

00384  
4  
LES



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

**FACULTAD DE CIENCIAS  
DIVISION DE ESTUDIOS DE POSGRADO**

**USO DE LA DISTRIBUCION PREDICTIVA EN  
BONDAD DE AJUSTE: GAUSSIANA INVERSA**

**FALLA DE ORIGEN**

**T E S I S**

**Que para obtener el grado Académico de**

**DOCTOR EN CIENCIAS**

**(MATEMATICAS)**

**p r e s e n t a**

**RAUL RUEDA DIAZ DEL CAMPO**

**DIRECTOR DE TESIS: DR. FEDERICO JORGE O' REILLY TOGNO**

1995



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

2007 ON 2007 0725  
ESTR DE LA UNIVERSIDAD  
ASOCIACION AF 10 01102

# USO DE LA DISTRIBUCION PREDICTIVA EN BONDAD DE AJUSTE: GAUSSIANA INVERSA

Goodness-of-fit of the Inverse Gaussian distribution when parameters are estimated is addressed, using the predictive distribution to transform the sample into the unit interval.

A parametric family of prior distributions is proposed, that results invariant to one-to-one transformations of the parameters, with the advantage that different parametrizations of the problem can be dealt with. Moreover, the usual non-informative prior distributions are reconstructed as particular members of this family.

A numerical study is done to compare the predictive transformations with the transformations obtained with maximum likelihood estimators and the Rao-Blackwell distribution function estimator and the generated empirical processes. These three transformations are quite similar yielding slightly different values. Nevertheless, the corresponding test statistics lead to the same conclusion. As another result, it can be said that the induced empirical processes converge to the same asymptotic process.

**A mis tres erres**

## INTRODUCCION

En muchos procedimientos de inferencia estadística la distribución de la cual se obtiene la muestra, se supone conocida en forma aunque con parámetros desconocidos. Las inferencias que se obtienen, con base en la muestra, en general dependen fuertemente de la suposición de la forma de la distribución mencionada. La verificación de este supuesto puede entonces ser considerado como uno de los problemas más importantes en la Estadística. El área de la Estadística que trata esta temática es conocida como Bondad de Ajuste y el libro de D'Agostino y Stephens (1986) es la referencia más reciente que puede darse sobre el tema.

De manera general, cualquier problema de Bondad de Ajuste es en si, un problema de prueba de hipótesis, por lo que debe tenerse una estadística que permita discriminar entre las hipótesis propuestas y conocer su distribución cuando la hipótesis nula se supone verdadera.

Así, hay dos aspectos fundamentales a tratar en cualquier procedimiento de Bondad de Ajuste: el primero, es la elaboración de una propuesta general para una "estadística de prueba", que contemple el hecho de que deben estimarse los parámetros desconocidos, y el segundo, la posibilidad de obtener la distribución de esta estadística bajo el supuesto de que la hipótesis nula es verdadera.

Sobre el primer punto, existen diferentes propuestas de entre las que sobresalen las basadas en la función de distribución empírica y dentro de éstas, una de las más utilizadas es la  $A^2$  de Anderson - Darling, con la que se han desarrollado procedimientos para verificar el ajuste de algunas distribuciones como la Exponencial y la Normal.

El caso en que la distribución especificada por la hipótesis nula es continua y sus parámetros son conocidos, ha sido ampliamente estudiado y se reduce al problema de pruebas sobre uniformidad. La situación se complica, cuando los parámetros son desconocidos. Si estos se estiman por máxima verosimilitud, se tienen resultados asintóticos y, en algunos casos, la distribución para  $n$  finita puede establecerse, aunque es complicada.

Sin embargo, existen otras propuestas de solución que consisten en estimar a toda la función de distribución usando el estimador de Rao-Blackwell o mediante la distribución predictiva, utilizada ampliamente en el contexto bayesiano de la Estadística. En estos casos, el problema principal es encontrar la distribución asintótica de la estadística de

prueba, pues dada la complejidad del problema pocos esfuerzos se han dedicado al caso finito. Existen resultados que permiten asegurar una equivalencia entre las distribuciones asintóticas que aparecen al estimar por máxima verosimilitud, por Rao-Blackwell o con la distribución predictiva, fundamentalmente cuando los parámetros son de escala y/o localización. Fuera de este caso, no existen resultados generales.

La distribución Gaussiana Inversa es un ejemplo en donde uno de los parámetros no es de escala ni de localización, por lo que surgió la inquietud de estudiar el comportamiento de la estadística  $A^2$  de Anderson-Darling, cuando se utiliza al estimador de Rao-Blackwell y a la distribución predictiva. El primer caso fué resuelto por O'Reilly & Rueda (1992) y el segundo es el tema de este trabajo.

Como es sabido, la distribución predictiva depende de la especificación de la distribución inicial, por lo que la obtención de distribuciones iniciales de referencia para los parámetros de la Gaussiana Inversa fué el primer aspecto que se tuvo que estudiar, obteniéndose una familia paramétrica que contiene como casos particulares a algunas de las distribuciones de referencia reportadas en la literatura. El siguiente paso, fué la obtención de las correspondientes distribuciones predictivas que servirán para transformar al intervalo unitario. La comparación de estos resultados con los obtenidos al utilizar al estimador de Rao-Blackwell y con los máximo verosímiles, fué el siguiente objetivo, así como estudiar el comportamiento de la correspondiente estadística  $A^2$ .

El trabajo se complementa con una descripción general del problema de Bondad de Ajuste y de las soluciones basadas en la función de distribución empírica, así como una breve introducción a las distribuciones predictivas.

Gracias a:

- Federico O'Reilly por su paciencia en la atinada dirección de este trabajo y sobre todo por su amistad.
- Belem, Silvia y Lety por su constante apoyo.
- A José Antonio, Francisco y Elida por su ayuda, sobre todo en los "bomberazos".
- Claudia y Marcelo por su ayuda en los engorrosos trámites.
- Todos los que apoyaron moralmente preguntando "¿ya para cuándo?": Sebastián, Carlos, Claudia, Nacho, Laura, José Antonio, Salvador, Rubén, Mauricio, Mari, Jaime, Toño y Pablo.

**- Nacho, Alfredo, Begoña, Belem, Silvia y Víctor por sus comentarios al revisar el trabajo.**

**El apoyo del IIMAS, en especial del Departamento de Probabilidad y Estadística, y del Sistema Nacional de Investigadores fue fundamental para el desarrollo del trabajo.**

**Muy en especial, gracias a Renata por haber decidido compartir su vida.**

## CAPITULO 1

### I. INTRODUCCION

La distribución Gaussiana Inversa aparece por primera vez en un trabajo de Schrödinger (1915) como la distribución de *primera llegada* en un proceso Gaussiano o movimiento browniano con deriva positiva. El nombre de Gaussiana Inversa fué dado por Tweedie (1956) después de haber notado la relación que existe entre la función generadora de cumulantes de la distribución Normal y la de esta distribución. En 1945, Tweedie demuestra que este tipo de relaciones se dan también en otras distribuciones como la Binomial con la Binomial Negativa y la Poisson con la Exponencial, por lo que acuña el nombre de *variables inversas*.

La distribución Gaussiana Inversa fue también generada por Wald (1947) como aproximación a la distribución del tamaño de muestra en las pruebas secuenciales de hipótesis.

En los últimos años, la distribución Gaussiana Inversa ha recibido mucha atención como un buen modelo de fenómenos de *tiempo de espera*. Una excelente revisión con abundantes ejemplos de aplicaciones, es el libro de Chhikara y Folks (1989). Además de presentar una gran cantidad de resultados relacionados con la distribución Gaussiana Inversa, tanto desde la perspectiva bayesiana como la clásica, incluye algunas aplicaciones en las áreas de confiabilidad, comunicación, cardiología, lingüística, agricultura, finanzas e ingeniería, entre otras, mostrando así la importancia de esta distribución. En todos los ejemplos que presentan los autores, se menciona que hubo una verificación del ajuste para determinar si la distribución Gaussiana Inversa explicaba razonablemente el conjunto de datos que se estudiaban. En algunos casos, la bondad del ajuste era verificada con pruebas de  $\chi^2$ , o simplemente con un ajuste gráfico de los datos.

Chhikara y Folks (1989) proponen, para probar el ajuste, un procedimiento basado en la estadística de Kolmogorov - Smirnov. Son bien conocidas las propiedades de potencia de estos procedimientos cuando los parámetros de la distribución son conocidos, sin embargo, en el caso de que los parámetros sean desconocidos, los valores críticos dependen de la distribución; y mas aún, posiblemente de los valores de los parámetros. En este caso, Chhikara y Folks (1989) calculan los valores críticos para esta prueba, mediante

simulación, para algunos valores de los parámetros y del tamaño de muestra, añadiendo algunos ajustes para cubrir todos los posibles valores de  $n$  y de los parámetros.

Una estadística muy usada en el área de bondad de ajuste es la  $A^2$  de Anderson - Darling (Anderson y Darling, 1954). Esta prueba tiene buenas propiedades de potencia (Stephens, 1974) y sus propiedades asintóticas han sido ampliamente estudiadas (Durbin, 1973a). Al igual que la mayoría de los procedimientos utilizados en bondad de ajuste, cuando los parámetros son desconocidos, los valores críticos dependen de la distribución especificada en la hipótesis y como ya se mencionó, posiblemente de los parámetros. Sin embargo, si los parámetros son de localización y/o escala, los valores críticos sólo dependen de la distribución que se trate pero no de los parámetros, permitiendo la construcción de tablas muy generales (e.g. Stephens, 1977, 1979; Lockhart y Stephens, 1985, entre otros).

En la siguiente sección, se describirá el problema de bondad de ajuste compuesto y algunas soluciones propuestas en la literatura.

## II. EL PROBLEMA DE BONDAD DE AJUSTE

Un problema común en Estadística es el siguiente: sea  $X$  una variable aleatoria cuya función de distribución es un elemento de  $\mathcal{F}$  una familia muy general de distribuciones (e.g. todas las distribuciones continuas). Supóngase que se tiene una muestra aleatoria de la distribución de  $X$ , es decir se tienen  $X_1, X_2, \dots, X_n$  variables aleatorias independientes con la misma distribución, y se desea probar la hipótesis

$$H_0 : F \in \mathcal{F}_0$$

donde  $F$  es la función de distribución de  $X$ ,  $\mathcal{F}_0 = \{F_0(X|\theta) : \theta \in \Theta\}$  es un subconjunto de  $\mathcal{F}$  y  $\Theta$  es un subconjunto de  $\mathbf{R}^k$ .

Existen diferentes procedimientos para probar esta hipótesis, de entre ellos, los basados en la función de distribución empírica han recibido mucha atención en la literatura. Si  $F_n(x)$  denota a la distribución empírica de la muestra, estos procedimientos consideran una medida de la discrepancia entre  $F_n(x)$  y  $F_0(x|\theta)$  y de acuerdo a la magnitud de esta discrepancia, se rechaza o no se rechaza la hipótesis  $H_0$ . El valor de esta discrepancia es comparado con los valores críticos de la distribución de la estadística de prueba, que son calculados suponiendo que  $H_0$  es verdadera.

Si la variable aleatoria  $X$  es continua y los parámetros de  $F$  son conocidos, llamado *caso simple* de bondad de ajuste, la hipótesis  $H_0$  puede ser llevada a una hipótesis equivalente, sobre uniformidad. Sea  $\{X_1, X_2, \dots, X_n\}$  una muestra aleatoria de la distribución de  $X$  y supóngase, sin pérdida de generalidad, que está ordenada de menor a mayor. Supóngase además que en la hipótesis la distribución  $F_0(X|\theta)$  está totalmente especificada, entonces es un hecho conocido que la transformación

$$U = U(X) = F_0(X|\theta) \quad (1)$$

transforma a la muestra  $\{X_1, X_2, \dots, X_n\}$  en una muestra ordenada  $\{U_1, U_2, \dots, U_n\}$  de variables aleatorias uniformes en el  $[0, 1]$ , de manera que la hipótesis  $H_0$  es equivalente a

$$H'_0 : F(u) = U(u|0, 1).$$

Así, dada la transformación (1), basta especificar una medida de la discrepancia entre la función de distribución empírica de la muestra  $\{U_1, U_2, \dots, U_n\}$ , que se denotará por  $G_n$ , y la propia distribución Uniforme. En la literatura se han propuesto una diversidad de funciones de discrepancia, de las que sobresalen dos clases: las del tipo supremo y las cuadráticas. Asociadas a la primera clase, se encuentran las estadísticas de Kolmogorov - Smirnov (Kolmogorov, 1933; Smirnov, 1939). La familia de Cramér - Von Mises (Cramér, 1928; Von Mises, 1931) se encuentra entre las del segundo tipo. Ambas medidas de discrepancia están fuertemente relacionadas con funciones de distancia o métricas definidas en espacios de funciones, la métrica del supremo y la métrica cuadrática. Específicamente y en términos de la transformación (1), estas medidas de discrepancia están definidas como

$$\begin{aligned} D_n &= \sup_{u \in [0,1]} |G_n(u) - u|, \\ \mathcal{W}_n^2 &= n \int_0^1 (G_n(u) - u)^2 \psi(u) du, \end{aligned} \quad (2)$$

donde  $\psi$  es una función no negativa y continua.

En ambos casos, la discrepancia es una función de

$$y_n(u) = \sqrt{n} \{G_n(u) - u\}, u \in [0, 1] \quad (3)$$

con  $u = F_0(x|\theta)$ . Para cada  $u \in [0, 1]$  y  $n \in \mathbb{N}$  la función  $y_n(u)$  es una variable aleatoria, por lo que

$$\{y_n\} = \{y_n(u) : u \in [0, 1]\} \quad (4)$$

define un proceso estocástico, llamado *proceso empírico*.

El principal problema es encontrar, para cada valor de  $n \in \mathbb{N}$ , las distribuciones de  $D_n$  y de  $\mathcal{W}_n^2$ , suponiendo  $H_0$  verdadera.

La distribución de  $D_n$  fue dada por Kolmogorov (1933) para algunos casos particulares y extendida al caso general por Durbin (1968). A pesar de tener una forma analítica para esta distribución, su aplicación no es sencilla pues requiere la construcción de una matriz de transición de dimensión  $2[n\alpha] + 1$ , donde  $\alpha$  es el punto donde se evalúa la distribución de  $D_n$ .

Relacionadas con  $D_n$ , están las estadísticas  $D_n^+$  y  $D_n^-$  definidas como

$$\begin{aligned} D_n^+ &= \sup_{u \in [0,1]} (G_n(u) - u), \\ D_n^- &= \sup_{u \in [0,1]} (u - G_n(u)), \end{aligned} \quad (5)$$

introducidas por Smirnov (1939, 1941). Es fácil ver que  $D_n = \max(D_n^+, D_n^-)$ . Las distribuciones de estas estadísticas para cualquier valor de  $n \in \mathbb{N}$ , fueron dadas por el mismo Smirnov (1941).

Con respecto a la distribución de  $\mathcal{W}_n^2$  para valores de  $n \in \mathbb{N}$ , poco puede decirse. Existen resultados para algunos valores de  $n$  y cuando  $\psi(u) = 1$ , para toda  $u$ . Marshall (1958) encontró la distribución de  $\mathcal{W}_n^2$  para  $n = 1, 2, 3$ . Stephens y Maag (1968) dan expresiones para los cuantiles inferiores extremos en este caso. En general, puede darse una ecuación diferencial en diferencias sobre la función característica, que puede resolverse numéricamente y que permite encontrar los cuantiles de  $\mathcal{W}_n^2$  (Durbin, 1973a).

Dadas las limitaciones para el caso  $\mathcal{W}_n^2$  y las complicaciones numéricas para  $D_n$ , se han dedicado grandes esfuerzos para derivar las distribuciones asintóticas de estas estadísticas. Como ambas dependen del proceso empírico  $\{y_n(u) : u \in [0, 1]\}$  definido en (4), lo natural

es encontrar la distribución de este proceso y posteriormente la de las transformaciones correspondientes que son

$$\begin{aligned}\sqrt{n}D_n &= \sup_{u \in [0,1]} |y_n(u)|, \\ \mathcal{W}_n^2 &= \int_0^1 y_n^2(u) \psi(u) du.\end{aligned}\tag{6}$$

Como

$$G_n(u) = \frac{1}{n} \sum_{i=0}^n I_{[0,u]}(U_i),\tag{7}$$

donde  $I_{[0,u]}$  denota a la función indicadora del intervalo  $[0, u]$  y  $\{U_1, \dots, U_n\}$  a la muestra ordenada, puede demostrarse fácilmente que

$$E(G_n(u)) = u \text{ y } \text{Cov}(G_n(u), G_n(u')) = \frac{u(1-u')}{n}, \quad u < u'\tag{8}$$

de donde puede deducirse que el proceso  $\{y_n(u) : u \in [0, 1]\}$  tiene media cero y función de covarianza dada por

$$\text{Cov}(y_n(u), y_n(u')) = \min(u, u') - uu'.\tag{9}$$

Usando el teorema central de límite, puede demostrarse que el vector  $(y_n(u_1), y_n(u_2), \dots, y_n(u_k))$  se distribuye asintóticamente como una Normal multivariada con media cero y matriz de varianzas calculada a partir de (9), para todo valor de  $k$ . Esto lo que significa, es que las distribuciones finito dimensionales del proceso  $\{y_n\}$  convergen a las distribuciones finito dimensionales de un proceso Gaussiano con media cero y función de covarianza dada por (9). La demostración de que un proceso con las características anteriores existe en  $C[0, 1]$  puede encontrarse en Billingsley (1968).

Denótese por  $\omega(u)$  al proceso browniano, es decir, al proceso definido en el intervalo  $[0, 1]$  que cumple con

- i.  $P[\omega(0) = 0] = 1$
- ii.  $p(\omega(u)) = N(\omega(u)|0, u)$  para cada  $u \in [0, 1]$
- iii.  $\omega(u_1) - \omega(u_0), \omega(u_2) - \omega(u_1), \dots, \omega(u_k) - \omega(u_{k-1})$  son independientes para cada  $0 \leq u_0 < u_1 < \dots < u_k \leq 1$ . (10)

Considérese ahora al proceso

$$y(u) = \omega(u) - u\omega(1),$$

llamado *browniano atado*, para el que puede demostrarse a partir de (10), que tiene media cero y función de covarianza dada por (9).

Sin embargo la convergencia de las distribuciones finito dimensionales no es suficiente para asegurar que, por ejemplo,  $\sup |y_n(u)| \rightarrow \sup |y(u)|$ , para esto es necesario el concepto de convergencia débil o convergencia en distribución (e.g. Billingsley, 1968).

Sea  $\Omega$  un espacio métrico con métrica  $d$  que contiene a las funciones  $\phi$  definidas en el  $[0, 1]$ , denótese con  $\mathcal{A}$  a la  $\sigma$ -álgebra generada por los abiertos en  $(\Omega, d)$  y considérese a  $(\Omega, \mathcal{A})$ . Sea  $\{\eta_n : n \in \mathbf{N}\}$  una sucesión de funciones medibles definidas sobre  $\Omega$  y  $\eta$  otra función medible definida también sobre  $\Omega$ . Se dice que  $\{\eta_n\}$  converge débilmente a  $\{\eta\}$  en  $(\Omega, \mathcal{A})$  si

$$\lim_{n \rightarrow \infty} E_{P_n}[f(\eta_n)] = E_P[f(\eta)]$$

para toda  $f$  continua y acotada en el espacio métrico  $(\Omega, d)$ , denotándolo por  $\{\eta_n\} \Rightarrow \{\eta\}$ , donde  $\{P_n : n \in \mathbf{N}\}$  y  $P$  son las correspondientes distribuciones de probabilidad definidas en  $(\Omega, \mathcal{A})$ . Esto es equivalente a decir que  $\{P_n : n \in \mathbf{N}\}$  converge débilmente a  $P$  y se denota con  $P_n \Rightarrow P$ .

Un resultado muy importante en convergencia débil, es el que asegura que si  $g$  es una función definida en  $\Omega$  que es continua con respecto a la métrica  $d$  y  $\{\eta_n\} \Rightarrow \{\eta\}$  entonces  $\{g(\eta_n)\} \Rightarrow \{g(\eta)\}$ .

Dado el resultado anterior, bastaría demostrar que las funciones definidas en  $\Omega$  como

$$\begin{aligned}
g_1(\eta) &= \sup_u |\eta(u)|, \\
g_2(\eta) &= \int_0^1 \eta(u) \psi(u) du,
\end{aligned}
\tag{11}$$

son continuas respecto a una métrica  $d$ .

Como el proceso empírico  $\{y_n\}$  no tiene trayectorias continuas en el intervalo  $[0, 1]$ , (mientras que las del proceso browniano atado si son continuas), el espacio natural de definición del proceso es  $D$ , el espacio de todas las funciones en el  $[0, 1]$  continuas por la derecha y cuyo límite izquierdo existe para todo  $u$  en  $[0, 1]$ . Este espacio puede ser dotado de diferentes métricas, por ejemplo, la métrica uniforme definida para toda  $\eta_1, \eta_2$  en  $D$  como

$$d_u(\eta_1, \eta_2) = \sup_{0 \leq u \leq 1} |\eta_1(u) - \eta_2(u)|. \tag{12}$$

El espacio métrico  $(D, d_u)$  es completo pero no separable (Billingsley, 1968), lo que presenta ciertas dificultades técnicas para demostrar la convergencia débil de ciertos procesos. Sin embargo, puede demostrarse (Skorokhod, 1956) que existe una métrica  $d_0$  definida sobre  $D \times D$  tal que el espacio métrico  $(D, d_0)$  es separable y completo. La introducción de esta métrica en el estudio de la convergencia débil, no es una condición necesaria, sin embargo facilita enormemente la teoría (Shorack y Wellner, 1986).

Si se considera al espacio  $(D, \mathcal{D})$ , donde  $\mathcal{D}$  es la  $\sigma$ -álgebra generada por los abiertos de  $(D, d_u)$  y se extiende el dominio de definición de los procesos browniano y browniano atado a  $(D, \mathcal{D})$ , puede demostrarse que el proceso  $\{y_n\}$  definido en (4) es un elemento de  $D$ . A partir de la convergencia de las distribuciones finito dimensionales de  $\{y_n\}$  a las de  $\{y\}$ , Doob (1949) mediante un argumento heurístico, demostró que  $\{y_n\} \Rightarrow \{y\}$ . Donsker (1952) dió una demostración formal de esta convergencia en  $(D, \mathcal{D}, d_u)$  (véase Pollard, 1984).

Por otro lado, puede demostrarse (Durbin, 1973a) que  $g_1$  y  $g_2$  definidas en (11) son continuas en  $(D, d_u)$ , por lo que

$$\sup_{0 \leq u \leq 1} |y_n(u)| \Rightarrow \sup_{0 \leq u \leq 1} |y(u)| \text{ y}$$

$$\int_0^1 y_n(u)\psi(u)du \Rightarrow \int_0^1 y(u)\psi(u)du. \quad (13)$$

Así, bastaría determinar las distribuciones de los términos de la derecha de la expresión anterior, para tener caracterizadas a las distribuciones asintóticas de interés.

Antes de continuar, vale la pena hacer un paréntesis en el tema de convergencia débil.

Una demostración interesante de que  $\{y_n\} \Rightarrow \{y\}$  en  $(D, \mathcal{D}, d_u)$ , es presentada por Pollard (1984): dada la convergencia de las distribuciones finito dimensionales, demuestra que las trayectorias del proceso  $\{y_n\}$  pueden aproximarse por procesos *locales* que implican la convergencia débil del proceso original. Esta condición es equivalente al concepto de *tensión* que se definirá un poco más adelante. Sin embargo, Pollard (1984) menciona que este tipo de condiciones para construir aproximaciones simples al proceso no son muy utilizadas debido, posiblemente, a que debe demostrarse en forma independiente la existencia del proceso límite. El tipo de demostración más utilizado requiere desarrollar otro tipo de herramientas (e.g. Billingsley, 1968).

Sea  $(S, \mathcal{L})$  con  $S$  un espacio métrico y  $\mathcal{L}$  la  $\sigma$ -álgebra generada por los abiertos de  $S$ . Sea  $\mathcal{P}$  una familia de medidas de probabilidad definidas sobre  $(S, \mathcal{L})$ . Se dice que la familia  $\mathcal{P}$  es *tensa* si para toda  $\varepsilon \in \mathbb{R}^+$  existe un conjunto compacto  $K \in \mathcal{L}$  tal que  $P(K) > 1 - \varepsilon$  para toda  $P \in \mathcal{P}$ .

La importancia de este concepto, queda establecida en el siguiente resultado, cuya demostración puede consultarse en Billingsley (1968): sea  $\{\eta_n : n \in \mathbb{N}\}$  una sucesión contenida en  $(D, \mathcal{D})$  y denótese por  $\{P_n : n \in \mathbb{N}\}$  la correspondiente sucesión de distribuciones de probabilidades. Si las distribuciones finito dimensionales del proceso convergen a las del proceso  $\{\eta\}$  y la familia  $\{P_n : n \in \mathbb{N}\}$  es *tensa*, entonces  $P_n \Rightarrow P$  donde  $P$  es la distribución de probabilidad asociada a  $\{\eta\}$ .

La condición de tensión puede relajarse, si el espacio de definición de los procesos es separable. Sea  $\mathcal{P}$  una familia de medidas de probabilidad definidas en el espacio métrico  $(S, \mathcal{L})$ . Se dice que la familia  $\mathcal{P}$  es *relativamente compacta* si para toda sucesión  $\{P_n : n \in \mathbb{N}\}$  contenida en  $\mathcal{P}$  existe una subsucesión  $\{P_{n'}\}$  que converge débilmente a una medida  $P$  definida en  $(S, \mathcal{L})$  (no necesariamente contenida en  $\mathcal{P}$ ).

Puede mostrarse que si  $\mathcal{P}$  es una familia relativamente compacta definida en  $(S, \mathcal{L})$  y  $S$  es separable y completo entonces  $\mathcal{P}$  es *tensa* (Billingsley, 1968).

Este último resultado, es la razón de que se use el espacio métrico separable  $(D, d_0)$ , pues en éste, basta demostrar la convergencia de las demostraciones finito dimensionales y que la familia de medidas de probabilidad asociadas sea relativamente compacta, para tener la convergencia débil.

Regresando al problema simple de Bondad de Ajuste, se tiene en resumen que, para probar la hipótesis  $H_0 : F(x) = F_0(x|\theta)$  con  $\theta$  conocida, primero hay que transformar la muestra ordenada  $\{X_1, X_2, \dots, X_n\}$ , mediante la transformación (1), encontrar la distribución del proceso  $\{y_n\}$  y finalmente la de la transformación correspondiente.

La distribución asintótica de  $D_n$  fue encontrada por Kolmogorov (1933) y una derivación de ella puede verse en Durbin (1973a). El resultado de Kolmogorov es

$$P[\sqrt{n}D_n \geq d] \rightarrow 2 \sum_{k=1}^{\infty} e^{-2k^2 d^2}. \quad (14)$$

La obtención de la distribución asintótica de  $\mathcal{W}_n^2$  para cada función  $\psi$  se esboza a continuación:

Si  $\xi_n(u) = \sqrt{\psi(u)y_n(u)}$ , entonces

$$\mathcal{W}_n^2 = \int_0^1 \xi_n^2(u) du$$

y además,

$$E(\xi_n(u)) = 0 \text{ y}$$

$$\kappa(u, u') = \text{cov}(\xi(u), \xi(u')) = \sqrt{\psi(u)\psi(u')}(\min(u, u') - uu'). \quad (15)$$

La idea es representar a  $\mathcal{W}_n^2$  como una suma ponderada de variables aleatorias no correlacionadas con media cero y varianza uno. El problema es análogo a cuando se tiene una suma finita de variables al cuadrado con media cero y cierta matriz de varianzas: diagonalizar la matriz usando los eigenvalores y eigenvectores correspondientes.

En el caso continuo, esto supone resolver la ecuación

$$\int_0^1 \kappa(u, u') \gamma(u') du' = \lambda \gamma(u), \quad (16)$$

donde  $\lambda_j$  y  $\gamma_j(u)$ ,  $j \in \mathbf{N}$  son los eigenvalores y las eigenfunciones del kernel  $\kappa(u, u')$  que cumplen con

$$\int_0^1 \gamma_j^2(u) du = 1 \text{ y } \int_0^1 \gamma_j(u) \gamma_k(u) du = 0 \text{ para } j \neq k.$$

De esta forma, si

$$\zeta_{nj} = \lambda_j^{-1/2} \int_0^1 \gamma_j(u) \xi_n(u) du,$$

entonces

$$\mathcal{W}_n^2 = \sum_{j=1}^{\infty} \lambda_j \zeta_{nj}^2, \quad (17)$$

donde  $\zeta_{n1}, \zeta_{n2}, \dots$  son variables aleatorias no correlacionadas con media cero y varianza unitaria (Durbin, 1973a).

Usando este resultado y como la distribución asintótica de  $\mathcal{W}_n^2$  es la misma que la de

$$\mathcal{W}^2 = \int_0^1 \psi(u) y^2(u) du,$$

donde  $\{y\}$  es el proceso browniano atado, se tiene que

$$\mathcal{W}^2 = \sum_{j=1}^{\infty} \lambda_j z_j^2, \quad (18)$$

con  $\lambda_j$  los eigenvalores del kernel  $\kappa(u, u') = \sqrt{\psi(u)\psi(u')}(\min(u, u') - uu')$  y en donde  $z_j$  son variables normales con media cero y varianza uno, que resultan de resolver (16). De

manera que  $\mathcal{W}^2$  es una serie ponderada de variables aleatorias independientes Ji-cuadrada con un grado de libertad, cuya función característica está dada por

$$\phi(t) = \prod_{j=1}^{\infty} (1 - 2it\lambda_j)^{-1/2}. \quad (19)$$

Así, habrá que invertir  $\phi(t)$  para encontrar los cuantiles de  $\mathcal{W}^2$ .

En general, resolver la ecuación (16) es complicado. En estos casos lo que se hace es calcular algunos cuantiles de la distribución asintótica haciendo una aproximación a (18) e invirtiendo numéricamente la correspondiente aproximación a (19). Específicamente, supóngase que

$$\mathcal{W}^2 \approx \frac{1}{N} \sum_{j=1}^N \psi(u_j) y^2(u_j) \quad u_j \in [0, 1], \quad j \in J_N \text{ y } N \in \mathbb{N}, \quad (20)$$

donde  $J_N = \{1, 2, \dots, n\}$ .

Si  $V$  denota a la matriz cuya entrada  $ij$ -ésima está dada por

$$v_{ij} = \frac{1}{N} \kappa(u_i, u_j) \quad i, j \in J_N, \quad (21)$$

con  $\kappa$  definido en (15), una aproximación a (18) está dada por

$$\mathcal{W}^2 \approx \sum_{j=1}^N \lambda_j \nu_j,$$

donde  $\lambda_j$ ,  $j \in J_N$  son los eigenvalores de la matriz (21) y  $\nu_j$  son variables aleatorias Ji-cuadrada independientes con un grado de libertad definidas para cada  $j$  como

$$\nu_j = \lambda_j^{-1/2} \sum_{k=1}^N \zeta_{jk} y(u_j),$$

con  $(\zeta_{j1}, \zeta_{j2}, \dots, \zeta_{jN})'$  el eigenvector asociado a  $\lambda_j$ .

Con esta aproximación, la expresión (19) puede aproximarse como

$$\phi(t) \approx \prod_{j=1}^N (1 - 2it\lambda_j)^{-1/2}. \quad (22)$$

Esta función característica puede invertirse numéricamente usando el siguiente resultado debido a Imhof (1961),

$$P[\mathcal{W}^2 > w] = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{sen}h(v)}{vg(v)} dv \quad (23)$$

donde

$$h(v) = \frac{1}{2} \left\{ \sum_{i=1}^N \tan^{-1} \lambda_i v - wv \right\} \text{ y}$$

$$g(v) = \prod_{i=1}^N (1 + \lambda_i^2 v^2)^{1/4}.$$

En resumen, para encontrar algún cuantil de la distribución asintótica de  $\mathcal{W}_n^2$  o equivalentemente, la distribución de  $\mathcal{W}^2$ , se calcula la matriz  $V$  definida en (21) para algún valor de  $N$ , se encuentran los eigenvalores y eigenvectores ortonormalizados asociados a  $V$  y se resuelve numéricamente la integral (23).

De esta manera, el problema simple de Bondad de Ajuste queda resuelto.

Sin embargo, el valor de  $\theta$  es desconocido en la mayoría de las aplicaciones de Bondad de Ajuste. En este caso la transformación (1) no puede aplicarse, pues la función de distribución no está totalmente especificada. Existen varias soluciones a este problema, en el fondo todas llevan a estimar a  $F_0(x|\theta)$ , ya sea estimando  $\theta$  y usar este estimador para transformar la muestra, o bien, estimando toda la función de distribución.

Supóngase entonces, que se tiene una muestra aleatoria  $\{X_1, X_2, \dots, X_n\}$  y que se quiere probar la hipótesis

$$H_0 : F(x) = F_0(x|\theta) \text{ con } \theta \text{ desconocido}$$

Sea  $\hat{\theta}_n = \hat{\theta}(X_1, X_2, \dots, X_n)$  un estimador de  $\theta$ . Siguiendo la lógica del caso  $\theta$  conocido, considérese la transformación

$$\hat{U}_i = \hat{U}(X_i) = F_0(X_i|\hat{\theta}) = \hat{F}_n(X_i) \text{ para toda } i \in J_n.$$

Si  $G_n$  denota a la distribución empírica de la muestra  $\{\hat{U}_1, \hat{U}_2, \dots, \hat{U}_n\}$ , el proceso empírico estimado asociado es

$$\hat{y}_n(\hat{u}) = \sqrt{n}(G_n(\hat{u}) - \hat{u}), \quad (24)$$

para el que habrá que encontrar su distribución asintótica.

Bajo ciertas condiciones, similares a las pedidas por Cramér (1946) en el desarrollo del método de máxima verosimilitud y puesto que  $\hat{u} \rightarrow u$  con  $u = F_0(x|\theta)$ , Durbin(1973b) demostró que  $\{\hat{y}_n(\hat{u})\} \Rightarrow \{\hat{y}(u)\}$  donde  $\{\hat{y}(u)\}$  es un proceso Gaussiano con media cero y función de covarianza dada por

$$E(\hat{y}(u), \hat{y}(u')) = \min(u, u') - uu' - h(u)'I^{-1}h(u') \text{ con } u, u' \in [0, 1] \quad (25)$$

y en donde  $h(u)$  es el vector cuya  $j$ -ésima entrada es  $\frac{\partial}{\partial \theta_j} F(x|\theta)$  cuando  $u = F(x|\theta)$  e  $I$  denota a la matriz de información de Fisher por unidad muestral.

En general, el estimador  $\hat{\theta}_n$  que se utiliza para generar (24) es el de máxima verosimilitud, pero puede ser cualquiera que cumpla con

$$\sqrt{n}(\hat{\theta}_n - \theta) = \frac{1}{\sqrt{n}} I^{-1} \sum_{j=1}^n \frac{\partial \log p(x_j|\theta)}{\partial \theta} + \varepsilon_n,$$

donde  $p(x|\theta)$  denota a la función de densidad y  $\varepsilon_n \rightarrow 0$  en probabilidad (Serfling, 1980).

Considerando las transformaciones  $g_1$  y  $g_2$  definidas en (11) se tiene que

$$\begin{aligned} \hat{W}_n^2 &= \int_0^1 \hat{y}_n^2(\hat{u})\psi(\hat{u})d\hat{u} \Rightarrow \int_0^1 \hat{y}(u)\psi(u)du \\ \hat{D}_n &= \sup_{0 \leq \hat{u} \leq 1} \sqrt{n}|\hat{y}_n(\hat{u})| \Rightarrow \sup_{0 \leq u \leq 1} |\hat{y}(u)|. \end{aligned} \quad (26)$$

Lilliefors (1967, 1969) encontró mediante simulación, la distribución asintótica de  $\hat{D}_n$  para el caso Normal con ambos parámetros desconocidos y para el caso Exponencial;

Durbin (1975) da la distribución asintótica para el caso Exponencial. Stephens (1986, y referencias incluidas), da los cuantiles asintóticos para algunas distribuciones, tanto de  $\hat{D}_n$  como de  $\hat{W}_n^2$ , para algunas formas específicas de la función  $\psi$ . Sin embargo una teoría asintótica general para  $\hat{D}_n$  no está dada (D'Agostino y Stephens, 1986).

Dos miembros de la familia Cramér-Von-Mises son de especial interés: la estadística Von-Mises, que se obtiene haciendo  $\psi(u) = 1$ ,  $u \in [0, 1]$  y la  $A^2$  de Anderson-Darling (Anderson y Darling, 1952, 1954), haciendo  $\psi(u) = [u(1-u)]^{-1}$ ,  $u \in [0, 1]$ . Ambas han sido ampliamente estudiadas y existe una teoría asintótica tanto para el caso de parámetros conocidos como cuando los parámetros se desconocen.

La prueba basada en la estadística  $A^2$  ha demostrado buena potencia, cuando los parámetros se estiman por máxima verosimilitud (Stephens, 1974).

Obsérvese que la función  $\psi(u) = [u(1-u)]^{-1}$  no es continua en  $\{0, 1\}$ , por lo que el método de prueba sobre la convergencia débil esbozado anteriormente, no puede ser aplicado. Este es un caso en el que el concepto de *tensión* es útil para encontrar la distribución asintótica del proceso empírico (véase Durbin, 1973b).

Nótese que la función de covarianza (25) es la del caso simple menos una función que depende de la forma de la función de distribución y, posiblemente, del parámetro. David y Johnson (1948) demostraron que la distribución de  $\hat{F}_n(x)$  no depende del parámetro si  $\theta$  es un parámetro de escala y/o localización, por lo tanto la distribución finita del proceso no depende de  $\theta$ , así que la distribución asintótica tampoco dependerá de  $\theta$  en estos casos. Por lo tanto, la generación de tablas con los cuantiles de la distribución asintótica resulta más sencilla, pues sólo dependerá de la forma de la función (véase Stephens, 1986). Esto es válido, tanto para la distribución asintótica de  $\hat{D}_n$  como para la de  $\hat{W}_n^2$ , incluyendo a la estadística  $\hat{A}_n^2$ .

Otro procedimiento de estimación es el que utiliza al estimador de Rao-Blackwell de la función de distribución, y fué introducido por Srinivasan (1970). Recuérdese que el estimador de Rao-Blackwell de la función de distribución es

$$\hat{F}_n(x) = E \left[ I_{[X_i \leq x]} | T_n \right], \quad (27)$$

donde  $T_n$  es una estadística suficiente minimal y  $X_i$  es cualquier elemento de la muestra. Srinivasan (1970) considera al proceso

$$\sqrt{n}\{F_n(x) - \hat{F}_n(x)\}, x \in \mathbf{R} \quad (28)$$

y encuentra, usando métodos de Monte Carlo, la distribución asintótica para el caso Normal y Exponencial, demostrando que estas distribuciones no dependen de  $\theta$ .

El proceso (28) puede escribirse como

$$\tilde{y}_n(\tilde{u}) = \sqrt{n}(G_n(\tilde{u}) - \tilde{u}), \quad (29)$$

donde  $\tilde{u} = \tilde{u}(x) = \hat{F}_n(x)$  y  $G_n$  denota ahora a la función de distribución empírica de la muestra transformada  $\{\tilde{U}_1, \tilde{U}_2, \dots, \tilde{U}_n\}$ .

El resultado que relaciona a las distribuciones asintóticas de los procesos empíricos estimados  $\{\hat{y}_n(\hat{u})\}$  y  $\{\tilde{y}_n(\tilde{u})\}$ , definidos en (24) y (29) respectivamente, fue establecido por Moore (1971) y es el siguiente:

Como

$$\sqrt{n}\{F_n(x) - \hat{F}_n(x)\} = \sqrt{n}\{F_n(x) - \tilde{F}_n(x)\} + \sqrt{n}\{\tilde{F}_n(x) - \hat{F}_n(x)\} \quad (30)$$

si

$$\sup |\tilde{F}_n(x) - \hat{F}_n(x)| \xrightarrow{P} 0 \quad (31)$$

entonces las distribuciones asintóticas de los procesos estimados  $\{\hat{y}_n(\hat{u})\}$  y  $\{\tilde{y}_n(\tilde{u})\}$  serán iguales.

Moore (1971) demostró que esto ocurría en los casos Normal, Exponencial y Uniforme con parámetro  $\theta$ . Este último caso es un ejemplo de que la propiedad (31) no es característica de la familia exponencial. Incluso Moore (1971) conjetura que (31) debe ser cierto en general, pero no exhibe una demostración.

Otro procedimiento de estimación de la función de distribución ha sido propuesto por O'Reilly y Villegas (1987). Estos autores proponen el uso de la distribución predictiva, que aparece en el contexto bayesiano, como un estimador de la función de distribución y demuestran que bajo ciertas condiciones, se cumple que

$$\tilde{F}(X_n|T_n) = F^\pi(X_n|T_{n-1}), \quad (32)$$

donde  $\pi$  es la medida de Haar invariante por la derecha y  $F^\pi$  la distribución predictiva obtenida a partir de  $\pi$ .

Así, si se considera al proceso empírico estimado

$$y_n^\pi(u) = \sqrt{n}\{G_n(u) - u\}, \quad (33)$$

donde  $G_n$  es ahora la función de distribución empírica de la muestra  $\{U_1^\pi, U_2^\pi, \dots, U_n^\pi\}$ , donde

$$U_i^\pi = U^\pi(X_i) = F^\pi(X_i|T_{(n-1)}) \quad (34)$$

y  $T_{(n-1)}$  la estadística suficiente minimal construida con las  $n-1$  observaciones restantes, se tendrá que la distribución asintótica es la misma que la del proceso  $\{\tilde{y}_n\}$ .

En el caso de que  $\theta$  sea un parámetro de localización y/o escala, la relación (34) se cumple y entonces las distribuciones asintóticas de los tres procesos empíricos estimados coinciden.

Es importante mencionar que la transformación  $\hat{U}$  no genera variables aleatorias Uniformes ni independientes (David y Johnson, 1948) y que la transformación  $\tilde{U}$  si genera variables Uniformes, aunque no independientes, (en O'Reilly y Quesenberry, 1973, se presenta un procedimiento basado en el estimador de Rao-Blackwell que transforma variables aleatorias a subconjuntos de variables aleatorias Uniformes e independientes), por lo que la transformación  $U^\pi$  también genera variables aleatorias Uniformes no independientes, si se cumple (34).

No existen condiciones necesarias y suficientes para saber cuándo los tres procesos tienen la misma distribución asintótica; en especial qué ocurre si el parámetro no es de localización ni de escala, como en el caso de la distribución Gaussiana Inversa.

En este trabajo se estudiará el problema de bondad de ajuste para la Gaussiana Inversa cuando se tienen que estimar los parámetros y se considerará únicamente a la estadística  $A^2$  de Anderson - Darling, debido a sus buenas propiedades de potencia reportadas en la literatura. En especial se tratará el proceso  $\{y_n^\pi\}$ .

En el siguiente capítulo se dará una revisión de las soluciones más relevantes reportadas en la literatura, se introducirá a la distribución predictiva y las propuestas de distribuciones *no informativas* que se han hecho sobre los parámetros de la distribución Gaussiana Inversa y se propone una familia de distribuciones iniciales *no informativas* que contiene a los casos más comunes.

En el capítulo 3 se encuentran las distribuciones finales asociadas a la familia propuesta y las correspondientes distribuciones predictivas. La obtención de estas últimas distribuciones no puede darse analíticamente, por lo que deben darse soluciones numéricas. Esto se hace en el capítulo 4, junto con un amplio estudio numérico sobre la sensibilidad de las soluciones encontradas, con respecto a los valores de las transformadas y de las estadísticas asociadas. Finalmente, en el capítulo 5 se dan algunos comentarios generales y conclusiones.

## CAPITULO 2

### I. INTRODUCCION

Como se mencionó en el capítulo anterior, el objetivo de este trabajo es estudiar el problema del ajuste de la distribución Gaussiana Inversa, usando a la distribución predictiva como estimador de la función de distribución. En este capítulo se introducirá a la distribución Gaussiana Inversa y algunas propiedades relevantes, y se describen dos soluciones ya propuestas al problema de Bondad de Ajuste para la distribución Gaussiana Inversa. Por otro lado, se verán los conceptos necesarios para la obtención de la distribución predictiva, y así obtener una propuesta adicional al problema de Bondad de Ajuste. En especial, se hará énfasis en el uso de *distribuciones finales de referencia* y de algunos métodos para obtenerlas. Finalmente, se hará una propuesta de una familia de distribuciones iniciales de referencia, que contiene como casos particulares a las distribuciones de referencia más comunes.

### II. LA DISTRIBUCION GAUSSIANA INVERSA

Se dice que la variable aleatoria  $X$  tiene una distribución Gaussiana Inversa, si su densidad es

$$IG(x|\mu, \lambda) = p(x|\mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp \left\{ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right\} \quad x, \mu, \lambda \in \mathbf{R}^+. \quad (1)$$

Existen diferentes parametrizaciones de esta densidad, sin embargo la descrita en (1) es la usual. Puede observarse que  $\lambda$  es un parámetro de escala, pero que  $\mu$  no es ni de escala ni de localización. Puede mostrarse que

$$E(X) = \mu \quad \text{y} \quad V(X) = \frac{\mu^3}{\lambda},$$

por lo que si  $\phi = \frac{\lambda}{\mu}$ , el coeficiente de variación puede escribirse como  $\sqrt{\phi^{-1}}$ . La reparametrización de (1) en términos de  $(\mu, \phi)$  o de  $(\lambda, \phi)$  tiene un interés especial, como se verá más adelante.

La función de distribución asociada a (1) está dada por (Chhikara y Folks, 1989)

$$F(x|\mu, \lambda) = \Phi \left[ \sqrt{\frac{\lambda}{x}} \left( \frac{x}{\mu} - 1 \right) \right] + e^{2\lambda/\mu} \Phi \left[ -\sqrt{\frac{\lambda}{x}} \left( 1 + \frac{x}{\mu} \right) \right], \quad (2)$$

donde  $\Phi$  denota a la función de distribución Normal estándar.

Si  $Z = \{X_1, X_2, \dots, X_n\}$  es una muestra aleatoria de una  $IG(x|\mu, \lambda)$ , la función de verosimilitud puede escribirse como

$$p(z|\mu, \lambda) = \left( \frac{\lambda}{2\pi} \right)^{\frac{n}{2}} \prod_{i=1}^n x_i^{-\frac{3}{2}} \exp \left\{ \frac{n\lambda}{\mu} \right\} \exp \left\{ -\frac{n\bar{x}\lambda}{2\mu^2} - \frac{n\bar{x}_h\lambda}{2} \right\}, \quad (3)$$

con  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  y  $\bar{x}_h = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}$ .

Entonces

$$\log p(z|\mu, \lambda) = -\frac{n}{2} \log(2\pi) + \frac{n}{2} \log \lambda - \frac{3}{2} \sum_{i=1}^n \log x_i + \frac{n\lambda}{\mu} - \frac{n\bar{x}\lambda}{2\mu^2} - \frac{n\bar{x}_h\lambda}{2},$$

de donde puede mostrarse que los estimadores máximo verosímiles de  $\mu$  y  $\lambda$  son

$$\hat{\mu} = \bar{x} \quad \text{y} \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^n \left( \frac{1}{x_i} - \frac{1}{\bar{x}} \right)}. \quad (4)$$

Chhikara y Folks (1989), presentan un procedimiento de Bondad de Ajuste basado en la estadística  $D_n$  de Kolmogorov-Smirnov. Como no existe una teoría asintótica general para encontrar los cuantiles de la estadística  $D_n$  cuando los parámetros son estimados, estos autores calculan, mediante simulación, la distribución de  $D_n$  para distintos valores de  $n$  y de  $\phi$ , utilizando los estimadores máximo verosímiles en la expresión de la función de distribución, y hacen una aproximación para poder extender sus resultados a cualquier  $n$  y cualquier  $\phi$ . Sin embargo las tablas que presentan, no muestran una estabilidad en el comportamiento de los cuantiles a medida que  $n$  crece, además de que hay un error en la expresión que presentan para extender los cuantiles a cualesquiera valores de  $n$  y  $\phi$ . Por otro lado, la potencia de la prueba de Kolmogorov - Smirnov es pobre (Stephens, 1974).

Otro procedimiento de prueba, fue desarrollado por O'Reilly y Rueda (1992), basado en la estadística  $A^2$  de Anderson-Darling y utilizando al estimador de Rao-Blackwell. Primero exhiben algunos cuantiles de la distribución asintótica de  $A^2$  cuando los parámetros se estiman por máxima verosimilitud. Esta distribución depende de los valores de  $\mu$  y  $\lambda$  sólo a través de  $\phi$ , lo que facilita la obtención de los cuantiles asintóticos usando el procedimiento descrito en el capítulo anterior. Por otro lado, el estimador de Rao-Blackwell de la función de distribución está dado por (Chhikara y Folks, 1989)

$$\tilde{F}_n(x) = G_{n-2}(W) + \frac{n-2}{n} \left[ 1 + \frac{4(n-1)\hat{\lambda}}{n^2\hat{\mu}} \right]^{\frac{n-3}{2}} G_{n-2}(-W'), \quad (5)$$

donde  $G_{n-2}$  denota a la distribución Student con  $n-2$  grados de libertad y

$$W = W(x; \hat{\mu}, \hat{\lambda}) = \frac{1}{C} \sqrt{n(n-2)} \left( \frac{x}{\hat{\mu}} - 1 \right),$$

$$W' = W'(x; \hat{\mu}, \hat{\lambda}) = \frac{1}{C} \sqrt{n(n-2)} \left( 1 + \frac{n-2}{n} \frac{x}{\hat{\mu}} \right),$$

con

$$C = \left\{ \frac{n}{\hat{\lambda}} \left( n - \frac{x}{\hat{\mu}} \right) x - n \left( 1 - \frac{x}{\hat{\mu}} \right)^2 \right\}^{\frac{1}{2}}.$$

Usando este estimador, O'Reilly y Rueda (1992) demuestran que

$$\sqrt{n} \sup_x \{ \tilde{F}_n(x) - \hat{F}_n(x) \} \rightarrow 0 \quad (6)$$

con probabilidad uno, mostrando así la conjetura de Moore (1973) para esta distribución.

De este resultado se sigue, que los procesos  $\{\hat{y}_n(\hat{u})\}$  y  $\{\tilde{y}_n(\hat{u})\}$  convergen al mismo proceso Gaussiano  $\{y(u)\}$ , con media cero y función de covarianza dada por

$$\rho(s, t) = \min(s, t) - st - h'(t) \mathcal{I}^{-1} h(s),$$

con  $\mathcal{I}$  la matriz de información de Fisher por unidad muestral y  $h'(t)$  el vector

$$h'(t) = \left( \frac{\partial F(x|\mu, \lambda)}{\partial \mu}, \frac{\partial F(x|\mu, \lambda)}{\partial \lambda} \right) \text{ evaluado en } t = F(x|\mu, \lambda). \quad (7)$$

Puede demostrarse que la expresión (7) depende de  $\mu$  y  $\lambda$  sólo a través de  $\phi$  (O'Reilly y Rueda, 1992).

Por otro lado, y en referencia al procedimiento propuesto por Chhikara y Folks (1989) comentado en párrafos anteriores, O'Reilly y Rueda (1992) analizaron cuatro conjuntos de datos presentados por Folks y Chhikara (1978) y obtuvieron conclusiones, en general, distintas a las obtenidas con una prueba Kolmogorov - Smirnov, en la que no se describe que tablas se usaron. Las tablas que aparecen en Chhikara y Folks (1989) fueron obtenidas por Stark y Chhikara (1988), por lo que no pudieron ser usadas en el trabajo de Folks y Chhikara (1978).

En resumen, O'Reilly y Rueda (1992) demostraron que las distribuciones asintóticas de la estadística  $A^2$  cuando se estiman los parámetros por máxima verosimilitud y la correspondiente a cuando se estima la función de distribución con el estimador de Rao - Blackwell, son iguales.

Es natural preguntarse si en este caso, el uso de la distribución predictiva como estimador de la función de distribución y usando a la estadística  $A^2$ , produciría la misma distribución asintótica. En la siguiente sección se darán los elementos necesarios para tratar este problema.

### III. ANALISIS BAYESIANO EN LA GAUSSIANA INVERSA

Desde la perspectiva bayesiana, puesto que el parámetro  $\theta = (\mu, \lambda)$  es desconocido, debe asignarse una distribución de probabilidad sobre  $\Theta$ , el espacio paramétrico, llamada *distribución inicial*, que exprese el conocimiento que se tiene sobre  $\theta$  y, en caso de tener una muestra aleatoria  $Z = \{X_1, X_2, \dots, X_n\}$  de  $IG(x|\mu, \lambda)$ , combinar esta información con la distribución inicial, mediante el teorema de Bayes, para actualizar el conocimiento sobre  $\theta$ . Es decir, si  $p(\theta)$  denota a la distribución inicial definida sobre  $\Theta$  y  $p(z|\theta)$  a la función de verosimilitud de  $\theta$  dada la muestra  $z$ , la distribución final,  $p(\theta|z)$ , se obtiene como

$$p(\theta|z) \propto p(\theta)p(z|\theta), \quad (8)$$

donde la constante de proporcionalidad es

$$p(z) = \int_{\Theta} p(\theta)p(z|\theta)d\theta. \quad (9)$$

A la distribución marginal de  $z$ , se le llama *distribución predictiva*. La razón del nombre queda más claro si se considera a  $X^*$ , una nueva observación generada por  $p(x|\theta)$ , que sea independiente de las  $n$  anteriores, dado  $\theta$ . Usando (8) y (9) se tendrá que la nueva distribución final es

$$p(\theta|z, x^*) = \frac{p(\theta|z)p(x^*|\theta)}{\int p(\theta|z)p(x^*|\theta)d\theta}. \quad (10)$$

En este caso, la distribución predictiva queda como

$$\begin{aligned} \int p(\theta|z)p(x^*|\theta)d\theta &= \int p(\theta|z)p(x^*|\theta, z) \\ &= \int p(\theta, x^*|z)d\theta \\ &= p(x^*|z), \end{aligned} \quad (11)$$

donde la primera igualdad se cumple por la independencia condicional (dado  $\theta$ ) entre  $X^*$  y  $Z$ . La expresión anterior permite hacer predicciones sobre la observación "futura"  $X^*$  dada la información de las  $n$  anteriores.

El problema principal es ¿cómo asignar  $p(\theta)$ ? En principio, la asignación de la distribución inicial depende del conocimiento que se tenga sobre  $\Theta$ . Una procedimiento muy común para asignar distribuciones iniciales, es usando *familias conjugadas*. Se dice que  $\mathcal{P}$  es una familia conjugada de distribuciones para  $\mathcal{F} = \{p(x|\theta) : \theta \in \Theta\}$  si  $p(\theta|z) \in \mathcal{P}$  para toda  $p(\theta) \in \mathcal{P}$  y  $p(x|\theta) \in \mathcal{F}$ . Un excelente tratado de este tema puede encontrarse en Raiffa y Schlaifer (1961).

Uno de los resultados más importantes sobre familias conjugadas involucra a las *familias exponenciales*.

Se dice que  $\mathcal{F}$  es una *familia exponencial regular* (e.g. Barndorff-Nielsen, 1978) si:

- i.  $\mathcal{F} = \{p(x|\theta) = a(\theta)b(x) \exp(\theta' t(x)) : \theta \in \Theta\}$  con  $\Theta \subset \mathbf{R}^k$ ,  $a(\theta)$  una función que sólo depende de  $\theta$  y  $b(x), t(x)$  funciones que sólo dependen de  $x$ .
- ii.  $\Theta = \{\theta \in \Theta : a(\theta) \in \mathbf{R}^+\}$  es un subconjunto abierto no vacío de  $\mathbf{R}^k$ .

Con esta definición, el resultado sobre familias conjugadas puede escribirse como sigue:

Si  $\mathcal{F}$  es una familia exponencial regular, entonces la familia conjugada de  $\mathcal{F}$  está dada por  $\mathcal{P} = \{p(\theta) \propto a^{n_0}(\theta) \exp(\theta' t) : t \in \mathbf{R}^k, n_0 \in \mathbf{R}, \theta \in \Theta\}$ , siempre que

$$\int_{\Theta} a^n(\theta) \exp\left(\sum_{i=1}^n \theta' t(x_i)\right) d\theta < \infty. \quad (12)$$

Nótese que  $\mathcal{P}$  es la familia exponencial correspondiente a  $\theta$ .

En el caso de la distribución Gaussiana Inversa, si  $\alpha = \lambda/\mu^2$ , (1) puede escribirse como

$$p(x|\lambda, \alpha) = \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} e^{\sqrt{\lambda\alpha} x - \frac{3}{2}} \exp\left\{-\frac{\alpha x}{2} - \frac{\lambda}{2x}\right\},$$

de donde se puede ver que es una familia exponencial con  $\theta' = (\lambda, \alpha)$ ,  $t(x) = -\frac{1}{2}(x^{-1}, x)$ ,  $b(x) = x^{-3/2}$  y  $a(\theta) = (\lambda/2\pi)^{1/2} \exp((\lambda\alpha)^{1/2})$ . Sin embargo, no es regular, pues el subconjunto  $\Theta$  de  $\mathbf{R}^2$  en donde  $a(\theta)$  es finito, es  $(0, \infty) \times [0, \infty)$ , que no es abierto. (La distribución que resulta al hacer  $\alpha = 0$  tiene como densidad a una distribución estable con exponente característico igual a  $1/2$  y parámetro de escala igual a  $\lambda^{-1}$ . Véase Barndorff-Nielsen, 1978).

Por otro lado el cumplimiento de la condición (12) depende de la parametrización usada. Por ejemplo, si a  $IG(x|\theta)$  se parametriza con  $\theta = (\mu, \lambda)$  – como en (1) –, la condición (12) no se cumple, pues la integral con respecto a  $\mu$  siempre diverge, para todo  $\lambda$ ; por lo que si  $\mu$  y  $\lambda$  son desconocidos la familia conjugada no existe. Pero si  $\mu$  es conocido, la familia conjugada para  $\lambda$  es una Gamma (Chhikara y Folks, 1989). Si la densidad se parametriza con  $(1/\mu, \lambda)$  entonces la familia conjugada para ambos parámetros es de la forma de una distribución Normal truncada en cero para  $\mu$  condicionado a  $\lambda$ , multiplicada por una distribución Gamma para  $\lambda$  (Banerjee y Bhattacharyya, 1979).

Por otro lado, el objetivo de tener una distribución inicial para  $\theta$  en el problema de Bondad de Ajuste, es comparar la solución que se obtiene al usar la distribución predictiva – que depende de la distribución inicial – como estimador de la función de distribución; por lo que sería deseable introducir una distribución inicial que permita que la distribución final esté basada fundamentalmente en la información muestral – como en el caso “clásico” –. Este tipo de distribuciones se conocen como *distribuciones no informativas* o *distribuciones de referencia*. La idea básica, es generar distribuciones iniciales que

“peturben” lo menos posible a la información obtenida por una muestra.

Existen diferentes procedimientos para obtener distribuciones de referencia, aunque la gran mayoría pueden clasificarse en tres grupos (Rueda, 1980): los basados en alguna caracterización del concepto de *información* (e.g. Bernardo, 1979; Jaynes, 1965; Lindley, 1961 y Zellner, 1977), los que utilizan familias conjugadas (e.g. DeGroot, 1970; Novick, 1969; Novick y Hall, 1965 y Raiffa y Schlaifer, 1961) y los que usan ciertos conceptos de *invarianza* como Jeffreys (1946), Hartigan (1964) y Villegas (1971, 1977) entre otros.

En este trabajo se usará el enfoque desarrollado en Bernardo (1979) y extendido posteriormente en Berger y Bernardo (1992). Hay dos razones principales por las que se seleccionó este procedimiento: la primera es porque da un tratamiento muy específico al caso en que se tienen *parámetros de ruido*, es decir, cuando no todo el parámetro es de interés, sino sólo una parte de él; y segundo, porque se fundamenta en un concepto muy claro de *información*. Este concepto de información fue introducido inicialmente por Lindley (1956), basado en un trabajo de Shannon (1948), y ampliamente estudiado en Bernardo (1976). El objetivo es obtener distribuciones finales que maximicen la cantidad de información contenida en un experimento.

En la siguiente sección se presentará una breve descripción de algunas de las propuestas de distribuciones de referencia existentes y una descripción del procedimiento de Berger y Bernardo (1992). En la sección V se encontrarán las distribuciones referencia para los parámetros de la distribución Gaussiana Inversa, basada en este procedimiento. Finalmente, en la última sección se propone una familia paramétrica de distribuciones iniciales, que contiene como casos particulares a las que se obtienen por el procedimiento de Berger y Bernardo (1992) y otras reportadas en la literatura.

#### IV. DISTRIBUCIONES DE REFERENCIA

En Chhikara y Folks (1989) se presentan algunas distribuciones de referencias basadas en la regla de Jeffreys (1946), para distintas parametrizaciones. Considérese la parametrización (1), es decir,

$$IG(x|\mu, \lambda) = (2\pi x^3)^{1/2} \left[ \lambda^{1/2} \exp \left( -\frac{\lambda x}{2\mu^2} + \frac{\lambda}{\mu} - \frac{\lambda}{2x} \right) \right],$$

entonces puede demostrarse que

$$\begin{aligned}\frac{\partial \log IG(x|\mu, \lambda)}{\partial \lambda} &= \frac{1}{2\lambda} - \frac{(x - \mu)^2}{2x\mu^2}, & \frac{\partial \log IG(x|\mu, \lambda)}{\partial \mu} &= \frac{\lambda(x - \mu)}{\mu^3}, \\ \frac{\partial^2 \log IG(x|\mu, \lambda)}{\partial \lambda^2} &= -\frac{1}{2\lambda^2}, & \frac{\partial^2 \log IG(x|\mu, \lambda)}{\partial \mu^2} &= -\frac{\lambda(3x - 2\mu)}{\mu^4}, \\ \frac{\partial^2 \log IG(x|\mu, \lambda)}{\partial \mu \partial \lambda} &= \frac{x - \mu}{\mu^3}.\end{aligned}$$

Por lo que la matriz de información de Fisher resulta ser

$$\mathcal{I}(\mu, \lambda) = \text{diag}\left(\frac{\lambda}{\mu^3}, \frac{1}{2\lambda^2}\right). \quad (13)$$

Puesto que la regla de Jeffreys establece que la *distribución inicial no informativa*, que se denotará con  $p_j$ , es proporcional a  $\{\det \mathcal{I}(\mu, \lambda)\}^{1/2}$ , se tiene que

$$p_j(\mu, \lambda) \propto (\lambda \mu^3)^{-1/2}, \quad (14)$$

o bien, si se usa la *regla modificada* de Jeffreys, es decir,  $p_{jm}(\mu, \lambda) = p_j(\lambda)p_j(\mu|\lambda)$ , se obtiene que

$$p_{jm}(\mu, \lambda) \propto \lambda^{-1} \mu^{-3/2}. \quad (15)$$

En cualquiera de los dos casos, es fácil obtener la distribución de Jeffreys en el caso de que uno de los dos parámetros sea conocido.

En el siguiente capítulo se demostrará que estas distribuciones iniciales producen distribuciones finales propias.

Ahora bien, si se considera la reparametrización  $(\eta, \lambda)$  con  $\eta = \mu^{-1}$ , es fácilmente demostrable que la matriz de información de Fisher es en este caso

$$\mathcal{I}(\eta, \lambda) = \text{diag}\left(\frac{\lambda}{\eta}, \frac{1}{2\lambda^2}\right),$$

por lo que las correspondientes distribuciones de Jeffreys son

$$p_j(\eta, \lambda) \propto (\eta \lambda)^{-1/2} \quad \text{y} \quad p_{jm}(\eta, \lambda) \propto \eta^{-1/2} \lambda^{-1}. \quad (16)$$

De acuerdo a Banerjee y Bhattacharyya (1979), la primera distribución no produce una distribución final propia. Dada esta limitación, proponen como distribución inicial no informativa a

$$p_{bb}(\eta, \lambda) \propto \lambda^{-1}. \quad (17)$$

Esta distribución no tiene mucho sustento, pues supone que  $p_{bb}(\eta|\lambda) \propto k$ , con  $k$  una constante positiva. Esta distribución no cumple ningún criterio de invarianza, pues  $\mu \in \mathbb{R}^+$  (veáse Jeffreys, 1946).

Berger y Bernardo (1992) proponen un método general para obtener distribuciones de referencia, basado en una propuesta de Bernardo (1979). A continuación se da una descripción de dicho método, manteniendo la notación original.

Supóngase que  $\Theta \subset \mathbb{R}^k$  y que las  $\theta_i$ 's puede separarse en  $m$  grupos de tamaños  $n_1, n_2, \dots, n_m$  como sigue

$$\begin{aligned} \theta_{(1)} &= (\theta_1, \theta_2, \dots, \theta_{n_1}) \\ \theta_{(2)} &= (\theta_{n_1+1}, \dots, \theta_{n_1+n_2}) \\ &\dots \\ \theta_{(i)} &= (\theta_{N_{i-1}+1}, \dots, \theta_{N_i}) \\ \theta_{(m)} &= (\theta_{N_{m-1}+1}, \dots, \theta_k) \end{aligned}$$

donde  $N_j = \sum_{i=1}^j n_i$ . Defínase a

$$\begin{aligned} \theta_{[j]} &= (\theta_{(1)}, \dots, \theta_{(j)}) = (\theta_1, \dots, \theta_{N_j}), \\ \theta_{[\sim j]} &= (\theta_{(j+1)}, \dots, \theta_{(m)}) = (\theta_{N_j+1}, \dots, \theta_k), \end{aligned}$$

con  $\theta_{[\sim 0]} = \theta_{[m]} = \theta$  y  $\theta_{[0]}, \theta_{[\sim m]}$  vacuos. Denótese a la *distancia de Kullback - Liebler* entre dos densidades  $p_1$  y  $p_2$  por

$$D(p_1, p_2) = \int_{\Theta} p_1(\theta) \log \frac{p_1(\theta)}{p_2(\theta)} d\theta. \quad (18)$$

El método general puede describirse en cuatro pasos:

**Primer paso.** Selecciónese una sucesión anidada  $\{\Theta^l\}$  de subconjuntos compactos de  $\Theta$  tal que  $\bigcup_{l=1}^{\infty} \Theta^l = \Theta$ .

**Segundo paso.** Divídase a  $\theta$  en  $m$  grupos  $\theta_{(1)}, \theta_{(2)}, \dots, \theta_{(m)}$ .

**Tercer paso.** Para  $j = m, m-1, \dots, 1$  calcúlese iterativamente a las densidades

$$\pi_j^l(\theta_{[\sim(j-1)]}|\theta_{[j-1]}) \propto \pi_{j+1}^l(\theta_{[\sim j]}|\theta_{[j]})h_j^l(\theta_{(j)}|\theta_{[j-1]}),$$

donde  $\pi_{m+1}^l = 1$  y  $h_j^l$  se calcula a partir de

**3a.** Defínase a  $p_t(\theta_{(j)}|\theta_{[j-1]})$  como

$$p_t(\theta_{(j)}|\theta_{[j-1]}) \propto \exp \left\{ \int p(z_t|\theta_{[j]}) \log p(\theta_{(j)}|z_t, \theta_{[j-1]}) dz_t \right\}, \quad (19)$$

donde

$$\begin{aligned} p(z_t|\theta_{[j]}) &= \int p(z_t|\theta) \pi_{j+1}^l(\theta_{[\sim j]}|\theta_{[j]}) d\theta_{[\sim j]}, \\ p(\theta_{(j)}|z_t, \theta_{[j-1]}) &\propto p(z_t|\theta_{(j)}|\theta_{[j-1]}). \end{aligned} \quad (20)$$

**3b.** Suponiendo que el límite existe, defínase a

$$h_j^l(\theta_{(j)}|\theta_{[j-1]}) = \lim_{l \rightarrow \infty} p_t(\theta_{(j)}|\theta_{[j-1]})$$

**Cuarto paso.** Defínase a una distribución inicial de referencia,  $\pi(\theta)$ , como cualquier inicial que cumple con

$$E_l^X D(\pi_1^l(\theta|X), \pi(\theta|X)) \rightarrow 0 \text{ cuando } l \rightarrow \infty \quad (21)$$

en donde  $E_l^X$  es el valor esperado con respecto a

$$p_l(x) = \int_{\Theta} p(x|\theta) \pi_1^l(\theta) d\theta$$

y  $D$  definida en (18).

Nótese que en el tercer paso se define a  $p_i$  en forma implícita, pues la densidad que aparece en el término del lado derecho también depende de  $p_i$  (véase (20)). Esto significa en la práctica que la obtención de la distribución inicial de referencia se complica. Sin embargo, en el caso regular, es decir cuando la verosimilitud  $p(z_i|\theta)$  es asintóticamente Normal, el tercer paso puede hacerse explícitamente.

Sea  $S(\theta) = \mathcal{I}^{-1}$ , con  $\mathcal{I}$  la matriz de información de Fisher. Supóngase que  $S = S(\theta)$  puede escribirse como

$$S = \begin{pmatrix} A_{11} & A'_{21} & \dots & A'_{m1} \\ A_{21} & A_{22} & \dots & A'_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{pmatrix}$$

en donde la dimensión de  $A_{ij}$  es  $n_i \times n_j$ . Defínase a  $S_j$  como la matriz de dimensión  $N_j \times N_j$  de la esquina superior izquierda de  $S$ , con  $S_m = S$  y  $H_j = S_j^{-1}$ . Las matrices  $h_j$  de  $n_j \times n_j$  de la esquina inferior derecha de  $H_j$  son relevantes. En particular se tiene que,  $h_1 = H_1 = A_{11}^{-1}$  y si  $S$  es diagonal a bloques (i.e.  $A_{ij} = 0$  para toda  $i \neq j$ ) entonces  $h_j = A_{jj}^{-1}$ .

Finalmente, sea  $\Theta^* \subset \Theta$ , y para cada  $\theta_{[j]}$  defínase a

$$\Theta^*(\theta_{[j]}) = \{\theta_{(j+1)} : (\theta_{[j]}, \theta_{(j+1)}, \theta_{[\sim(j+1)]}) \in \Theta^* \text{ para algún } \theta_{[\sim(j+1)]}\}$$

Con estos elementos, el tercer paso del algoritmo puede reescribirse como sigue

Tercer paso. Defínase a

$$\pi_m^l(\theta_{[\sim(m-1)]}|\theta_{[m-1]}) = \pi_m^l(\theta_{(m)}|\theta_{[m-1]}) = \frac{|h_m(\theta)|^{1/2} I_{\Theta^l(\theta_{[m-1]})}(\theta_{(m)})}{\int_{\Theta^l(\theta_{[m-1]})} |h_m(\theta)|^{1/2} d\theta_{(m)}} \quad (22)$$

Para  $j = m-1, m-2, \dots, 1$  defínase a

$$\pi_j^1(\theta_{[\sim(j-1)]}|\theta_{[j-1]}) = \frac{\pi_{j+1}^1(\theta_{[\sim j]}|\theta_{[j]}) \exp\{\frac{1}{2}E_j^1[(\log|h_j(\theta)|)|\theta_{[j]}]I_{\Theta^1(\theta_{[j-1]})}(\theta_{[j]})\}}{\int_{\Theta^1(\theta_{[j-1]})} \exp\{\frac{1}{2}E_j^1[(\log|h_j(\theta)|)|\theta_{[j]}]d\theta_{[j]}\}}, \quad (23)$$

donde

$$E_j^1[g(\theta)|\theta_{[j]}] = \int_{\{\theta_{[\sim j]}: (\theta_{[j]}, \theta_{[\sim j]}) \in \Theta^1\}} g(\theta) \pi_{j+1}^1(\theta_{[\sim j]}|\theta_{[j]}) d\theta_{[\sim j]}. \quad (24)$$

Algunos comentarios son pertinentes. El argumento central de este método es encontrar una distribución inicial,  $\pi(\theta)$  que maximice la cantidad de información que proporciona un experimento sobre  $\theta$ . La caracterización de *cantidad de información* que se utiliza es debida a Lindley (1956) y extendida por Bernardo (1976), y se define para una muestra  $Z_t = \{X_1, X_2, \dots, X_n\}$  como

$$I_t^\theta = \int \int p(z_t|\theta) \pi(\theta) \log \frac{\pi(\theta|z_t)}{\pi(\theta)} d\theta dz_t. \quad (25)$$

El objetivo es maximizar (25) y pasar al límite en  $t$  para encontrar  $\pi(\theta)$ . Usando un argumento variacional se puede demostrar, bajo condiciones muy generales (Bernardo, 1979), que

$$\pi_t(\theta) \propto \exp \left\{ \int p(z_t|\theta) \log \pi(\theta|z_t) dz_t \right\}. \quad (26)$$

La expresión (19) es una extensión al caso multivariado de esta ecuación.

En el caso univariado y bajo condiciones que garanticen la convergencia de la distribución final a la distribución Normal (Walker, 1969), puede demostrarse que este procedimiento reproduce a la *Regla de Jeffreys*.

El orden en que se forman los grupos de  $\theta$ 's puede ser relevante, así como la construcción de la sucesión de compactos. Berger y Bernardo (1992) recomiendan que, a menos que se tenga una idea muy clara sobre el papel que juegan los parámetros, cada grupo contenga sólo un elemento y el orden en que se arreglen tome en cuenta cuáles parámetros son de

ruido. Para la selección de la sucesión recomiendan que se tomen intervalos cerrados en  $\mathbf{R}$  o sus equivalentes en  $\mathbf{R}^k$ .

## V. DISTRIBUCIONES DE REFERENCIA EN LA GAUSSIANA INVERSA

Usando el procedimiento descrito en la sección pasada, se encontrarán las distribuciones de referencia para los parámetros de la distribución Gaussiana Inversa, para tres parametrizaciones: la usual,  $\theta = (\mu, \lambda)$  y, haciendo  $\phi = \lambda/\mu$ , considerando a  $\theta = (\phi, \lambda)$  y  $\theta = (\phi, \mu)$ . Para cada parametrización se tratarán tres casos: cuando todo  $\theta$  es de interés y los dos casos en donde uno de los parámetros es de interés y el otro es de ruido.

Las parametrizaciones en términos de  $\phi$  tienen un interés especial, debido a que la distribución asintótica de los procesos estimados por máxima verosimilitud y Rao-Blackwell, depende de los parámetros sólo a través de  $\phi$ . Por otro lado, en la construcción de intervalos de probabilidad para  $\phi$ , Rueda (1988) exhibe una distribución de referencia para  $\phi$ , considerando al otro parámetro como de ruido, que produce resultados muy similares a los intervalos de confianza clásicos.

Usando la notación de la sección anterior, considérese a la siguiente sucesión de subconjuntos de  $\mathbf{R}^+ \times \mathbf{R}^+$ . Para cada  $l \in \mathbf{N}$  sea

$$\Theta^l = \left\{ [a_l, b_l] \times [a_l, b_l] \subset \mathbf{R}^+ \times \mathbf{R}^+ : \{a_l : l \in \mathbf{N}\} \downarrow 0 \text{ y } \{b_l : l \in \mathbf{N}\} \uparrow \infty \right\}.$$

(Por ejemplo,  $a_l = \frac{1}{l}$  y  $b_l = l$  para toda  $l \in \mathbf{N}$ ). Por definición,  $\Theta^l$  es compacto para cada  $l \in \mathbf{N}$  y además  $\bigcup_{l=1}^{\infty} \Theta^l = \Theta$ .

Denótese con  $\pi_\eta(\theta)$  a la distribución inicial de referencia para  $\theta$  cuando  $\eta$  es el parámetro de interés.

La obtención de las distribuciones iniciales de referencia se dan en las siguientes proposiciones. La primera se refiere a la parametrización usual.

**Proposición 1.** Supóngase a la distribución Gaussiana Inversa en la parametrización usual, es decir,  $\theta = (\mu, \lambda)$ . Entonces

$$i. \quad \pi_\theta(\mu, \lambda) \propto (\mu^3 \lambda)^{-1/2},$$

$$\text{ii. } \pi_{\mu}(\mu, \lambda) \propto (\lambda \mu^{3/2})^{-1},$$

$$\text{iii. } \pi_{\lambda}(\mu, \lambda) \propto (\lambda \mu^{3/2})^{-1}.$$

**Demostración.** Como  $\theta = (\mu, \lambda)$ , entonces

$$p(x|\mu, \lambda) = IG(x|\mu, \lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, \quad x, \mu, \lambda \in \mathbf{R}^+$$

De acuerdo a (13), la matriz de información de Fisher está dada por

$$\mathcal{I}(\mu, \lambda) = \text{diag}\left(\frac{\lambda}{\mu^3}, \frac{1}{2\lambda^2}\right).$$

i. Como ambos parámetros son de interés se tiene que  $\theta_{(1)} = \theta_{[1]} = (\mu, \lambda)$ . Así,

$$h_1(\mu, \lambda) = H_1(\mu, \lambda) = S^{-1}(\mu, \lambda) = \mathcal{I}(\mu, \lambda).$$

Usando (22),

$$\pi_1^t(\mu, \lambda) = \frac{\left(\frac{1}{2\mu^3\lambda}\right)^{1/2} I_{\Theta^t}(\theta_{(1)})}{\int_{\Theta^t} (2\mu^3\lambda)^{-1/2} d\theta_{(1)}},$$

y resolviendo la integral, se obtiene

$$\pi_1^t(\mu, \lambda) = \frac{4(b_t^{1/2} - a_t^{1/2})^2}{\sqrt{a_t b_t}} (\mu^3 \lambda)^{-1/2}.$$

**Finalmente**

$$\pi_{\theta}(\mu, \lambda) \propto (\mu^3 \lambda)^{-1/2}.$$

ii. Como  $\lambda$  se considera parámetro de ruido, se tiene ahora que  $\theta_{(1)} = \mu$  y  $\theta_{(2)} = \lambda$ , y entonces  $\theta_{[1]} = \mu$ ,  $\theta_{[\sim 1]} = \lambda$  y  $\theta_{[2]} = (\mu, \lambda)$ . Por otro lado,

$$S_1(\mu, \lambda) = \frac{\mu^3}{\lambda}, S_2(\mu, \lambda) = S(\mu, \lambda) = \text{diag}\left(\frac{\mu^3}{\lambda}, 2\lambda^2\right),$$

de donde

$$H_1(\mu, \lambda) = \mathcal{I}(\mu, \lambda) = \text{diag}\left(\frac{\lambda}{\mu^3}, \frac{1}{2\lambda^2}\right),$$

y, por lo tanto

$$h_1(\mu, \lambda) = \frac{\lambda}{\mu^3} \text{ y } h_2(\mu, \lambda) = \frac{1}{2\lambda^2}.$$

Nuevamente, de (22) se sigue que

$$\begin{aligned} \pi_2^I(\lambda|\mu) &= \frac{(2\lambda^2)^{-1/2} I_{[a_1, b_1]}(\lambda)}{\int_{[a_1, b_1]} (2\lambda^2)^{-1/2} d\lambda} \\ &= \frac{1}{\log b_1 - \log a_1} \lambda I_{[a_1, b_1]}(\lambda). \end{aligned}$$

De acuerdo a (23),

$$\pi_1^I(\mu, \lambda) = \frac{\pi_2^I(\lambda|\mu) \exp\left\{\frac{1}{2} E_1^I\left[\log \frac{\lambda}{\mu^3} | \mu\right]\right\} I_{[a_1, b_1]}(\mu)}{\int_{[a_1, b_1]} \exp\left\{\frac{1}{2} E_1^I\left[\log \frac{\lambda}{\mu^3} | \mu\right]\right\} d\mu},$$

y como

$$\begin{aligned} E_1^I\left[\log \frac{\lambda}{\mu^3} | \mu\right] &= \int_{[a_1, b_1]} \log \frac{\lambda}{\mu^3} \pi_2^I(\lambda|\mu) d\lambda \\ &= K_0 - 3 \log \mu \end{aligned}$$

con  $K_0 = \log b_1 + \log a_1$ , entonces

$$\pi_1^I(\mu, \lambda) = K(\lambda\mu^{3/2})^{-1} I_{[a_1, b_1]}$$

con  $K^{-1} = 4\left(\frac{1}{\sqrt{b_i}} - \frac{1}{\sqrt{a_i}}\right) \log \frac{b_i}{a_i}$ .

Por lo tanto,

$$\pi_{\mu}(\mu, \lambda) \propto (\lambda \mu^{3/2})^{-1}.$$

iii. En este caso,  $\theta_{(1)} = \lambda$  y  $\theta_{(2)} = \mu$  y la demostración es análoga al caso anterior.

La siguiente proposición tienen que ver con la parametrización  $\theta = (\mu, \phi)$ , donde  $\phi = \lambda/\mu$ .

**Proposición 2.** Supóngase que la densidad Gaussiana Inversa es parametrizada con  $\theta = (\mu, \phi)$ , donde  $\phi = \lambda/\mu$ . Entonces

i.  $\pi_{\theta}(\mu, \phi) \propto \mu^{-1} \phi^{-1/2}$ ,

ii.  $\pi_{\mu}(\mu, \phi) \propto (\phi \mu)^{-1}$ ,

iii.  $\pi_{\phi}(\mu, \phi) \propto \mu^{-1} (\phi + 2\phi^2)^{-1}$ .

**Demostración.** En este caso, la densidad Gaussiana Inversa se escribe como

$$IG(x|\mu, \phi) = \left(\frac{\phi \mu}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left\{-\frac{\phi(x-\mu)^2}{2\mu x}\right\}.$$

De aquí, la matriz de información de Fisher es

$$\mathcal{I}(\mu, \phi) = \frac{1}{2} \begin{pmatrix} \frac{1+2\phi}{\mu^2} & \frac{1}{\mu\phi} \\ \frac{1}{\mu\phi} & \frac{1}{\phi^2} \end{pmatrix}.$$

i. Como  $\theta_{(1)} = (\mu, \phi)$ , se tiene que

$$\pi_1^1(\mu, \phi) = \frac{\mu^{-1} \phi^{-1/2} I_{\Theta^1}(\mu, \phi)}{\int_{\Theta^1} \mu^{-1} \phi^{-1/2} d\mu d\phi},$$

por lo que

$$\pi_1^i(\mu, \lambda) = \frac{\mu^{-1} \phi^{-1/2} I_{\Theta^i}(\mu, \phi)}{2(\log b_i - \log a_i)(\sqrt{b_i} - \sqrt{a_i})}.$$

Así que la distribución inicial de referencia, cuando todo  $\theta$  es de interés, es

$$\pi_\theta(\mu, \phi) \propto \mu^{-1} \phi^{-1/2}.$$

Nótese que esta distribución se puede obtener a partir de la especificada en el inciso i de la proposición anterior, haciendo el cambio de variable  $\phi = \lambda/\mu$ .

ii. Como  $\phi$  es parámetro de ruido, se sigue que  $\theta_{(1)} = \mu$  y  $\theta_{(2)} = \phi$ .

Como

$$S(\mu, \phi) = \mathcal{I}^{-1} = \begin{pmatrix} \frac{\mu^2}{\phi} & -\mu \\ -\mu & \phi(1 + 2\phi) \end{pmatrix}, \quad (27)$$

se tiene que

$$\begin{aligned} S_1(\mu, \phi) &= \frac{\mu^2}{\phi} \quad \text{y} \quad S_2(\mu, \phi) = S(\mu, \phi), \\ H_1(\mu, \phi) &= \frac{\phi}{\mu^2} \quad \text{y} \quad H_2(\mu, \phi) = \mathcal{I}(\mu, \phi), \\ h_1(\mu, \phi) &= \frac{\phi}{\mu^2} \quad \text{y} \quad h_2(\mu, \phi) = \frac{1}{2\phi^2}, \end{aligned}$$

y entonces

$$\pi_2^i(\mu, \phi) = \frac{\left(\frac{1}{2\phi^2}\right)^{1/2} I_{[a_i, b_i]}(\phi)}{\int_{[a_i, b_i]} \left(\frac{1}{2\phi^2}\right)^{1/2} d\phi},$$

de donde

$$\pi_2^i(\mu, \phi) = \frac{\phi^{-1} I_{[a_i, b_i]}(\phi)}{\log b_i - \log a_i}.$$

Usando esta distribución y como

$$\begin{aligned} E_1^1 \left[ \log \frac{\phi}{\mu^2} \mid \phi \right] &= \left( \log \frac{b_1}{a_1} \right)^{-1} \int_{[a_1, b_1]} \phi^{-1} \log \frac{\phi}{\mu^2} d\phi \\ &= K_0 - \log \mu^2 \end{aligned}$$

con  $K_0 = \frac{1}{2} \log b_1 a_1$ , se obtiene finalmente que

$$\pi_\mu(\mu, \phi) \propto (\mu\phi)^{-1}.$$

iii. En este caso se supone que  $\phi$  es ahora el parámetro de interés y que  $\mu$  el de ruido.

Es fácil verificar que  $S(\phi, \mu)$  está dada por (27), intercambiando las entradas de la diagonal principal para que el orden de los parámetros coincida, obteniéndose las siguientes relaciones

$$\begin{aligned} S_1(\phi, \mu) &= \phi(1 + 2\phi) \text{ y } S_2(\phi, \mu) = S(\phi, \mu), \\ H_1(\phi, \mu) &= \frac{1}{\phi(1 + 2\phi)} \text{ y } H_2(\phi, \mu) = S^{-1}(\phi, \mu), \\ h_1(\phi, \mu) &= \frac{1}{\phi(1 + 2\phi)} \text{ y } h_2(\phi, \mu) = 2 \frac{(1 + 2\phi)}{\mu^2}. \end{aligned}$$

De donde puede verificarse fácilmente que

$$\pi_\phi(\phi, \mu) \propto \mu^{-1}(\phi + 2\phi)^{-1/2}.$$

El último grupo de distribuciones se refiere a la parametrización  $(\lambda, \phi)$ . La demostración de esta proposición es completamente análoga a la anterior, por lo que se omite.

**Proposición 3.** Supóngase que la densidad Gaussiana Inversa se parametriza con  $\theta = (\lambda, \phi)$ , con  $\phi = \lambda/\mu$ . Entonces

i.  $\pi_\theta(\lambda, \phi) \propto \lambda^{-1}\phi^{-1/2}$ ,

ii.  $\pi_\lambda(\lambda, \phi) \propto \lambda^{-1}\phi^{-1/2}$ ,

$$\text{iii. } \pi_{\phi}(\phi, \lambda) \propto \lambda^{-1}(\phi + 2\phi^2)^{-1/2}.$$

Nótese que estas distribuciones pueden obtenerse a partir de la proposición 2, haciendo el cambio de variable  $\lambda = \mu$ .

## VI. UNA FAMILIA DE DISTRIBUCIONES DE REFERENCIA PARA LA DISTRIBUCION GAUSSIANA INVERSA

En la sección anterior se obtuvieron nueve distribuciones iniciales de referencia que podrían usarse, además de las presentadas en la tercera sección, para obtener a la distribución predictiva que estimará a la función de distribución en el problema de Bondad de Ajuste.

Sin embargo, no todas son adecuadas. Por un lado, Bernardo y Berger (1992) mencionan que si el objetivo final es hacer predicciones, las que se harían con la distribución predictiva, entonces el parámetro de interés debe ser  $E(X)$ , que en la parametrización usual es igual a  $\mu$ . De acuerdo a esto, la distribución inicial de la Proposición 2(iii) no es adecuada, pues supone a  $\phi$  como el parámetro de interés y a  $\mu$  como el de ruido, y como la distribución de la Proposición 3(iii) puede obtenerse a partir de la anterior, el mismo comentario se aplica a ella.

Por otro lado, como se demostrará en el siguiente capítulo, la distribución encontrada en 2(ii), no produce distribuciones finales propias, por lo que la distribución predictiva correspondiente no existe.

Puesto que en la distribución de la proposición 1(ii), se trata a  $\mu$  como el parámetro de interés y a  $\lambda$  como parámetro de ruido, esta distribución es la más adecuada.

El objetivo principal de este trabajo, es proponer una familia de distribuciones iniciales que, entre otras ventajas, contiene como casos particulares a las dos distribuciones mencionadas anteriormente.

Específicamente, se propone la siguiente familia paramétrica de distribuciones:

$$\mathcal{P}_{IG} = \{\pi(\mu, \lambda) \propto \lambda^{-r} \mu^{-3/2} : r \in \mathbf{R}^+\} \quad (28)$$

Esta familia tiene las siguientes características:

- i. Al ser una familia paramétrica, es más sencilla la asignación de una distribución inicial.
- ii. Como depende de un sólo parámetro, permite trabajar en forma general diferentes distribuciones iniciales propuestas en la literatura. Por ejemplo, las distribuciones que se obtienen por la Regla de Jeffreys, incluyendo la que se obtiene con la parametrización  $(1/\mu, \lambda)$ , están incluidas en esta familia (ver expresiones 14, 15 y 16).
- iii. Para un amplio rango de valores de  $r$ , produce distribuciones finales propias, teniendo así que las distribuciones predictivas correspondientes existen.
- iv. Contiene a las distribuciones encontradas en la proposición 1 y las de las proposiciones 2(i), 3(i) y 3(ii), lo que permite un análisis, a través de diferentes valores de  $r$ , del efecto de considerar a uno de los parámetros como parámetro de ruido y a distintas parametrizaciones.
- v. Puede extenderse, considerando a

$$\mathcal{P}_{IG}^* = \left\{ \pi(\mu, \lambda) \propto \lambda^{-r} \pi(\mu) : r \in \mathbf{R}^+ \right\} \quad (29)$$

donde  $\pi(\mu)$  es cualquier distribución para  $\mu$ , posiblemente impropia.

Los primeros cuatro puntos serán tratados en el siguiente capítulo. El punto *v* no será tratado, pues como se verá en el capítulo 4, la familia propuesta en (28) da soluciones satisfactorias al problema de Bondad de Ajuste, que es finalmente, el que se quiere resolver.

## CAPITULO 3

### I. INTRODUCCION

En el capítulo anterior se hizo una propuesta de una familia de distribuciones iniciales para los parámetros de la distribución Gaussiana Inversa, que contiene como casos particulares a las distribuciones de referencia más comunes reportadas en la literatura. En este capítulo se encontrará la familia correspondiente de distribuciones finales y se establecerán condiciones que permitirán definir a una familia de distribuciones finales propias. Se darán expresiones para la familia de distribuciones predictivas que resultan de la familia de distribuciones finales propias.

### II. DISTRIBUCIONES FINALES PARA LA DISTRIBUCION GAUSSIANA INVERSA

La siguiente familia de distribuciones iniciales para los parámetros de una distribución Gaussiana Inversa, fue propuesta en el capítulo anterior,

$$\mathcal{P}_{IG} = \left\{ \pi(\mu, \lambda) \propto \lambda^r \mu^{-3/2} : r \in \mathbf{R}^+ \right\}. \quad (1)$$

Si  $Z = \{X_1, X_2, \dots, X_n\}$  una muestra aleatoria de  $IG(x|\mu, \lambda)$ , la verosimilitud puede escribirse como

$$p(z|\mu, \lambda) = \left( \frac{\lambda}{2\pi} \right)^{n/2} \prod_{i=1}^n x_i^{-3/2} \exp \left\{ - \frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} \right\}.$$

En virtud del teorema de Bayes y usando (1), la distribución final de  $(\mu, \lambda)$ , para cada  $r \in \mathbf{R}^+$ , es

$$\pi_r(\mu, \lambda|z) \propto \mu^{-3/2} \lambda^{n/2-r} \exp \left\{ - \frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} \right\},$$

o bien, desarrollando el exponente,

$$\pi_r(\mu, \lambda | z) \propto \mu^{-3/2} \lambda^{n/2-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{n\lambda\bar{x}}{2\mu^2} - \frac{n\lambda\tilde{x}}{2}\right\}, \quad (2)$$

donde  $\bar{x}$  es la media aritmética y  $\tilde{x}$  es el inverso de la media armónica, es decir,  $\tilde{x} = \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}$ .

Esta distribución será propia siempre que

$$\int_0^{\infty} \int_0^{\infty} \mu^{-3/2} \lambda^{n/2-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{n\lambda\bar{x}}{2\mu^2} - \frac{n\lambda\tilde{x}}{2}\right\} d\mu d\lambda \quad (3)$$

sea finita para toda  $n, \bar{x}, \tilde{x}$  y  $r \in \mathbf{R}^+$ .

Defínase a

$$g(\mu, \lambda) = \mu^{-3/2} \lambda^a \exp\left\{\frac{b\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda c_1}{\mu^2} - c_2 \lambda\right\} \text{ para toda } \mu, \lambda \in \mathbf{R}^+, \quad (4)$$

con  $a, b, c_1$  y  $c_2 \in \mathbf{R}$ , y denótese por  $\Lambda(g)$  a la doble integral de  $g$  respecto a  $\mu$  y  $\lambda$ , es decir,

$$\Lambda(g) = \int_0^{\infty} \int_0^{\infty} g(\mu, \lambda) d\mu d\lambda.$$

Finalmente, considérese el siguiente cambio de variable

$$u = \mu \quad \text{y} \quad v = \frac{\lambda}{\mu},$$

cuyo jacobiano es  $u$ , entonces

$$\Lambda(g) = \int_0^{\infty} \int_0^{\infty} v^a u^{a-1/2} \exp\{bv\} \exp\left\{-\frac{c_1 v}{u} - c_2 uv\right\} dudv. \quad (5)$$

De acuerdo a Gradshteyn y Ryzhik (1980), se tiene que

$$\int_0^{\infty} x^{\nu-1} \exp\left\{-\frac{\beta}{x} - \gamma x\right\} dx = 2\left(\frac{\beta}{\gamma}\right)^{\frac{\nu}{2}} \mathbf{K}_{\nu}(2\sqrt{\beta\gamma}) \text{ si } \beta, \gamma \in \mathbf{R}^+ \text{ y } \nu \in \mathbf{R},$$

donde  $K_\nu$  denota a la función de Bessel modificada de orden  $\nu$ .

Usando este resultado y haciendo  $\nu = a + \frac{1}{2}$ , al integrar (5) respecto a  $u$  y suponiendo que  $c_1, c_2 \in \mathbf{R}^+$ , se obtiene

$$\Lambda(g) = 2 \left( \frac{c_1}{c_2} \right)^{2a + \frac{1}{4}} \int_0^\infty v^a e^{bv} K_{2a + \frac{1}{2}}(2v\sqrt{c_1 c_2}) dv. \quad (6)$$

Si  $2a \in \mathbf{N}$ , la función de Bessel tiene una forma cerrada (McLachlan, 1955) y por lo tanto la expresión (6) podría integrarse analíticamente. En general esto no ocurre, por lo que (6) debe obtenerse numéricamente, si es que la integral existe. Sin embargo, si  $\nu$  es fija (Abramowitz y Stegun, 1972),

$$\begin{aligned} K_\nu(t) &= \sqrt{\frac{\pi}{2t}} e^{-t} \left\{ 1 + \frac{4\nu^2 - 1}{8t} + O(t^{-2}) \right\} && \text{para valores grandes de } t \text{ y} \\ K_\nu(t) &= \frac{1}{2} \Gamma(\nu) \left( \frac{1}{2t} \right)^{-\nu} && \text{para valores pequeños de } t, \end{aligned}$$

y entonces

$$\begin{aligned} K_\nu(t) &= O(t^{-5/2} e^{-t}) && \text{para valores grandes de } t \text{ y} \\ K_\nu(t) &= O(t^{-\nu}) && \text{para valores pequeños de } t. \end{aligned}$$

Así, el comportamiento de la función de Bessel en los extremos del intervalo de integración, hacen que la aproximación numérica sea inestable.

Con el fin de tener una forma sencilla de evaluar numéricamente a la integral (3) y al mismo tiempo, encontrar condiciones para que sea finita, se demostrará que puede escribirse como la suma de dos integrales finitas sobre el intervalo  $(0, 1)$ . Primero se mostrará la descomposición en la suma de integrales y posteriormente se demostrará que son finitas, teniéndose así, condiciones para la existencia de las correspondientes distribuciones predictivas.

**Proposición 1.** Si  $a, b, c_1$  y  $c_2 \in \mathbf{R}^+$  son tales que  $\frac{c_1}{c_2} + c_2 u - b \in \mathbf{R}^+$ , entonces

$$\Lambda(g) = \Gamma(a + 1) \left\{ \int_0^1 \frac{u^{2a + \frac{1}{2}} du}{(c_2 u^2 - bu + c_1)^{a+1}} + 2 \int_0^1 \frac{du}{(c_1 u^4 - bu^2 + c_2)^{a+1}} \right\}. \quad (7)$$

**Demostración.** Sean  $a, b, c_1$  y  $c_2 \in \mathbf{R}^+$  tales que  $\frac{c_1}{u} + c_2u - b \in \mathbf{R}^+$ . Intercambiando el orden de integración en (5), se sigue que

$$\begin{aligned}\Lambda(g) &= \int_0^{\infty} u^{a-\frac{1}{2}} \left\{ \int_0^{\infty} v^a \exp \left\{ -v \left( \frac{c_1}{u} + c_2u - b \right) \right\} dv \right\} du \\ &= \Gamma(a+1) \int_0^{\infty} \frac{u^{a-\frac{1}{2}} du}{\left( \frac{c_1}{u} + c_2u - b \right)^{a+1}} \\ &= \Gamma(a+1) \int_0^{\infty} \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}}\end{aligned}$$

La penúltima igualdad es válida por las hipótesis sobre  $a, c_1, c_2$  y  $b$ . Partiendo la integral, se tiene

$$\Lambda(g) = \Gamma(a+1) \left\{ \int_0^1 \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}} + \int_0^{\infty} \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}} \right\}.$$

Haciendo el cambio de variable  $t = \frac{1}{\sqrt{u}}$  en la segunda integral, se obtiene

$$\Lambda(g) = \left\{ \int_0^1 \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}} - \int_{t(1)}^{t(\infty)} \frac{2t^{-(4a+1)} dt}{t^3 [t^{-4}(c_2 - bt^2 + c_1t^4)]^{a+1}} \right\}$$

y finalmente

$$\Lambda(g) = \Gamma(a+1) \left\{ \int_0^1 \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}} + 2 \int_0^1 \frac{du}{(c_1u^4 - bu^2 + c_2)^{a+1}} \right\}.$$

Una de las ventajas de escribir así a la doble integral (3), es que ninguna de las dos integrales que aparecen en (7) tiene singularidades en los extremos del intervalo de integración. Por otro lado, es obvio que la segunda integral de la expresión (7) converge, pues el integrando es una función continua y acotada en  $[0, 1]$ . Para verificar que la primera también es convergente, se demostrará una proposición más general.

**Proposición 2.** Sean  $\gamma, \nu, v_1, v_2, v_3 \in \mathbf{R}^+$ . Defínase

$$f(t) = \frac{t^\gamma}{(v_1 t^2 - 2v_2 t + v_3)^{\frac{\nu+1}{2}}}, \quad t \in \mathbf{R}^+.$$

Si  $v_1 v_3 \geq v_2^2$  y  $\nu > \gamma$ , entonces

$$\int_0^\infty f(t) dt < \infty.$$

**Demostración.** Sean

$$g: \mathbf{R}^+ \cup \{0\} \rightarrow \mathbf{R} \text{ tal que } g(t) = t^\gamma \text{ y}$$

$$h: \mathbf{R}^+ \cup \{0\} \rightarrow \mathbf{R} \text{ tal que } h(t) = (v_1 t^2 - 2v_2 t + v_3)^{\frac{\nu+1}{2}},$$

entonces, puede mostrarse, dadas las hipótesis, que

$$g(t) = O(t^\gamma) \text{ en } \infty \text{ y } g(0) = 0$$

y además

$$h(t) = O(t^{\nu+1}) \text{ en } \infty \text{ y } h(t) \in \mathbf{R}^+ \forall t \in \mathbf{R}^+.$$

Para demostrar que la imagen de  $h$  es  $\mathbf{R}^+$ , defínase a  $r(t) = v_1 t^2 - 2v_2 t + v_3$ . Si  $D_t^n r(t)$  denota a la derivada de orden  $n$  de  $r$  con respecto a  $t$ , se tiene que

$$D_t r(t) = 2v_1 t - v_2,$$

por lo que

$$t^* = \frac{v_2}{v_1}$$

es un extremo de  $r$ ; pero

$$D_t^2 r(t) \in \mathbf{R}^+, \quad \text{para toda } t \in \mathbf{R}^+,$$

así que  $t^*$  es un mínimo global de  $r$ .

Por otro lado, como

$$\begin{aligned} r(t^*) &= v_1 \left( \frac{v_2}{v_1} \right)^2 - 2v_2 \left( \frac{v_2}{v_1} + v_3 \right) \\ &= v_1(v_1 v_3 - v_2^2), \end{aligned}$$

y por hipótesis,  $v_1 \in \mathbf{R}^+$ , entonces  $r(t^*) \in \mathbf{R}^+$ , por lo que  $h(t) \in \mathbf{R}^+$  para toda  $t \in \mathbf{R}^+$ .

Entonces

$$f(t) = \frac{g(t)}{h(t)} \quad \forall t \in \mathbf{R}^+ \cup \{0\},$$

y

$$f(0) = 0 \text{ y } f(t) = O(t^{-(\nu-\gamma)-1}).$$

Así,

$$\int_0^{\infty} f(t) dt < \infty \text{ si y sólo si } \int_{\varepsilon}^{\infty} t^{-(\nu-\gamma)-1} dt < \infty,$$

para cualquier  $\varepsilon \in \mathbf{R}^+$ .

Pero

$$\int t^{-(\nu-\gamma)-1} dt = \begin{cases} \log t, & \text{si } \nu = \gamma; \\ t^{-(\nu-\gamma)} & \text{si } \nu \neq \gamma. \end{cases}$$

por lo que si  $\nu = \lambda$ , entonces la integral no existe. Si  $\nu \neq \lambda$ , la integral existe, pero sólo es finita cuando  $\nu - \gamma > 0$ .

**Corolario 1.** Sean  $a, c_1, c_2$  y  $b \in \mathbf{R}^+$ . Si  $4c_1c_2 \geq b^2$  entonces  $\Lambda(g) < \infty$ .

**Demostración.** Por la proposición 1 basta demostrar que

$$\int_0^1 \frac{u^{2a+\frac{1}{2}} du}{(c_2u^2 - bu + c_1)^{a+1}} < \infty.$$

Sea  $\gamma = 2a + \frac{1}{2}$  y  $\nu = 2a + 1$ , se tiene que  $\gamma, \nu \in \mathbf{R}^+$  y obviamente  $\nu > \gamma$ . El resultado se obtiene de la proposición 2, haciendo  $c_2 = v_1, c_1 = v_3$  y  $v_2 = b/2$ .

Finalmente, pueden darse condiciones para que la integral (3) exista.

**Proposición 3.** Si  $r < \frac{n}{2}$  entonces

$$\int_0^\infty \int_0^\infty \mu^{-3/2} \lambda^{n/2-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{n\lambda\bar{x}}{2\mu^2} - \frac{n\lambda\bar{x}}{2}\right\} d\mu d\lambda < \infty.$$

**Demostración.** Denótese con  $\Lambda_0$  a la doble integral

$$\int_0^\infty \int_0^\infty \mu^{-3/2} \lambda^{n/2-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{n\lambda\bar{x}}{2\mu^2} - \frac{n\lambda\bar{x}}{2}\right\} d\mu d\lambda.$$

Defínase a  $a = \frac{n}{2} - r, b = n, c_1 = \frac{n\bar{x}}{2}$  y  $c_2 = \frac{n\bar{x}}{2}$ . Como

$$\frac{c_1}{u} + c_2u - b = \frac{1}{u}(c_2u^2 - bu + c_1),$$

con un razonamiento similar al de la demostración de la proposición 2, se puede probar que

$$\frac{c_1}{u} + c_2u - b \in \mathbf{R}^+,$$

ya que  $c_1c_2 \geq b^2$  pues  $\bar{x}\bar{x} \geq 1$  como puede mostrarse fácilmente usando la desigualdad de Cauchy - Schwartz.

Por lo tanto, de la proposición 1 se sigue que

$$\Lambda_0 = \Gamma\left(\frac{n}{2} - r\right) \left\{ \int_0^1 \frac{u^{n-2r+\frac{1}{2}} du}{\left(\frac{n\bar{x}u^2}{2} - nu + \frac{n\bar{x}}{2}\right)^{\frac{n}{2}-r+1}} + 2 \int_0^1 \frac{du}{\left(\frac{n\bar{x}u^4}{2} - nu^2 + \frac{n\bar{x}}{2}\right)^{\frac{n}{2}-r+1}} \right\},$$

y del corolario 1 se sigue que la primera integral es finita, pues  $n - 2r + \frac{1}{2} > n - 2r$ . Por lo tanto  $\Lambda_0$  es finita.

Resumiendo, si  $r < \frac{n}{2}$  la familia de distribuciones iniciales propuesta, esto es,

$$\mathcal{P}_{IG} = \left\{ \pi_r(\mu, \lambda) \propto \lambda^r \mu^{-3/2} : r \in \mathbf{R}^+ \right\},$$

produce una familia de distribuciones finales propias, dada por

$$\mathcal{P}_{IG}^* = \left\{ \pi_r(\mu, \lambda|z) \propto \mu^{-3/2} \lambda^{n/2-r} \exp \left\{ -\frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} \right\} \right\},$$

así que la correspondientes distribuciones predictivas existen.

A partir de la proposición 2 también puede demostrarse que la distribución inicial obtenida en la proposición 2.2(ii), no produce una distribución final propia.

**Corolario 2.** Si  $\pi(\mu, \lambda) \propto (\mu\phi)^{-1}$  entonces  $\pi(\mu, \lambda|z)$  no es propia para todo  $z$ .

**Demostración.** Por el teorema de Bayes, la distribución final asociada a esta distribución inicial es

$$\pi(\mu, \lambda|z) \propto (\mu\lambda)^{-1} (\mu\lambda)^{n/2} \exp \left\{ -\frac{\phi}{2} \sum_{i=0}^n \frac{(x_i - \mu)^2}{\mu x_i} \right\}.$$

Esta distribución es propia, si para toda  $z$ , la integral

$$\int_0^\infty \int_0^\infty \phi^{\frac{n}{2}-1} \mu^{\frac{n}{2}-1} \exp \left\{ -\frac{\phi}{2} \sum_{i=0}^n \frac{(x_i - \mu)^2}{\mu x_i} \right\} d\mu d\lambda$$

es finita. Si se denota por  $\Lambda_0$  a esta integral y se integra primero con respecto a  $\phi$ , se obtiene que

$$\begin{aligned}
\Lambda_0 &= \Gamma\left(\frac{n}{2}\right) \int_0^{\infty} \left(\frac{\bar{x}}{\mu} - 2n + \mu\bar{x}\right)^{-\frac{n}{2}} \mu^{\frac{n}{2}-1} d\mu \\
&= \Gamma\left(\frac{n}{2}\right) \int_0^{\infty} \frac{\mu^{\frac{n}{2}-1} d\mu}{\left[\frac{1}{\mu}(\bar{x} - 2n\mu + \bar{x}\mu^2)\right]^{\frac{n}{2}}} \\
&= \Gamma\left(\frac{n}{2}\right) \int_0^{\infty} \frac{\mu^{n-1} d\mu}{(\bar{x}\mu^2 - 2n\mu + \bar{x})^{\frac{n}{2}}}
\end{aligned}$$

De acuerdo a la proposición 2, y haciendo  $\gamma = n - 1$  y  $\frac{\nu+1}{2} = \frac{n}{2}$ , es decir  $\nu = n - 1$ , se tiene que la integral no existe, pues  $\gamma = \nu$ , por lo tanto la distribución final no es propia.

4

Con respecto a las otras distribuciones iniciales encontradas en el capítulo anterior, sólo las de la proposición 2.1 y las de las proposiciones 2.2(i), 2.3(i) y 2.3(ii) están en la familia. En efecto, en el caso de las distribuciones propuestas en la proposición 2.1, si  $r = 1/2$  se obtiene 2.1(i) y si  $r = 1$  se obtienen los otros dos incisos. Por otro lado, 2.2(i) se obtiene de 2.1(i) si se considera a  $\phi = \lambda/\mu$ , por lo que pertenece a  $\mathcal{P}_{IG}$  con  $r = 1/2$ . Finalmente, en el caso de 2.3(i) y 2.3(ii), haciendo  $\mu = \lambda/\phi$  se puede mostrar que pertenecen a  $\mathcal{P}_{IG}$  con  $r = 3/2$ . En todos los casos, basta que  $n$  sea mayor a tres, para que todas las distribuciones finales asociadas, sean propias y las correspondientes predictivas existan.

Un último comentario relativo a las distribuciones iniciales presentadas en el capítulo anterior, es relevante. Banerjee y Bhattacharyya (1979) proponen una distribución inicial para  $(\eta, \lambda)$  diferente a la obtenida mediante la regla de Jeffreys, argumentando que esta última producía una final impropia. Esta aseveración es falsa. Recuérdese que la inicial de Jeffreys está dada por

$$p_j(\eta, \lambda) \propto (\eta\lambda)^{-1/2},$$

y al hacer el cambio de variable  $\mu = 1/\eta$  se obtiene

$$p_j(\mu, \lambda) \propto \mu^{-3/2} \lambda^{-1/2},$$

que pertenece a  $\mathcal{P}_{IG}$  y para la que ya se demostró que la distribución final asociada es propia.

En esta sección se han dado todos los elementos necesarios para la obtención de las distribuciones predictivas que servirán para el problema de Bondad de Ajuste. La obtención de estas distribuciones, es el tema de la siguiente sección.

### III. DISTRIBUCIONES PREDICTIVAS

En el capítulo anterior se introdujo a la distribución predictiva de una observación "futura"  $X^*$  dada la información contenida en una muestra  $Z = \{X_1, X_2, \dots, X_n\}$  independiente, dado  $\theta$ , de  $X^*$ . Específicamente, si con  $Z_n$  se denota a la muestra de tamaño  $n$ , la distribución predictiva de  $X^*$  dado  $Z_n$  es

$$p(x^*|z_n) = \int p(\theta|z_n)p(x^*|\theta)d\theta. \quad (8)$$

El objetivo de encontrar una distribución predictiva de la distribución Gaussiana Inversa, es para usarla como un estimador de la función de distribución en el problema de Bondad de Ajuste. Desde este punto de vista, lo que se requiere es la distribución predictiva de cada  $X_i$  de la muestra dadas las  $n - 1$  restantes. En este caso, la distribución final asociada es *distinta* para cada  $X_i$ , pues los parámetros de esta distribución dependen de la muestra observada.

Con el propósito de no complicar la notación, y sin pérdida de generalidad, se supondrá que  $X_n$  es la "variable a predecir" dadas  $X_1, X_2, \dots, X_{n-1}$ . Las expresiones que se encuentren bajo este supuesto, serán totalmente generales y sólo cambiarán al evaluarlas en una muestra específica.

El objetivo es entonces, dada una muestra de tamaño  $n$ , encontrar las distribuciones predictivas de  $X_n$  dadas  $X_1, X_2, \dots, X_{n-1}$ , asociadas a la familia de distribuciones finales,  $\mathcal{P}_{IG}^*$ , definida en la sección anterior.

Primero se dará una expresión general que permite calcular numéricamente y de manera eficiente a la distribución predictiva. Posteriormente, se dará la forma específica para el caso de la distribución Gaussiana Inversa.

Sea  $Z_n = \{X_1, X_2, \dots, X_n\}$ , una muestra aleatoria de  $p(x|\theta)$ , entonces

$$\begin{aligned}
p(x_n|z_{n-1}) &= \int p(\theta|z_{n-1})p(x_n|\theta)d\theta \\
&= \int \frac{p(\theta)p(z_{n-1}|\theta)}{p(z_{n-1})}p(x_n|\theta)d\theta \quad \text{por el teorema de Bayes,} \\
&= \frac{\int p(\theta)p(z_{n-1}, x_n|\theta)d\theta}{\int p(\theta)p(z_{n-1}|\theta)d\theta}. \tag{9}
\end{aligned}$$

La última igualdad es debida a la independencia, dado  $\theta$ , entre  $X_n$  y  $Z_{n-1}$ , y la definición de  $p(z_{n-1})$ . Finalmente, la función de distribución predictiva queda como

$$\begin{aligned}
F_{n-1}^p(x) &= P[X_n \leq x|z_{n-1}] \\
&= \frac{\int \int_x^{\infty} p(\theta)p(z_{n-1}, x_n|\theta)d\theta dx_n}{\int p(\theta)p(z_{n-1}|\theta)d\theta}. \tag{10}
\end{aligned}$$

Esta expresión tiene la ventaja de no requerir del cálculo específico de la distribución final.

La ecuación (9) y los resultados de la sección anterior, permiten dar una expresión para la distribución predictiva, que puede ser evaluada numéricamente para cada  $r \in \mathbf{R}^+$ . Para hacer esto, defínase para cada  $k \in \mathbf{N}$  a

$$\begin{aligned}
\bar{x}_k &= \frac{1}{k} \sum_{i=1}^k x_i \quad \text{y} \\
\tilde{x}_k &= \frac{1}{k} \sum_{i=1}^k \frac{1}{x_i},
\end{aligned}$$

la media aritmética y el inverso de la media armónica respectivamente, de una muestra de tamaño  $k$ .

Análogamente,  $Z_k = \{X_1, X_2, \dots, X_k\}$  denota a una muestra de tamaño  $k$  de una distribución Gaussiana Inversa con parámetros  $(\mu, \lambda)$ . Entonces

$$p(z_{n-1}|\mu, \lambda) = \left(\frac{\lambda}{2\pi}\right)^{\frac{n-1}{2}} \prod_{i=1}^{n-1} x_i^{-3/2} \exp\left\{\frac{n-1\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_{n-1}}{2\mu^2} - \frac{\lambda\tilde{x}_{n-1}}{2}\right\} \quad \text{y}$$

$$p(z_n|\mu, \lambda) = \left(\frac{\lambda}{2\pi}\right)^{\frac{n}{2}} \prod_{i=1}^n x_i^{-3/2} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_n}{2\mu^2} - \frac{\lambda\bar{x}_n}{2}\right\}, \quad (11)$$

son las verosimilitudes de  $(\mu, \lambda)$  dadas  $Z_{n-1}$  y  $Z_n$ , respectivamente.

Si se supone que la distribución inicial de  $(\mu, \lambda)$  es un elemento de  $\mathcal{P}_{IG}$ , entonces

$$\pi_r(\mu, \lambda)p(z_{n-1}|\mu, \lambda) \propto \lambda^{\frac{n-1}{2}-r} \exp\left\{\frac{(n-1)\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_{n-1}}{2\mu^2} - \frac{\lambda\bar{x}_{n-1}}{2}\right\} \quad \text{y}$$

$$\pi_r(\mu, \lambda)p(z_n|\mu, \lambda) \propto \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \lambda^{\frac{1}{2}-r} x_n^{-3/2} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_n}{2\mu^2} - \frac{\lambda\bar{x}_n}{2}\right\}$$

por lo que

$$p(x_n|z_{n-1}) = \frac{\left(\frac{1}{2\pi}\right)^{\frac{1}{2}} x_n^{-3/2} \iint \lambda^{\frac{n}{2}-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_n}{2\mu^2} - \frac{\lambda\bar{x}_n}{2}\right\} d\mu d\lambda}{\iint \lambda^{\frac{(n-1)}{2}-r} \exp\left\{\frac{(n-1)\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_{n-1}}{2\mu^2} - \frac{\lambda\bar{x}_{n-1}}{2}\right\} d\mu d\lambda}. \quad (12)$$

Finalmente,

$$F_{n-1}^{\pi}(x) = \frac{\left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \int_0^x u^{-3/2} \iint \lambda^{\frac{n}{2}-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_n(u)}{2\mu^2} - \frac{\lambda\bar{x}_n(u)}{2}\right\} d\mu d\lambda du}{\iint \lambda^{\frac{(n-1)}{2}-r} \exp\left\{\frac{(n-1)\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda\bar{x}_{n-1}}{2\mu^2} - \frac{\lambda\bar{x}_{n-1}}{2}\right\} d\mu d\lambda}. \quad (13)$$

Esta última expresión, define a la familia de estimadores de la función de distribución que serán usados en el problema de Bondad de Ajuste, en donde se ha denotado con  $\bar{x}_n(u)$  y  $\bar{x}_{n-1}(u)$ , para hacer notar la dependencia en el proceso de integración.

Nótese que las integrales de la expresión (12) cumplen las hipótesis de la proposición 1, por lo que cada una puede escribirse como la suma de dos integrales finitas sobre el intervalo  $[0, 1]$ , y podrán ser integradas numéricamente.

En el siguiente capítulo, se resolverá (13) numéricamente y se harán comparaciones con otros estimadores de la función de distribución propuestos en la literatura.

## CAPITULO 4

### I. INTRODUCCION

El objetivo de este capítulo es evaluar numéricamente a las distribuciones predictivas encontradas en el capítulo anterior. Al mismo tiempo, se estimará a la función de distribución de la Gaussiana Inversa usando a los estimadores máximo verosímil y el de Rao-Blackwell, con el fin de hacer una comparación entre los tres. Por otro lado, las distribuciones predictivas encontradas dependen de un parámetro  $r \in \mathbb{R}^+$ , por lo que se hará un estudio numérico para ver el efecto de este parámetro en la estimación. Finalmente, se realizará un estudio numérico para comparar las estadísticas de prueba que se generan con cada una de las transformadas.

### II. DISTRIBUCIONES PREDICTIVAS: SOLUCION NUMERICA

Dada la familia de distribuciones iniciales  $\mathcal{P}_{IG}$ , en el capítulo anterior se encontró la siguiente expresión para la función de distribución predictiva en la Gaussiana Inversa,

$$F_{n-1,r}^{\pi}(x) = \frac{\int_0^x u^{-3/2} \iint \mu^{-3/2} \lambda^{\frac{n}{2}-r} \exp\left\{\frac{n\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda \bar{x}_n(u)}{2\mu^2} - \frac{\lambda \tilde{x}_n(u)}{2}\right\} d\mu d\lambda du}{\sqrt{2\pi} \iint \mu^{-3/2} \lambda^{\frac{(n-1)}{2}-r} \exp\left\{\frac{(n-1)\lambda}{\mu}\right\} \exp\left\{-\frac{\lambda \bar{x}_{n-1}}{2\mu^2} - \frac{\lambda \tilde{x}_{n-1}}{2}\right\} d\mu d\lambda}, \quad (1)$$

en donde los índices denotan que se está calculando la distribución predictiva dada una muestra de tamaño  $n-1$  y se está usando una distribución inicial perteneciente a  $\mathcal{P}_{IG}$  de parámetro  $r$ . Además, se hace énfasis en la dependencia del argumento  $u$  en las primeras dos integrales del numerador, al usar  $\bar{x}_n(u)$  y  $\tilde{x}_n(u)$ .

Denótese con  $\Lambda_1(u)$  y  $\Lambda_2$  a las integrales con respecto a  $\mu$  y  $\lambda$  del numerador y denominador de (1) respectivamente. Entonces

$$F_{n-1,r}^{\pi}(x) = \frac{1}{\sqrt{2\pi}\Lambda_2} \int_0^x u^{-3/2} \Lambda_1(u) du.$$

Para cada  $u \in \mathbf{R}^+$ , tanto  $\Lambda_1$  como  $\Lambda_2(u)$  son de la forma

$$\iint g(\mu, \lambda) d\mu d\lambda,$$

donde

$$g(\mu, \lambda) = \mu^{-3/2} \lambda^a \exp\left\{\frac{b\lambda}{\mu}\right\} \exp\left\{-\frac{c_1\lambda}{\mu^2} - c_2\lambda\right\},$$

con  $a, b, c_1$  y  $c_2 \in \mathbf{R}^+$  tales que  $4c_1c_2 \geq b^2$ , por lo que si  $r < \frac{n-1}{2}$ , las hipótesis de la proposición 3.3 se cumplen, teniéndose entonces que

$$\begin{aligned} \Lambda_1(u) &= \Gamma\left(\frac{n}{2} - r\right) \left\{ \int_0^1 \frac{t^{n-2r+\frac{1}{2}} dt}{\left(\frac{n\bar{x}_n(u)t^2}{2} - nt + \frac{n\bar{x}_n(u)}{2}\right)^{\frac{n}{2}-r+1}} \right. \\ &\quad \left. + 2 \int_0^1 \frac{dt}{\left(\frac{n\bar{x}_n(u)t^4}{2} - nt^2 + \frac{n\bar{x}_n(u)}{2}\right)^{\frac{n}{2}-r+1}} \right\} \quad y \\ \Lambda_2 &= \Gamma\left(\frac{n-1}{2} - r\right) \left\{ \int_0^1 \frac{t^{n-2r-\frac{1}{2}} dt}{\left(\frac{(n-1)\bar{x}_{n-1}(u)t^2}{2} - (n-1)t + \frac{(n-1)\bar{x}_{n-1}(u)}{2}\right)^{\frac{n+1}{2}-r}} \right. \\ &\quad \left. + 2 \int_0^1 \frac{dt}{\left(\frac{(n-1)\bar{x}_{n-1}(u)t^4}{2} - (n-1)t^2 + \frac{(n-1)\bar{x}_{n-1}(u)}{2}\right)^{\frac{n+1}{2}-r}} \right\}. \end{aligned} \quad (2)$$

Dada una muestra observada  $z_n = \{x_1, x_2, \dots, x_n\}$ , lo que interesa es obtener una muestra de una distribución cuyas marginales son uniformes, denotada con  $\{u_1^{\pi}, u_2^{\pi}, \dots, u_n^{\pi}\}$ , que como se vió en el capítulo 2, se obtienen mediante la transformación

$$U_i^{\pi}(x_1, x_2, \dots, x_n) = F_{n-1,r}^{\pi}(x_i | x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \quad \text{para toda } i \in J_n. \quad (3)$$

Esto es, para obtener la  $i$ -ésima transformada, se supone que  $X_i$  es la observación futura y se encuentra su distribución predictiva evaluada en el valor observado  $x_i$  dada la información de  $\{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ .

Sin pérdida de generalidad y con el fin de no complicar la notación, se supondrá que  $x_1, x_2, \dots, x_n$  se ordenan de manera tal que  $x_n$  es siempre el valor donde se evaluará a la distribución predictiva de la observación futura, denotada con  $X_n$ , dada la información de las primeras  $n - 1$  observaciones. Bajo este supuesto,  $\Lambda_2$  puede ser evaluada numéricamente si se conocen los valores de  $n, r, \bar{x}_{n-1}$  y  $\tilde{x}_{n-1}$ .

Por otro lado,  $\Lambda_1(u)$  puede evaluarse para cada valor de  $u$ , dados  $n, r, \bar{x}_{n-1}$  y  $\tilde{x}_{n-1}$ , ya que

$$\begin{aligned} n\bar{x}_n(u) &= (n-1)\bar{x}_{n-1}(u) + u & \text{y} \\ n\tilde{x}_n(u) &= (n-1)\tilde{x}_{n-1}(u) + \frac{1}{u}, \end{aligned} \quad (4)$$

es lo único que depende de  $u$ , y posteriormente calcular

$$\int_0^{x_n} u^{-3/2} \Lambda_1(u) du. \quad (5)$$

En esta última expresión, se tiene un proceso de doble integración, por lo que numéricamente debe realizarse al revés, es decir, primero debe integrarse (5) y en cada llamada de la función a integrar, resolver la integral  $\Lambda_1(u)$ .

Como el objetivo es ver cómo se comporta la distribución predictiva como un estimador de la función de distribución en el problema de Bondad de Ajuste, y dado que no existe una expresión exacta para la primera que permita estudiar analíticamente el comportamiento teórico de los procesos empíricos estimados involucrados, las únicas comparaciones que pueden hacerse son mediante simulación.

Los primeros problemas a resolver son: ¿cuál método de integración numérica debe utilizarse? y ¿cuáles valores de  $r$  se usan en las comparaciones?

En el primer caso, al haber transformado las integrales respecto a  $\mu$  y  $\lambda$  que aparecen en (1), a las correspondientes expresiones dadas en (2), se obtuvieron dos puntos favorables para la solución numérica: primero, se eliminaron todas las singularidades de las

funciones involucradas en la integración. En efecto, puede observarse fácilmente de (2) que todas las funciones que se van a integrar, toman valores finitos en los extremos del intervalo de integración, y dada la continuidad de las mismas, se asegura la inexistencia de singularidades en todo el intervalo. Segundo, al tener intervalos con extremos finitos, las aproximaciones son mejores.

Al conjuntarse estas dos características – ausencia de singularidades y extremos finitos en los intervalos de integración –, se pueden usar métodos *cerrados* de integración numérica, es decir, métodos que a diferencia de los llamados *abiertos*, permiten la evaluación de la función a integrar en los extremos del intervalo, produciendo soluciones más precisas.

Sin embargo, la integral del numerador con respecto a  $u$ , tiene una singularidad en el cero que no puede removerse, pues la integral depende tanto de  $u$  como de su inverso, como se observa en (4).

Podría realizarse la integración en (2) usando un método *cerrado* y resolver (5) con un método *abierto*, pero la precisión que se gana al hacerlo no cambia significativamente los resultados. La única ventaja en este caso, es que al usar un método *cerrado* se realizarían menos evaluaciones de la función al integrar.

Dado lo anterior, se optó por usar un mismo método *abierto* de integración numérica para resolver todas las integrales.

El procedimiento usado es una adaptación del método del trapecio presentado por Press *et. al.* (1989). Se hicieron varias corridas, variando la precisión requerida en la aproximación, y los resultados fueron esencialmente los mismos. El programa usado, que se presenta en el apéndice, requiere de dos llamadas independientes a la subrutina de integración, debido a que en fortran, las llamadas recursivas no son válidas.

El segundo problema que debe resolverse antes de hacer las comparaciones entre los diferentes estimadores de la función de distribución, es sobre los valores de  $r$  que deben usarse.

Para medir el efecto del parámetro  $r$  en la obtención de la distribución predictiva, se generaron muestras de la distribución Gaussiana Inversa para distintos valores de los parámetros y variando  $n$ , el tamaño de muestra y  $r$ . Dado que el efecto de  $r$  disminuye a medida que  $n$  crece, las simulaciones se hicieron sólo para valores pequeños de  $n$ .

En la tabla 1, se presentan los resultados para  $n = 5$ . La tabla contiene a la muestra generada, así como las transformaciones resultantes al usar el estimador de Rao-Blackwell ( $\hat{U}$ ) y el estimador predictivo ( $U_r^*$ ) para cinco valores de  $r$ . Los valores de  $r$  fueron elegidos, tomando en cuenta que para  $r = 0.5$  y  $r = 1$ , se obtienen las distribuciones de referencia más comunes en la literatura.

**TABLA 1. Transformadas predictivas y Rao-Blackwell**

$\phi$	$x$	$U_{0.3}^*$	$U_{0.5}^*$	$U_{0.8}^*$	$U_1^*$	$U_{1.5}^*$	$\hat{U}$
$2^{-4}$	0.019728	0.030775	0.039059	0.056194	0.071811	0.135735	0.080238
	0.056671	0.302375	0.324744	0.362707	0.391337	0.479029	0.346465
	0.127622	0.526360	0.545171	0.575967	0.598319	0.663696	0.504617
	0.203806	0.632728	0.648059	0.673519	0.690919	0.742597	0.575609
	3.569357	0.987144	0.989356	0.988582	0.988085	0.987753	0.953238
$2^{-3}$	0.050848	0.094068	0.109673	0.138585	0.162488	0.245695	0.149949
	0.073208	0.234807	0.254824	0.289499	0.316128	0.400767	0.289879
	0.113325	0.429274	0.446086	0.474502	0.495968	0.562260	0.455291
	0.135579	0.508958	0.522862	0.546548	0.564533	0.620726	0.520333
	0.531820	0.992089	0.989874	0.985453	0.981768	0.970082	0.980621
$2^{-2}$	0.044924	0.037474	0.046847	0.065776	0.082769	0.149730	0.084169
	0.071839	0.157193	0.176352	0.210427	0.237564	0.335548	0.222317
	0.628771	0.792503	0.798680	0.809296	0.817392	0.842320	0.731443
	0.706974	0.816497	0.821506	0.830240	0.837000	0.858252	0.753832
	0.840848	0.850400	0.853736	0.859794	0.864711	0.881004	0.787600
$2^{-1}$	0.145042	0.026724	0.034727	0.051576	0.067253	0.131627	0.060687
	0.300961	0.318120	0.336858	0.368774	0.393168	0.469388	0.362379
	0.497219	0.571178	0.580708	0.597608	0.610943	0.655063	0.573163
	0.507720	0.581416	0.590452	0.606577	0.619365	0.661997	0.581659
	1.258897	0.958086	0.952745	0.943936	0.937826	0.923575	0.932466

**TABLA 1. Transformadas (continuación)**

$\phi$	$x$	$U_{0.3}^x$	$U_{0.5}^x$	$U_{0.8}^x$	$U_1^x$	$U_{1.5}^x$	$\tilde{U}$
$2^0$	0.551358	0.009540	0.013477	0.022834	0.032675	0.082120	0.029379
	0.822983	0.336960	0.350460	0.373549	0.391447	0.449895	0.372765
	1.051484	0.638940	0.637453	0.636152	0.636350	0.645047	0.621111
	1.131942	0.725590	0.719860	0.711007	0.706722	0.700849	0.693967
	1.334350	0.890150	0.880259	0.863766	0.851855	0.820760	0.845764
2	0.442933	0.061904	0.074788	0.099713	0.121120	0.199620	0.109462
	0.548644	0.147697	0.165829	0.198198	0.224024	0.309527	0.204823
	1.226724	0.583654	0.591227	0.605112	0.616433	0.655646	0.583870
	1.864437	0.799701	0.796412	0.792909	0.791963	0.797455	0.770355
	2.245694	0.882918	0.876755	0.867935	0.862681	0.855164	0.849695
$2^2$	0.511665	0.030411	0.038852	0.056535	0.073002	0.141448	0.065763
	0.681035	0.232065	0.248996	0.278269	0.301132	0.375412	0.282182
	0.999988	0.633646	0.632986	0.633567	0.635177	0.648358	0.618417
	1.070021	0.703063	0.699093	0.693733	0.691216	0.692925	0.676680
	1.389765	0.924309	0.915269	0.899676	0.888086	0.856916	0.883748
$2^3$	0.831586	0.183156	0.199120	0.227143	0.249311	0.323369	0.235930
	0.866230	0.237443	0.252856	0.279453	0.300179	0.368254	0.285612
	0.889488	0.275048	0.289587	0.314498	0.333789	0.396988	0.318680
	1.089198	0.592690	0.592547	0.593239	0.594711	0.606162	0.580710
	1.594133	0.996947	0.995470	0.991730	0.987635	0.966303	0.987261
$2^4$	0.671149	0.010940	0.015274	0.025423	0.035975	0.087580	0.032294
	0.996663	0.327882	0.341815	0.365640	0.384064	0.444227	0.365154
	1.234511	0.586145	0.587539	0.590969	0.594793	0.613139	0.577835
	1.546424	0.835148	0.825781	0.810924	0.800745	0.777753	0.792102
	1.549617	0.837103	0.827696	0.812754	0.802497	0.779214	0.793941

Como puede observarse, independientemente del valor del parámetro  $\phi$ , las transformadas predictivas para  $r$  en el rango de 0.5 a 1, son más parecidas a la correspondiente transformada de Rao-Blackwell, que los casos  $r = 0.3$  y  $r = 1.5$ . Estos resultados se mantienen para valores de  $n$  menores a 25.

Como consecuencia de esto, se decidió usar para las estadísticas de Bondad de Ajuste a las transformaciones predictivas para  $r$  igual a 0.5, 0.8 y 1, para hacer las comparaciones con las correspondientes transformadas usando máxima verosimilitud y Rao-Blackwell.

Así, el siguiente paso fué calcular el valor de la estadística  $A^2$ , para cada uno de los cinco casos mencionados en el párrafo anterior y para distintos valores del parámetro  $\phi$ . Para cada combinación se generaron muestras de tamaño 10(10)150. La tabla 2 muestra estos resultados, para algunos valores de  $n$ .

**TABLA 2. Valores de la estadística  $A^2$  para las distintas transformadas.**

$n$	$\phi$	$A^2_{\pi,0.5}$	$A^2_{\pi,0.8}$	$A^2_{\pi,1}$	$\bar{A}^2$	$\hat{A}^2$
10	$2^{-4}$	0.470929	0.466217	0.472638	0.236095	0.459760
	$2^{-3}$	0.733298	0.770320	0.804549	0.284564	0.593229
	$2^{-2}$	0.265825	0.227236	0.210078	0.257592	0.290800
	$2^{-1}$	0.278170	0.246681	0.233733	0.246524	0.313868
	$2^0$	0.327172	0.300314	0.289018	0.286038	0.363973
	$2^1$	0.364235	0.326485	0.307473	0.325709	0.361786
	$2^2$	0.205134	0.178799	0.166200	0.173425	0.194297
	$2^3$	0.854168	0.807816	0.781645	0.796949	0.777774
	$2^4$	0.605148	0.600386	0.600332	0.587859	0.418187
	50	$2^{-4}$	0.450755	0.424169	0.426312	0.422787
$2^{-3}$		0.623420	0.650981	0.670693	0.617069	0.647619
$2^{-2}$		0.207761	0.204369	0.202659	0.214079	0.204050
$2^{-1}$		0.357988	0.346557	0.340082	0.343881	0.375814
$2^0$		0.300821	0.312944	0.322017	0.310655	0.286301

**TABLA 2. Valores de la estadística  $A^2$**   
(continuación)

$n$	$\phi$	$A_{\pi,0.5}^2$	$A_{\pi,0.8}^2$	$A_{\pi,1}^2$	$\tilde{A}^2$	$\hat{A}^2$
100	$2^1$	0.435217	0.426718	0.421813	0.438692	0.420219
	$2^2$	0.442768	0.431923	0.426011	0.420999	0.459009
	$2^3$	0.317106	0.310399	0.306653	0.310964	0.309026
	$2^4$	1.071499	1.080562	1.086185	1.079968	1.041673
	$2^{-4}$	0.629421	0.617332	0.602816	0.616756	0.680595
	$2^{-3}$	0.237000	0.245866	0.245928	0.240362	0.242229
	$2^{-2}$	0.216715	0.231531	0.211517	0.222864	0.208873
	$2^{-1}$	0.375978	0.388457	0.393532	0.374097	0.364126
	$2^0$	0.767501	0.770306	0.772356	0.780695	0.749987
	$2^1$	0.457140	0.451756	0.447032	0.457140	0.451256
150	$2^2$	0.768466	0.760284	0.756150	0.749542	0.780388
	$2^3$	0.267995	0.265938	0.263677	0.268065	0.264970
	$2^4$	1.122895	1.130301	1.130105	1.138900	1.108407
	$2^{-4}$	1.093217	1.071676	1.057766	1.078664	1.134091
	$2^{-3}$	0.242282	1.110900	1.312236	0.230612	0.224459
	$2^{-2}$	0.167151	0.169476	0.868756	0.172720	0.159016
	$2^{-1}$	0.562603	0.773900	0.778380	0.551756	0.578618
	$2^0$	0.838626	0.845924	0.851070	0.855560	0.816253
	$2^1$	0.670732	0.663044	0.657956	0.669992	0.665857
	$2^2$	0.913787	0.905062	0.899400	0.895908	0.925390
	$2^3$	0.287084	0.286070	0.285424	0.289860	0.283874
	$2^4$	0.565807	0.572071	0.576800	0.581772	0.558631

$A_{\pi,r}^2$  :  $A^2$  usando la transformada predictiva de parámetro  $r$

$\tilde{A}^2$  :  $A^2$  usando la transformada de Rao-Blackwell

$\hat{A}^2$  :  $A^2$  usando la transformada máximo verosímil

Los valores de  $A^2$  usando las distintas transformaciones son muy parecidos, exceptuando para valores pequeños tanto de  $n$  como de  $\phi$ . A medida que  $n$  crece, la similitud entre la  $A^2$  basada en la transformación de Rao-Blackwell y la  $A^2$  usando la transformación predictiva con  $r = 0.8$ , es mayor.

Sólo se generó una muestra para cada combinación  $n$  y  $\phi$ , pues se tiene teoría cuando se usan las transformadas de Rao-Blackwell y máxima verosimilitud. Así, las comparaciones son con el fin de ver cómo se comportan las transformadas predictivas para los diferentes valores de  $r$ , al contrastarlas con resultados bien establecidos.

El hecho de que los valores de la estadística  $A^2$  correspondientes a las diferentes transformaciones sean parecidos, no asegura que se tenga que llegar a las mismas conclusiones cuando éstas se usan. Es decir, las inferencias que se obtengan pueden ser sensitivas al tipo de transformación que se utilice. Para analizar este aspecto, se calcularon los niveles de significancia descriptivos para cada caso. El cálculo de estos niveles de significancia descriptivos, se hizo interpolando en la tabla de valores asintóticos presentada en O'Reilly y Rueda (1992) y que se anexa en el apéndice. Los resultados se presentan en la tabla 3.

**TABLA 3. Niveles de significancia descriptivos (nsd) para algunos casos de la Tabla 2.**

$n$	$\phi$	$p(A_{\pi,0.5}^2)$	$p(A_{\pi,0.8}^2)$	$p(A_{\pi,1}^2)$	$p(\tilde{A}^2)$	$p(\hat{A}^2)$
10	$2^{-3}$	0.001000	0.001000	0.001000	0.001000	0.001000
10	$2^2$	0.001000	0.001000	0.001000	0.001000	0.001000
10	$2^3$	0.033814	0.042728	0.047761	0.044818	0.048505
10	$2^4$	0.129043	0.133012	0.133057	0.143451	0.159844
50	$2^{-2}$	0.001000	0.001000	0.001000	0.001000	0.001000
50	$2^2$	0.001000	0.001000	0.001000	0.001000	0.001000
100	$2^0$	0.084166	0.083231	0.082548	0.079768	0.090004
100	$2^2$	0.058282	0.061429	0.063019	0.065561	0.053697

**TABLA 3. Niveles de significancia descriptivos (nsd) para algunos casos de la Tabla 2. (continuación)**

$n$	$\phi$	$p(A_{\pi,0.5}^2)$	$p(A_{\pi,0.8}^2)$	$p(A_{\pi,1}^2)$	$p(\bar{A}^2)$	$p(\hat{A}^2)$
150	$2^{-4}$	0.083997	0.088139	0.090814	0.086795	0.076136
150	$2^0$	0.060458	0.058025	0.056310	0.054813	0.067916
150	$2^1$	0.106620	0.112111	0.115746	0.107149	0.110102
150	$2^2$	0.026195	0.027873	0.028962	0.029633	0.024524
150	$2^4$	0.161828	0.156608	0.152667	0.148523	0.167808

$p(A_{\pi,r}^2)$  : nsd usando la transformación predictiva de parámetro  $r$

$p(\bar{A}^2)$  : nsd usando la transformación rao-blackwell

$p(\hat{A}^2)$  : nsd usando la transformación máximo verosímil

Como puede observarse, los niveles de significancia descriptivos son muy parecidos en los cinco casos, por lo que puede decirse que no hay diferencias en el uso de  $A^2$  bajo las distintas transformaciones.

Finalmente, dada la dificultad de encontrar la distribución asintótica del proceso

$$y_n^\pi(u) = \sqrt{n}\{G_n(u) - u\},$$

definido en (2.33), se hizo un estudio numérico sobre su comportamiento asintótico.

Tomando en cuenta que, para cualesquier par de transformaciones  $F^*(X)$  y  $F^{**}(X)$  se tiene

$$\sqrt{n}\{F_n(X) - F^*(X)\} = \sqrt{n}\{F_n(X) - F^{**}(X)\} + \sqrt{n}\{F^{**}(X) - F^*(X)\},$$

entonces, si se demuestra que

$$\sup_X \sqrt{n}|F^*(X) - F^{**}(X)| \rightarrow 0$$

en probabilidad, entonces puede concluirse que la distribución asintótica de los procesos

$$\sqrt{n}\{F_n(X) - F^*(X)\} \text{ y } \sqrt{n}\{F_n(X) - F^{**}(X)\}$$

es la misma.

En O'Reilly y Rueda (1992) se demostró que el proceso

$$\sup_X \sqrt{n}|\tilde{F}_n(X) - \hat{F}_n(X)| \rightarrow 0 \quad (6)$$

cuando  $n \rightarrow \infty$ , con probabilidad 1, por lo que el comportamiento de los procesos construidos con la diferencia entre las transformaciones predictivas y cualquiera de las otras dos - máxima verosimilitud y Rao-Blackwell -, deberá compararse con (6).

Para hacer esto se generó una muestra de tamaño 150 de una distribución Gaussiana Inversa para cada valor de  $\phi$  en el rango  $2^{-4}$  a  $2^4$ . En cada caso se fué transformando la muestra, aumentando 10 observaciones en cada etapa, con las cinco propuestas: tres distribuciones predictivas, Rao-Blackwell y máxima verosimilitud. Se calculó el máximo de las diferencias entre las distribuciones empíricas de las muestras obtenidas con cada una de las transformadas predictivas y la obtenida usando Rao-Blackwell, y las correspondientes diferencias usando la muestra transformada con máxima verosimilitud.

Para cada proceso se obtuvieron 15 puntos para cada valor del parámetro  $\phi$  y se graficaron para estudiar la convergencia de los diferentes procesos. Se presentan aquí algunas gráficas y las restantes se anexan en el apéndice.

En todas las gráficas, la línea roja representa al proceso

$$\max_x \sqrt{n}\{\tilde{F}_n(x) - \hat{F}_n(x)\}, \quad (7)$$

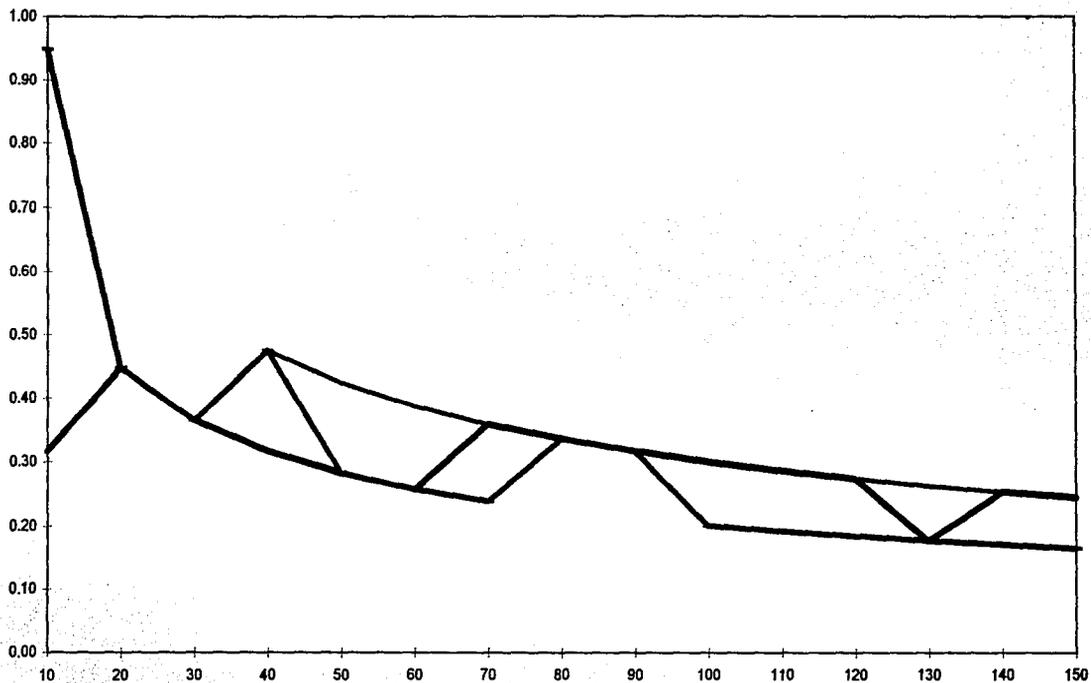
la línea azul, al proceso

$$\max_x \sqrt{n}\{F_n^\pi(x) - \hat{F}_n(x)\} \quad (8)$$

y la negra al proceso

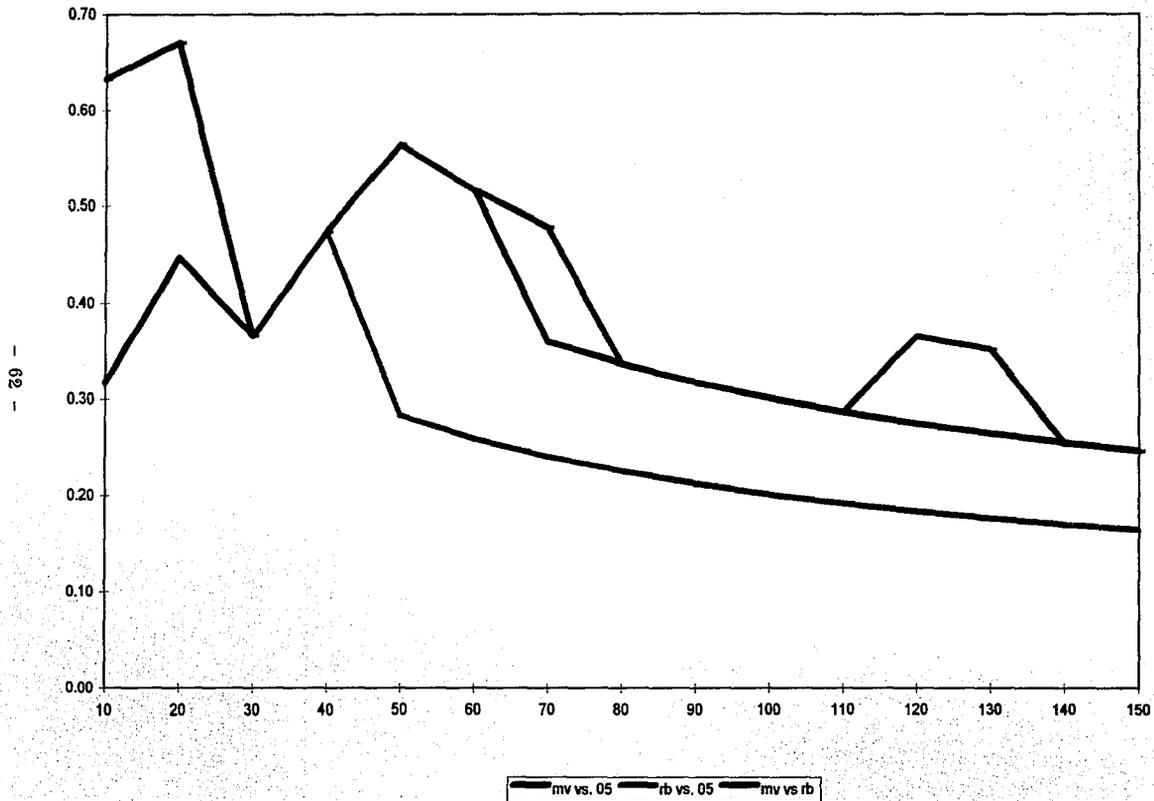
$$\max_x \sqrt{n}\{F_n^*(x) - \hat{F}_n(x)\}, \quad (9)$$

GRAFICA 1. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.0625)

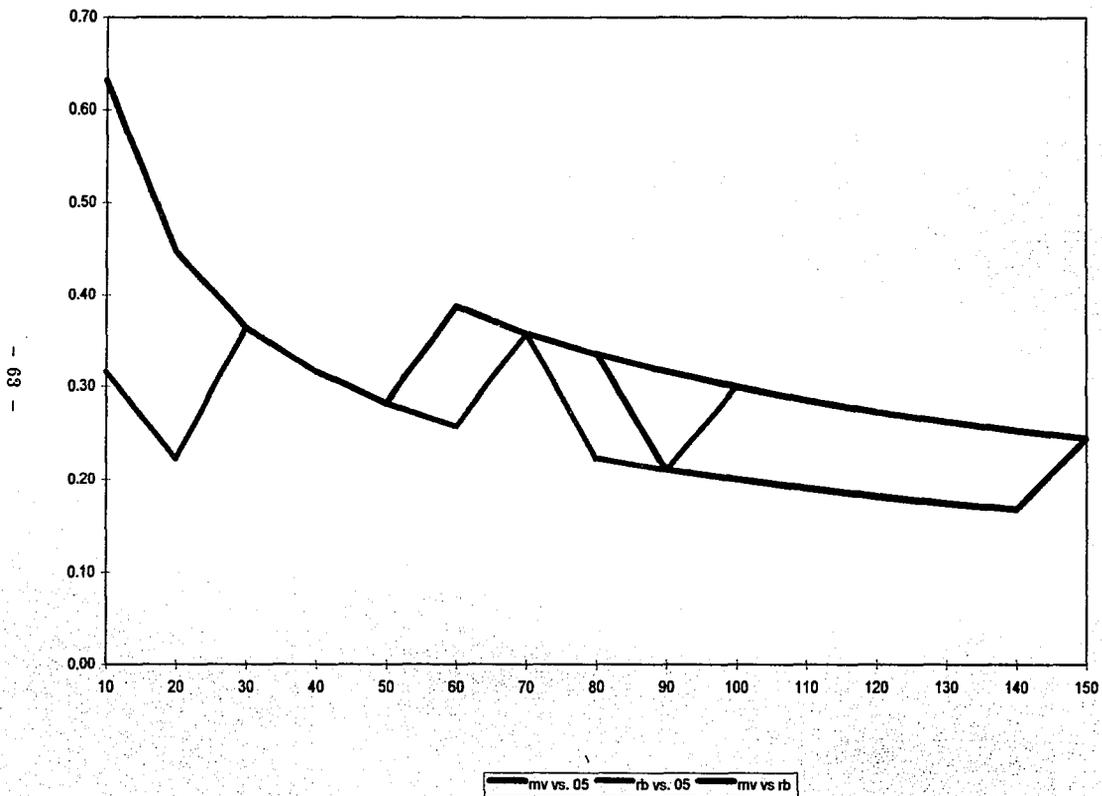


mv vs. 05 rb vs. 05 mv vs. rb

GRAFICA 2. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.125)



GRAFICA 3. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.25)



donde  $F_n^r(x)$ ,  $\tilde{F}_n(x)$  y  $\hat{F}_n(x)$  denotan las funciones de distribución empíricas de las transformadas predictivas, Rao-Blackwell y máximo verosímil, respectivamente.

A medida que  $n$  aumenta, todos los procesos van acercándose a cero. En general, el proceso (7) lo hace más lentamente que los otros dos y el (9) más rápido que los restantes. Esto permite conjeturar que el proceso

$$\sup_X \sqrt{n} |\tilde{F}_n(X) - F_{\pi,r}(X)|$$

converge a cero en probabilidad, lo que llevaría a que la distribución asintótica del proceso *predictivo* es la misma que la obtenida transformando con Rao-Blackwell o con máxima verosimilitud.

En resumen, de las simulaciones puede conjeturarse que el uso de la distribución predictiva como estimador de la función de distribución en el problema de Bondad de Ajuste para la distribución Gaussiana Inversa, da los mismos resultados que se obtienen al transformar con el estimador de Rao-Blackwell.

## CAPITULO 5

La primera referencia que se conoce sobre el uso de la *distribución predictiva* en Bondad de Ajuste es la de Guttman (1967), quien transforma el problema al de probar que la muestra fué generada por la distribución predictiva para una observación futura  $X_{n+1}$ , obtenida al usar el modelo especificado en la hipótesis original y una distribución inicial de referencia para  $\theta$ . Construye una prueba basada en la estadística  $X^2$  de Pearson, en donde los valores esperados son calculados utilizando a la predictiva ya mencionada.

El trabajo que aqui se presenta usa también la distribución predictiva para transformar a la muestra original, pero transformando a cada una de las observaciones de la muestra con su correspondiente predictiva, esto es, la calculada para esa observación. Con lo anterior se obtiene a una muestra (transformada) de variables aleatorias Uniformes en el  $(0, 1)$ . Esta transformación se basa en el trabajo de O'Reilly y Villegas (1987), quienes suponen un *modelo estructural* en el sentido de Fraser (1961, 1968), para el resultado central de su trabajo.

Un modelo estructural consiste en lo siguiente:

Sea  $(\mathcal{X}, G, \mathcal{P}, \Theta)$ , donde  $\mathcal{X}$ , el espacio muestral, es un subconjunto de  $\mathbf{R}$ ,  $G$  es un grupo topológico localmente compacto que actúa sobre  $\mathcal{X}$  y  $\mathcal{P}$  es una familia invariante de distribuciones de probabilidad indexada por  $\Theta$ .

La estructura anterior implica, entre otras cosas, lo siguiente:

- i. Existe  $\tilde{G}$  y un isomorfismo  $\phi : G \rightarrow \tilde{G}$ , tal que  $\tilde{G}$  es un grupo que actúa sobre  $\Theta$ ,
- ii. Si la distribución de  $X$  es  $P_\theta$  entonces la de  $gX$  es  $P_{\tilde{g}X}$  para toda  $X \in \mathcal{X}$ ,  $g \in G$  y  $\tilde{g} \in \tilde{G}$ .
- iii. Si además se supone que  $G = \Theta$ , de manera que  $\theta X \in \mathcal{X}$ , entonces existe  $\theta^{-1}$  tal que la distribución de  $\theta^{-1}X$  es  $P_e$ , con  $e$  el elemento unitario del grupo. A  $\theta^{-1}X$  se le llama *variable pivotal*, pues su distribución es independiente de  $\theta$  y por lo tanto conocida.

Un ejemplo sencillo de un modelo estructural como el presentado, es cuando  $\theta$  es un parámetro de localización y el grupo de transformaciones está formado por todas las

traslaciones definidas en  $\mathbf{R}$ .

La ventaja de esta estructura es que a partir de la variable pivotal y el grupo de transformaciones, puede generarse cualquier otro elemento de  $\mathcal{X}$  y por tanto, su distribución asociada.

Con esta estructura y algunos supuestos adicionales, O'Reilly y Villegas (1987) demuestran que

$$\tilde{F}(X_n|T_n) = F_{\pi}(X_n|T_{n-1}),$$

donde  $T_n$  es la estadística suficiente minimal.

El punto central de este procedimiento, es la suposición de que el modelo estadístico puede ser dotado de un estructura de grupo, situación que parece no cumplirse en el caso de la distribución Gaussiana Inversa.

Aunque en Chhikara y Folks (1989) se presentan algunas transformaciones que generan variables cuyas distribuciones son independientes de los parámetros, y que son usadas para calcular intervalos de confianza y realizar pruebas de hipótesis, estas distribuciones no pertenecen a la familia de distribuciones Gaussiana Inversa, y por tanto no son *variables pivotaes* en el sentido expuesto. Esto no demuestra que no exista un grupo de transformaciones con las características mencionadas anteriormente, y a pesar de no tener una demostración formal de este hecho, parece ser reconocida esta limitación en esta distribución.

Por otro lado, O'Reilly y Rueda (1992) demostraron que la distribución asintótica de los procesos empíricos inducidos por usar los estimadores de máxima verosimilitud o utilizar el estimador Rao-Blackwell en el caso de la distribución Gaussiana Inversa para transformar al  $(0,1)$ , es la misma. La pregunta inmediata fué: ¿qué pasa si se utiliza a la distribución predictiva como estimador?

Lo primero que se hizo fué encontrar una distribución inicial cuya distribución predictiva asociada resultara útil para el problema que se trataba de resolver. Se propuso una familia de distribuciones iniciales que contenía, entre otras, a las distribuciones de referencia reportadas en la literatura. Para algunas de estas distribuciones de referencia se demostró que no tenían todas las propiedades que se les atribuían. Por otro lado, la familia propuesta resultó invariante ante transformaciones uno a uno de los parámetros, lo

que permite su uso en distintas parametrizaciones de la distribución Gaussiana Inversa. Además, bajo condiciones muy generales, la familia de distribuciones finales de referencia que se induce, está constituida sólo por distribuciones propias, lo que asegura la existencia de las distribuciones predictivas correspondientes.

Independientemente de su utilización en el área de Bondad de Ajuste, estas distribuciones predictivas pueden ser útiles en problemas propios de la estadística Bayesiana.

Aunque no se pudo encontrar una forma analítica cerrada para las distribuciones predictivas ya mencionadas, no se asevera la imposibilidad de que ésta exista para alguna combinación de  $r$  y  $n$ , pero no se hicieron más intentos por encontrarla.

Dado lo anterior, sólo pudieron realizarse estudios, respaldados numéricamente, del comportamiento de la transformada predictiva y del proceso empírico asociado. Como resultado de estos estudios, se encontró que la transformada predictiva fué muy similar a la transformada de Rao-Blackwell, para algunos valores de  $r$ , y a medida que  $\phi$  y  $n$  aumentaban, el parecido era mucho mayor.

Sin embargo, lo relevante no es el parecido entre las diferentes estimaciones. Lo importante es que los resultados numéricos realizados indican que la distribución asintótica del proceso predictivo coincide con la correspondiente al estimar por máxima verosimilitud o Rao-Blackwell, lo que permite usarla como una transformación alternativa.

¿Cuál es la ventaja de tener otro estimador detrás de una nueva transformación? Un primer argumento a favor, es la existencia de transformadas predictivas en casos en donde el estimador de Rao-Blackwell no existe, como O'Reilly y Villegas (1987) muestran para la distribución  $U(X|0, \theta)$ , y desde luego para la distribución  $U(X|\theta_1, \theta_2)$ , en las que la predictiva asociada a la distribución de referencia usual si existe. En este caso, resulta relativamente sencillo constatar que las transformadas tienen una distribución (frecuencial) Uniforme en el  $(0, 1)$ , lo que resulta interesante.

Por otro lado, puede ocurrir que la convergencia del proceso empírico asociado usando un estimador sea más rápida que el correspondiente proceso empírico inducido por otra transformación, lo que permite mejores aproximaciones cuando el tamaño de muestra es pequeño. Finalmente, la potencia del procedimiento de prueba puede ser mejor.

Estos dos últimos puntos se sugieren como temas a tratar para tener más elementos en relación al uso de la transformada predictiva en el caso específico de la distribución

**Gaussiana Inversa.** Por último, aunque parece ser difícil, está el reto de proporcionar una demostración formal de la convergencia del proceso empírico utilizando la predictiva.

## REFERENCIAS

- Abramowitz, M. y Stegun, I.A. (1972). *Handbook of Mathematical Functions*. New York: Dover.
- Anderson, T.W. y Darling, D.A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.*, 23, 193-212.
- Anderson, T.W. y Darling, D.A. (1954). A test of goodness of fit. *J. Amer. Statist. Assoc.*, 49, 765-769.
- Banerjee, A.K. y Bhattacharyya, G.K. (1979). Bayesian results for the inverse Gaussian distribution with an application. *Technometrics*, 21, 247-251.
- Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Chichester: Wiley.
- Berger, J.O. y Bernardo, J.M. (1992). On the development of reference priors (con discusión). En *Bayesian Statistics 4*. (J.M. Bernardo, J.O. Berger, A.P. Dawid y A.F.M. Smith, eds.) Oxford: University Press, 35-60.
- Bernardo, J.M. (1976). The use of information in the design and analysis of scientific experimentation. *Ph. D. Thesis*. University of London.
- Bernardo, J.M. (1979). Reference posterior distributions for bayesian inference (con discusión). *J. R. Statist. Soc., B* 41, 113-147.
- Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.
- Box, G.E.P. y Muller, M.E. (1958). A note on the generation of random normal deviates. *Ann. Math. Statist.*, 29, 610-611.
- Cramér, H. (1928). On the composition of elementary errors. Second paper: Statistical applications. *Skand. Aktuarietidskr.*, 11, 141-180.
- Cramér, H. (1946). *Mathematical methods in statistics*. Princeton: University Press.
- Chhikara, R.S. y Folks, L. (1989). *The Inverse Gaussian Distribution*. New York: Marcel Dekker, Inc.
- D'Agostino, R.B. y Stephens, M.A. (eds.) (1986). *Goodnes-of-fit techniques*. New York: Marcel Dekker, Inc.

- David, F.N. y Johnson, N.L. (1948). The probability integral transformation when parameters are estimated from the sample. *Biometrika*, *35*, 182-192.
- DeGroot, M.H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.
- Doob, J.L. (1949). Heuristic approach to the Kolmogorov - Smirnov theorems. *Ann. Math. Statist.*, *20*, 393-403.
- Donsker, M.D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov - Smirnov theorems. *Ann. Math. Statist.*, *23*, 277-281.
- Durbin, J. (1968). The probability that the sample distribution function lies between two parallel straight lines. *Ann. Math. Statist.*, *39*, 398-411.
- Durbin, J. (1973a). Distribution theory for tests based on the sample distribution function. *Reg. Conf. Ser. Appl. Math.*, *9*. Philadelphia: SIAM.
- Durbin, J. (1973b). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* *1*, 279-290.
- Durbin, J. (1975). Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests of spacing. *Biometrika*, *62*, 5-22.
- Folks, L. y Chhikara, R.S. (1978). The inverse Gaussian distribution and its statistical application - a review. *J. R. Statist. Soc., B*, *40*, 263-289.
- Fraser, D.A.S. (1961). The fiducial method and invariance. *Biometrika*, *48*, 261-280.
- Fraser, D.A.S. (1968). *The structure of inference*. New York: Wiley.
- Gradshteyn, I.S. y Ryzhik, I.M. (1980). *Table of integrals, series and products*. New York: Academic Press.
- Guttman, I. (1967). The use of the concept of a future observation in goodness-of-fit problems. *J. R. Statist. Soc., B*, *29*, 83-100.
- Hartigan, J.A. (1964). Invariant prior distributions. *Ann. Math. Statist.*, *35*, 836-845.
- Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, *48*, 419-426.
- Jaynes, E.T. (1965). Prior probabilities. *IEEE Trans. Systems, Science & Cybernetics*, *SCC-4*, 227-291.

- Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proc. R. Soc. London, A 186*, 453-461.
- Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge distribuzione. *Giorn. Ist. Ital. Attuari.*, 4, 83-91.
- Lilliefors, H.W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Amer. Statist. Assoc.*, 62, 399-402.
- Lilliefors, H.W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *J. Amer. Statist. Assoc.* 64, 387-389.
- Lindley, D.V. (1956). On a measure of the information provided by an experiment. *Ann. Math. Statist.*, 27, 986-1005.
- Lindley, D.V. (1961). The use of prior probability distributions in statistical inference and decisions. *Proc. 4th. Berkeley Symp.*, 1, 436-468.
- Lockhart, R.A. y Stephens, M. (1985). Goodness-of-fit tests for the Gamma distribution. *Technical Report*. Department of Mathematics and Statistics, Simon Fraser University.
- Marshall, A.W. (1958). The small sample distribution of  $nw_n^2$ . *Ann. Math. Statist.*, 29, 307-309.
- McLachlan, N.W. (1955). *Bessel functions for engineers*. Oxford: Clarendon Press.
- Michael, J.R.; Schucany, W.R. y Haas, R.W. (1976). Generating random variables using transformations with multiple roots. *Amer. Statist.*, 30, 88-90.
- Moore, D.S. (1973). A note on Srinivasan's goodness-of-fit test. *Biometrika*, 60, 209-211.
- Novick, M.R. (1969). Multiparameter bayesian indifference procedures (con discusión). *J.R. Statist. Soc.*, B 31, 29-64.
- Novick, M.R. y Hall, W.J. (1965). A bayesian indifference procedure. *J. Amer. Statist. Assoc.* 60, 1104-1117.
- O'Reilly, F. y Quesenberry, C.P. (1973). The conditional probability integral transformations and applications to obtain chi-square goodness-of-fit tests. *Ann. Statist.*, 1, 74-83.
- O'Reilly, F. y Rueda, R. (1992). Goodness of fit for the inverse Gaussian distribution.

*Can. J. Statist.*, 20, 387-397.

O'Reilly, F. y Villegas, C. (1987). Conditional and predictive probability integral transforms. *Can. J. Statist.*, 15, 151-158.

Press, W.H.; Flannery, B.P.; Teukolsky, S.A. y Vetterling, W.T. (1989). *Numerical Recipes*. Cambridge: University Press.

Pollard, D. (1984). *Convergence of Stochastic Processes*. New York: Springer-Verlag.

Raiffa, H. y Schlaifer, R. (1961). *Applied Statistical Decision Theory*. Boston: Harvard Business School.

Rueda, R. (1980). Distribuciones de Referencia. *Tesis de Maestría en Estadística*. IIMAS - CCH, UNAM.

Rueda, R. (1988). Intervalos de confianza y de probabilidad en la distribución Gaussiana Inversa. *Aportaciones Matemáticas, Com.*, 5, 369-375.

Schrödinger, E. (1915). Zur theorie der fall- und steigversuche an teilchen mit Brownscher bewegung. *Phys. Ze.*, 16, 289-295.

Serfling, R.J. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Tech. J.*, 27, 379-423, 623-656.

Shorack, G.L. y Wellner, J.A. (1986). *Empirical processes with applications to statistics*. New York: Wiley.

Skorokhod, A.V. (1956). Limit theorems for stochastic processes. *Theor. Prob. Appl.* 1, 261-290.

Smirnov, N.V. (1939). Sur les écarts de la courbe de distribution empirique. *Mat. Sbornik*, 48, 3-26.

Smirnov, N.V. (1941). Approximate laws of distribution of random variables from empirical data. *Uspekhi Mat. Nauk.*, 10, 179-206. (En ruso).

Srinivasan, R. (1970). An approach to testing the goodness of fit of incompletely specified distributions. *Biometrika*, 57, 605-611.

- Stark, G.E. y Chhikara, R.S. (1988). Kolmogorov - Smirnov statistics for the inverse Gaussian distribution. *Conference Proceedings: International Conference on Applied Simulation and Modelling*. Galveston, Texas, Mayo 18-20, 1988.
- Stephens, M. (1974). EDF statistics for goodness-of-fit and some comparisons. *J. Amer. Statist. Assoc.*, 69, 730-737.
- Stephens, M. (1977). Goodness-of-fit for the extreme value distribution. *Biometrika*, 64, 583-588.
- Stephens, M. (1979). Test of fit for the logistic distribution based on the empirical distribution function. *Biometrika*, 66, 591-595.
- Stephens, M. (1986). Tests based on EDF statistics. En *Goodnes-of-fit techniques*. (R.B. D'Agostino y M. Stephens, eds.) New York: Marcel Dekker, Inc. 97-193.
- Stephens, M. y Maag, U.R. (1968). Further percentage points for  $W_n^2$ . *Biometrika*, 55, 428-430.
- Strecok, A.J. (1968). On the calculation of the inverse of the error function. *Mathematics of Computation*, 22, 144-158.
- Tweedie, M.C.K. (1945). Inverse statistical variates. *Nature*, 155, 453.
- Tweedie, M.C.K. (1956). Some statistical properties of inverse Gaussian distributions. *Virginia J. Sci.*, 7, 160-163.
- Villegas, C. (1971). On Haar priors. En *Foundations of statistical inference*. (V.P. Godambe y D.A. Sprott, eds.) Toronto: Holt, Rinehart & Winston. 409-414.
- Villegas, C. (1977). On the representation of ignorance. *J. Amer. Statist. Assoc.* 72, 651-654.
- Von Mises, R. (1931). *Wahrscheinlichkeitsrechnung*. Leipzig: Wein.
- Wald, A. (1947). *Sequential Analysis*. New York: Wiley.
- Walker, D.M. (1969). On the asymptotic behaviour of a posterior distribution. *J. R. Statist. Soc.*, B 31, 80-88
- Wichmann, B.A. y Hill, I.D. (1982). An efficient and portable pseudo-random number generator. *Applied Statistics*, 31, 188-190.

Zellner, A. (1977). Maximal data information prior distributions. En *New Developments in the Applications of Bayesian Methods*. (A. Aykaç y C. Brumat, eds.) Amsterdam: North-Holland. 114-132.

## APENDICE

**TABLA A1. DISTRIBUCION ASINTOTICA PARA  $A^2$   
EN LA DISTRIBUCION GAUSSIANA INVERSA**

$\phi/\alpha$	0.25	0.20	0.15	0.10	0.05	0.025	0.01
0	0.80	0.90	1.02	1.19	1.50	1.83	2.27
$2^{-10}$	0.80	0.89	1.01	1.18	1.49	1.82	2.26
$2^{-9}$	0.80	0.89	1.00	1.18	1.48	1.81	2.24
$2^{-8}$	0.79	0.88	0.99	1.17	1.47	1.79	2.22
$2^{-7}$	0.77	0.86	0.98	1.15	1.45	1.76	2.19
$2^{-6}$	0.75	0.84	0.95	1.12	1.41	1.71	2.13
$2^{-5}$	0.73	0.81	0.92	1.07	1.35	1.64	2.03
$2^{-4}$	0.69	0.77	0.87	1.01	1.27	1.53	1.90
$2^{-3}$	0.65	0.72	0.81	0.93	1.16	1.40	1.73
$2^{-2}$	0.60	0.66	0.74	0.85	1.05	1.26	1.54
$2^{-1}$	0.56	0.61	0.68	0.78	0.95	1.13	1.37
$2^0$	0.52	0.57	0.63	0.72	0.87	1.02	1.23
$2^1$	0.50	0.54	0.60	0.68	0.82	0.96	1.14
$2^2$	0.49	0.53	0.58	0.66	0.79	0.92	1.09
$2^3$	0.48	0.52	0.57	0.65	0.77	0.90	1.07
$2^4$	0.48	0.52	0.57	0.64	0.76	0.89	1.05
$2^5$	0.47	0.51	0.56	0.64	0.76	0.88	1.05
$\infty$	0.47	0.51	0.56	0.63	0.76	0.88	1.04

Tomada de O'Reilly y Rueda, 1992

$\alpha$  = Nivel de significancia

$$\phi = \frac{\lambda}{\mu}$$

```

c*****
c  programa ppf01.for
c  se generan muestras de variables aleatorias gaussiana inversa,
c  para distintos valores de los parámetros y tamaños de muestra.
c  los resultados se escriben en archivos definidos por el usuario
c  subrutinas que usa: ginv generador de va gaussianas inversas
c                      dnor generador de va normales (0,1)
c                      dun generador de va uniformes (0,1)
c*****
c  implicit double precision ( a - h , o - z )
c  dimension z ( 170 ) , gi ( 170 )
c  character*10 narch ( 10 ) , esc
c*****
c  se fija el número de observaciones a generar, se hace  $\mu = 1$  y
c  se piden las semillas iniciales, para generar n variables aleato-
c  rias gaussiana inversa de parámetros  $\phi$ , l
c*****
c  l = 0
c  1 write ( * , 70 )
c    read ( * , * ) n
c    write ( * , 80 )
c    read ( * , * ) m
c    write ( * , 90 )
c    read ( * , * ) dla
c    if ( l .eq. 0 ) then
c      write ( * , 100 )
c      read ( * , * ) ix , iy , iz
c    endif
c    dmua = 1
c*****
c  se pregunta los nombres de los archivos donde se guardarán las
c  observaciones generadas
c*****
c  do 2 i = 1 , m

```

```

write ( * , 110 ) i
read ( * , 120 ) narch ( i )
2 continue
c*****
c comienza el loop principal. primero se abre el archivo
c*****
do 4 j = 1 , m
open( unit=5, file=narch(j), status='new' )
c*****
c se generan variables aleatorias gaussiana inversa y se guardan
c en el archivo designado
c*****
call ginv (ix , iy , iz , n , dmU , dla , gi )
write ( 5 , 130 ) n , dla
do 3 i = 1 , n
z ( i ) = gi ( i )
3 write ( 5 , 140 ) z ( i )
c*****
c se cierra el archivo de escritura , se incrementa en potencias
c de dos el valor de lambda y se termina el loop principal
c*****
close ( unit = 5 , err = 600 , iostat = nerr )
dla = 2d0 * dla
4 continue
c*****
c se abre un archivo donde se escribe el valor de m y los nombres
c de los archivos generados
c*****
write ( * , 115 )
read ( * , 120 ) esc
open ( 10 , file = esc , status = 'new ' )
write ( 10 , 150 ) m
do 5 k = 1 , m
write ( 10 , 160 ) narch ( k )

```

```

5 continue
  write ( * , 170 )
  read ( * , * ) 1
  if ( 1.eq. 1 ) goto 1
c*****
c  formatos
c*****
70 format ( 1x , ' dar el valor de n : ' , $ )
80 format ( 1x , ' cuántos valores de los parámetros ? ' , $ )
90 format ( 1x , ' dar el primer valor de lambda : ' , $ )
100 format ( 1x , ' dar las semillas iniciales : ' , $ )
110 format ( 1x , ' nombre del archivo número ', i3 , ' : ' , $ )
115 format ( 1x , ' nombre del archivo de escritura : ' , $ )
120 format ( a10 )
130 format ( 1x , i3 , 2x , f10.6 )
140 format ( 1x , f10.6 )
150 format ( 1x , i3 )
160 format ( 1x , a10 )
170 format ( 1x , ' quieres otro caso ? ' , $ )
600 write(*,610) nerr
610 format ( 1x , 'error número ',i2,' en el cierre de un archivo ')
c*****
  stop
  end
c*****
c*****
c  subrutinas
c*****
c*****
  subroutine dnoe(i,j,n,z)
c  generador de variables aleatorias normales estandar
c  con el metodo de box y muller (1958)
c  entrada: n    número de variables a generar
c             de  de  semillas iniciales

```

5 continue

write ( \* , 170 )

read ( \* , \* ) 1

if ( 1 .eq. 1 ) goto 1

c\*\*\*\*\*

c     formatos

c\*\*\*\*\*

70 format ( 1x , ' dar el valor de n : ' , \$ )

80 format ( 1x , ' cuántos valores de los parámetros ? ' , \$ )

90 format ( 1x , ' dar el primer valor de lambda : ' , \$ )

100 format ( 1x , ' dar las semillas iniciales : ' , \$ )

110 format ( 1x , ' nombre del archivo número ' , i3 , ' : ' , \$ )

115 format ( 1x , ' nombre del archivo de escritura : ' , \$ )

120 format ( a10 )

130 format ( 1x , i3 , 2x , f10.6 )

140 format ( 1x , f10.6 )

150 format ( 1x , i3 )

160 format ( 1x , a10 )

170 format ( 1x , ' quieres otro caso ? ' , \$ )

600 write(\*,610) nerr

610 format ( 1x , 'error número ',i2,' en el cierre de un archivo ' )

c\*\*\*\*\*

stop

end

c\*\*\*\*\*

c\*\*\*\*\*

c     subrutinas

c\*\*\*\*\*

c\*\*\*\*\*

subroutine dnor(ix,iy,iz,n,z)

c     generador de variables aleatorias normales estandar

c     con el metodo de box y muller (1950)

c     entradas: n        número de variables a generar

c         ix,iy,iz semillas iniciales

c salida: z vector de dimensión n con las va-  
 c riables aleatorias  
 c subrutina que usa: dun(ix,iy,iz,u) generador de va-  
 c riables uniformes

c\*\*\*\*\*

```

implicit double precision(a-h,o-z)
dimension z(n),un(2)
pi = 1d0
pi = 4d0 * datan ( pi )
pi = 2d0 * pi
l = mod ( n , 2 )
if ( l.eq. 0 ) k = n
if ( l.eq. 1 ) k = n - 1
do 2 j = 1 , k , 2
do 1 j1 = 1 , 2
call dun ( ix , iy , iz , u )
1 un ( j1 ) = u
z ( j ) = ( dsqrt ( ( - 2d0 ) * dlog ( un ( 1 ) ) ) ) * dcos
* ( pi * un ( 2 ) )
2 z ( j + 1 ) = ( dsqrt ( ( - 2d0 ) * dlog ( un ( 1 ) ) ) ) *
* dsin ( pi * un ( 2 ) )
if ( l.eq. 0 ) goto 4
do 3 j1 = 1 , 2
call dun ( ix , iy , iz , u )
3 un ( j1 ) = u
if ( un ( 1 ) .lt. un ( 2 ) ) then
z ( n ) = ( dsqrt ( ( - 2d0 ) * dlog ( un ( 1 ) ) ) ) * dcos
* ( pi * un ( 2 ) )
else
z ( n ) = ( dsqrt ( ( - 2d0 ) * dlog ( un ( 1 ) ) ) ) *
* dsin ( pi * un ( 2 ) )
endif
4 return
end

```

```

c*****
c*****
  subroutine dun(ix,iy,iz,u)
c  generador de variables aleatorias uniformes
c  en el intervalo (0,1), usando el algoritmo
c  as 183 de applied statistics (1982) vol 31
c  entrada: ix,iy,iz semillas iniciales
c  salida:      u número aleatorio en (0,1)
c*****
  implicit double precision(a-h,o-z)
  ix = 171 * mod ( ix , 177 ) - 2 * ( ix / 177 )
  iy = 172 * mod ( iy , 176 ) - 35 * ( iy / 176 )
  iz = 170 * mod ( iz , 178 ) - 63 * ( iz / 178 )
  if ( ix .le. 0 ) ix = ix + 30269
  if ( iy .le. 0 ) iy = iy + 30307
  if ( iz .le. 0 ) iz = iz + 30323
  u = dmod ( dble ( ix ) / 30269d0 + dble ( iy ) / 30307d0
  *+ dble ( iz ) / 30323d0 , dble ( 1 ) )
  if ( u .ge. 1d0 ) u = 0.999999d0
  if ( u .le. 0d0 ) u = 1d-6
  return
  end
c*****
c*****
  subroutine ginv (ix,iy,iz,n,dmu,dla,gi)
c  generador de variables aleatorias de la distribución gaussiana
c  inversa, usando el algoritmo de michael et al (1976)
c  entradas: ix,iy,iz semillas iniciales
c           n      número de observaciones que se desean generar
c           dmu, dla parámetros de la distribución
c  salida:  gi      vector de dimensión n, conteniendo a las va-
c           riables aleatorias gi(dmu,dla)
c  subrutinas que usa: dun(ix,iy,iz,u) generador de variables alea-
c           torias uniformes en el (0,1)

```

```

c          dnor(ix,iy,iz,n,z) generador de variables a-
c          leatorias normales
c*****
      implicit double precision (a-h,o-z)
      dimension gi(n) , z(250) , y(250)
      call dnor ( ix , iy , iz , n , z )
      do 1 i = 1 , n
      gi ( i ) = 0d0
      call dun ( ix , iy , iz , u )
1 y ( i ) = u
      const = dmU / ( 2d0 * dla )
      do 2 j = 1 , n
      z ( j ) = dmU * z ( j ) * z ( j )
      x1 = z ( j ) - dsqrt ( z ( j ) * ( 4d0 * dla + z ( j ) ) )
      x1 = dmU + const * x1
      p1 = dmU / ( dmU + x1 )
      x = x1
      if ( y ( j ) .ge. p1 ) x = dmU * dmU / x1
2 gi ( j ) = x
      return
      end
c*****

```

```

c*****
c  programa ppf02.for
c  programa que calcula la distribución predictiva de la dis-
c  tribución gaussiana inversa. como subrutina de integración
c  se usa una modificación a la regla del trapecio
c  subrutinas que usa:
c      arit(n,x,a) suma de n observaciones
c      armo(n,x,b) suma de los inversos de n obser-
c                   caciones
c      dest ( ... ) estimadores máximo verosímiles
c      fac(x)      función logaritmo del factorial
c                   de x-1 (logaritmo de la función
c                   gamma evaluada en x)
c      fden(t)     función integrando del denomina-
c                   dor
c      ffnu(u)     función integrando del numerador,
c                   que depende de fnum(t)
c      fnum(t)     función integrando del numerador
c      qtrap (...) subrutina de integración basada
c                   en la regla del trapecio
c      sort ( ... ) ordenamiento de observaciones
c      tmidpt (..) regla abierta del trapecio
c*****
c*****
c      implicit double precision ( a - h , o - z )
c      dimension pr ( 170 ) , coc ( 170 )
c      dimension z ( 170 ) , erre ( 3 )
c      character*10 narch1 , narch2 ( 10 ) , narch3 ( 50 ) , ent
c      external fden,fnum,ffnu
c      common /eme/ rn,rm
c      common /par/ a,b
c      common /parn/ an,bn
c      data pi /3.14159265358979d0/
c      data erre / 0.5d0 , 0.8d0 , 1d0 /

```

```

open ( 12 , file = 'basura.dat', status = 'old' )
c*****
c lee el nombre del archivo donde están los nombres de los
c archivos que contienen la información y los asigna
c*****
11 write ( * , 100 )
   read ( * , 110 ) narch1
   open (10, file = narch1 , status = 'old' )
   read ( 10 , 90 ) m
   read ( 10 , 80 ) ( narch2 ( i ) , i = 1 , m )
   close ( unit = 10 , err = 600 , iostat = nerr )
   write ( * , 100 )
   read ( * , 110 ) ent
   open ( 50 , file = ent , status = 'old' )
   read ( 50 , 90 ) na
   read ( 50 , 110 ) ( narch3 ( j ) , j = 1 , na )
   close ( unit = 50 , err = 600 , iostat = nerr )
c*****
c inicia preguntando el índice de erre, que es lo que define al
c valor de r
c*****
   write ( * , 105 )
   read ( * , * ) ir
   r = erre ( ir )
c*****
c inicia do principal ( sobre el número de archivos abiertos )
c*****
do 3000 inc = 1 , m
c*****
c abre el archivo correspondiente y lee la información relevante
c*****
open ( 20 , file = narch2 ( inc ) , status = 'old' )
read ( 20 , 120 ) nin , dla
read ( 20 , 130 ) ( z ( i ) , i = 1 , nin )

```

```

close ( unit = 20 , err = 600 , iostat = nerr )
n = 10
c*****
c   cálculo de los estimadores de lambda y mu, para muestras de ta-
c   maño 10 ( 10 ) 150. inicia loop de acumulación de información.
c   lee las primeras nin observaciones y las ordena
c*****
do 2000 is = 1 , 15
ik = ( inc - 1 ) * 15 + is
write ( * , 160 ) ik
call sort ( n , z )
call dest ( n , z , dmue , dlae )
c*****
c   cálculo del estimador predictivo
c   cálculo de constantes
c*****
rn = dble ( n )
rm = rn - 2d0 * r
nn = n - 1
rrm = ( rm + 1d0 ) / 2d0
rrm = fac ( rrm )
rrm = dexp ( rrm )
dpi = 2d0 * pi
dpi = dsqrt ( dpi )
dpi = 1d0 / dpi
gar = rm / 2d0 + 1d0
gar = fac ( gar )
gar = dexp ( gar )
c*****
c   comienza el loop principal, para cada i se define a Xn , se
c   calculan la suma de las n - 1 observaciones y la de sus in-
c   versos
c*****
do 1000 j = 1 , n

```

```

c*****
c  generación de la observación que se va a transformar
c*****
  pr ( n ) = z ( j )
  if ( j .eq. 1 ) then
    do 1 k = 2 , n
1   pr ( k - 1 ) = z ( k )
    else
    do 2 k = 1 , j - 1
2   pr ( k ) = z ( k )
    do 3 kk = j + 1 , n
3   pr ( kk - 1 ) = z ( kk )
  endif
c*****
c  se calculan las estadísticas suficientes con las n - 1 observa-
c  ciones. estas estadísticas son la suma de las observaciones (a)
c  y la suma de sus inversos (b)
c*****
  call arit ( nn , pr , a )
  call armo ( nn , pr , b )
c*****
c  se calcula la predictiva. primero se calcula la integral del
c  numerador, después la del denominador y por último el cociente
c  entre ellas
c*****
  dlim = pr ( n )
  call qtrap ( fnu , 0d0 , dlim , su )
  dnum = su * gar * dpi
  call qtrapss ( fden , 0d0 , 1d0 , suu )
  dden = suu * rrm
  coc ( j ) = dnum / dden
  write ( * , 170 ) j
c*****
c  termina loop principal

```

```

c*****
1000 continue
c*****
c  escritura de resultados : primero abre al archivo donde va a es-
c  cribir los resultados, después escribe n, lambda y las transfor-
c  madas
c*****
      open ( 30, file = narch3 ( ik ) , status = 'new' )
      write ( 30 , 140 ) n , dla , dmue , dlae
      write ( 30 , 150 ).( coc ( kk ) , kk = 1 , n )
      close ( unit = 30 , err = 600 , iostat = nerr )
c*****
c  terminan los do's
c*****
      n = n + 10
      write ( * , 160 ) ik
2000 continue
c*****
c  termina do principal
c*****
3000 continue
c*****
c  se pregunta si se quiere otra 'erre'
c*****
      close ( unit = 12 , err = 600 , iostat = nerr )
      write ( * , 190 )
      read ( * , * ) l
      if ( l .eq. 1 ) goto 11
c*****
c  formatos
c*****
      80 format ( 1x , a10 )
      90 format ( 1x , i3 )
      100 format ( 1x , ' nombre del archivo : ' , $ )

```

```

105 format ( 1x , ' cuánto vale r ? ' , $ )
110 format ( a10 )
120 format ( 1x , i3 , 2x , f10.6 )
130 format ( 1x , f10.6 )
140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
150 format ( 1x , f8.6 )
160 format ( '+ caso ' , i3 )
170 format ( '+ observación ' , i4 )
190 format ( 1x , ' quieres otro caso ? ' , $ )
600 write ( * , 610 ) nerr
610 format(1x,' error número ' ,i2, ' en el cierre de un archivo ')
c*****
c   fin del programa
c*****
c   stop
c   end
c*****
c*****
c   subrutinas
c*****
c*****
c   subroutine arit(n,z,a)
c   cálculo de la suma de n observaciones
c*****
c   implicit double precision (a-h,o-z)
c   dimension z(n)
c   sum = 0d0
c   do 1 i=1,n
1 sum = sum + z(i)
c   a = sum
c   return
c   end
c*****
c*****

```

```

subroutine armo(n,z,b)
c cálculo de la suma de los inversos de n observaciones
c*****
implicit double precision (a-h,o-z)
dimension z(n)
sum = 0d0
do 1 i = 1 , n
1 sum = sum + 1 / z(i)
b = sum
return
end
c*****
c*****
subroutine dest (n,z,dmue,dlae)
c cálculo de los estimadores máximo verosímiles de los parámetros
c de una distribución gaussiana inversa
c entradas: n número de observaciones
c z vector de dimensión n, con las observaciones
c salidas: dmue estimador máximo verosímil de mu
c dlae estimador máximo verosímil de lambda
c*****
implicit double precision (a-h,o-z)
dimension z(n)
sum=0d0
sumi=0d0
do 1 i=1,n
sum=sum+z(i)
1 sumi=sumi+1d0/z(i)
dmue=sum/dble(n)
dlae=dble(n)/(sumi-dble(n)/dmue)
return
end
c*****
c*****

```

function fac(x)

- c cálculo del logaritmo del factorial de x-1. se usa
- c la función gamma si x es menor que 150 y la
- c aproximación de sterling para x mayor que 150.
- c entrada: x argumento donde se evalúa la función
- c salida: fac logaritmo del factorial de x - 1

c\*\*\*\*\*

implicit double precision (a-h,o-z)

data a1,a2,a3,a4,a5/0.0833333333333333d0,0.0027777777777778d0,  
\*0.000793650793651d0,0.000595238095238d0,0.918938533204673d0/

pi = 1d0

pi = 4d0 \* datan ( pi )

if ( x .eq. 0d0 ) then

fac = 1d0

return

endif

if ( x .gt. 150d0 ) goto 1

y1 = ( x - 0.5d0 ) \* dlog ( x ) - x + a5

y2 = 1d0 / x

y3 = y2 \* y2

r = ( a1 - y3 \* ( a2 - y3 \* ( a3 - a4 \* y3 ) ) ) \* y2

fac = r + y1

return

1 spi = 2d0 \* pi

x = x - 1d0

fac = ( x + 0.5d0 ) \* dlog ( x ) - x + 0.5d0 \* dlog ( spi )

return

end

c\*\*\*\*\*

c\*\*\*\*\*

function fden ( t )

- c función integrando del denominador

c\*\*\*\*\*

implicit double precision ( a - h , o - z )

```

common /eme/ rn,rm
common /par/ a,b
t2 = t * t
a2 = a / 2d0
b2 = b / 2d0
espd = ( rm + 1d0 ) / 2d0
espn = rm - 0.5d0
f1n = espn * dlog ( t )
f1d = t * ( b2 * t - ( rn - 1d0 ) ) + a2
f1d = espd * dlog ( f1d )
f1 = f1n - f1d
f1 = dexp ( f1 )
f2d = t2 * ( a2 * t2 - ( rn - 1d0 ) ) + b2
f2d = espd * dlog ( f2d )
f2 = dlog ( 2d0 ) - f2d
f2 = dexp ( f2 )
fden = f1 + f2
return
end

```

c\*\*\*\*\*

c\*\*\*\*\*

```
function ffnu ( u )
```

c función integrando del numerador

c\*\*\*\*\*

```
implicit double precision ( a - h , o - z )
```

```
external fnum
```

```
common /eme/ rn,rm
```

```
common /parn/ an,bn
```

```
common /par/ a,b
```

```
an = a + u
```

```
bn = b + 1d0 / u
```

```
call qtraps ( fnum , 0d0 , 1d0 , ssu )
```

```
ep = 3d0 / 2d0
```

```
v = 1d0 / u
```

```

v = v ** ep
ffnu = v * ssu
return
end

```

```

c
c*****
c*****

```

```

function fnum ( t )

```

```

c primera función integrando del numerador

```

```

c*****

```

```

implicit double precision ( a - h , o - z )

```

```

common /éme/ rn,rm

```

```

common /parn/ an,bn

```

```

t2 = t * t

```

```

a2 = an / 2d0

```

```

b2 = bn / 2d0

```

```

espd = ( rm + 2d0 ) / 2d0

```

```

espn = rm + 0.5d0

```

```

fln = espn * dlog ( t )

```

```

fld = t * ( b2 * t - rn ) + a2

```

```

fld = espd * dlog ( fld )

```

```

f1 = fln - fld

```

```

f1 = dexp ( f1 )

```

```

f2d = t2 * ( a2 * t2 - rn ) + b2

```

```

f2d = espd * dlog ( f2d )

```

```

f2 = dlog ( 2d0 ) - f2d

```

```

f2 = dexp ( f2 )

```

```

fnum = f1 + f2

```

```

return

```

```

end

```

```

c*****

```

```

c*****

```

```

subroutine qtrap ( func , a , b , s )

```

```

c subrutina de integración basada en la regla del trapecio

```

```

c*****
  implicit double precision ( a - h , o - z )
  external func
  parameter ( eps = 1d-3 , jmax = 20 )
  olds = - 1d30
  do 11 j= 1 , jmax
    call tmidpt ( func , a , b , s , j )
    if ( dabs ( s - olds ) .lt. eps * dabs ( olds ) ) return
    olds = s
11  continue
  pause 'número máximo de iteraciones alcanzado'
  end
c*****
c*****
  subroutine qtraps ( func , a , b , s )
c  subrutina de integración basada en la regla del trapecio
c*****
  implicit double precision ( a - h , o - z )
  external func
  parameter ( eps = 1d-3 , jmax = 20 )
  olds = - 1d30
  do 11 j= 1 , jmax
    call tmidptss ( func , a , b , s , j )
    if ( dabs ( s - olds ) .lt. eps * dabs ( olds ) ) return
    olds = s
11  continue
  pause 'número máximo de iteraciones alcanzado'
  end
c*****
c*****
  subroutine qtrapss ( func , a , b , s )
c  subrutina de integración basada en la regla del trapecio
c*****
  implicit double precision ( a - h , o - z )

```

```

external func
parameter ( eps = 1d-3 , jmax = 20 )
olds = - 1d30
do 11 j= 1 , jmax
    call tmidpts ( func , a , b , s , j )
    if ( dabs ( s - olds ) .lt. eps * dabs ( olds ) ) return
    olds = s
11 continue
pause 'número máximo de iteraciones alcanzado'
end
c*****
c*****
    subroutine sort ( n,z)
c    cálculo de las estadísticas de orden de una muestra de n
c    observaciones, usando el algoritmo de heapsort (press et
c    al,1985).
c    entradas: n número de observaciones en la muestra
c             z vector de dimensión n, que contiene a la muestra
c             original
c    salida:  z vector de dimensión n, formado con las estadísti-
c            cas de orden
c*****
    implicit double precision (a-h,o-z)
    dimension z(n)
    l = n / 2 + 1
    ir = n
10 continue
    if ( l .gt. 1 ) then
        l = l - 1
        rra = z ( l )
    else
        rra = z ( ir )
        z ( ir ) = z ( l )
        ir = ir - 1

```

```

        if ( ir .eq. 1 ) then
            z ( 1 ) = rra
            goto 100
        endif
    endif
    i = 1
    j = 1 + 1
20    if ( j .le. ir ) then
        if ( j .lt. ir ) then
            if ( z ( j ) .lt. z ( j + 1 ) ) j = j + 1
        endif
        if ( rra .lt. z ( j ) ) then
            z ( i ) = z ( j )
            i = j
            j = j + j
        else
            j = ir + 1
        endif
        goto 20
    endif
    z ( i ) = rra
    goto 10
100 return
end

```

c\*\*\*\*\*

c\*\*\*\*\*

subroutine tmidpt ( func , a , b , s , n )

c método del trapecio, como subrutina de integración

c\*\*\*\*\*

implicit double precision ( a - h , o - z )

external func

if ( n .eq. 1 ) then

s = ( b - a ) \* func ( 0.5d0 \* ( a + b ) )

it = 1

```

else
    tnm = it
    del = ( b - a ) / ( 3d0 * tnm )
    ddel = del + del
    x = a + 0.5d0 * del
    sum = 0d0
    do 11 j = 1 , it
        sum = sum + func ( x )
        x = x + ddel
        sum = sum + func ( x )
        x = x + del
11    continue
    s = ( s + ( b - a ) * sum / tnm ) / 3d0
    it = 3 * it
endif
return
end

c*****
c*****
subroutine tmidpts ( func , a , b , s , n )
c    método del trapecio, como subrutina de integración
c*****
implicit double precision ( a - h , o - z )
external func
if ( n .eq. 1 ) then
    s = ( b - a ) * func ( 0.5d0 * ( a + b ) )
    it = 1
else
    tnm = it
    del = ( b - a ) / ( 3d0 * tnm )
    ddel = del + del
    x = a + 0.5d0 * del
    sum = 0d0
    do 11 j = 1 , it

```

```

        sum = sum + func ( x )
        x = x + ddel
        sum = sum + func ( x )
        x = x + del
11  continue
    s = ( s + ( b - a ) * sum / tnm ) / 3d0
    it = 3 * it
endif
return
end
c*****
c*****
    subroutine tmidptss ( func , a , b , s , n )
c  método del trapecio, como subrutina de integración
c*****
    implicit double precision ( a - h , o - z )
    external func
    if ( n .eq. 1 ) then
        s = ( b - a ) * func ( 0.5d0 * ( a + b ) )
        it = 1
    else
        tnm = it
        del = ( b - a ) / ( 3d0 * tnm )
        ddel = del + del
        x = a + 0.5d0 * del
        sum = 0d0
        do 11 j = 1 , it
            sum = sum + func ( x )
            x = x + ddel
            sum = sum + func ( x )
            x = x + del
11  continue
        s = ( s + ( b - a ) * sum / tnm ) / 3d0
        it = 3 * it

```

```
endif  
return  
end
```

```
c*****
```

```

c*****
c  programa pf03.for
c  transformaciones al (0,1) usando el estimador de rao-blackwell,
c  en la gaussiana inversa. en el archivo nom.dat se leen los nom-
c  bres de los archivos donde se encuentran las observaciones que
c  se van a transformar.
c*****
c*****
c  subrutinas que usa: dest ( ... ) estimación de parámetros
c                      erfc(x) función complemento error
c                      rab(.....) estimador rao-blackwell para x(n)
c                      sort ( ... ) subrutina de ordenación
c                      tstu(gl,x) distribución student
c*****
c*****
c  implicit double precision (a-h,o-z)
c  dimension z ( 170 ) , frb ( 170 )
c  character*10 narch1 ; narch2(10) , narch3(50) , ent
c  common /val/ dmue,dlae
c  open ( 12 , file = 'basura.dat', status = 'old' )
c*****
c  lee el nombre del archivo donde están los nombres de los
c  archivos que contienen la información y los asigna
c*****
c  write ( * , 100 )
c  read ( * , 110 ) narch1
c  open (10, file = narch1 , status = 'old' )
c  read ( 10 , 90 ) m
c  read ( 10 , 80 ) ( narch2 ( i ) , i = 1 , m )
c  close ( unit = 10 , err = 600 , iostat = nerr )
c  write ( * , 100 )
c  read ( * , 110 ) ent
c  open ( 50 , file = ent , status = 'old' )
c  read ( 50 , 90 ) na

```

```

      read ( 50 , 110 ) ( narch3 ( j ) , j = 1 , na )
      close ( unit = 50 , err = 600 , iostat = nerr )
c*****
c  inicia do principal ( sobre el número de archivos )
c*****
      do 2000 inc = 1 , m
c*****
c  abre el archivo correspondiente y lee la información relevante
c*****
      open ( 20 , file = narch2 ( inc ) , status = 'old' )
      read ( 20 , 120 ) nin , dla
      read ( 20 , 130 ) ( z ( i ) , i = 1 , nin )
      close ( unit = 20 , err = 600 , iostat = nerr )
      n = 10
c*****
c  cálculo de los estimadores de lambda y mu, para muestras de ta-
c  maño 10 ( 10 ) 150. inicia loop de acumulación de información.
c  lee las primeras nin observaciones y las ordena
c*****
      do 1000 is = 1 , 15
      ik = ( inc - 1 ) * 15 + is
      write ( * , 160 ) ik
      call sort ( n , z )
      call dest ( n , z , dmue , dlae )
c*****
c  cálculo del estimador rao-blackwell
c*****
      call rab ( n , z , frb )
c*****
c  escritura de resultados : primero abre al archivo donde va a es-
c  cribir los resultados, después escribe n, lambda y las transfor-
c  madas
c*****
      open ( 30 , file = narch3 ( ik ) , status = 'new' )

```

```

write ( 30 , 140 ) n , dla , dmue , dlae
write ( 30 , 150 ) ( frb ( kk ) , kk = 1 , n )
close ( unit = 30 , err = 600 , iostat = nerr )
c*****
c termina do de acumulación de información
c*****
n = n + 10
1000 continue
c*****
c termina do principal
c*****
2000 continue
close ( unit = 12 , err = 600 , iostat = nerr )
c*****
c formatos
c*****
80 format ( 1x , a10 )
90 format ( 1x , i3 )
100 format ( 1x , ' nombre del archivo : ' , $ )
110 format ( a10 )
120 format ( 1x , i3 , 2x , f10.6 )
130 format ( 1x , f10.6 )
140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
150 format ( 1x , f8.6 )
160 format ( '+ , caso ' , i3 )
600 write ( * , 610 ) nerr
610 format(1x,' error número ', i2, ' en el cierre de un archivo ')
stop
end
c*****
c*****
c subrutinas
c*****
c*****

```

```

function erfc(x)
c cálculo de la función erfc -complemento de la función
c error-, usando el algoritmo 7.1.26 de abramowitz y ste-
c gun (1972)
c entrada: x valor donde se evalua la función
c salida: erfc valor de la función en el punto x
c*****
implicit double precision (a-h,o-z)
data a1,a2,a3,a4,a5,p/.254829592d0,-.284496736d0,1.421413741d0,
*-1.453152027d0,1.061405429d0,.3275911d0/
te = 1d0 / ( 1d0 + p * x )
ex = x * x
erfc = ( te * ( a1 + te * ( a2 + te * ( a3 + te * ( a4 + a5 * te )
* ) ) ) * dexp ( -ex )
return
end
c*****
c*****
subroutine rab (n,z,frb)
c cálculo del estimador por rao-blackwell de la distribu-
c ción gaussiana inversa
c entradas: n número de observaciones en la muestra
c z vector de longitud n, que contiene a la
c muestra ordenada
c salida: frb vector de dimensión n conteniendo a la
c función de distribución estimada por
c rao-blackwell y evaluada en z(n)
c subrutinas que usa: function tstu (gl,x) función de dis-
c tribución student estándar con
c gl grados de libertad, evalua-
c da en x
c*****
implicit double precision (a-h,o-z)
common /val/dmue,dlae

```

```

dimension z(n),frb(n)
external tstu,erfc
rn = dble ( n )
rn1 = rn - 1d0
rn2 = rn - 2d0
rn3 = rn - 3d0
rns = dsqrt ( rn * rn2 )
rnc = rn3 / 2d0
pr = 1d0 + 4d0 * rn1 * dlae / ( rn * rn * dmue )
pr1 = ( rn2 / rn ) * ( pr ** rnc )
do 1 i = 1 , n
co = ( rn / dlae ) * ( rn - z ( i ) / dmue ) * z ( i )
co = co - rn * ( 1d0 - z ( i ) / dmue ) * ( 1d0 - z ( i ) / dmue )
co = dsqrt ( co )
w1 = rns * ( z ( i ) / dmue - 1d0 ) / co
w2 = rns * ( 1d0 + rn2 * z ( i ) / ( rn * dmue ) )
w2 = - w2 / co
frb(i) = tstu ( rn2 , w1 ) + pr1 * tstu ( rn2 , w2 )
100 format ( ' + voy en la observación ', i3 )
1 continue
return
end

```

```

c*****
c*****

```

function tstu(gl,x)

```

c cálculo de la función de distribución student estándar con
c gl grados de libertad, evaluada en el punto x
c entradas: gl grados de libertad
c x valor donde se evalua la función
c salida: tstu valor de la función con gl grados de liber-
c tad, en el punto x
c subrutina que usa: function erfc(x) complemento de la fun-
c ción error

```

```

c*****

```

```

implicit double precision (a-h,o-z)
external erfc
t = dabs(x)
an = gl
m = idint(an)
t = t*t
y = t/an
b = 1d0+y
if ( ( an .ne. m .or. an .ge. 20d0 ) .and. t .lt. an .or. an .gt.
* 200d0 ) goto 4
if ( an .lt. 20d0 .and. t .lt. 4d0 ) goto 6
a = 1d0
y = an
xj = 0d0
z = 0d0
1 if ( a .eq. z ) goto 2
xj = xj + 2d0
z = a
y = y * ( xj - 1d0 ) / ( b * xj )
a = a + y / ( an + xj )
goto 1
2 if ( an .le. 1d0 .or. a .lt. 1d-30 ) goto 3
a = ( an - 1d0 ) / ( b * an ) * a
an = an - 2d0
goto 2
3 if ( an .ne. 0d0 ) a = dsqrt ( b ) * 0.6366198d0 * a / b
q = a
goto 9
4 w = b - 1d0
if ( w .ne. 0d0 ) y = y * ( dlog ( b ) / w )
a = an - 0.5d0
b = 48d0 * a * a
y = y * a
y = ( ( ( ( ( -0.4d0 * y - 3.3d0 ) * y - 24d0 ) * y - 85.5d0 ) /

```

```

* ( 0.8d0 * ( y * y ) + 100d0 + b ) + y + 3d0 ) / b + 1d0 ) *
* dsqrt ( y )
if ( y .lt. 18.8125d0 ) goto 5
q = 0d0
goto 9
5 q = erfc ( y * 0.7071068d0 )
goto 9
6 y = dsqrt ( y )
a = y
if ( an .eq. 1d0 ) a = 0d0
7 an = an - 2d0
if ( an .le. 1d0 ) goto 8
a = ( an - 1d0 ) / ( b * an ) * a + y
goto 7
8 if ( an .eq. 0d0 ) a = a / dsqrt ( b )
if ( an .ne. 0d0 ) a = ( datan ( y ) + a / b ) * 0.6366198d0
q = 1d0 - a
9 tstu = 1d0 - q / 2d0
if ( x .lt. 0d0 ) tstu = 1d0 - tstu
return
end
c*****

```

```

c*****
c  programa ppf04.for
c  transformaciones al (0,1) usando el estimador máximo verosímil,
c  en la gaussiana inversa. en el archivo nom.dat se leen los nom-
c  bres de los archivos donde se encuentran las observaciones que
c  se van a transformar.
c  para calcular la distribución gaussiana inversa se usará la re-
c  lación que existe con la distribución normal.
c*****
c*****
c  subrutinas que usa:
c          dest (...) estimadores máximo verosímiles de
c          los parámetros
c          fmv1(...) estimador máximo verosímil de IG,
c          usando la relación con la normal
c          sort ( ... ) subrutina de ordenación
c          tnor ( ... ) distribución normal
c*****
c*****
c  implicit double precision (a-h,o-z)
c  dimension z ( 170 ) , femv ( 170 )
c  character*10 narch1 , narch2(20) , narch3(50) , ent
c  common /val/ dmue,dlae
c*****
c  lee el nombre del archivo donde están los nombres de los
c  archivos que contienen la información y los asigna
c*****
c  write ( * , 100 )
c  read ( * , 110 ) narch1
c  open (10, file = narch1 , status = ' old ' )
c  read ( 10 , 90 ) m
c  read ( 10 , 110 ) ( narch2 ( i ) , i = 1 , m )
c  close ( unit = 10 , err = 600 , iostat = nerr )
c  write ( * , 100 )

```

```

read ( * , 110 ) ent
open ( 50 , file = ent , status = 'old' )
read ( 50 , 90 ) na
read ( 50 , 110 ) ( narch3 ( j ) , j = 1 , na )
close ( unit = 50 , err = 600 , iostat = nerr )
c*****
c  inicia do principal ( sobre el número de archivos )
c*****
do 2000 inc = 1 , m
c*****
c  abre el archivo correspondiente y lee la información relevante
c*****
open ( 20 , file = narch2 ( inc ) , status = 'old' )
read ( 20 , 120 ) nin , dla
read ( 20 , 130 ) ( z ( i ) , i = 1 , nin )
close ( unit = 20 , err = 600 , iostat = nerr )
n = 10
c*****
c  cálculo de los estimadores de lambda y mu, para muestras de ta-
c  maño 10 ( 10 ) 150. inicia loop de acumulación de información.
c  lee las primeras nin observaciones y las ordena
c*****
do 1000 is = 1 , 15
ik = ( inc - 1 ) * 15 + is
write ( * , 160 ) ik
call sort ( n , z )
call dest ( n , z , dmue , dlae )
c*****
c  cálculo del estimador máximo verosímil
c*****
call fmv1 ( n , z , femv )
c*****
c  escritura de resultados : primero abre al archivo donde va a es-
c  cribir los resultados, después escribe n, lambda y las transfor-

```

```

c   madas
c*****
  open ( 30, file = narch3 ( ik ) , status = 'new' )
  write ( 30 , 140 ) n , dla , dmue , dlae
  write ( 30 , 150 ) ( femv ( kk ) , kk = 1 , n )
  close ( unit = 30 , err = 600 , iostat = nerr )
c*****
c   termina do de acumulación de información
c*****
  n = n + 10
1000 continue
c*****
c   termina do principal
c*****
2000 continue
  close ( unit = 12 , err = 600 , iostat = nerr )
c*****
c   formatos
c*****
  80 format ( 1x , a10 )
  90 format ( 1x , i3 )
  100 format ( 1x , ' nombre del archivo : ' , $ )
  110 format ( a10 )
  120 format ( 1x , i3 , 2x , f10.6 )
  130 format ( 1x , f10.6 )
  140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
  150 format ( 1x , f8.6 )
  160 format ( '+          voy en el caso ' , i3 )
  600 write ( * , 610 ) nerr
  610 format(1x,' error número ' , i2, ' en el cierre de un archivo ' )
c*****
c   fin del programa
c*****
  stop

```

```

end
c*****
c*****
c  subrutinas
c*****
c*****
  subroutine fmv1 ( n , z , femv )
c  cálculo del estimador máximo verosímil de
c  la distribución gaussiana inversa, usando
c  la relación con la distribución normal
c  entradas: n número de observaciones muestrales
c           z vector de dimensión n con la muestra ordenada
c  salida: femv vector de dimensión n, que contiene a la
c          función de distribución gaussiana inversa
c          con parámetros (dmue,dlae), evaluada en
c          el vector z(n)
c  subrutinas que usa: tnor ( x ) distribución normal
c*****
  implicit double precision (a-h,o-z)
  external tnor
  common /val/ dmue,dlae
  dimension z(n), femv(n)
  co = dexp ( 2d0 * dlae / dmue )
  do 1 i = 1 , n
    fr1 = dlae * z ( i )
    fr1 = dsqrt ( fr1 )
    fr2 = dlae / z ( i )
    fr2 = dsqrt ( fr2 )
    rr = - fr2 + fr1 / dmue
    rl = - fr2 - fr1 / dmue
    femv ( i ) = tnor ( rr ) + tnor ( rl ) * co
1 continue
  return
end

```

```

c*****
c*****
function tnor(x)
c  cálculo de la función de distribución normal usando
c  el algoritmo de Strecok (1968), para la función error.
c*****
implicit double precision ( a - h , o - z )
dimension a1 ( 19 ) , a2 ( 19 )
a1(1) = 0.7032250027437753888217370d0
a1(2) = 0.3305015219166062231414836d0
a1(3) = 0.2013397472647062732206869d0
a1(4) = 0.1086302450227406906873823d0
a1(5) = 0.0467755234324848608231226d0
a1(6) = 0.0153985726157101996635576d0
a1(7) = 0.0038015076798529871151801d0
a1(8) = 0.0006971837924080287357368d0
a1(9) = 0.0000944909268810454980951d0
a1(10) = 0.0000094328116983836678903d0
a1(11) = 0.0000006919275203251400557d0
a1(12) = 0.0000000372252349369107964d0
a1(13) = 0.0000000014666061423380014d0
a1(14) = 0.000000000422616144318049d0
a1(15) = 0.000000000008897865267233d0
a1(16) = 0.000000000000136760444757d0
a1(17) = 0.000000000000001533423425d0
a1(18) = 0.000000000000000012536751d0
a1(19) = 0.000000000000000000074517d0
a2(1) = 0.2472551681400521469013060d0
a2(2) = 0.1442272263615747171484869d0
a2(3) = 0.0869894549959345538475537d0
a2(4) = 0.0439773381940833680821082d0
a2(5) = 0.0172439625886622621167913d0
a2(6) = 0.0050790696122025702812471d0
a2(7) = 0.0011086064534234065370328d0

```

```

a2(8) = 0.0001782280162548616973045d0
a2(9) = 0.0000210404583073251381878d0
a2(10) = 0.0000018206631636434076704d0
a2(11) = 0.0000001153309909443694132d0
a2(12) = 0.0000000053427502760308268d0
a2(13) = 0.0000000001808485878095127d0
a2(14) = 0.00000000000044696822924881d0
a2(15) = 0.0000000000000806068838945d0
a2(16) = 0.0000000000000010601364636d0
a2(17) = 0.000000000000000101649277d0
a2(18) = 0.00000000000000000710005d0
a2(19) = 0.00000000000000000000000d0
pi = 3.141592653589793238462643d0
t = dabs ( x )
y = t / dsqrt ( 2d0 )
ck = 2d0 * y / 5d0
c = dcos ( ck )
s = dsin(ck)
alf = 2d0 * c * c - 1d0
d = ck / pi
sum = 0d0
do 1 i = 1 , 19
  ii = i - 1
  rp = ( a1 ( i ) + 2d0 * c * a2 ( i ) ) * (alf** ii)
1  sum = sum + rp
  erf = d + s * sum
  res = 0.5d0 * ( 1d0 + erf )
  if ( x .le. 0d0 ) res = 1d0 - res
  tnor = res
return
end

```

c\*\*\*\*\*

```

c*****
c  programa ppf4b.for
c  transformaciones al (0,1) usando el estimador máximo verosímil,
c  en la gaussiana inversa. en el archivo nom.dat se leen los nom-
c  bres de los archivos donde se encuentran las observaciones que
c  se van a transformar.
c*****
c*****
c  subrutinas que usa:
c          dgausin (...) densidad IG
c          dest (...) estimadores máximo verosímiles de
c          los parámetros
c          fmv(...) estimador máximo verosímil de IG
c          gausin (...) integral de la densidad IG
c          qtrap (...) subrutina de integración basada
c          en la regla del trapecio
c          sort ( ... ) subrutina de ordenación
c          tmidpt (..) regla abierta del trapecio
c*****
c*****
c          implicit double precision (a-h,o-z)
c          dimension z ( 170 ) , femv ( 170 ) , x ( 170 )
c          character*10 narch1 , narch2(10) , narch3(50) , ent
c          common /val/ dmue,dlae
c          open ( 12 , file = 'basura.dat',status = 'old')
c*****
c  lee el nombre del archivo donde están los nombres de los
c  archivos que contienen la información y los asigna
c*****
c          write ( * , 100 )
c          read ( * , 110 ) narch1
c          open (10, file = narch1 , status = ' old ' )
c          read ( 10 , 90 ) m
c          read ( 10 , 110 ) ( narch2 ( i ) , i = 1 , m )

```

```

close ( unit = 10 , err = 600 , iostat = nerr )
write ( * , 100 )
read ( * , 110 ) ent
open ( 50 , file = ent , status = 'old' )
read ( 50 , 90 ) na
read ( 50 , 110 ) ( narch3 ( j ) , j = 1 , na )
close ( unit = 50 , err = 600 , iostat = nerr )
c*****
c  inicia do principal ( sobre el número de archivos )
c*****
do 2000 inc = 1 , m
c*****
c  abre el archivo correspondiente y lee la información relevante
c*****
open ( 20 , file = narch2 ( inc ) , status = 'old' )
read ( 20 , 120 ) nin , dla
read ( 20 , 130 ) ( x ( i ) , i = 1 , nin )
close ( unit = 20 , err = 600 , iostat = nerr )
n = 10
c*****
c  cálculo de los estimadores de lambda y mu, para muestras de ta-
c  maño 10 ( 10 ) 200. inicia loop de acumulación de información.
c  lee las primeras nin observaciones y las ordena
c*****
do 1000 is = 1 , 15
ik = ( inc - 1 ) * 15 + is
write ( * , 160 ) ik
do 1 i = 1 , n
1 z ( i ) = x ( i )
call sort ( n , z )
call dest ( n , z , dmue , dlae )
c*****
c  cálculo del estimador máximo verosímil
c*****

```

```

    call fmv ( n , z , fmv )
c*****
c  escritura de resultados : primero abre al archivo donde va a es-
c  cribir los resultados, después escribe n, lambda y las transfor-
c  madas
c*****
    open ( 30, file = narch3 ( ik ) , status = 'new' )
    write ( 30 , 140 ) n , dla , dmue , dlae
    write ( 30 , 150 ) ( fmv ( kk ) , kk = 1 , n )
    close ( unit = 30 , err = 600 , iostat = nerr )
c*****
c  termina do de acumulación de información
c*****
    n = n + 10
    1000 continue
c*****
c  termina do principal
c*****
    2000 continue
    close ( unit = 12 , err = 600 , iostat = nerr )
c*****
c  formatos
c*****
    80 format ( 1x , a10 )
    90 format ( 1x , i3 )
    100 format ( 1x , ' nombre del archivo : ', $ )
    110 format ( a10 )
    120 format ( 1x , i3 , 2x , f10.6 )
    130 format ( 1x , f10.6 )
    140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
    150 format ( 1x , f8.6 )
    160 format ( '+          voy en el caso ', i3 )
    600 write ( * , 610 ) nerr
    610 format(1x,' error numero ',i2, ' en el cierre de un archivo ')

```

```

c*****
c  fin del programa
c*****
  stop
  end
c*****
c*****
c  subrutinas
c*****
c*****
  function dgausin(x)
c  función de densidad gaussiana inversa
c  con parámetros dmue,dlae
c  entradas: dmue,dlae parámetros de la función (por common)
c          x valor donde se evalua la función
c  salida: dgausin valor de la función
c*****
  implicit double precision (a-h,o-z)
  common /val/dmue,dlae
  pi = 3.141592653589793238462643d0
  x3 = x * x * x
  c= dlae / ( 2d0 * pi * x3 )
  c = dsqrt ( c )
  ot = dlae * ( x - dmue ) * ( x - dmue ) /
*          ( 2d0 * x * dmue * dmue )
  ot = dexp ( -ot )
  dgausin = c * ot
  return
  end
c*****
c*****
  subroutine fmv(n,z,femv)
c  cálculo del estimador máximo verosímil de la distribución
c  gaussiana inversa, usando integración directa de la función

```

```

c   de densidad
c   entradas: n número de observaciones muestrales
c           z vector de dimensión n con la muestra ordenada
c   salida: femv vector de dimensión n, que contiene a la
c           función de distribución gaussiana inversa
c           con parámetros (dmue,dlae), evaluada en
c           el vector z(n)
c   subrutinas que usa: gausin(finf,fsup,fint)
c*****
      implicit double precision ( a - h , o - z )
      common /val/ dmue , dlae
      dimension z ( n ) , femv ( n )
      do 1 i = 1 , n
      finf = 1d-8
      fsup = z ( i )
      call gausin ( finf , fsup , fint )
      femv ( i ) = fint
1 continue
      return
      end
c*****
c*****
      subroutine gausin ( finf , fsup , fint )
c   cálculo de la función de distribución gaussiana inversa, por
c   integración directa de la función de densidad
c   entradas: finf,fsup límites de integración
c   salida: fint valor de la integral
c   subrutinas que usa: dgausin(x,dmu,dla) función de densi-
c           dad gaussiana inversa con paráme-
c           tros dmue, dla
c*****
      implicit double precision (a-h,o-z)
      common /val/ dmue , dlae
      external dgausin

```

```
call qtrap ( dgausin , finf , fsup , s )  
fint = s  
return  
end
```

```
c*****
```

```

c*****
c  programa ppf05.for
c  cálculo de la estadística  $A^2$ . las observaciones las lee de archi-
c  vos cuyos nombres se proporcionan.
c  subrutinas que usa: ander (n,z,a) cálculo de  $A^2$ 
c*****
      implicit double precision ( a - h , o - z )
      dimension z ( 170 )
      character*10 narch,nomb(60),nsal
c*****
c  se abre el archivo donde están los nombres de los archivos que
c  contienen a los datos. se pide también el nombre del archivo de
c  salida
c*****
      1 write ( * , 100 )
        read ( * , 110 ) narch
        open ( 5 , file = narch , status = 'old' )
        read ( 5 , 80 ) m
        read ( 5 , 90 ) ( nomb ( i ) , i = 1 , m )
        close ( unit = 5 , err = 600 , iostat = nerr )
        write ( * , 100 )
        read ( * , 110 ) nsal
        open ( 10 , file = nsal , status = 'new' )
c*****
c  comienza el loop principal. lee la información contenida en los ar-
c  chivos, calcula la estadística de anderson - darling y escribe el
c  resultado, junto con el valor de n y el valor de lambda
c*****
      do 1000 inc = 1 , m
        open ( 15 , file = nomb ( inc ) , status = 'old' )
        read ( 15 , 140 ) n , dla , dmue , dlae
        read ( 15 , 150 ) ( z ( kk ) , kk = 1 , n )
        close ( unit = 15 , err = 600 , iostat = nerr )
        call ander ( n , z , a )

```

```

write ( 10 , 160 ) n , dla , a
write ( * , 170 ) inc
1000 continue
c*****
c termina el loop principal
c*****
write ( * , 190 )
read ( * , * ) l
if ( l.eq. 1 ) goto 1
c*****
c formatos
c*****
80 format ( 1x , i3 )
90 format ( 1x , a10 )
100 format ( 1x , ' nombre del archivo : ' , $ )
110 format ( a10 )
140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
150 format ( 1x , f8.6 )
160 format ( 1x , i3 , 2 ( 2x , f10.6 ) )
170 format ( '+ voy en el caso ' , i3 )
190 format ( 1x , ' quieres otro caso ? ' , $ )
600 write ( * , 610 ) nerr
610 format(1x,' error número ',i2,' en el cierre de un archivo ')
c*****
c fin del programa
c*****
stop
end
c*****
c*****
c subrutinas
c*****
c*****
subroutine ander (n,z,a)

```

- c cálculo de la estadística a cuadrada de anderson y darling
- c para una muestra ordenada de n observaciones
- c entradas: n número de observaciones
- c z vector de dimensión n con la muestra ordenada
- c salida: a valor de la estadística a cuadrada

```

c*****
implicit double precision ( a - h , o - z )
dimension z ( n )
a = 0d0
do 1 i = 1 , n
if ( z ( i ) .lt. 1d-7 ) z ( i ) = 1d-7
if ( z ( n + 1 - i ) .gt. 0.9999999d0)
*           z ( n + 1 - i ) = 0.9999999d0
di = dble ( i )
a = a - ( 2d0 * di - 1d0 ) * ( dlog ( z ( i ) )
1 a = a + dlog ( 1d0 - z ( n - i + 1 ) ) )
rn = dble ( n )
a = a / rn - rn
return
end
c*****

```

```

c*****
c  programa ppf06.for
c  cálculo del proceso  $\sqrt{n} \{Fn(x) - Gn(x)\}$ , donde F y G
c  son transformadas al (0,1). la información la lee de archivos
c  proporcionados por el usuario
c  subrutinas que usa: ddist ( x , n , z ) cálculo de la función
c                      de distribución empírica
c                      sort ( n , z ) estadísticas de orden
c*****
c  implicit double precision ( a - h , o - z )
c  dimension x ( 170 ) , y ( 170 ) , w1 ( 170 ) , w2 ( 170 )
c  character*10 arch1 , arch2 , sal , narch1 ( 50 ) , narch2 ( 50 )
c*****
c  se preguntan los nombres de los archivos que contienen los nombres
c  de los archivos donde están las transformadas
c*****
c  1 write ( * , 100 )
c    read ( * , 110 ) arch1
c    write ( * , 100 )
c    read ( * , 110 ) arch2
c*****
c  se lee la información y se verifica que se tenga el mismo número
c  de archivos
c*****
c  open ( 10 , file = arch1 , status = 'old' )
c  read ( 10 , 90 ) m1
c  read ( 10 , 80 ) ( narch1 ( i ) , i = 1 , m1 )
c  close ( unit = 10 , err = 600 , iostat = nerr )
c  open ( 20 , file = arch2 , status = 'old' )
c  read ( 20 , 90 ) m2
c  read ( 20 , 80 ) ( narch2 ( i ) , i = 1 , m2 )
c  close ( unit = 20 , err = 600 , iostat = nerr )
c  if ( m1 .ne. m2 ) then
c    write ( * , 190 ) arch1 , arch2

```

```

    goto 2000
endif
c*****
c se pregunta el nombre del archivo de escritura
c*****
    write ( * , 105 )
    read ( * , 110 ) sal
    open ( 50 , file = sal , status = 'new' )
c*****
c inicia do principal
c*****
    do 1000 inc = 1 , m1
c*****
c abre los archivos donde están las transformadas, lee la informa-
c ción y verifica que el número de observaciones sea el mismo y que
c se hayan generado con el mismo valor de lambda
c*****
    open ( 30 , file = narch1 ( inc ) , status = 'old' )
    read ( 30 , 140 ) n1 , dla1 , dlae1 , dmuel
    read ( 30 , 150 ) ( x ( i ) , i = 1 , n1 )
    close ( unit = 30 , err = 600 , iostat = nerr )
    open ( 40 , file = narch2 ( inc ) , status = 'old' )
    read ( 40 , 140 ) n2 , dla2 , dlae2 , dimue2
    read ( 40 , 150 ) ( y ( i ) , i = 1 , n2 )
    close ( unit = 40 , err = 600 , iostat = nerr )
    if ( n1 .ne. n2 .or. dla1 .ne. dla2 ) then
        write ( * , 180 ) narch1 ( inc ) , narch2 ( inc )
        goto 2000
    endif
c*****
c calcula las diferencias entre las distribuciones empíricas evaluadas
c en la primera muestra
c*****
    call sort ( n1 , x )

```

```

dn = dble ( n1 )
rn = dsqrt ( dn )
do 2 i = 1 , n1
  di = dble ( i )
  di1 = di - 1d0
  cua = di / dn
  cual = di1 / dn
  p = x ( i )
  oca = ddist ( p , n2 , y )
  w1 ( i ) = rn * dabs ( cua - oca )
  w2 ( i ) = rn * dabs ( cual - oca )
2 continue
c*****
c  calcula el máximo de las diferencias
c*****
  call sort ( n1 , w1 )
  call sort ( n2 , w2 )
  if ( w1 ( n1 ) .ge. w2 ( n2 ) ) then
    res = w1 ( n1 )
  else
    res = w2 ( n1 )
  endif
c*****
c  escribe los resultados
c*****
  write ( 50 , 145 ) n1 , dlal , res
  write ( * , 160 ) inc , ml
c*****
c  termina el do principal , sobre el número de archivos, y pregunta
c  si se quiere repetir el proceso para otro conjunto de datos
c*****
1000 continue
2000 write ( * , 170 )
  read ( * , * ) l

```

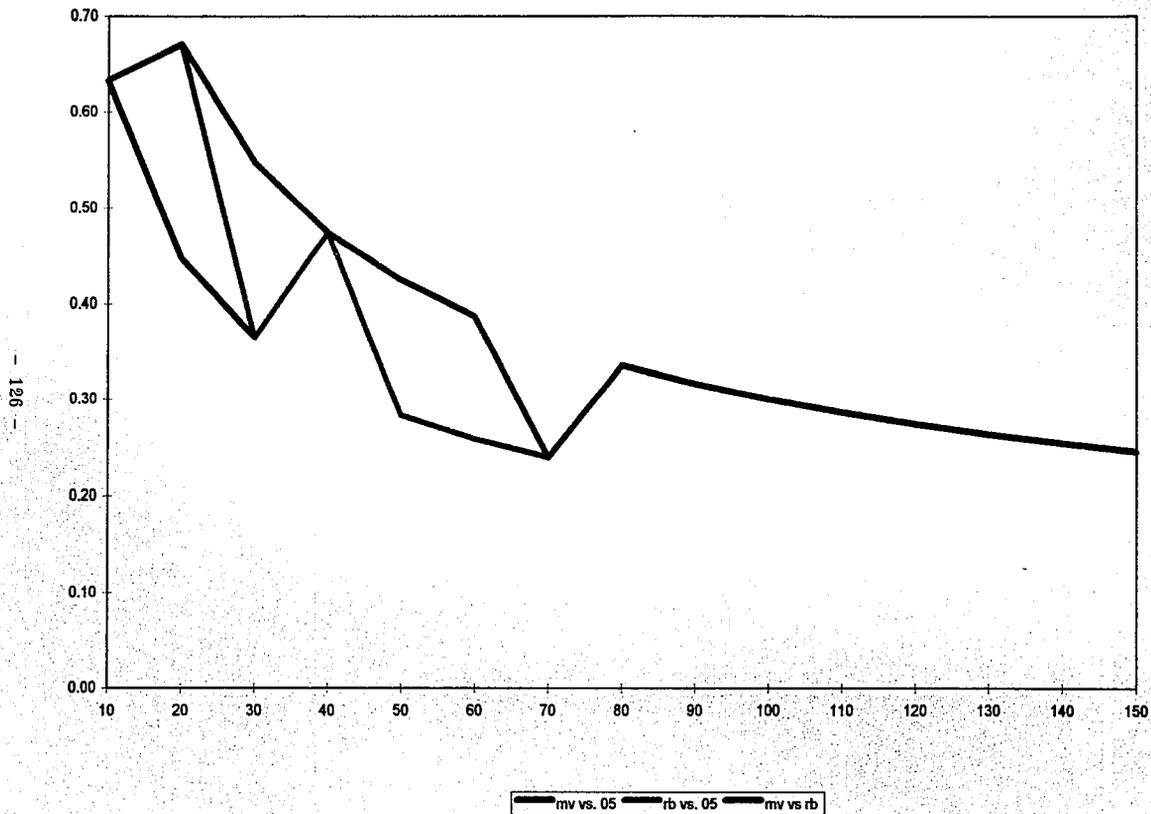
```

    if ( l.eq. 1 ) goto 1
c*****
c   formatos
c*****
    80 format ( 1x , a10 )
    90 format ( 1x , i3 )
    100 format ( 1x , ' nombre del archivo : ', $ )
    105 format ( 1x , ' nombre del archivo de escritura : ', $ )
    110 format ( a10 )
    120 format ( 1x , i3 , 2x , f10.6 )
    140 format ( 1x , i3 , 3 ( 2x , f10.6 ) )
    145 format ( 1x , i3 , 2 ( 2x , f10.6 ) )
    150 format ( 1x , f8.6 )
    160 format ( '+ voy en el caso ', i4 , ' de un total de ', i4 )
    170 format ( 1x , ' quieres tratar otro caso ? ', $ )
    180 format ( 1x , ' los archivos ', 2 ( 1x , a10 ) , / ,
      *1x , ' tienen n ó Ú distinto ' )
    190 format ( 1x , ' los archivos ', 1x , a10 , 1x , a10 , ' son de di
      *stinto tamaño ' )
    600 write ( * , 610 ) nerr
    610 format ( 1x , ' error numero ', i2 , ' en el cierre de un archivo ' )
c*****
    stop
    end
c*****
c   subrutinas
c*****
c*****
function ddist ( x , n , z )
c   cálculo de la distribución empírica de la muestra z de tamaño n
c   en el punto x
c*****
    implicit double precision ( a - h , o - z )

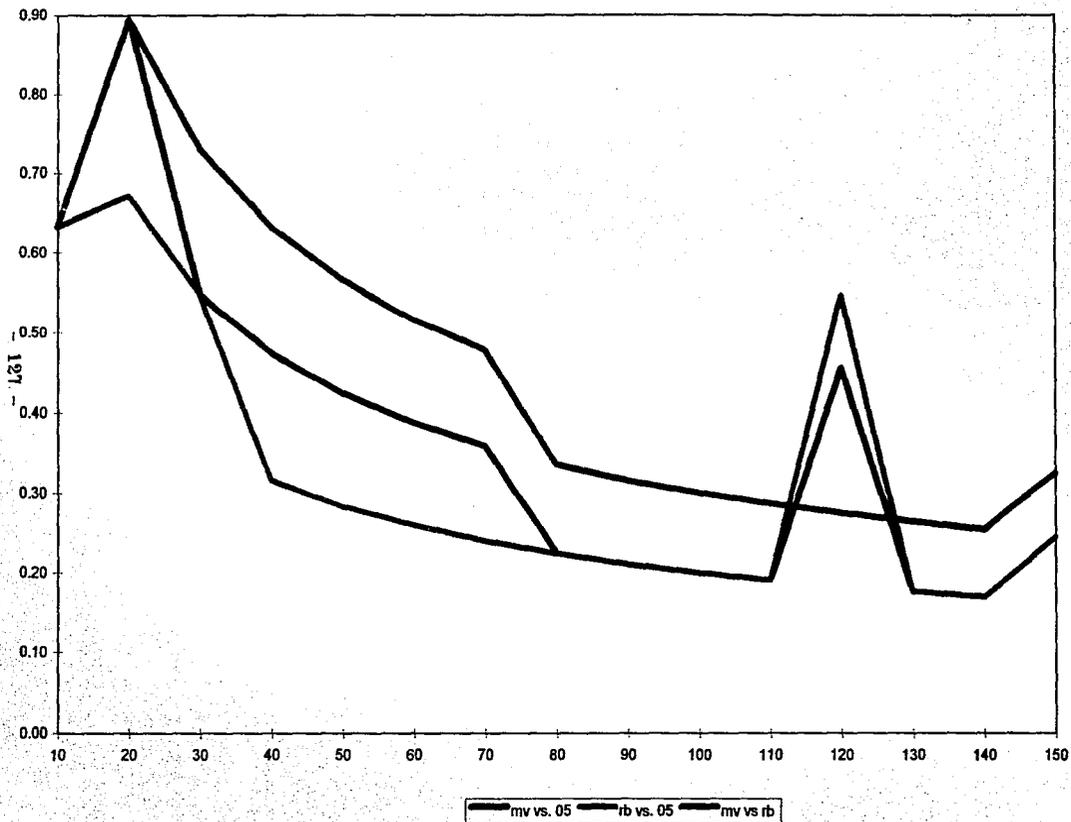
```

```
dimension z ( n )
l = 0
do 1 i = 1 , n
  s = z ( i )
  if ( s .le. x ) l = l + 1
1 continue
dl = dble ( l )
dn = dble ( n )
ddist = dl / dn
return
end
c*****
```

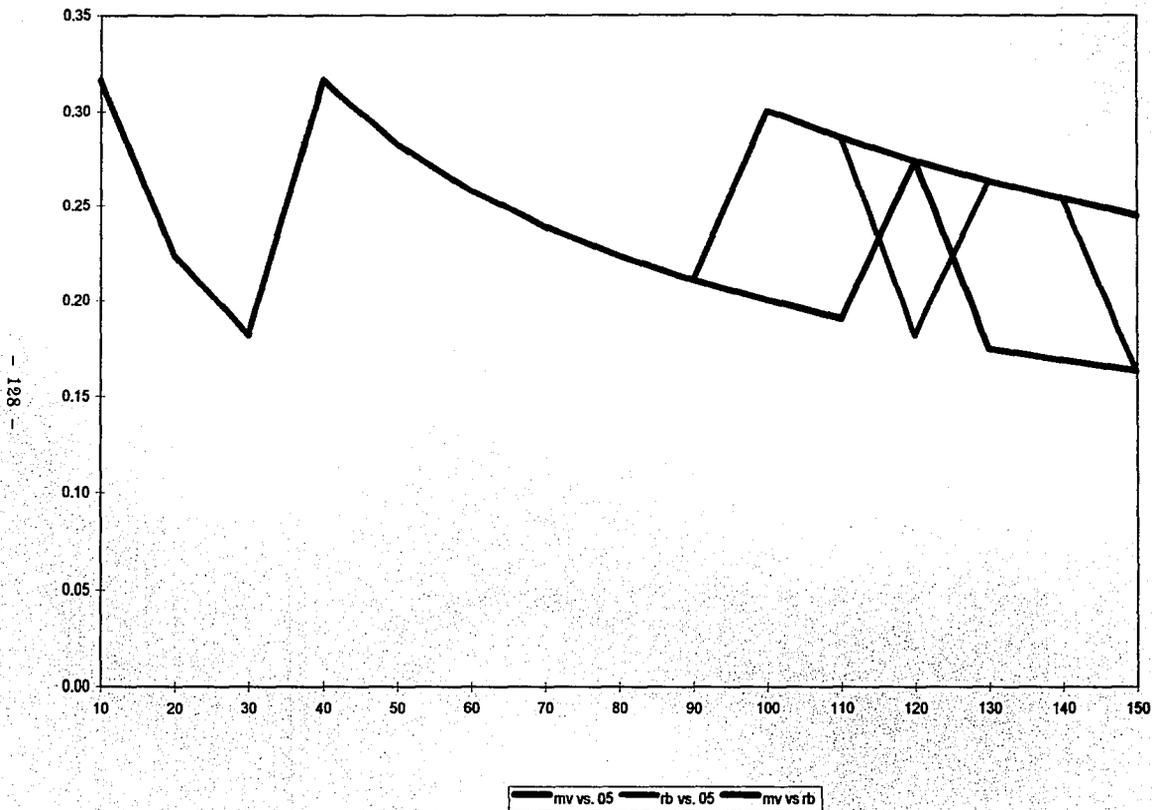
GRAFICA 4. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.50)



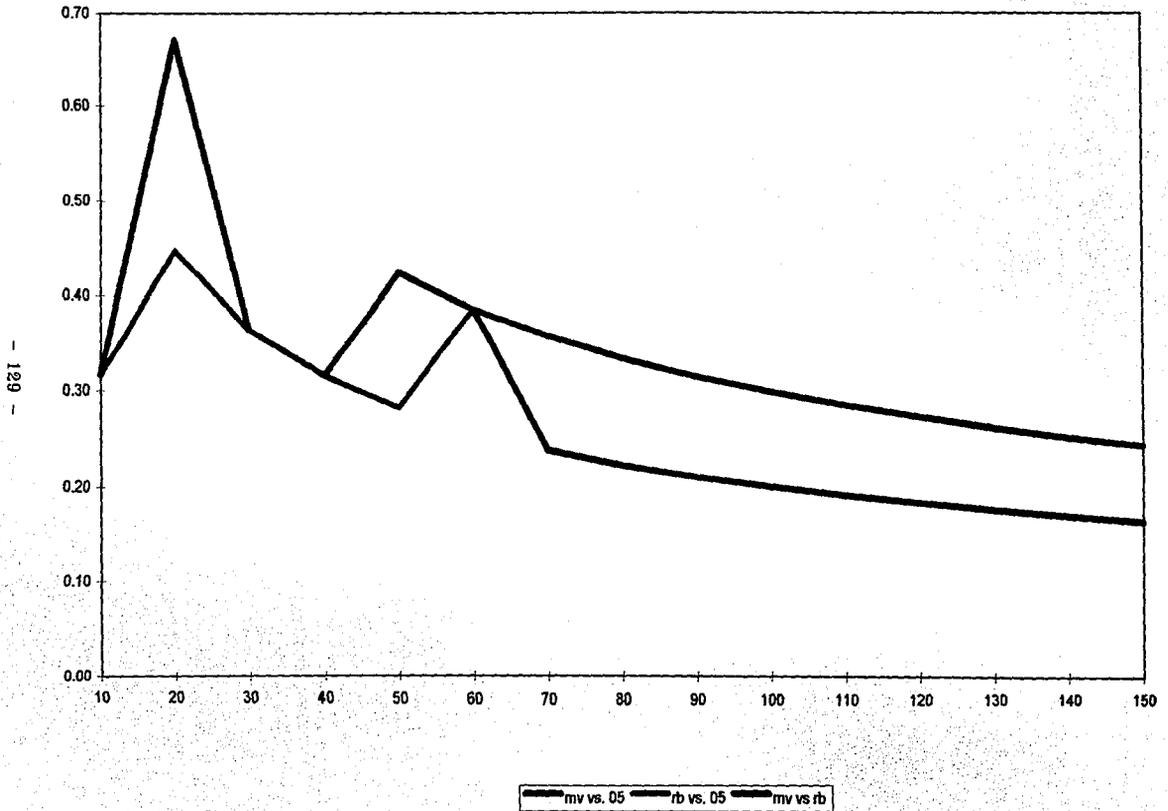
GRAFICA 5. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 1)



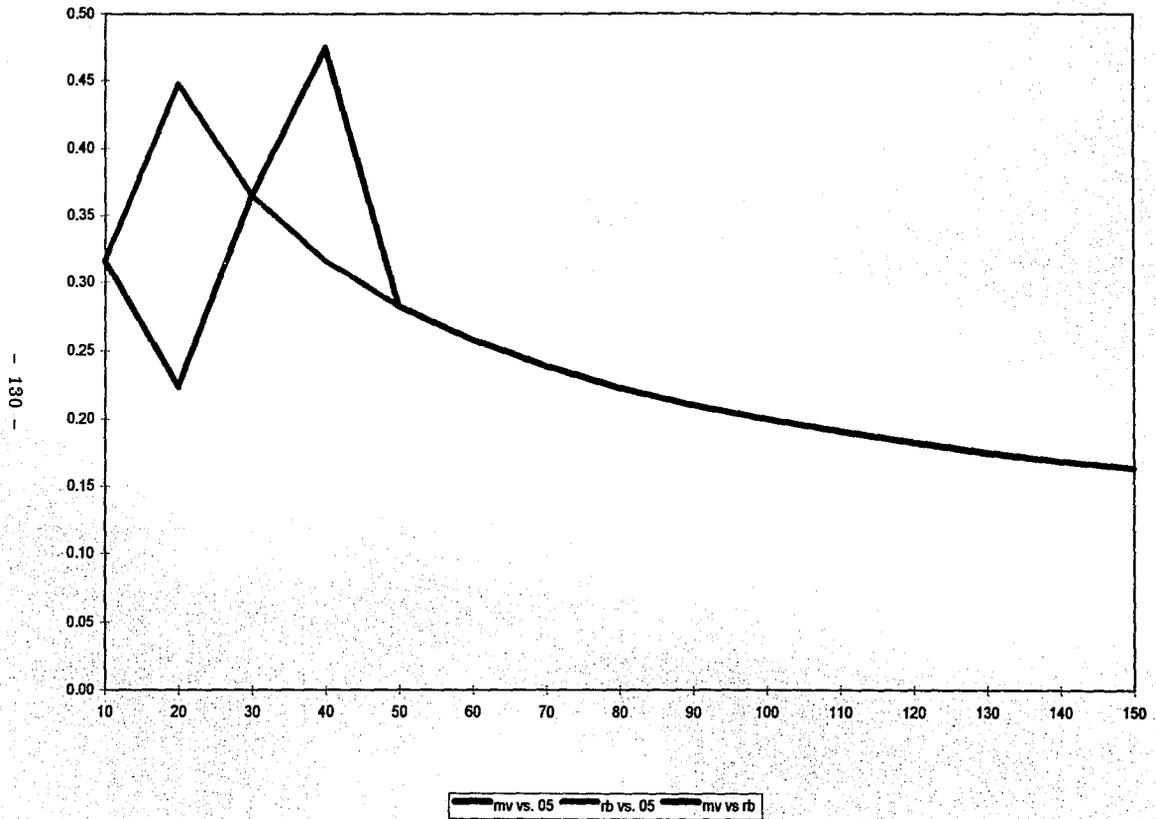
GRAFICA 6. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 2)



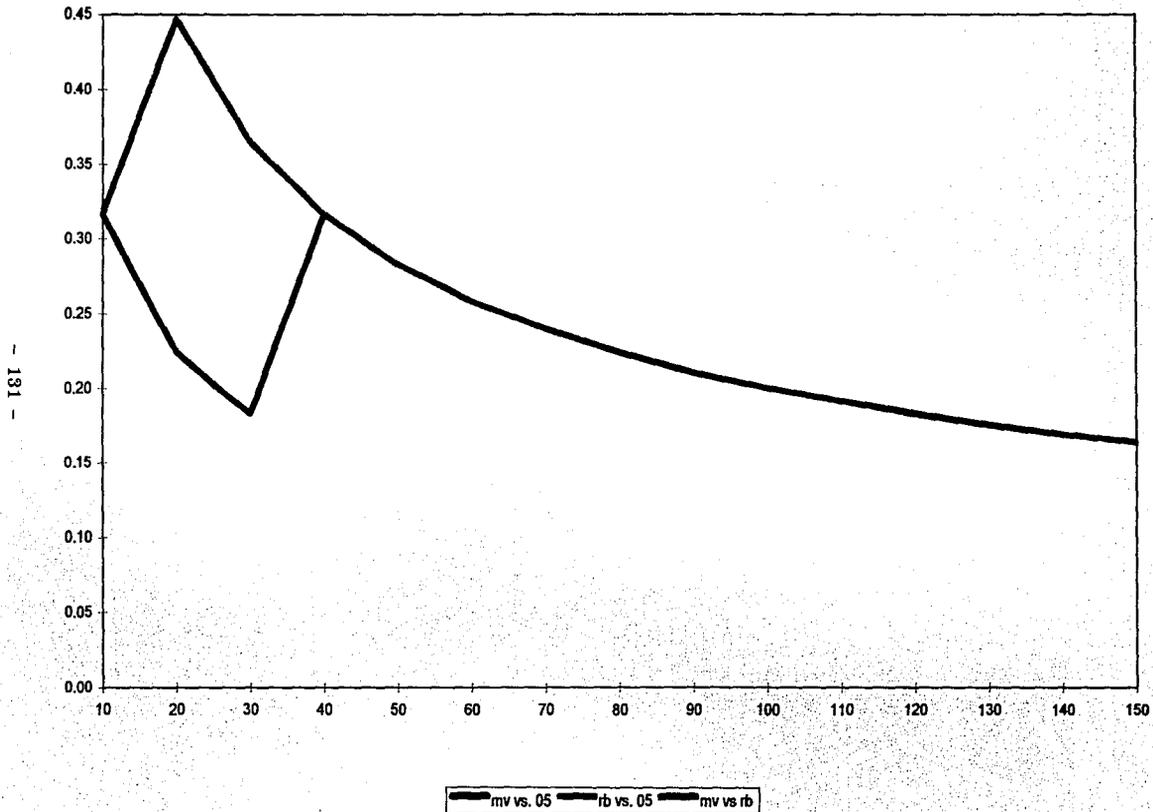
GRÁFICA 7. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 4)



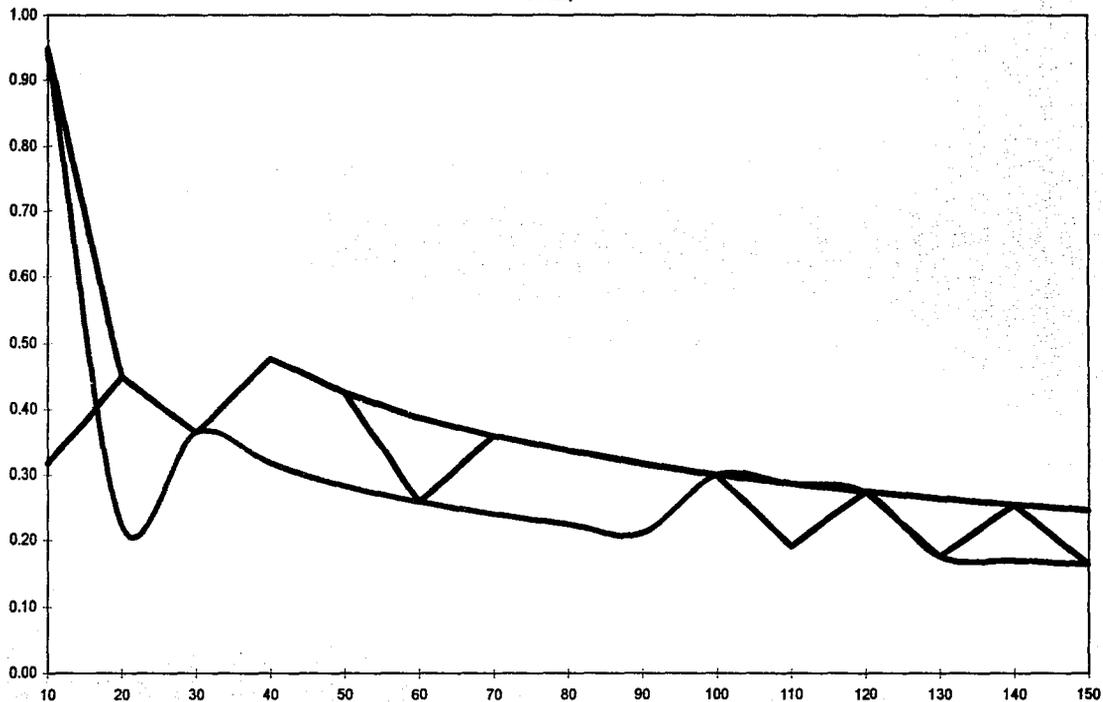
GRAFICA 8. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 8)



GRAFICA 9. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 16)

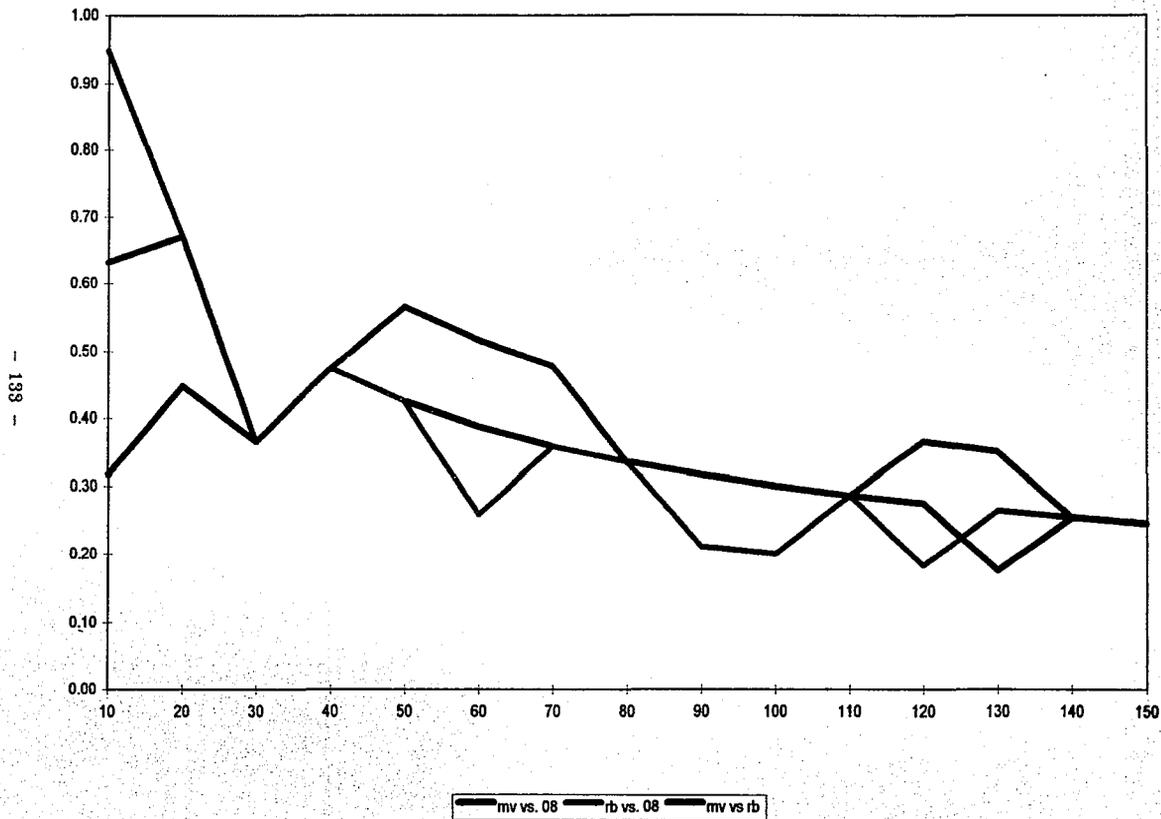


GRAFICA 10. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.0625)

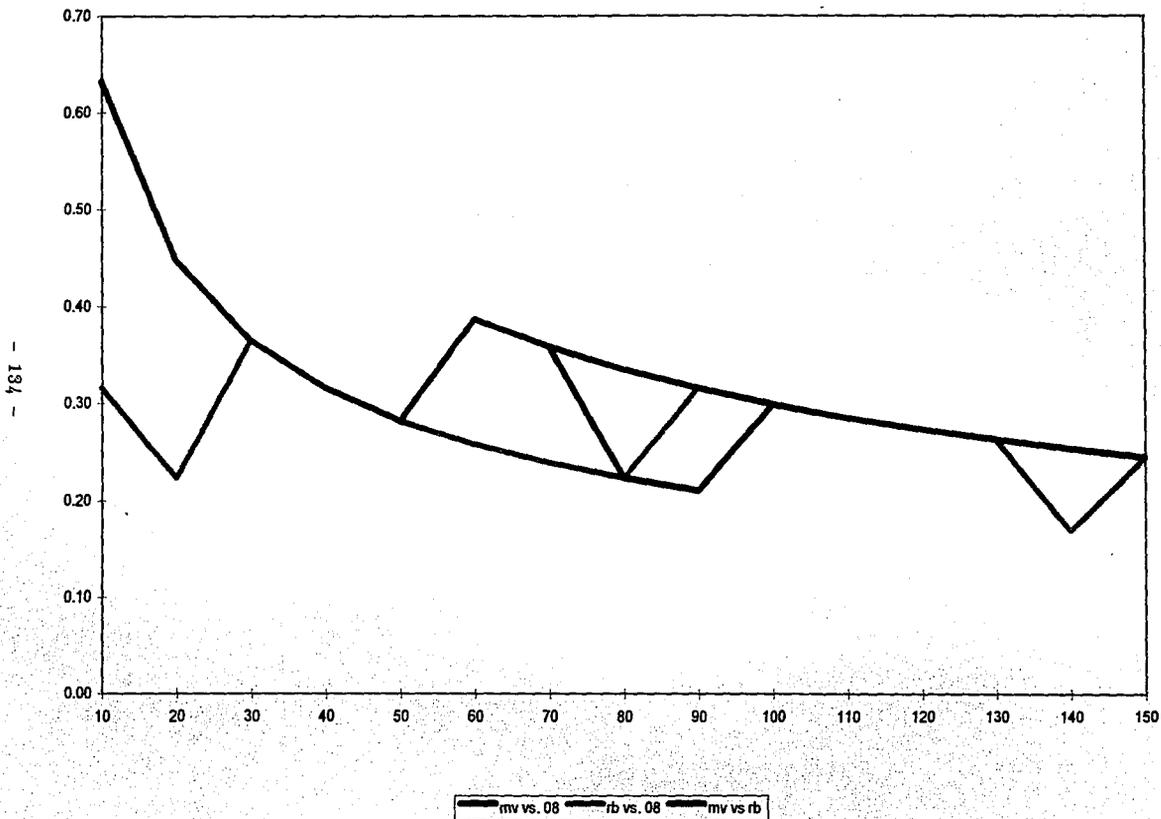


— mv vs. 08 — rb vs. 08 — mv vs rb

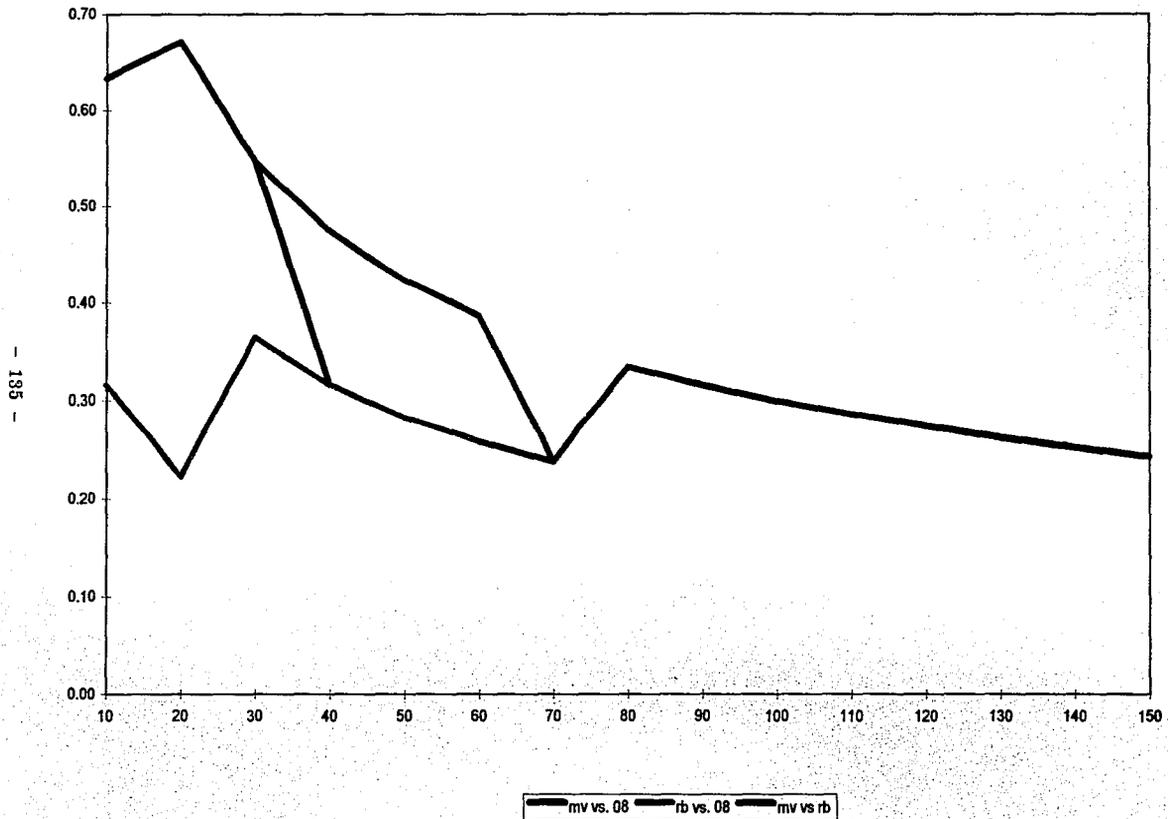
GRAFICA 11. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.125)



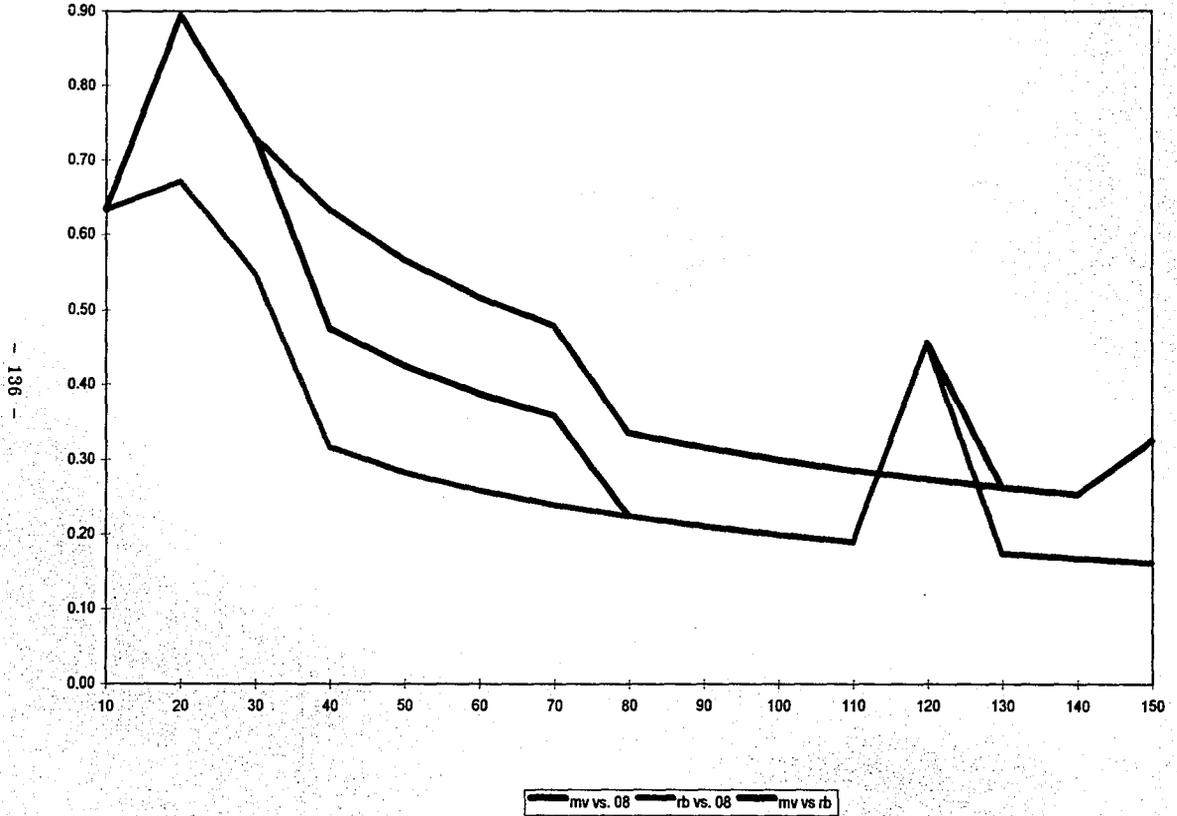
GRAFICA 12. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.250)



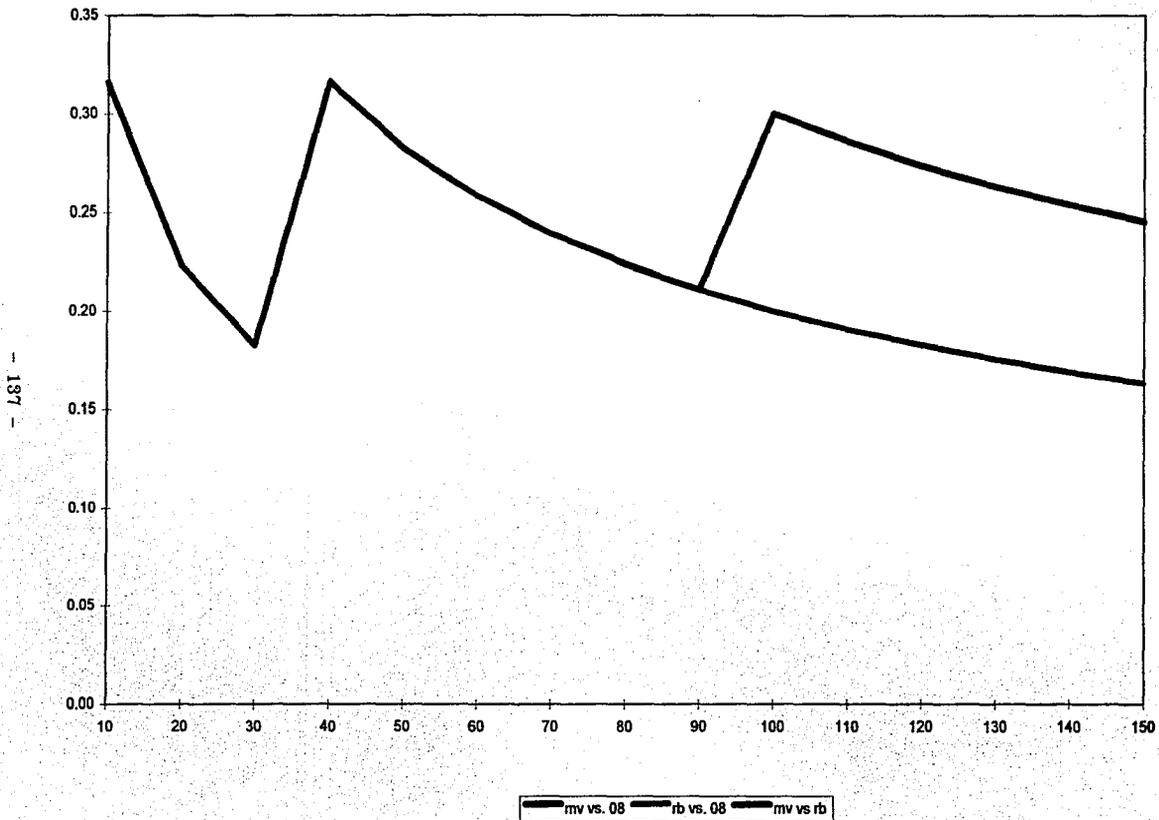
GRAFICA 13. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.50)



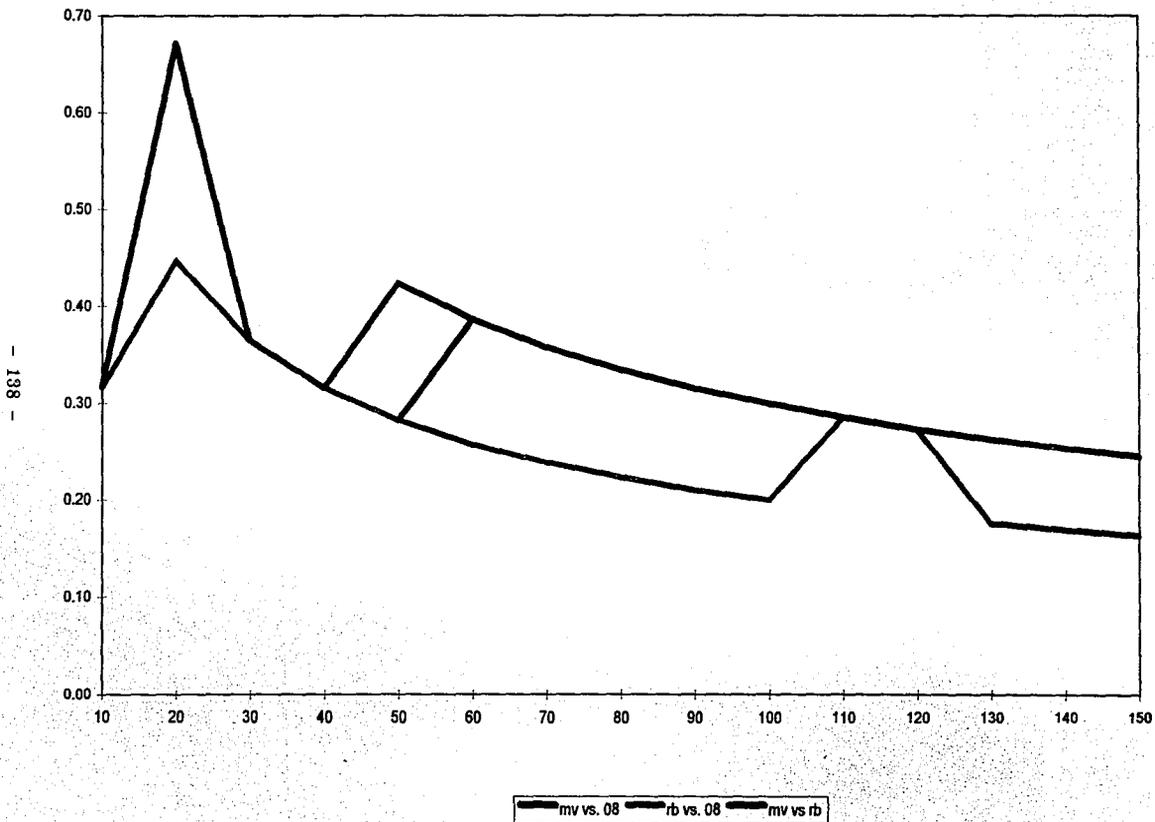
GRAFICA 14. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 1)



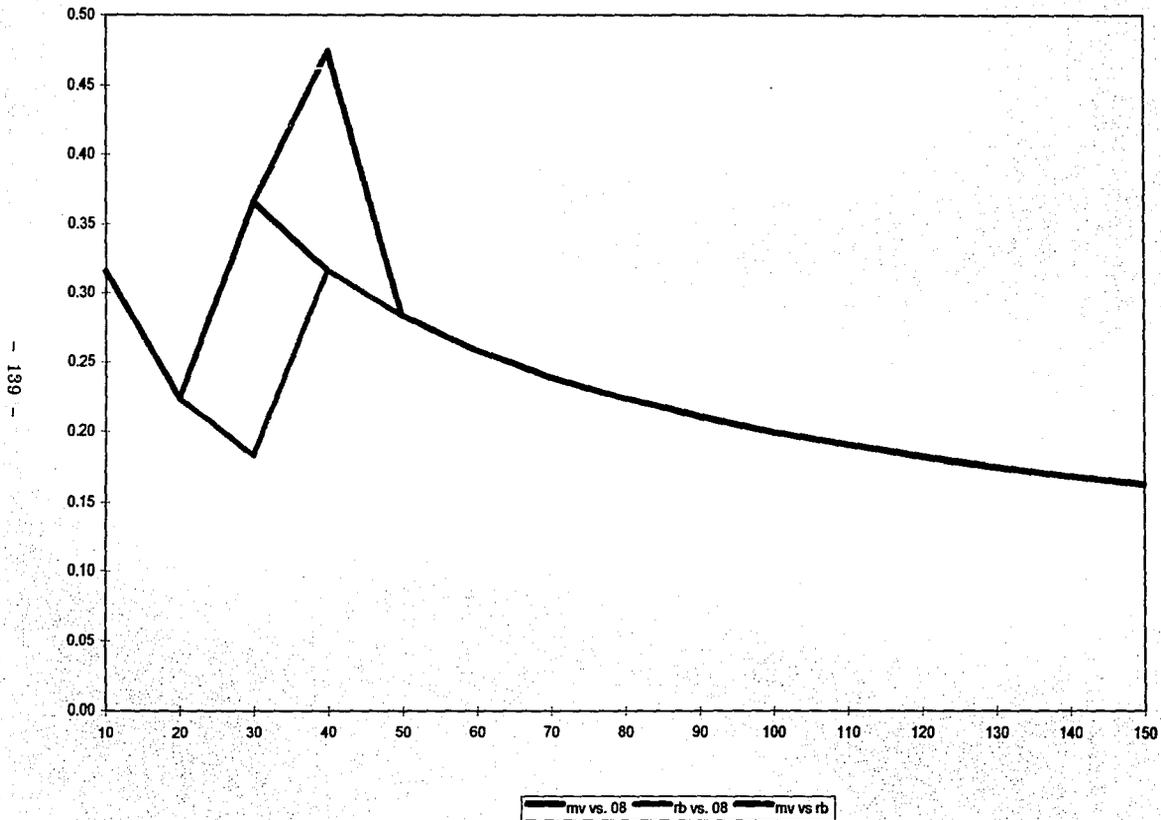
GRAFICA 15. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 2)



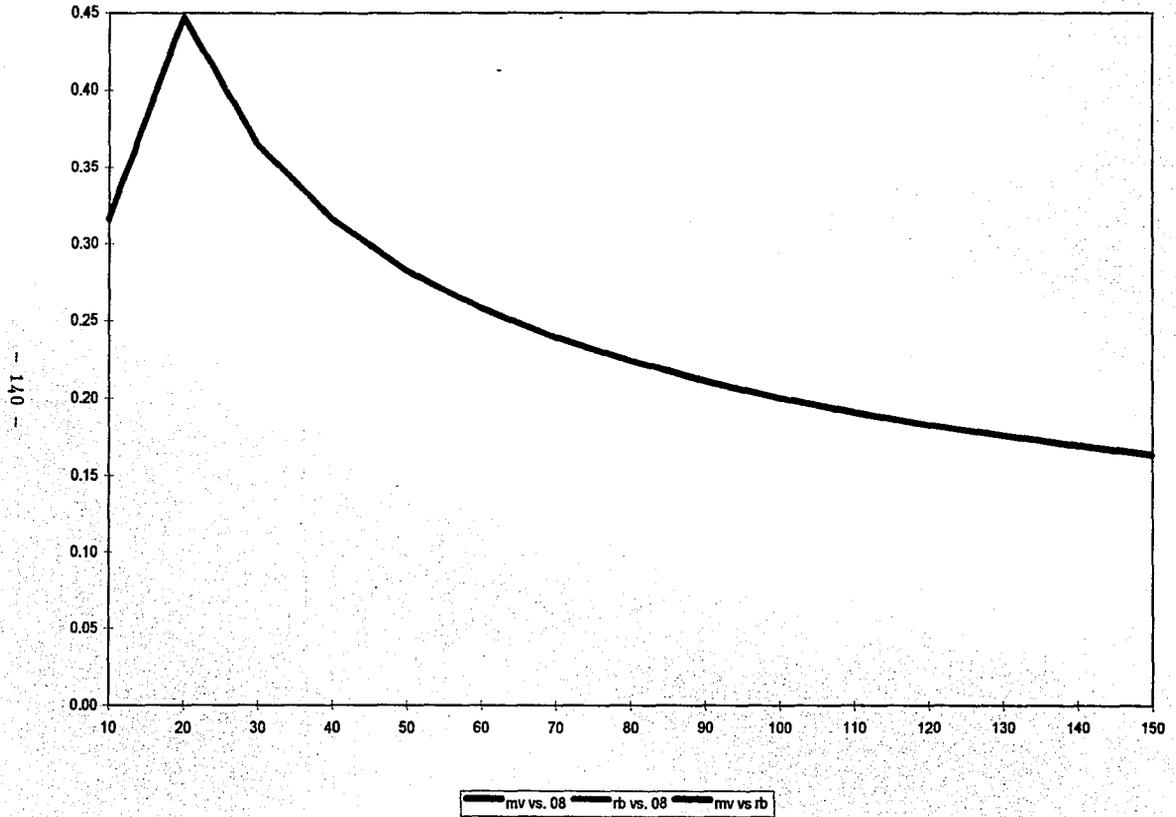
GRAFICA 16. SUPREMO DE LAS IDFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 4)



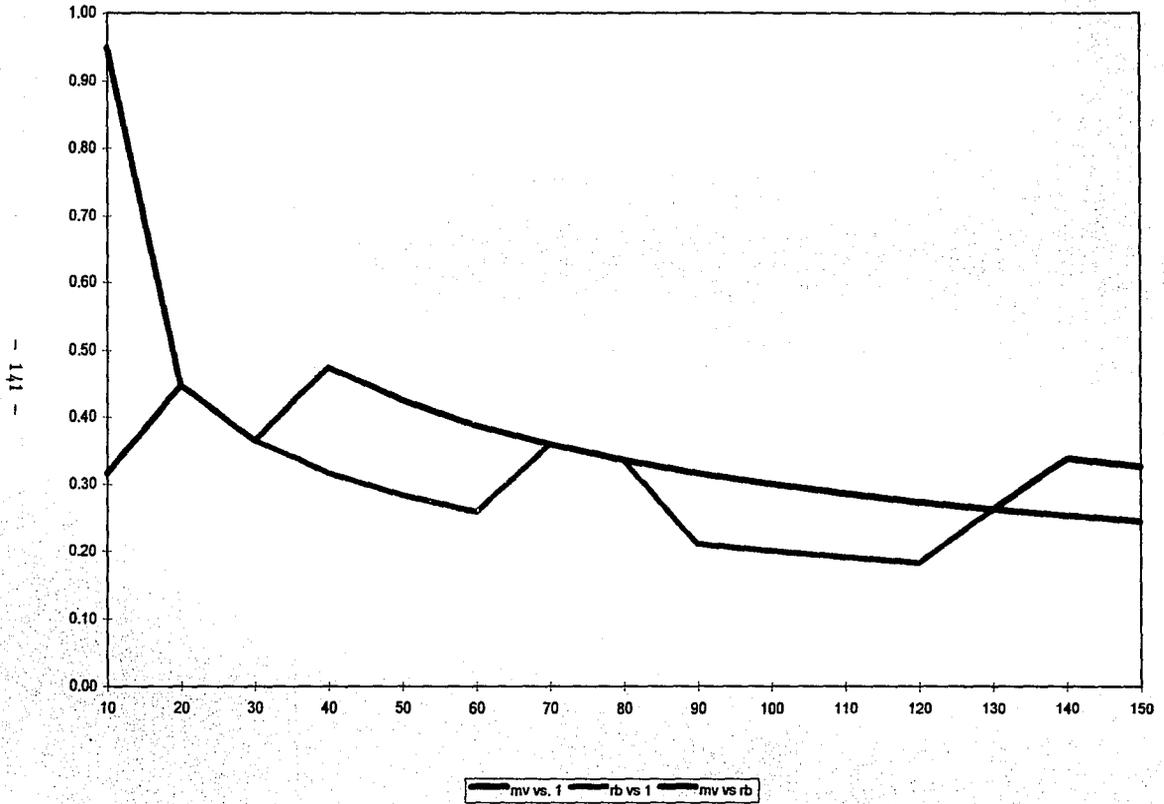
GRAFICA 17. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 8)



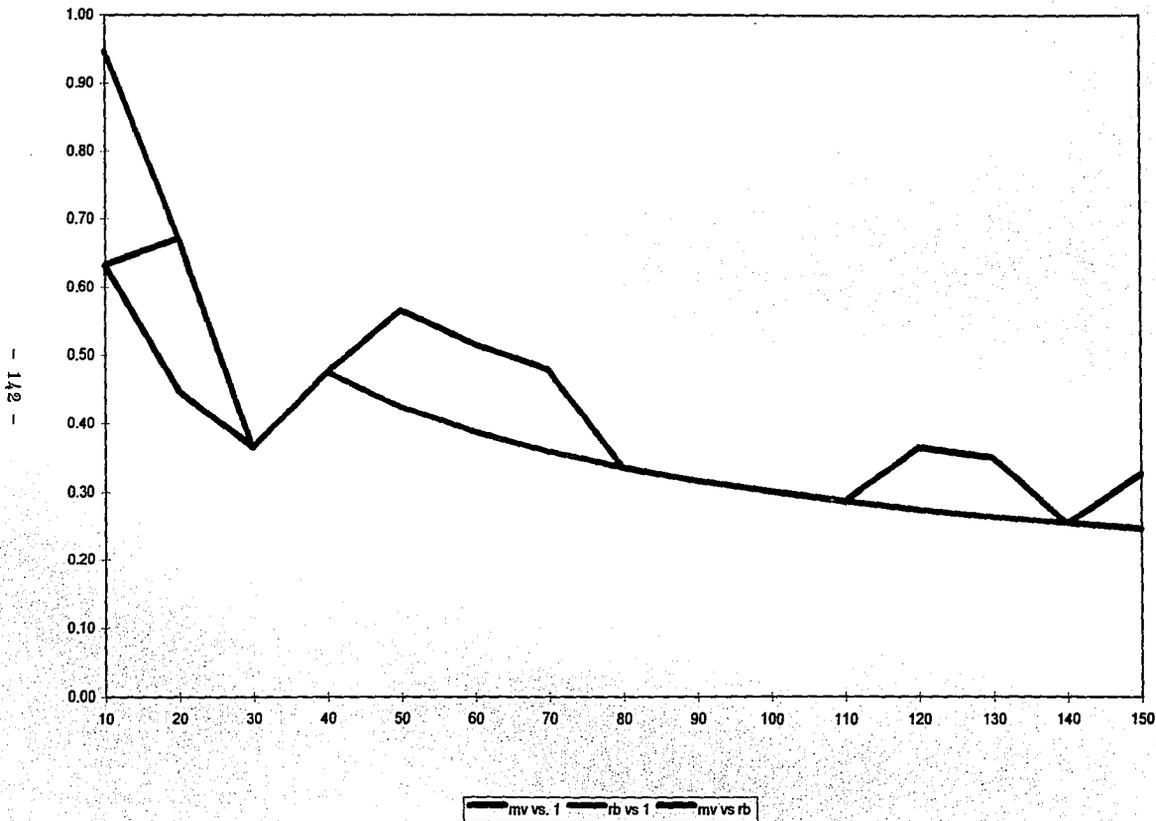
GRAFICA 18. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 16)



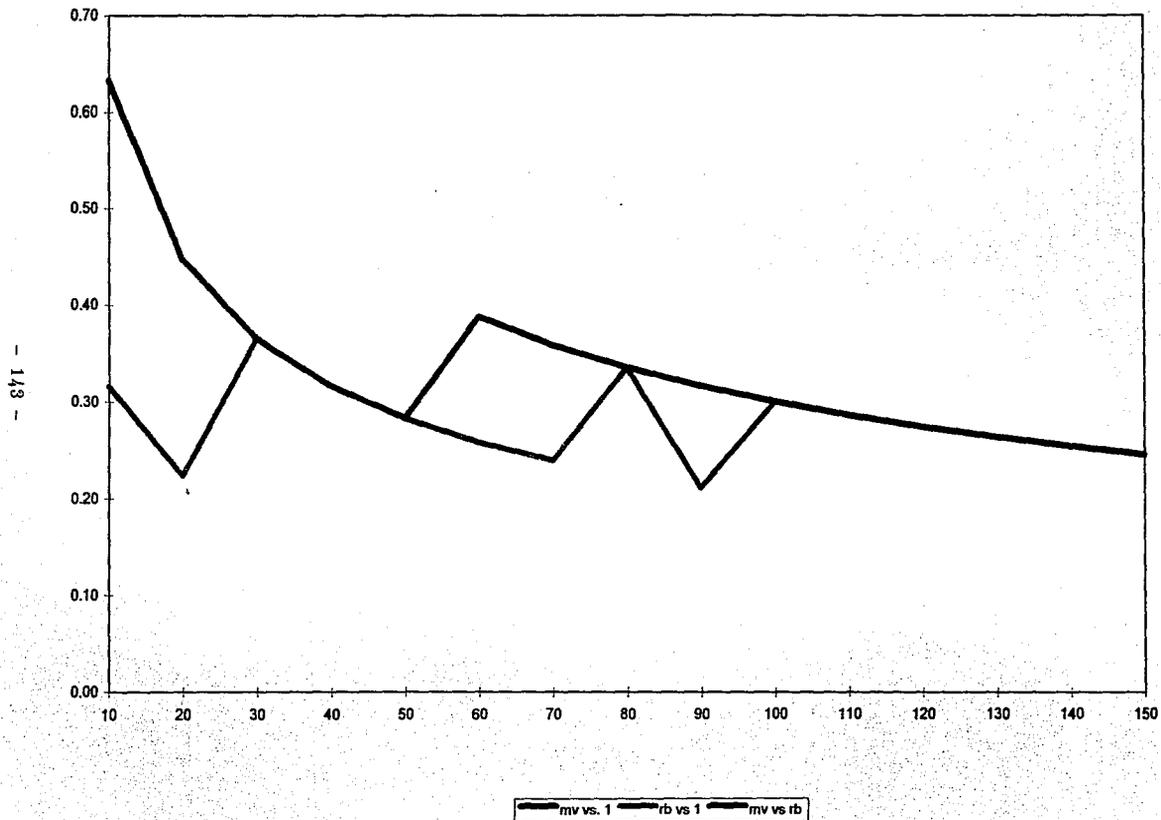
GRAFICA 19. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.0625)



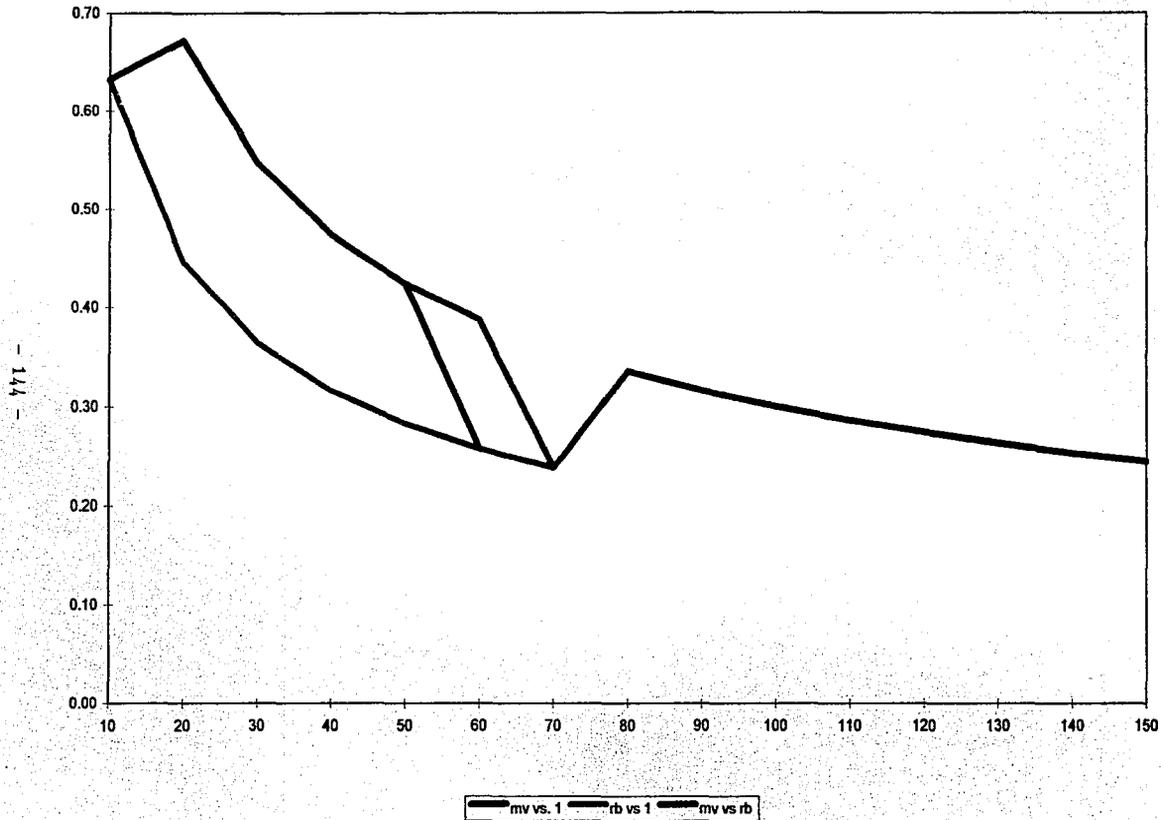
GRAFICA 20. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.125)



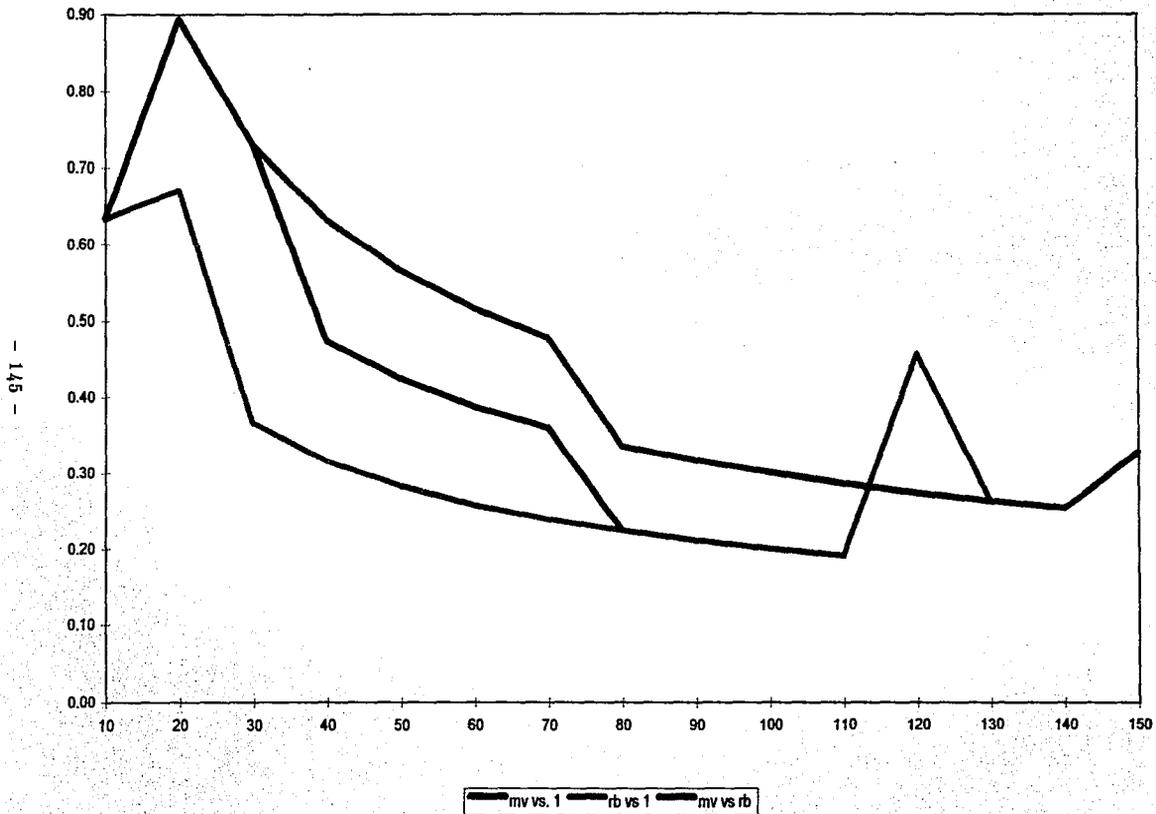
GRAFICA 21. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.25)



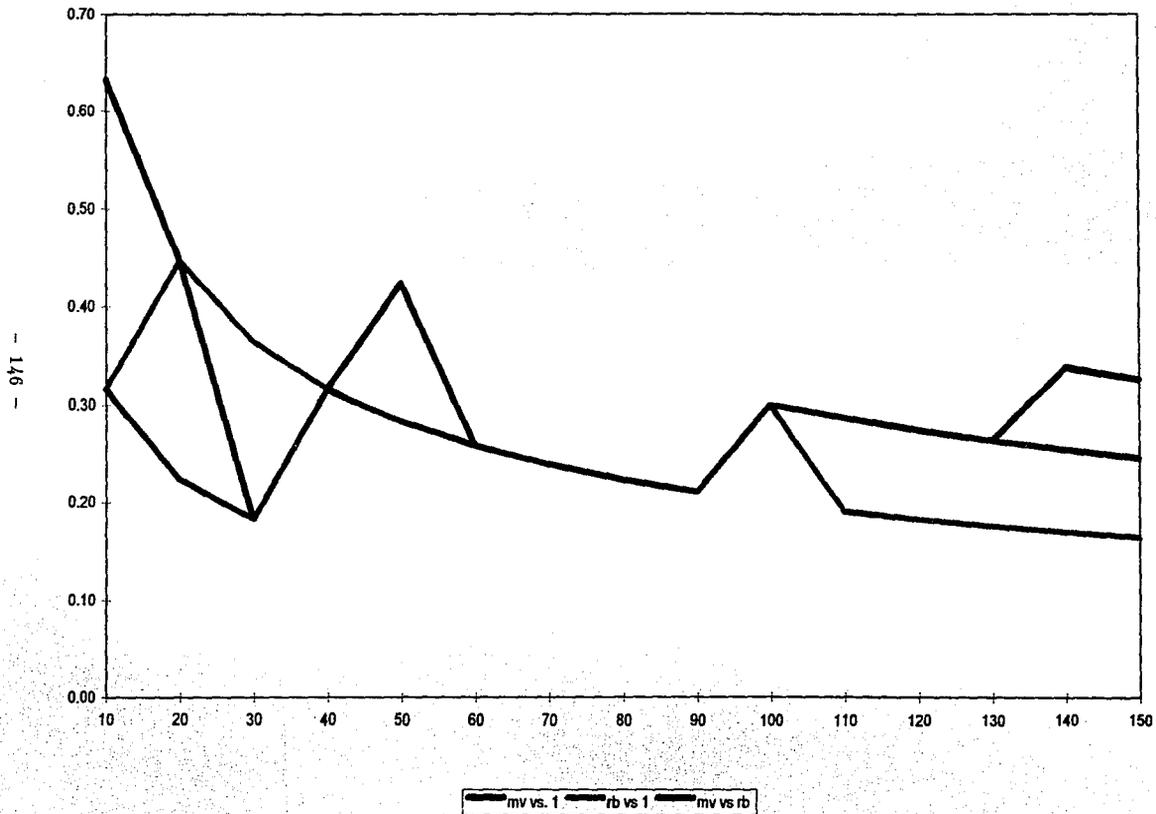
GRAFICA 22. SUPREMO DE LAS IDFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 0.5)



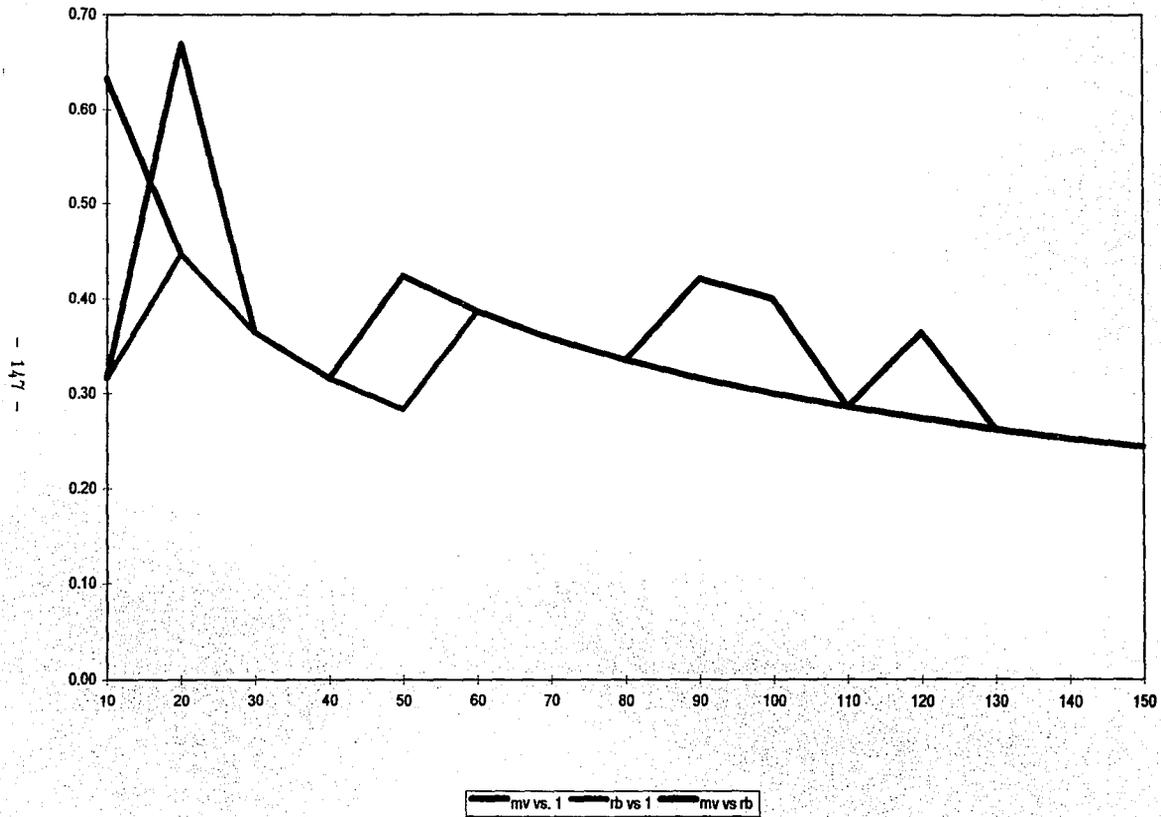
GRAFICA 23. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 1)



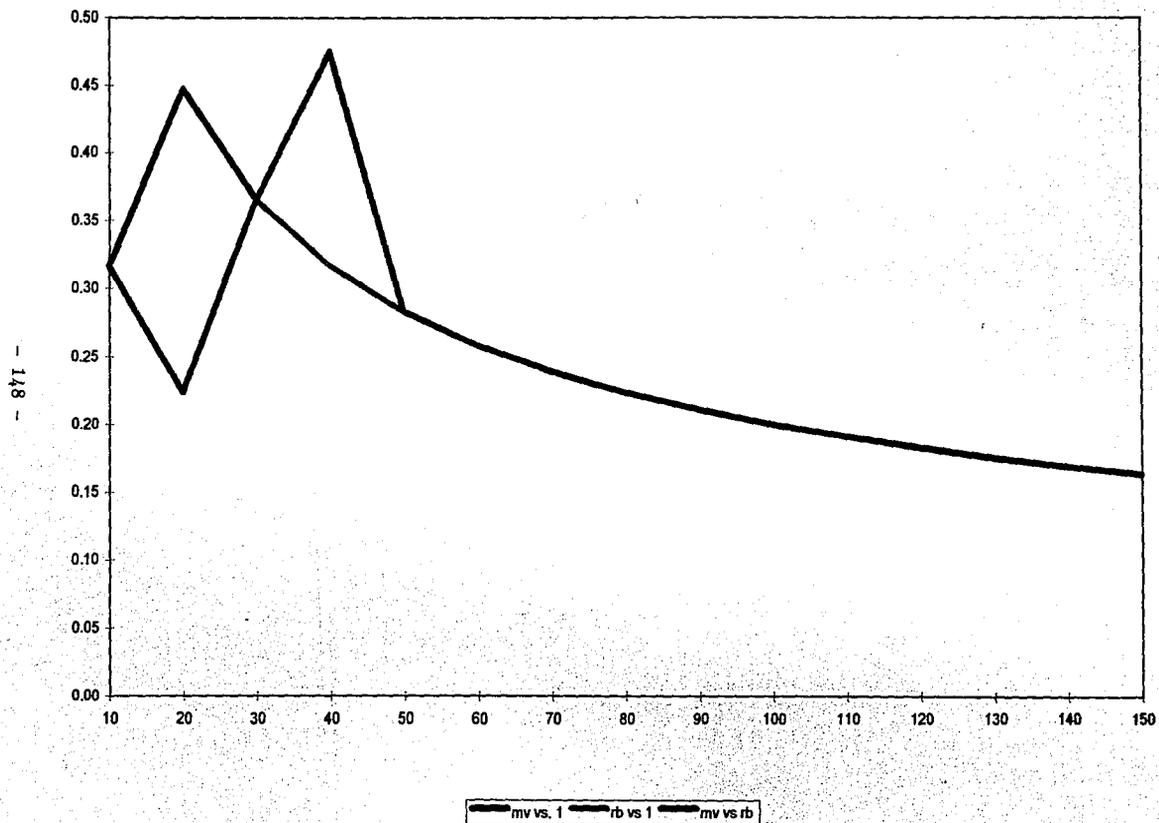
GRAFICA 24. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 2)



GRAFICA 25. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 4)



GRAFICA 26. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 8)



GRAFICA 27. SUPREMO DE LAS DIFERENCIAS ENTRE FUNCIONES DE DISTRIBUCION EMPIRICAS (PARAMETRO IGUAL A 16)

