ln f$  es el logaritmo de la función de verosimilitud,  $n$  el número de observaciones y  $k$  el número de parámetros estimados.

<sup>16</sup>Información de corte transversal. Finalmente, se trabajó con 990 observaciones y se eliminaron nueve inconsistentes.

<sup>17</sup>Financiado con recursos de la Universidad de Maryland, bajo la dirección del profesor Darrell Hueth con el apoyo de la Federación Nacional de Cafeteros de Colombia y la valiosa colaboración de Diego Pizano, Julián García, Alfonso Ángel Uribe, Óscar Jaramillo García y Omar Acevedo Chamorro.

<sup>18</sup>La encuesta se aplicó en 2004, recolectando la información de los cafeteros para 2003. Hasta el momento, no se cuenta con más información microeconómica de este tipo en las mismas fincas entrevistadas y en otros periodos. Ninguna entidad cafetera la recolecta para otra muestra en Colombia.

<sup>19</sup>Socioeconómicas, producción del grano y otras actividades en la finca, financieras, relacionadas con la asistencia técnica, geográficas, ambientales, propias y entorno de la finca o vivienda cafetera.

<sup>20</sup>Pequeños (entre 0 y 2,1 hectáreas productivas en café; 662 observaciones en la encuesta equivale a 66,87% de la muestra total), medianos (entre más de 2,1 y 6,9 hectáreas productivas en café; 250 observaciones en la encuesta equivale a 25,25% de la muestra total) y grandes (más de 6,9 hectáreas productivas en café; 78 observaciones en la encuesta equivale a 7,88% de la muestra total). Lo anterior guarda las mismas proporciones poblacionales puntualizadas por la Comisión de Ajuste de la Institucionalidad (2002), descritas en los pies de páginas 1, 2 y 3 de la sección 1.

<sup>21</sup>Incluye todas las actividades de la finca relacionadas con arado, fertilización, fumigación con químicos, cosechas, riego, poda de maleza, transporte, administración y otros servicios.

nitrógenos, fósforo, potasio y otros; medido en *Kg*) y maquinaria (cantidad de herramientas).

De acuerdo con lo anterior y con Greene (2002), es incorrecto en funciones no lineales o flexibles, como la Translogarítmica Minflex Laurent (tabla 3), analizar la relevancia de los estimadores, interpretar sus resultados directamente e inferir algún tipo de afirmación sobre si el comportamiento y la intensidad de los insumos cafeteros son complementarios, sustitutos o exhiben productividad marginal decreciente según los signos (negativos o positivos) parciales de los parámetros en los términos cuadráticos y las interacciones del modelo estimado.

La tabla 2 presenta las elasticidades y economías a escala derivadas de la función ajustada Translogarítmica Minflex Laurent (tabla 3) para los pequeños caficultores, que deben calcularse con el fin de entender apropiadamente la incidencia de los principales insumos empleados (tierra, trabajo, fertilizantes y maquinaria) en la producción del grano y su rendimiento.

TABLA 2: Elasticidades y economías a escala de los pequeños productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada
<b>Variables Indep.</b>	<b>CGBC</b>	<b>TLBC</b>	<b>LGBC</b>	<b>CGBC</b>	<b>CRC</b>
Tierra	0,019	0,86	0,83	0,8	151
Trabajo	-0,0064	0,47	0,44	0,32	57
Fertilizantes	-0,0073	0,026	0,12	0,14	22,3
Maquinaria	0,0031	0,026	0,04	0,07	10
<b>Economías a escala</b>	<b>0,0084</b>	<b>1,382</b>	<b>1,43</b>	<b>1,33</b>	<b>240,3</b>

Indica relevancia estadística en los niveles del (\*)10%, (\*\*)5% y (\*\*\*)1% de importancia

	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
<b>Variables Indep.</b>	<b>ESC</b>	<b>CD</b>	<b>TLML</b>	<b>LG</b>	<b>C</b>
Tierra	0,02	0,89	0,81***	2,73	24,88
Trabajo	0,01	0,47	(-0,30)***	-0,04	0,17
Fertilizantes	0,0024	0,03	0,69***	5,61	24,21
Maquinaria	0,0014	0,03	0,13***	0,03	0,003
<b>Economías a escala</b>	<b>0,0338</b>	<b>1,42</b>	<b>1,33</b>	<b>8,33</b>	<b>49,263</b>

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

En este sentido, la elasticidad del factor tierra lo revela como el insumo más importante (1% de importancia) en las pequeñas unidades cafeteras, que un incremento de 0,81% en la producción del grano cuando las hectáreas cultivadas de café aumentan 1%; seguido por la intensidad de fertilizantes y maquinaria (1% de importancia), porque ampliando su uso parcialmente en 1% la producción crece 0,69% y 0,13%, respectivamente.

La cantidad de trabajo es la menos importante (1%). El rendimiento del cultivo disminuye 0,30% al incrementar el factor en 1%. Este grupo de caficultores exhibe rendimientos crecientes a escala en su actividad productiva dado que aumentar simultáneamente en 1% la tierra, el trabajo, los fertilizantes y la maquinaria incrementa la producción en 1,33%. Con base en el método de frontera estocástica,

estimado mediante máxima verosimilitud, la tabla 3 presenta las estimaciones<sup>22</sup> de las funciones de la tabla 1 para los pequeños caficultores en la zona. Puede observarse que la forma funcional de producción cafetera afecta el valor y la relevancia estadística del parámetro de eficiencia ( $\gamma$ ), de promedio de eficiencia técnica (ET) y de las economías a escala (tabla 2), debido a sus diferentes resultados bajo cada frontera.

Por consiguiente, la función Translogarítmica Minflex Laurent se ajusta a los pequeños caficultores porque el valor del criterio Akaike (1.178) fue el menor, comparado con el de las otras funciones convencionales y flexibles. Asimismo, el parámetro de eficiencia ( $\gamma = 0.993$ ) y las varianzas ( $\sigma_v^2$  y  $\sigma_u^2$ ) son determinantes (al 5 % y 1 % de importancia), indicando ineficiencia técnica ( $u_i$ ) y aleatoria ( $v_i$ ) con efectos negativos sobre la producción del grano de los pequeños cafeteros. En otras palabras si  $u_i$  y  $v_i$  aumentan individualmente 1 %, en promedio la producción de ellos disminuye anualmente 1,74 % y 1,73 % respectivamente. Igualmente lo muestra el promedio de eficiencia técnica (71 %) logrado por este grupo de cultivadores de café, inferior al 100 %.

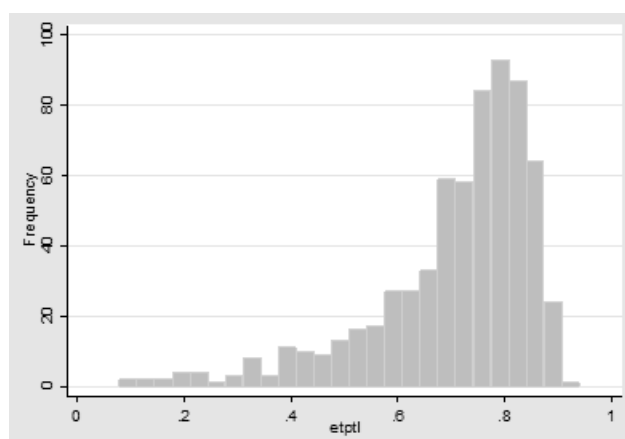


FIGURA 2: Eficiencia técnica de los pequeños productores de café.

El eje X representa el porcentaje de ET y el Y la frecuencia (el número de caficultores) a la cual le corresponde el valor de ET en el eje X. Estas mismas relaciones se aprecian más adelante en las gráficas de medianos y grandes productores en el sector general de productores.

A partir de la función de producción Translogarítmica Minflex Laurent, fue obtenido el nivel de eficiencia técnica ( $ET_i$ ), de acuerdo con la ecuación (3), por cada pequeño caficultor y su distribución, como puede apreciarse en la figura 2. Se observa sesgo negativo, lo cual implica que la mayor parte de los pequeños caficultores están ubicados por encima del promedio de ET (71 %) y debajo de la eficiencia máxima (100 %). Únicamente entre 80 y 100 productores de todo el grupo consiguen un máximo de eficiencia técnica del 80 % con un 20 % restante a mejorar, pero ninguno alcanza el 100 % de ET.

<sup>22</sup>En Stata 9 (frontera estocástica) y Eviews 4.1 (Box-Cox), utilizando errores estándar robustos y mínimos cuadrados generalizados para eliminar la heteroscedasticidad.

De acuerdo con los criterios examinados en los pequeños productores de café, se continuará en este sentido el análisis de los resultados para los medianos y grandes cultivadores cafeteros del sector en este estudio. Así, la tabla 4 presenta las elasticidades y las economías a escala derivadas de la función ajustada Translogarítmica Minflex Laurent (tabla 5) para los medianos caficultores, calculadas con el fin de entender apropiadamente la incidencia de los principales insumos empleados (tierra, trabajo, fertilizantes y maquinaria) en la producción del grano y su rendimiento.

La elasticidad del factor tierra lo revela como el insumo más importante (1 % de importancia) en las medianas unidades cafeteras, que reflejado un incremento de 0,71 % en la producción del grano cuando las hectáreas cultivadas de café aumentan 1 %, seguido por la intensidad de trabajo, fertilizantes y maquinaria (1 % de importancia), porque ampliando su uso parcialmente en 1 % la producción crece 0,20 %, 0,18 % y 0,12 % respectivamente.

Este grupo de caficultores exhibe rendimientos crecientes a escala en su actividad productiva dado que aumentar simultáneamente en 1 % la tierra, el trabajo, los fertilizantes y la maquinaria incrementa la producción en 1,21 %. Con base en el método de frontera estocástica estimado mediante máxima verosimilitud, la tabla 5 presenta las estimaciones de las funciones en la tabla 1 para los medianos caficultores en la zona. Puede observarse que la forma funcional de producción cafetera afecta el valor y la relevancia estadística del parámetro de eficiencia ( $\gamma$ ), del promedio de eficiencia técnica ( $ET$ ) y de las economías a escala (tabla 4), debido a sus diferentes resultados bajo cada frontera.

Por consiguiente, la función Translogarítmica Minflex Laurent se ajusta a los medianos caficultores porque el valor del criterio Akaike (358) fue el menor, comparado con el de las otras funciones convencionales y flexibles. Asimismo, el parámetro de eficiencia ( $\gamma = 1,93$ ) y las varianzas ( $\sigma_v^2$  y  $\sigma_u^2$ ) son determinantes (5 % y 1 % de importancia), lo que indica ineficiencia técnica ( $u_i$ ) y aleatoria ( $v_i$ ) con efectos negativos sobre la producción del grano de los medianos cafeteros. En otras palabras, si  $u_i$  y  $v_i$  aumentan individualmente 1 %, en promedio su producción disminuye anualmente 1,05 % y 2,36 %, respectivamente. Igualmente, lo muestra el promedio de eficiencia técnica (66 %) logrado por este grupo de cultivadores de café, inferior al 100 %.

A partir de la función de producción Translogarítmica Minflex Laurent fue obtenido el nivel de eficiencia técnica ( $ET_i$ ), de acuerdo con la ecuación (3), por cada mediano caficultor. Su distribución puede apreciarse en la figura 3, donde se observa sesgo negativo, implicando que la mayor parte de los medianos caficultores están ubicados por encima del promedio de  $ET$  (66 %) y debajo de la eficiencia máxima (100 %). Únicamente, entre 40 y 45 productores del grupo consiguen un máximo de eficiencia técnica del 70 %; el 30 % restante debe mejorar. Pero ninguno alcanza el 100 % de  $ET$ .

Prosiguiendo el análisis, la tabla 6 presenta las elasticidades y economías a escala derivadas de la función ajustada Translogarítmica Minflex Laurent (tabla 7) para los grandes caficultores, calculadas con el fin de entender apropiadamente la incidencia de los principales insumos empleados (tierra, trabajo, fertilizantes

TABLA 3: FE de pequeños productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variables independientes	CGBC	TLBC	LGBC	CGBC	CRG	ESC	CD	TLML	LG	C
Constante	5,1238723	3,96868***	29,96749	56,97914	2306,745	5,268704	4,0669***	3,41999***	49,219	-16,84038
Tierra	2,14048***	0,89183***	19,56354	64,77872***	1434,546	2,26902***	0,8879***	0,77886***	104,295***	31,75793
Trabajo	0,70433***	0,44935***	15,34484***	3,69829***	323,1948	0,645277***	0,4685***	0,01575***	-1,084029	1,889196
Fertilizantes	0,04820***	0,0113032	0,686391**	0,04809***	8,83554***	0,04745***	0,0339***	0,650582	0,0016066	0,06284***
Maquinaria	0,37186***	0,10397***	8,698136	5,32886*	411,5735	0,12872***	0,0276***	0,20984***	1,374729	8,807975
Tierra <sup>2</sup>	-	-	-	-	-	-	-	0,008945	(-84,911)***	14,54999
Trabajo <sup>2</sup>	-	-	-	-	-	-	-	-0,099681	-7,461878	(-0,0492)***
Fertilizantes <sup>2</sup>	-	-	-	-	-	-	-	0,01336***	0,3538563	(-0,00001)***
Maquinaria <sup>2</sup>	-	-	-	-	-	-	-	0,03656***	-1,648577	-0,7354321
Tierra*Trabajo	0,21584	0,0245723	26,23024***	(-7,9683)***	(-1594,944)	-	-	(-0,022)***	41,20967	2,220996
Tierra*Fertilizantes	0,00012***	(-0,0332)***	1,09310***	(-0,0525)***	2,53878	-	-	0,052587	1,02229***	0,009729*
Tierra*Maquinaria	(-0,0391)***	0,035372	10,74741	81,4351***	16346,94***	-	-	0,012529	11,82213	0,7294544
Trabajo*Fertilizantes	0,1719587	0,01297	0,0686377	0,01742***	4,08138***	-	-	0,002566	0,344339	0,001091
Trabajo*Maquinaria	-	(-0,07148)**	2,720545	(-8,00703)***	(-1215,691)	-	-	(-0,036)***	6,458284	0,63910***
Fertilizantes*Maqui.	-0,0092981	-0,002622	(-0,47405)**	(-0,10203)***	(-34,7112)***	-	-	-0,00230	(-1,03002)***	(-0,0109)***
Theta ( $\Theta$ , Box-Cox)	0,10440845	0	0,5	0,5	1	-	0	-	-	-
Lambda ( $\lambda$ , Box-Cox)	0,2797122	0	0,5	1	1	-	0	-	-	-
Logaritmo natural de V	0,65077***	(-1,6562)***	8,67406***	8,66456***	20,57648***	0,66490***	(-1,78)***	(-1,730)***	8,66267***	8,65828***
Logaritmo natural de U	(-5,0773)***	(-1,7475)***	(-5,30543)	(-5,305384)	(-5,48595)	(-5,169036)	(-0,60)***	(-1,745)***	-5,305431	-5,305433
Desviación estándar de V ( $\sigma_v$ )	1,38457	0,43688***	76,76480***	76,11774***	29385,06***	1,39438***	0,4105***	0,42099***	76,04575***	74,8820***
Desviación estándar de U ( $\sigma_u$ )	0,07897	0,41737***	0,0704597	0,0704596	0,0643784	0,0754324	0,7401***	0,41796***	0,0704596	0,0704596
Varianza del modelo $\sigma_s^2$	1,92326	0,63506***	5849,211***	5793,197***	$8,63 \times 10^8$ ***	1,9499***	0,72***	0,35**	5783***	5607***
Parámetro de eficiencia ( $\gamma$ )	0,05704	0,9553***	0,0009213	0,00093	0,0000022	0,054	1,80***	0,993***	0,000926	0,000941
Logaritmo función de verosimilitud del modelo	-1155,13	-1337,165	-3810,4517	-3807,000	-7750,000	-1160,000	-599	-572,000	-3806	-3796
Razón de verosimilitud ( $\gamma = 0$ )	0,00	35,40***	0	0	0	0	24,44**	40,20**	0	0
Eficiencia técnica (ET, promedio)	0,9399795	0,705911	0,9461736	0,9461736	0,95	0,94	0,61	0,71	0,95	0,95
Wald	1845,59***	1669,25	937,49***	953***	299***	1811***	1540***	1787***	956***	1006
Akaike (AIC)	2334	1199	7647	7641	15526	2334	1211	1178	7647	7627
Scharwz (SIC)	2388	1258	7705	7699	15584	2365	1243	1255	7724	7703
Número de observaciones (n)	662	662	662	662	662	662	662	662	662	662

Indica relevancia estadística en los niveles del (\*) 10 %, (\*\*) 5 % y (\*\*\*) 1 % de importancia

TABLA 4: Elasticidades y economías a escala de los medianos productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada
Variables Indep.	CGBC	TLBC	LGBC	CGBC	CRC
Tierra	0,0023	0,74	0,74	0,71	0,99
Trabajo	0,0016	0,62	0,50	0,27	-0,01
Fertilizantes	-0,0003	-0,011	0,05	0,15	0,19
Maquinaria	0,0001	-0,003	0,03	0,13	0,25
Economías a escala	0,0037	1,346	1,32	1,26	1,42

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variables Indep.	ESC	CD	TLML	LG	C
Tierra	0,003	0,77	0,71***	12,46	39,40
Trabajo	0,002	0,59	0,20***	-1,16	0,12
Fertilizantes	0	0,004	0,18***	21,44	40
Maquinaria	-0,0001	-0,03	0,12***	0,11	0,04
Economías a escala	0,0049	1,334	1,21	32,85	79,56

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

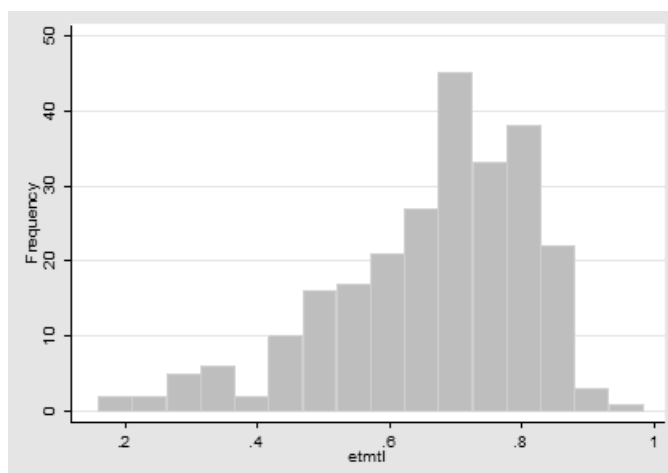


FIGURA 3: Eficiencia técnica de los medianos productores de café.

y maquinaria) en la producción del grano y su rendimiento. En este sentido, la elasticidad del factor trabajo lo revela como el insumo más importante (1% de importancia) en las grandes unidades cafeteras. Esto refleja un incremento de 0,78% en la producción del grano cuando la mano de obra aumenta 1%, seguido por la intensidad de fertilizantes (1% de importancia), aunque el rendimiento del cultivo disminuye 0,58% por incrementar el factor en 1%.

Sin embargo, la cantidad de maquinaria y tierra resultaron irrelevantes (no significativas en el 1%, 5% y 10% de importancia) en el cultivo de los grandes caficultores. Este grupo exhibe rendimientos decrecientes a escala en su actividad

TABLA 5: FE de medianos productores de café.

Variables independientes	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
	CGBC	TLBC	LGBC	CGBC	CRC	ESC	CD	TLML	LG	C
Constante	4.84***	3.87***	85,10	109,28	82220,93	3.93***	3.78***	3.59***	-548,91	-307,27
Tierra	0,32	0,63	-50,22	32,12	10946,72	1,29***	0,77***	1,06	-304,73	207,26***
Trabajo	0,55	0,41**	-4,81	-0,48	(-15280,77)***	0,99***	0,59***	0,036	(-13,19)***	5,93
Fertilizantes	0,09**	0,03	3,46***	0,11***	103,90***	0,006	0,004	0,33	0,04*	0,10***
Maquinaria	0,09	0,05	-59,44	-10,63	-16378,52	-0,04	-0,03	0,21	73,53***	-37,02
Tierra <sup>2</sup>	-	-	-	-	-	-	-	-0,61	409,53	(-25,37)*
Trabajo <sup>2</sup>	-	-	-	-	-	-	-	-0,03	(-71,81)*	(-0,10)***
Fertilizantes <sup>2</sup>	-	-	-	-	-	-	-	0,008	7,26***	-0,00000046
Maquinaria <sup>2</sup>	-	-	-	-	-	-	-	0,05***	(-193,42)**	2,36
Tierra*Trabajo	0,87	0,24	49,40**	4,18**	7314,82***	-	-	0,25	(-5,92)**	-0,005
Tierra*Fertilizantes	(-0,14)*	(-0,1)***	-1,68	(-0,03)*	-18,21	-	-	(-0,06)**	177,55***	2,68***
Tierra*Maquinaria	0,04	0,04	138,83**	37,52**	4979,18	-	-	-0,02	55,40	8,44
Trabajo*Fertilizantes	-0,06	0,03	0,13	0,0004	-1,43	-	-	0,008	-0,29	-0,0007
Trabajo*Maquinaria	-	-0,04	8,25	0,46	5105,52***	-	-	-0,004	31,49	0,7
Fertilizantes*Maquinaria	-0,022	-0,012	(-1,98)**	(-0,024)***	(-27,29)***	-	-	(-2,36)***	11,08**	11,02***
Theta ( $\Theta$ , Box-Cox)	0,04	0	0,5	0,5	1	0,04	0	-	-	-
Lambda ( $\lambda$ , Box-Cox)	-0,03	0	0,5	1	1	-0,03	0	-	-	-
Logaritmo natural de V	(-1,34)**	(-2,22)***	11,18***	11,15***	25***	(-1,28)***	(-2,33)***	(-2,36)***	11,08***	11***
Logaritmo natural de U	-0,14	(-2,05)***	-5,21	-5,22	-5,84	-0,14	(-0,91)***	(-1,05)***	-5,22	-5,22
Desviación estándar de V ( $\sigma_v$ )	0,51***	0,33***	268***	265***	281926*	0,53***	0,31***	0,31***	254***	245***
Desviación estándar de U ( $\sigma_u$ )	0,93***	0,36***	0,07	0,07	0,05	0,93***	0,63***	0,60***	0,07	0,07
Varianza del modelo $\sigma_s^2$	1,13***	0,24***	71839***	70210***	$795 \times 10^7$ ***	1,15***	0,50***	0,44***	64756***	60024***
Parámetro de eficiencia ( $\gamma$ )	1,82***	1,09***	0,0003	0,003	0,0000002	1,77***	2,03***	1,93***	0,0003	0,0003
Logaritmo función de verosimilitud del modelo	-283,00	-165	-1753	-1750	-3492	-287	-173	-162	-1740	1730
Razón de verosimilitud ( $\gamma = 0$ )	10,85***	23,17***	0	0	0	10,70***	15,47***	11,50**	0	0
Eficiencia técnica (ET, promedio)	0,55	0,74	0,94	0,94	0,96	0,55	0,65	0,66	0,94	0,95
Wald	325***	371***	302	315***	383	321***	324***	365***	363***	411
Akaike (AIC)	589	356	3531	3525	7010	587	360	358	3513	3494
Schwarz (SIC)	631	402	3576	3571	7056	612	385	418	3573	3554
Número de observaciones (n)	250	250	250	250	250	250	250	250	250	250

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

TABLA 6: Elasticidades y economías a escala de los grandes productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada
Variables Indep.	CGBC	TLBC	LGBC	CGBC	CRC
Tierra	-	0,01	0,41	0,11	0,04
Trabajo	-	0,61	0,58	0,68	0,84
Fertilizantes	-	-1,12	0,06	0,14	0,15
Maquinaria	-	0,14	-0,02	0,04	0,14
Economías a escala	-	-0,36	1,03	0,97	1,17

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variables Indep.	ESC	CD	TLML	LG	C
Tierra	0,0008	0,84	0,01	-0,23	-41,56
Trabajo	0,002	0,07	0,78***	0,68	1,13
Fertilizantes	0,0007	0,04	(-0,58)***	-10,73	-41,95
Maquinaria	-0,0001	-0,009	0,05	-0,03	-0,07
Economías a escala	0,0034	0,941	0,26	-10,31	-82,45

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

productiva dado que aumentar simultáneamente en 1% la tierra, el trabajo, los fertilizantes y la maquinaria incrementa la producción en 0,26%.

Con base en el método de frontera estocástica, estimado mediante máxima verosimilitud, la tabla 7 presenta las estimaciones de las funciones en la tabla 1 para los grandes caficultores en la zona. Puede observarse que la forma funcional de producción cafetera afecta el valor del parámetro de eficiencia ( $\gamma$ ), del promedio de eficiencia técnica ( $ET$ ) y de las economías a escala (tabla 6), debido a sus diferentes resultados bajo cada frontera. Por consiguiente, la función Translogarítmica Minflex Laurent se ajusta a los grandes caficultores porque el valor del criterio Akaike (84) fue el menor, comparado con el de las otras funciones convencionales y flexibles. Asimismo, el parámetro de eficiencia ( $\gamma = 0.03$ ) y las varianzas ( $\sigma_v^2$  y  $\sigma_u^2$ ) no son determinantes (al 5% y 1% de importancia).

Lo anterior indica su eficiencia técnica ( $u_i$ ); igualmente, lo evidencia el promedio de eficiencia técnica ubicado en 100%. Esto significa que este grupo opera con eficiencia técnica, aunque no aleatoria ( $v_i$ ). Esta última tiene efectos negativos sobre la producción del grano de los grandes cafeteros; en otras palabras, si aumentan 1%, en promedio su producción disminuye 2,20% anualmente.

A partir de la función de producción Translogarítmica Minflex Laurent fue obtenido el nivel de eficiencia técnica ( $ET_i$ ), de acuerdo con la ecuación (3), para cada grande caficultor. Su distribución puede apreciarse en la figura 4, donde se observa que todos los grandes productores logran el máximo de eficiencia técnica del 100%. Nuevamente, determina la ausencia de ineficiencia técnica en este grupo.

Una vez estimada las funciones de producción por tamaño de caficultor, son agrupadas las tres muestras para obtener los resultados del sector general. Así, la tabla 8 presenta sus elasticidades y economías a escala derivadas de la función ajustada Translogarítmica Minflex Laurent (tabla 9), calculadas con el fin de en-



TABLA 7: FE de grandes productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variables independientes	CGBC	TLBC	LGBC	CGBC	CRC	ESC	CD	TLML	LG	C
Constante	16.37***	7.01***	1477.98	622.99	900683.5	16.51**	5.44***	6.80***	2073.55	718.13
Tierra	(-1.01)**	(-2.36)**	(-686.80)***	(-102.53)***	(-985550)***	0.41***	0.84***	(-1.21)*	65.30	(-134.12)**
Trabajo	0.57***	0.68*	50.36	22.37***	85226.89**	0.31***	0.07*	-0.06	32.52***	27.47***
Fertilizantes	0.009**	-0.17	3.36	0.08***	477.83*	0.0067***	0.04**	0.002	0.15**	0.05
Maquinaria	0.46	1.15***	73	-9.91	-78787.93	-0.10	-0.009	0.66*	-71.48	-19.80
Tierra <sup>2</sup>	-	-	-	-	-	-	-	-0.30	(-1026.97)**	2.11
Trabajo <sup>2</sup>	-	-	-	-	-	-	-	0.06***	147.15	0.07***
Fertilizantes <sup>2</sup>	-	-	-	-	-	-	-	0.009	3.33	0,0000029
Maquinaria <sup>2</sup>	-	-	-	-	-	-	-	-0.0008	1.45	-0,0009
Tierra*Trabajo	-0,00022	0.62***	57.57***	0.007	7801.62***	-	-	0.37***	1.45	-0,0009
Tierra*Fertilizantes	0.0011	0.02	-0.53	0.004**	25.56*	-	-	0.0006	-6.52	(-0.78)**
Tierra*Maquinaria	0.44***	1.09***	256.9***	26.75***	116396.3**	-	-	0.28**	640.04***	19.70***
Trabajo*Fertilizantes	(-0.08)***	0.07***	0.60	0.0003	(-5.96)*	-	-	0.01	-1.30	0.0005***
Trabajo*Maquinaria	0	(-1.24)***	(-58.79)***	(-1.26)***	(-7092.53)***	-	-	(-0.36)****	(-160.46)***	(-1.78)***
Fertilizantes*Maquinaria	(-0.004)*	0.043	-3.04	(-0.021)***	-44.84	-	-	0.02	-6.70	(-0.001)***
Theta ( $\Theta$ , Box-Cox)	0.04	0	0.5	0.5	1	0.04	0	-	-	-
Lambda ( $\lambda$ , Box-Cox)	-0.03	0	0.5	1	1	-0.03	0	-	-	-
Logaritmo natural de V	2.40***	(-2.07)***	14.06***	13.84***	32***	2.70***	(-1.92)***	(-2.20)***	13.86***	13.50***
Logaritmo natural de U	-5,49	-9,18	-5,17	-5,21	(-5.12)	-5,48	-0,61	-9,29	-5,21	-5,22
Desviación estándar de V ( $\sigma_v$ )	3.31***	0.35***	1134.32***	1010.24***	8876673***	3.87***	0.38***	0.33***	1022.50***	853.19**
Desviación estándar de U ( $\sigma_u$ )	0.06	0.01	0.08	0.07	0.08	0.06	0.73	0.01	0.07	0.07
Varianza del modelo $\sigma_s^2$	11***	0.13***	1286673***	1020588***	8.8E+12	14.95***	0.68***	0.11***	1045503***	727926***
Parámetro de eficiencia ( $\gamma$ )	0,02	0,03	0,000067	0,000073	8,69E-08	0,02	1,92**	0,03	0,00007	0,00009
Logaritmo función de verosimilitud del modelo	-204,00	-30	-659	-650	-1359	-216	-67	-25	-651	-637
Razón de verosimilitud ( $\gamma = 0$ )	0	0	0	0	0	0	0	0	0	0
Eficiencia técnica (ET, promedio)	0,95	0,96	0,94	0,94	0,94	0,95	0,61	0,99	0,94	0,94
Wald	461***	404**	715***	921	1109***	319***	105***	470***	897***	1323***
Akaike (AIC)	432	86	1345	1327	2741	446	149	84	1336	1308
Schwarz (SIC)	461	116	1375	1357	2769	463	165	124	1376	1348
Número de observaciones (n)	78	78	78	78	78	78	78	78	78	78

Indica relevancia estadística en los niveles del (\*) 10%, (\*\*) 5% y (\*\*\*) 1% de importancia

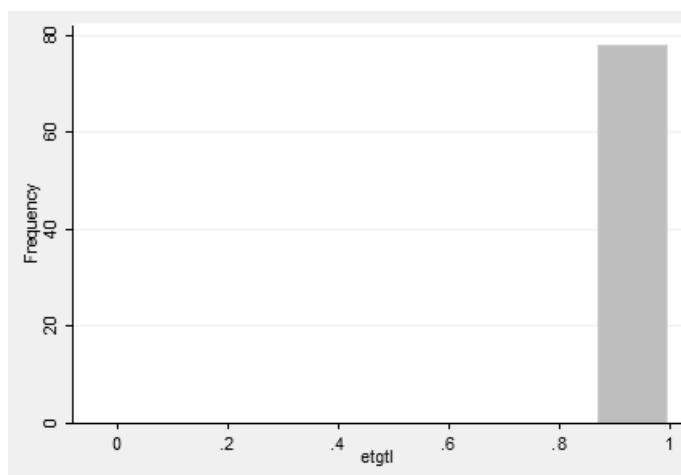


FIGURA 4: Eficiencia técnica de los grandes productores de café.

TABLA 8: Elasticidades y economías a escala del sector general de productores de café.

	Cuadrática generalizada Box-Cox	Translogarítmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada
Variables Indep.	CGBC	TLBC	LGBC	CGBC	CRC
Tierra	0,004	0,88	0,46	0,11	-4,53
Trabajo	-0,002	0,41	0,57	0,72	4,14
Fertilizantes	-0,005	-0,004	0,12	0,24	1,98
Maquinaria	0,002	0,02	0,04	0,05	-0,24
Economías a escala	-0,001	1,306	1,19	1,12	1,35

Indica relevancia estadística en los niveles del (\*) 10 %, (\*\*) 5 % y (\*\*\*) 1 % de anualmente

	Elasticidad de sustitución constante	Cobb-Douglas	Translogarítmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variables Indep.	ESC	CD	TLML	LG	C
Tierra	0,004	0,91	0,69***	5,97	1,21
Trabajo	0,002	0,30	0,41***	-0,43	0,57
Fertilizantes	0,0004	0,03	0,26***	-3,68	0,84
Maquinaria	0,0002	0,02	0,11	0,09	0,1
Economías a escala	0,0066	1,26	1,47	1,95	2,72

Indica relevancia estadística en los niveles del (\*) 10 %, (\*\*) 5 % y (\*\*\*) 1 % de anualmente

tender apropiadamente la incidencia de los principales insumos empleados (tierra, trabajo, fertilizantes y maquinaria) en la producción del grano y su rendimiento.

En este sentido, la elasticidad del factor tierra lo revela como el insumo más importante (1 % de importancia) en el sector general cafetero, lo cual refleja un incremento de 0,69 % en la producción del grano cuando las hectáreas cultivadas de café aumentan 1 %, seguido por la intensidad de trabajo y fertilizantes (relevantes al 1 % de importancia), porque ampliando su uso parcialmente en 1 % la producción crece 0,41 % y 0,26 % respectivamente.

La cantidad de maquinaria resultó irrelevante (no significativa en el 1 %, 5 % y 10 % de importancia) en el cultivo del sector general. Este grupo exhibe rendimientos crecientes a escala en su actividad productiva, dado que aumentar simultáneamente en 1 % la tierra, el trabajo, los fertilizantes y la maquinaria incrementa la producción en 1,47 %.

Con base en el método de frontera estocástica, estimado mediante máxima verosimilitud, la tabla 9 presenta las estimaciones de las funciones en la tabla 1 para el sector general de caficultores en la zona, donde puede observarse que la forma funcional de producción cafetera afecta el valor y la relevancia estadística del parámetro de eficiencia ( $\gamma$ ), del promedio de eficiencia técnica ( $ET$ ) y de las economías a escala (tabla 8), debido a sus diferentes resultados bajo cada frontera.

Por consiguiente, la función Translogarítmica Minflex Laurent se ajusta al sector general de caficultores porque el valor del criterio Akaike (1.625) fue el menor, comparado con el de las otras funciones convencionales y flexibles. Asimismo, el parámetro de eficiencia ( $\gamma = 0.99$ ) y las varianzas ( $\sigma_v^2$  y  $\sigma_u^2$ ) son determinantes (al 5 % y 1 % de importancia).

Lo anterior indica ineficiencia técnica ( $u_i$ ) y aleatoria ( $v_i$ ) con efectos negativos sobre la producción del grano. En otras palabras, si  $u_i$  y  $v_i$  aumentan individualmente 1 %, en promedio la producción disminuye anualmente 1,87 % y 1,85 % respectivamente. Igualmente, lo muestra el promedio de eficiencia técnica (72 %) logrado por este grupo de cultivadores de café, inferior al 100 %.

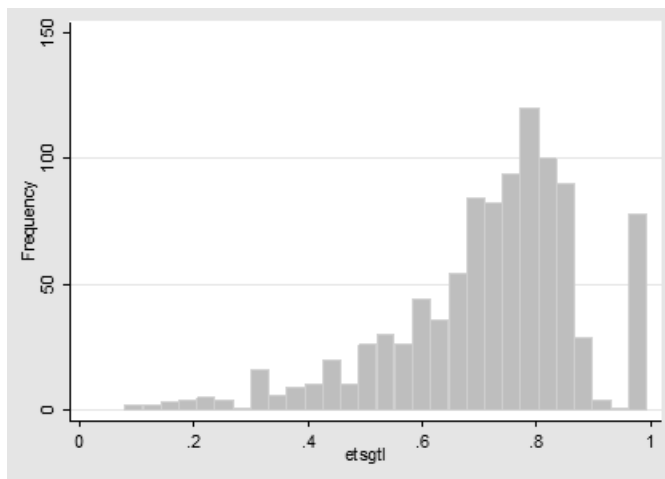


FIGURA 5: Eficiencia técnica del sector general de productores de café.

El nivel de eficiencia técnica ( $ET_i$ ) del sector general puede apreciarse en la figura 5, obtenido agrupando este valor de los resultados expuestos para cada unidad productiva analizada, con el objetivo de no sobrestimar o subestimar esta medida, dada la heterogeneidad de producción en cada rango de tamaño cafetero. También porque no son directamente comparables la eficiencia técnica del pequeño

TABLA 9: FE del sector general de productores de café.

	Cuadrática generalizada Box-Cox	Translogaritmica (Translog desde Box-Cox)	Leontief generalizada desde Box-Cox	Cuadrática generalizada desde Box-Cox	Cuadrática, raíz cuadrada	Elasticidad de sustitución constante	Cobb-Douglas	Translogaritmica (Minflex Laurent Translog)	Leontief generalizada (Diewert)	Cuadrática
Variab. indep.	CGBC	TLBC	LGBC	CGBC	CRC	ESC	CD	TLML	LG	C
Constante	5.14***	4.14***	-107,87	-43,99	(-457418.2)***	4.69***	4.36***	3.80***	36,07	-95,70
Tierra	1.61***	0.99***	(-106.32)***	-3,35	(-576025.4)***	1.65***	0.91***	0.79***	(-102)***	42.99***
Trabajo	0.81***	0.43***	39.19***	16.54***	53213.72***	0.68***	0.30***	0.02***	15.41***	13.83***
Fertilizantes	0.05***	0,02	1.93***	0.008***	339.91***	0.05***	0.03***	0.27***	0.12***	0.03***
Maquinaria	0.36***	0.17***	72.14***	10,58	-25891.59	0.07**	0.02***	0.22***	39.14**	14,99
Tierra <sup>2</sup>	-	-	-	-	-	-	-	(-0.10)***	(-149.44)***	(-2.75)***
Trabajo <sup>2</sup>	-	-	-	-	-	-	-	0.10***	7,33	0.02***
Fertilizantes <sup>2</sup>	-	-	-	-	-	-	-	0.01***	1.37***	0.0000042***
Maquinaria <sup>2</sup>	-	-	-	-	-	-	-	0.03***	-0,59	1,21
Tierra*Trabajo	0,04	(-0.08)**	48.16***	0,05	8163.86***	-	-	(-0.02)***	97.93***	0,08
Tierra*Fertilizantes	(-0.02)*	(-0.02)**	(-0.59)*	0.00085*	14.74***	-	-	-0.00009	0,82	0.002**
Tierra*Maquinaria	0.18***	0.08***	106.41***	18.80***	80672.76***	-	-	0,01	255.14***	10.51***
Trabajo*Fertilizantes	(-0.22)***	0,01	0,15	0.000183*	(-5.57)***	-	-	-0,001	(-1.61)***	0,00006
Trabajo*Maquinaria	-	(-0.14)***	(-27.56)***	(-0.92)***	(-5574.9)***	-	-	(-0.05)***	(-2.52)***	(-0.01)***
Fertilizantes*Maquinaria	-0,003	-0,002	(-0.94)**	(-0.01)***	-14,52	-	-	-0,002	(-2.52)***	(-0.01)***
Theta ( $\Theta$ , Box-Cox)	0.08***	0	0,5	0,5	1	0.08***	0	-	-	-
Lambda ( $\lambda$ , Box-Cox)	0.15***	0	0,5	1	1	0.15***	0	-	-	-
Logaritmo natural de V	0,12	(-1.80)***	11.86***	11.71**	29.54***	0.43***	(-1.61)***	(-1.85)***	11.72***	11.47***
Logaritmo natural de U	0,03	(-0.72)***	-5,30	-5,30	-5,12	-7,63	(-0.79)***	(-1.87)***	-5,21	-5,21
Desviación estándar de V ( $\sigma_v$ )	1,06	0.41***	377***	350***	2598910***	1.24***	0.45***	0.40***	350***	310***
Desviación estándar de U ( $\sigma_u$ )	0,03	(-0.72)***	-5,31	-5,31	-5,13	-7,63	(-0.80)***	(-1.88)***	-5,21	-522
Varianza del modelo $\sigma_s^2$	2,15	0.65***	141864.5***	122352***	47500000000	1.54***	0.65***	0.31***	122422.7***	96126.67***
Parámetro de eficiencia ( $\gamma$ )	0,96	1.72***	0,0002	0,0002	0,00000003	0,02	1.50***	0.99***	0,002	0,0002
Logaritmo función de verosimilitud del modelo	-1603	-860	-7277	-7204	-16028	-1617	-897	-795	-7203	-7084
Razón de verosimilitud ( $\gamma = 0$ )	3.59***	33.31***	0	0	0	0	24***	65***	0	0
Eficiencia técnica (ET, promedio)	0,52	0,62	0,95	0,95	0,94	0,98	0,63	0,72	0,94	0,94
Wald	6545***	5416***	10906***	12803***	14221***	6379***	4592***	0.06***	12796***	16566***
Akaike (AIC)	3232	1727	14579	14433	32079	3249	1807	1625	14442	14202
Schwarz (SIC)	3291	1791	14643	14497	32138	3283	1841	1708	14525	14285
Número de observaciones ( $n$ )	990	990	990	990	990	990	990	990	990	990

Indica relevancia estadística en los niveles del (\*) 10 %, (\*\*) 5 % y (\*\*\*) 1 % de importancia

con la de un grande o mediano caficultor, debido a las amplias diferencias en sus condiciones tecnológicas en el cultivo del grano.

En la figura 5 se observa sesgo negativo, implicando que la mayor parte de los caficultores está ubicadas por encima del promedio (72 %) y por debajo del 80 % de *ET*. Esto indica ineficiencia técnica para el sector cafetero en la zona de estudio colombiana; únicamente entre 50 y 80 productores son eficientes técnicamente por alcanzar el 100 % de *ET*.

## 5. Conclusiones y sugerencias

De acuerdo con el objetivo planteado y los resultados obtenidos en la sección anterior, mediante frontera estocástica, se presentan las principales conclusiones y sugerencias del caso para pequeños, medianos y grandes productores de café ubicados en la zona cafetera colombiana (Caldas, Quindío y Risaralda). Este estudio determinó la función de producción cafetera, las economías a escala y la eficiencia técnica por tipo de productor.

En este sentido, y con información estadística microeconómica sobre el cultivo, recolectada en 2004 en 999 fincas cafeteras, pudo establecerse una función de producción cafetera flexible Translog minflex Laurent para pequeños, medianos y grandes cultivadores del grano, determinada mediante frontera estocástica, empleando los resultados del criterio Akaike en las estimaciones de máxima verosimilitud y mínimos cuadrados ordinarios, una vez estimadas las distintas funciones de producción cafeteras convencionales y flexibles en la tabla 1.

La metodología de frontera estocástica exige previamente una función de producción adecuada o correctamente especificada para evaluar y estimar las economías a escala y la eficiencia técnica en determinada actividad (en este caso el cultivo de café), porque el parámetro de eficiencia ( $\gamma$ ) y los estimadores del modelo son susceptibles a la forma funcional, como se apreció en las tablas 2, 4, 6 y 8.

De acuerdo con Greene (2002, p. 104), es incorrecto en funciones como la Translogarítmica Minflex Laurent analizar la relevancia de los estimadores, interpretar sus resultados directamente e inferir algún tipo de afirmación sobre si el comportamiento y la intensidad de los insumos cafeteros son complementarios, sustitutos o exhiben productividad marginal decreciente según los signos (negativos o positivos) parciales de los parámetros en los términos cuadráticos y las interacciones del modelo estimado.

Por tanto, deben calcularse las elasticidades y las economías a escala derivadas de la función ajustada (Translogarítmica Minflex Laurent), como fueron presentadas en las tablas 1, 3, 5 y 7 con el fin de entender apropiadamente la incidencia de los principales insumos empleados (tierra, trabajo, fertilizantes y maquinaria) en la producción del grano y su rendimiento por unidad productiva cafetera.

Con los valores encontrados de las elasticidades y economías a escala puede concluirse que el factor tierra es el insumo más importante para desarrollar la actividad ejercida por los pequeños cultivadores, seguido por la intensidad de fertilizantes y maquinaria. La cantidad de trabajo, aunque es relevante, mantiene

una relación inversa con la producción, dado que este grupo genera gran parte de mano de obra consumida en las explotaciones cafeteras campesinas y empresariales, y sus hectáreas cultivadas son de poca extensión. Esta situación no les obliga a demandar jornaleros, sustituidos por el trabajo propio de los minifundistas propietarios.

Igualmente, este grupo de caficultores exhibe rendimientos crecientes a escala en su actividad productiva, aunque son ineficientes técnicamente. Así, el cultivo del grano para ellos es afectado negativamente por la ineficiencia aleatoria y técnica, con 20 % en promedio por mejorar en esta última. Esto indica que no están asignando ni empleando adecuadamente los principales insumos en el área productiva en café, mano de obra, cantidad de químicos y maquinaria, los cuales pueden controlar para no incurrir en costos más altos de producción, baja productividad y competitividad en el sector. Lo anterior también sucede en las explotaciones cafeteras campesinas o medianas, y en el sector general, pero con un 30 % promedio de eficiencia técnica a mejorar.

El caso contrario ocurre en las grandes unidades cafeteras o empresariales, las cuales exhiben rendimientos decrecientes a escala, eficiencia técnica e ineficiencia estocástica. Esta última, señala existencia de factores externos no controlables por este grupo de cafeteros que afectan negativamente el desempeño adecuado de su producción; igualmente sucede en la actividad de los pequeños, medianos y sector general cafetero.

Entre estos factores pueden resaltarse los climáticos, el control de plagas en fincas vecinas, las medidas institucionales (adoptadas por el gobierno, la Federación Nacional de Cafeteros en Colombia y los organismos internacionales de café) la incertidumbre de la producción internacional de café y la volatilidad de los precios externos del grano. Los resultados señalaron también que la cantidad de mano de obra es el factor más importante para las explotaciones empresariales, dado que este grupo necesita la mayor proporción laboral ofrecida en el sector, porque su gran extensión en las hectáreas cultivadas de café obliga a demandar el trabajo requerido para atender la producción. Por esta razón, el factor tierra no es significativo para ellos y se le resta importancia en el análisis de la actividad. Por consiguiente, los resultados obtenidos en este estudio ayudan a entender cuáles deberían ser las políticas orientadas al aumento de productividad y competitividad del cultivo en la región cafetera de Colombia.

Las instituciones encargadas de prestar asesoría de eficiencia técnica a los productores de café en Colombia deben fortalecerse y apoyar principalmente a los pequeños y medianos productores, porque a que estos grupos son la mayor parte de caficultores del país y constituyen los sectores más vulnerables a cambios estructurales del mercado internacional por la ineficiencia técnica y estocástica presentada en su actividad.

Para los pequeños y medianos caficultores se requiere incentivar el acceso a tierras productivas en café; y para los empresariales, desincentivar el uso de químicos y promover la mano de obra. Los resultados encontrados son importantes para formulación de políticas cafeteras en Colombia. Así, las entidades encargadas de orientarlas pueden continuar recolectando este tipo de información microeconómi-

co, en el tiempo por unidad de producción y continuar con este tipo de estudios bajo distintas metodológicas analíticas que lleven a deducciones cuantitativas para implementar y fortalecer la política cafetera colombiana.

## Agradecimientos

Agradecemos al profesor Andrés L. Medaglia (profesor asociado al Departamento de Ingeniería Industrial de la Universidad de Los Andes) por los comentarios brindados a nuestro trabajo. Igualmente a los cuatro árbitros que con su trabajo voluntario, cuidadoso y observaciones constructivas contribuyeron a mejorar significativamente el resultado final del presente artículo. Finalmente a Tania Paola Barraza González (Matemática de la Universidad del Atlántico, Barranquilla) por realizar voluntariamente la edición final del documento en L<sup>A</sup>T<sub>E</sub>X.

[Recibido: agosto de 2010 — Aceptado: abril de 2011]

## Referencias

- Aigner, D., Lovell, C. & Schmidt, P. (1977), 'Formulation and estimation of stochastic frontier production model', *Journal of Econometrics* **6**, 21–37.
- Banco Mundial (2002), 'Estudio del sector cafetero en Colombia: resumen ejecutivo', *Ensayos sobre Economía Cafetera* **15**(18), 27–32.
- Battese, G. & Broca, S. (1991), 'Functional forms of stochastic frontier production functions and models for technical inefficiency effects: A comparative study for wheat farmers in Pakistan', *Journal of Productivity Analysis* **8**, 395–414.
- Battese, G. & Coelli, T. (1988), 'Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data', *Journal of Econometrics*, **38**(3), 387–399.
- Behr, A. & Tente, S. (2008), 'Stochastic frontier analysis by means of maximum likelihood and the method of moments', *Banking and Financial Studies* (19).  
\*<http://econpapers.repec.org/RePEc:zbw:bubdp2:200819>
- Cárdenas, G., Vedenov, D. & Houston, J. (2008), Analysis of production efficiency of Mexican coffee-producing districts, Technical Report 19470, American Agricultural Economics Association, 2005 Annual meeting, July 24-27, Providence, RI.  
\*<http://ageconsearch.umn.edu/handle/19470>
- Chiang, A. (1984), *Fundamental Methods of Mathematical Economics*, 3rd edn, McGraw-Hill, New York.
- Coelli, T. (1995), 'Estimators and hypothesis test for a stochastic frontier function: a Monte Carlo analysis', *Journal of Productivity Analysis* **6**(4), 247–268.

- Coelli, T. J. & Fleming, E. M. (2003), Diversification economies and specialisation efficiencies in a mixed food and coffee smallholder farming system in Papua New Guinea, Technical Report 25841, International Association of Agricultural Economists, 2003 Annual Meeting, August 16-22, Durban, South Africa. \*<http://ideas.repec.org/p/ags/iaae03/25841.html>
- Diewert, W. E. (1974), Applications of duality theory, *in* M. D. Intriligator & D. A. Kendrick, eds, 'Frontiers of Quantitative Economics', Vol. II, Amsterdam, North-Holland, pp. 106–171.
- Donnell, C. & Griffiths, W. (2006), 'Estimating state-contingent production frontiers', *American Journal of Agricultural Economics* **88**(1), 249–266.
- García, J. & Ramírez, J. (2002), 'Sostenibilidad económica de las pequeñas explotaciones cafeteras colombianas', *Ensayos sobre Economía Cafetera* **15**(18), 73–89.
- Greene, W. (1998), *Análisis Econométrico*, 3rd edn, Prentice Hall, New York.
- Greene, W. (2002), *Econometric Analysis*, 5th edn, Prentice Hall, New York.
- Guhl, A. (2004), 'Café y cambio de paisaje en la zona cafetera colombiana', *Ensayos sobre Economía Cafetera* (20), 137–153.
- Joachim, B., Kalilou, S., Ibrahim, D. & Gwendoline, N. (2003), 'Factors affecting technical efficiency among coffee farmers in Côte D'Ivoire: Evidence from the centre west region', *African Development Review* **15**, 1–66.
- Konstantinos, G., Kien, T. & Vangelis, T. (2003), 'On the choice of functional form in stochastic frontier modeling', *Empirical Economics* **28**, 75–100.
- Kumbhakar, S. (1993), 'Short-Run Returns to Scale, Farm-Size, and Economic Efficiency', *Journal Review of Economics and Statistics* **75**, 336–341.
- Lohr, L. & Park, T. (2006), 'Technical efficiency of U.S. organic farmers: The complementary roles of soil management techniques and farm experience', *Agricultural and Resource Economics Review* **35**, 327–338.
- Meesen, W. & Van den Broeck, J. (1977), 'Efficiency estimation from Cobb Douglas production functions with composed error', *International Economic Review* **18**, 435–444.
- Mendieta, J. & Perdomo, J. (2008), *Fundamentos de Economía del Transporte: Teoría, Metodología y Análisis de Política*, 1 edn, Ediciones Uniandes, Bogotá, Colombia.
- Mosheim, R. (2002), 'Organizational type and efficiency in the Costa Rican coffee processing sector', *Journal of Comparative Economics* **30**, 296–316.
- Nicholson, W. (2002), *Microeconomic Theory*, South-Western Thomson Learning, United States.



- Ornelas, F., Shumway, R. & Ozuna, T. (1994), 'Using quadratic Box-Cox for flexible functional form selection and unconditional variance computation', *Empirical Economics* **19**, 639–645.
- Perdomo, J. (2006), Estimación de funciones de producción y eficiencia técnica en el eje cafetero colombiano: una aplicación con fronteras estocásticas vs DEA, Tesis de maestría, Universidad de los Andes, Facultad de Economía, Bogotá, Colombia.
- Perdomo, J. (2010), 'Una propuesta metodológica para estimar los cambios sobre el valor de la propiedad: estudio de caso para Bogotá aplicando *propensity score matching* y precios hedónicos espaciales', *Lecturas de Economía* **73**, 49–65.
- Perdomo, J. (2011), 'A methodological proposal to estimate changes of residential property value: Case study developed in Bogota', *Applied Economics Letters* **18**(7), 605–610.
- Perdomo, J. & Hueth, D. (2010), Funciones de producción y eficiencia técnica en el eje cafetero colombiano: una aproximación con frontera estocástica, Technical Report 007606, Universidad de los Andes.  
\*<http://ideas.repec.org/p/col/000089/007606.html>
- Perdomo, J., Hueth, D. & Mendieta, J. (2007), 'Factores que afectan la eficiencia técnica en el sector cafetero de colombiano: una aplicación con análisis envolvente de datos', *Ensayos sobre Economía Cafetera* **22**, 121–140.
- Perdomo, J. & Mendieta, J. (2007), 'Factores que afectan la eficiencia técnica y asignativa en el sector cafetero de colombiano: una aplicación con análisis envolvente de datos', *Revista Desarrollo y Sociedad* **60**, 1–45.
- Ramírez, L., Silva, G., Valenzuela, L., Villegas, A. & Villegas, L. (2002), 'Comisión de ajuste de la institucionalidad cafetera', *Ensayos sobre Economía Cafetera* (17).
- Reinhard, S., Lovell, K. & G., T. (1999), 'Econometric estimation of technical and environmental efficiency: An application to Dutch dairy farms', *American Journal in Agriculture* **81**, 44–60.
- Rios, A. R. & Shively, G. E. (2005), Farm size and nonparametric efficiency measurements for coffee farms in Vietnam, Technical Report 19159, American Agricultural Economics Association, 2005 Annual meeting, July 24-27, Providence, RI.  
\*<http://ideas.repec.org/p/ags/aaea05/19159.html>
- Rosales, R., Perdomo, J., Morales, C. & Urrego, A. (2010), *Fundamentos de Econometría Intermedia: Teoría y Aplicaciones*, number 1, Universidad de Los Andes, Bogotá, Colombia. Apuntes de Clase CEDE. ISBN/ISSN: 1909-4442.
- Saravia, S. (2007), A Stochastic Frontier Model of the Nicaraguan Coffee Sector: Analyzing Efficiency and Performance Under Changing Political Environment, II International Symposium on Economic Theory, Policy and Application, Athens, Greece.

Thanda, K. & Matthias von, O. (1999), Stochastic Frontier Production Function and Technical Efficiency Estimation: A Case Study on Irrigated Rice in Myanmar, Deutscher Tropentag, Berlin.

Wollni, M. (2007), Productive and Efficiency of Specialty and Conventional Coffee Farmers in Costa Rica: Accounting for the Use of Different Technologies and Self-Selection, Technical Report 9956, Department of Applied Economics and Management, 2007 Annual Meeting, July 29-August 1, Portland, Oregon TN.  
\*<http://ageconsearch.umn.edu/handle/9956>

## Apéndice

### Varianza de los estimadores mediante máxima verosimilitud en la frontera estocástica

De acuerdo con Behr & Tente (2008, 19-20), Greene (1998, 2002), la ecuación (2) y características del componente estocástico ( $\eta_i = v_i + u_i$ ), la varianza de los estimadores mediante máxima verosimilitud en la frontera estocástica está determinada por las condiciones de primer orden en el logaritmo de la función de verosimilitud ( $\ln f$ ) de la siguiente manera:

$$\ln f(\sigma_s^2, \hat{\beta}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma_s^2) + \sum_{i=1}^n \ln[1 - \varphi(z_i)] - \frac{n}{2\sigma_s^2} \sum_{i=1}^n [q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta})]^2$$

$$\frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \hat{\beta}} = \frac{n}{\sigma_s^2} \sum_{i=1}^n [q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta})] x_i + \frac{\gamma}{\sigma_s} \sum_{i=1}^n \left( \frac{\varphi^*}{1 - \varphi(z_i)} \right) x_i,$$

$$\varphi^* = \ln[1 - \varphi(z_i)]$$

$$\frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \sigma_s^2} = -\frac{n}{\sigma_s^2} + \frac{1}{2\sigma_s^4} \sum_{i=1}^n [q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta})]^2 + \frac{1}{2\sigma_s^3} \sum_{i=1}^n \left( \frac{\varphi^*}{1 - \varphi(z_i)} \right) [q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta})]$$

$$\frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \gamma} = -\frac{1}{\sigma_s} + \frac{1}{2\sigma_s^4} \sum_{i=1}^n \left( \frac{\varphi^*}{1 - \varphi(z_i)} \right) [q_i - f(x_{1i}, x_{2i}, x_{3i}, x_{4i}, \hat{\beta})]$$

Por consiguiente, la matriz  $\mathbf{W}$  contiene las condiciones de primer orden, y la matriz de varianza y covarianza de los estimadores de máxima verosimilitud ( $Var[\hat{\beta}]$ ) se estima consistentemente como

$$\mathbf{W} = \begin{bmatrix} \frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \beta} \\ \frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \sigma_s^2} \\ \frac{\partial \ln f(\sigma_s^2, \hat{\beta})}{\partial \gamma} \end{bmatrix} \quad Var[\hat{\beta}] = (\widehat{\mathbf{W}}' \widehat{\mathbf{W}})^{-1}$$

## Información para los autores

La **Revista Colombiana de Estadística** publica artículos originales de carácter teórico o aplicado en cualquiera de las ramas de la estadística. Se consideran también artículos divulgativos de gran calidad de exposición sobre metodologías o técnicas estadísticas aplicadas en diferentes campos del saber. Únicamente se publican artículos en español e inglés, si el autor escribe en una lengua diferente a la nativa debe enviar un certificado de un traductor oficial o de un corrector de estilo que haya revisado el texto.

El Comité Editor únicamente acepta trabajos para evaluación que no han sido publicados previamente y que no están siendo propuestos simultáneamente para publicación en otros medios, ni lo serán sin previo consentimiento del Comité, a menos que, como resultado de la evaluación, se decida no publicarlos en la Revista. Se supone además que cuando los autores hacen entrega de un documento con fines de publicación en la **Revista Colombiana de Estadística**, conocen las condiciones anteriores y que están de acuerdo con ellas.

### Material

Los artículos remitidos a la **Revista Colombiana de Estadística** deben ser presentados en archivo PDF o PS, con textos, gráficas y tablas en color negro y, además, los autores deben agregar una versión del artículo sin nombres ni información de los autores, que se utilizará para el arbitraje. Se debe enviar una carta firmada por cada uno de los autores, donde manifiesten estar de acuerdo con someter el artículo y con las condiciones de la Revista. Si un artículo es aceptado, los autores deben poner a disposición del Comité Editorial los archivos: fuente en  $\text{\LaTeX}$  y de gráficas en formato EPS en blanco y negro.

Para facilitar la preparación del material publicado se recomienda utilizar  $\text{MiKTeX}^1$ , usando los archivos de la plantilla y del estilo *revcoles* disponibles en la página Web de la Revista<sup>2</sup> y siguiendo las instrucciones allí incorporadas.

Todo artículo debe incluir:

- Título en español y su traducción al inglés.
- Los nombres completos y el primer apellido, la dirección postal o electrónica y la afiliación institucional de cada autor.
- Un resumen con su versión en inglés (*abstract*). El resumen en español no debe pasar de 200 palabras y su contenido debe destacar el aporte del trabajo en el tema tratado.
- Palabras clave (*Key words*) en número entre 3 y 6, con su respectiva traducción al inglés, siguiendo las recomendaciones del *Current Index to Statistics* (CIS)<sup>3</sup>.

---

<sup>1</sup><http://www.ctan.org/tex-archive/systems/win32/miktex/>

<sup>2</sup><http://www.estadistica.unal.edu.co/revista>

<sup>3</sup><http://www.statindex.org/CIS/homepage/keywords.html>

- Cuando el artículo se deriva de una tesis o trabajo de grado debe indicarse e incluirse como una referencia.
- Si se deriva de un proyecto de investigación, se debe indicar el título del proyecto y la entidad que lo patrocina.
- Referencias bibliográficas, incluyendo solamente las que se hayan citado en el texto.

### **Referencias y notas al pie de página**

Para las referencias bibliográficas dentro del texto se debe utilizar el formato autor-año, dando el nombre del autor seguido por el año de la publicación dentro de un paréntesis. La plantilla L<sup>A</sup>T<sub>E</sub>X suministrada utiliza, para las referencias, los paquetes BibT<sub>E</sub>X y Harvard<sup>4</sup>. Se recomienda reducir el número de notas de pie de página, especialmente las que hacen referencia a otras notas dentro del mismo documento y no utilizarlas para hacer referencias bibliográficas.

### **Tablas y gráficas**

Las tablas y las gráficas, con numeración arábica, deben aparecer referenciadas dentro del texto mediante el número correspondiente. Las tablas deben ser diseñadas en forma que se facilite su presentación dentro del área de impresión de la Revista. En este sentido, los autores deben considerar en particular la extensión de las tablas, los dígitos representativos, los títulos y los encabezados. Las gráficas deben ser visualmente claras y debe ser posible modificar su tamaño. Cuando el artículo sea aceptado para su publicación, los autores deben poner la versión definitiva a disposición del Comité Editorial. Todos los elementos como barras, segmentos, palabras, símbolos y números deben estar impresos en color negro.

### **Responsabilidad legal**

Los autores se hacen responsables por el uso de material con propiedad intelectual registrada como figuras, tablas, fotografías, etc.

### **Arbitraje**

Los artículos recibidos serán revisados por el Comité Editorial y sometidos a arbitraje por pares especializados en el tema respectivo. El arbitraje es “doble ciego” (árbitros anónimos para los autores y viceversa). El Comité Editorial decide aceptar, rechazar o solicitar modificaciones a los artículos con base en las recomendaciones de los árbitros.

---

<sup>4</sup><http://tug.ctan.org/tex-archive/macros/latex/contrib/harvard>