

Distr.
RESTRINGIDA

LC/R.460
12 de octubre de 1985

ORIGINAL: ESPAÑOL

C E P A L

Comisión Económica para América Latina y el Caribe



METODO DE ESTIMACION DE LOS INGRESOS MEDIOS
EN DISTRIBUCIONES DE FRECUENCIAS AGRUPADAS

Este documento fue preparado por la División de Estadística y Análisis
Cuantitativo de la CEPAL.

85-10-1470

INDICE

| | <u>Página</u> |
|--|---------------|
| 1. INTRODUCCION ;..... | 1 |
| 2. METODOS DE ESTIMACION | 4 |
| 2.1 INGRESOS MEDIOS : GENERALIDADES | 4 |
| 2.2 ESTIMACION DEL INGRESO MEDIO DE LOS INTERVALOS INTERMEDIOS | 5 |
| 2.3 ESTIMACION DEL INGRESO MEDIO DEL PRIMER INTERVALO . | 5 |
| 2.3.1 Punto medio del intervalo | 5 |
| 2.3.2 Ajuste de una función cúbica | 6 |
| 2.4 ESTIMACION DEL INGRESO MEDIO DEL ULTIMO INTERVALO MEDIANTE EL AJUSTE DE UNA FUNCION DE PARETO | 8 |
| 2.4.1 La distribución de Pareto | 8 |
| 2.4.2 La estimación del ingreso medio del último intervalo | 8 |
| 2.4.3 Estimación de parámetros mediante cuantilas . | 10 |
| 2.4.4 Estimación de parámetros por máxima verosi- militud | 10 |
| 2.4.5 Estimación de parámetros por mínimos cuadrados | 11 |
| 2.4.6 Estimación por el método de los momentos | 12 |
| 3. RESULTADOS DE LOS ENSAYOS | 14 |
| 3.1 LA BASE DE DATOS UTILIZADA | 14 |
| 3.2 EL AJUSTE PARA LOS INTERVALOS INTERMEDIOS | 15 |
| 3.3 ESTIMACION DEL INGRESO MEDIO DEL PRIMER INTERVALO . | 15 |
| 3.4 ESTIMACION DEL INGRESO MEDIO DEL INTERVALO SUPERIOR | 17 |
| 3.4.1 Ajuste de una función de Pareto por el método de cuantilas | 17 |
| 3.4.2 Ajuste de una función de Pareto por el método de máxima verosimilitud..... | 18 |
| 3.4.3 Ajuste de una función de Pareto por el método de mínimos cuadrados | 19 |
| 4. EL PROCEDIMIENTO ELEGIDO | 20 |
| CUADROS ESTADÍSTICOS | 23 |

1. INTRODUCCION

Desde hace bastante tiempo economistas y estadísticos han venido trabajando en proponer una especificación funcional que describa apropiadamente la forma que sigue la distribución de los ingresos entre las personas y/o unidades familiares. A este fin se han dedicado ingentes esfuerzos teóricos y empíricos y la vasta literatura sobre el tema da cuenta de ello.

Sin embargo, en la actualidad no se puede hablar aún de que exista un consenso en esta materia, de suerte que cada investigación de los problemas vinculados a la distribución del ingreso se ve en la necesidad de satisfacer sus inevitables requerimientos de estimación estadística en base a particulares hipótesis y opciones metodológicas. Esto se explica además, en alguna medida, porque el fenómeno mismo de la distribución del ingreso presenta una realidad diversa entre países -según sus rasgos estructurales- y entre sectores sociales; a lo que debe sumarse los sesgos característicos de los instrumentos de medición, sesgos que suelen diferir tanto en magnitud como en la forma en que se distribuyen según los distintos tipos de ingreso y categoría ocupacional de los receptores.

De allí entonces que se ha considerado conveniente explicitar el método de estimación de la distribución del ingreso utilizado en

los estudios que, sobre el tema, viene realizando la División de Estadística y Análisis Cuantitativo de la CEPAL para los países de América Latina. Con tal objeto en esta nota se describe dicho método, junto a una breve reseña de otros procedimientos metodológicos alternativos, posibilitando así una cierta comparación a nivel de sus resultados. No obstante, estos últimos se incluyen sólo a manera de ilustración y no representan un examen exhaustivo sobre el particular. Igualmente, se debe señalar que los métodos aquí presentados se aplicaron en muchas ocasiones a antecedentes derivados de encuestas que constituyeron primeros esfuerzos de los países de la región en este campo. Tanto este hecho como la naturaleza de las desigualdades en la distribución del ingreso que mostraban esas encuestas tienen estrecha relación con el instrumental analítico utilizado.

Por otra parte, cabe observar que las mencionadas estimaciones son necesarias en todos aquellos casos en que la información disponible no resulta suficiente para estructurar una distribución de los ingresos. Aún cuando es preciso reconocer que esto ha venido cambiando en los últimos años, lo más común de encontrar todavía en las publicaciones realizadas por los países es que los datos de ingreso se presenten agrupados por tramos, sin incluir los ingresos agregados o ingresos medios de cada tramo; de modo que sólo se cuenta con la frecuencia de observaciones y el valor de los ingresos que acotan cada intervalo. Asimismo, cuando se desea con posterioridad cierta interpolación que reagrupe las unidades receptoras llevándolas a grupos decílicos o percentílicos cualesquiera, con el propósito por ejemplo de uniformar

la presentación de las distribuciones, también se requiere de alguna hipótesis en cuanto a la forma funcional que sigue la distribución de los ingresos.

Para tales efectos, en nuestro caso, la distribución se ha dividido en tres áreas: primer intervalo (o tramo de menores ingresos), intervalos intermedios e intervalo superior. Para cada uno de ellos se propone ajustar una función distinta, las cuales se han descrito teóricamente en la sección 2. La sección 3 exhibe los resultados de algunas pruebas empíricas, en tanto que la sección 4 sintetiza el método adoptado.

2. METODOS DE ESTIMACION

2.1 INGRESOS MEDIOS; GENERALIDADES

Conviene clarificar y precisar la terminología. Cuando se habla de "ingreso medio" se está aludiendo a la "marca de clase", o valor representativo de un intervalo de clase, en una distribución de frecuencias de ingresos con datos agrupados. Este es un concepto muestral.

Su contrapartida poblacional, cuando se conoce o se postula una distribución probabilística, es la media de la "distribución truncada" sobre el intervalo:

$$\mu (x', x'') = \frac{\int_{x'}^{x''} x f(x) dx}{\int_{x'}^{x''} f(x) dx} = \frac{\int_{x'}^{x''} x f(x) dx}{F(x'') - F(x')}, \quad (a)$$

donde $F(x)$ es la función de distribución y $f(x) = F'(x)$ la función de densidad.

El procedimiento de estimación más simple para el ingreso medio de un intervalo consiste, en principio, en asignar el punto medio del intervalo. Esta suposición, como se indica más arriba, equivale a suponer que todas las unidades dentro del intervalo tienen el mismo ingreso y que éste es igual al punto medio del intervalo.

Matemáticamente lo anterior se traduce en que se postula una función de densidad constante sobre el intervalo:

$$F(x) = \frac{n (x - x')}{N (x'' - x')}, \quad x \in [x', x''] \quad (b)$$

$$f(x) = \frac{n}{N (x'' - x')}$$

donde n es la frecuencia absoluta del primer intervalo y N el tamaño de la muestra. En efecto, la marca de clase es

$$\begin{aligned} \bar{x}(x', x'') &= \frac{\int_{x'}^{x''} x \frac{n}{N(x'' - x')} dx}{\frac{n}{N} - 0} = \frac{N}{n} \cdot \frac{n}{N(x'' - x')} \cdot \frac{x^2}{2} \Big|_{x'}^{x''} = \\ &= \frac{1}{x'' - x'} \cdot \frac{1}{2} (x''^2 - x'^2) = \frac{x' + x''}{2} \quad (d) \end{aligned}$$

Antes de exponer otros métodos de estimación, dejemos establecido que el estimador de la media poblacional será, para mayor claridad, denominado "ingreso medio general", para indicar que nos referimos a la distribución entera.

2.2. ESTIMACION DEL INGRESO MEDIO DE LOS INTERVALOS INTERMEDIOS

Los resultados empíricos, expuestos en la sección 3.2, indican que la estimación de los ingresos medios de los intervalos intermedios mediante el respectivo punto medio resulta muy satisfactoria. Por esta razón no se ha intentado otro procedimiento, dedicando esfuerzo teórico y computacional al primer y último intervalo. Las fórmulas aparecen en el párrafo precedente.

2.3. ESTIMACION DEL INGRESO MEDIO DEL PRIMER INTERVALO

2.3.1 Punto medio del intervalo

Contrariamente a lo señalado para el caso de los intervalos intermedios, los resultados empíricos indican que este supuesto es poco verosímil para el primer intervalo, en el que las unidades tienden sistemáticamente a agruparse más a medida que se alejan del valor mínimo (usualmente cero). O sea, el punto medio del intervalo es un estimador fuertemente sesgado que subestima al ingreso medio del primer intervalo.

2.3.2 Ajuste de una función cúbica

Dado que los resultados ofrecidos por el punto medio del primer intervalo subestiman considerablemente (ver sección 3.3) el ingreso medio observado, se plantea la idea de sustituir el tramo rectilíneo de la curva acumulada por una parábola cúbica:

$$y = F(x) = a + b(x-x_0) + c(x-x_0)^2 + d(x-x_0)^3, \quad x \in [x_0, x_1]$$

Para simplificar los desarrollos, pongamos $z = x-x_0$, $z_1 = x_1 - x_0$:

$$y = F(z) = a + bz + cz^2 + dz^3$$

Se imponen las siguientes condiciones:

- (i) La curva es una interpolación, o sea, pasa por los puntos $(0,0)$.

y (z_1, y_1) :

$$0 = a + b \cdot 0 + c \cdot 0 + d \cdot 0, \quad \text{o sea } a = 0 \quad (a)$$

$$y_1 = bz_1 + cz_1^2 + dz_1^3 \quad (b)$$

- (ii) Su pendiente en $(0,0)$ es nula:

$$y' = b + 2cz + 3dz^2 \quad (c)$$

$$0 = b + 2c \cdot 0 + 3d \cdot 0, \quad \text{o sea } b = 0 \quad (d)$$

- (iii) Su pendiente en (z_1, y_1) es igual a

$$F'(z_1) = h = \frac{1}{2} \left[\frac{y_1}{z_1} + \frac{y_2 - y_1}{z_2 - z_1} \right] = \frac{1}{2} \left[\frac{y_1}{x_1 - x_0} + \frac{y_2 - y_1}{x_2 - x_1} \right] \quad (e)$$

donde z_2 e y_2 corresponden al extremo derecho del segundo intervalo.

Como resumen de (a), (b), (c) y (d) se tiene

$$y = cz^2 + dz^3$$

$$y' = 2cz + 3dz^2$$

Evaluando (b) y (e) en (z_1, y_1) resulta el sistema de ecuaciones

$$\begin{cases} cz_1^2 + dz_1^3 = y_1 \\ 2cz_1 + 3dz_1^2 = h \end{cases} \quad (f)$$

cuyas soluciones son

$$c = \frac{3y_1 - hz_1}{z_1^2}$$

$$d = \frac{hz_1 - 2y_1}{z_1^3}$$

Pasemos al cálculo de \bar{x}_1 . Para empezar,

$$\bar{x}_1 = \bar{z}_1 + x_0$$

Usando 2.1 (a),

$$F(z_1) - F(z_0) = y_1$$

$$\begin{aligned} \int_0^{z_1} f(z) dz &= \int_0^{z_1} z(2cz + 3dz^2) dz = \frac{2}{3} cz_1^3 + \frac{3}{4} dz_1^4 = \\ &= \frac{2}{3} \cdot \frac{3y_1 - hz_1}{z_1^2} \cdot z_1^3 + \frac{3}{4} \cdot \frac{hz_1 - 2y_1}{z_1^3} \cdot z_1^4 = \\ &= \frac{z_1(6y_1 + hz_1)}{12} \end{aligned}$$

Finalmente,

$$\bar{x}_1 = \bar{z}_1 + x_0 = \frac{z_1(6y_1 + hz_1)}{12y_1} + x_0 = \frac{x_0 + x_1}{2} + \frac{h(x_1 - x_0)^2}{12y_1}$$

Se observará que este estimador es siempre mayor que el punto medio del intervalo (que es $(x_0 + x_1)/2$) y el efecto de esta corrección puede apreciarse en el cuadro 3.

2.4 ESTIMACION DEL INGRESO MEDIO DEL ULTIMO INTERVALO MEDIANTE EL AJUSTE DE UNA FUNCION DE PARETO

2.4.1 La distribución de Pareto

La función de distribución de una variable aleatoria según la ley de Pareto es

$$F(x) = 1 - \left(\frac{k}{x}\right)^\alpha, x \geq k > 0, \alpha > 0 \quad (a)$$

Por lo tanto su función de densidad será

$$f(x) = F'(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}} \quad (b)$$

Si $\alpha > 1$, la media de la distribución será

$$\mu = E(x) = \int_k^{+\infty} x \cdot \frac{\alpha k^\alpha}{x^{\alpha+1}} dx = \frac{\alpha k}{\alpha-1} \quad (c)$$

y su varianza

$$\begin{aligned} V(x) = E(x - \mu)^2 &= \int_k^{+\infty} (x - \mu)^2 \frac{\alpha k^\alpha}{x^{\alpha+1}} dx \\ &= \frac{\alpha k^2}{(\alpha-1)^2 (\alpha-2)} \end{aligned}$$

lo que exige que $\alpha > 2$. Esta condición no es fácilmente cumplida en la práctica, por lo que los momentos de orden superior a 1 son poco útiles en trabajos de estimación.

2.4.2 Estimación del ingreso medio del último intervalo

Estudiemos la aplicación a un intervalo cualquiera $[x', x'']$ de la fórmula 2.1.1(c):

$$\bar{x} = \frac{\int_{x'}^{x''} x f(x) dx}{F(x'') - F(x')} \quad (a)$$

que expresa la "marca de clase" de dicho conjunto.

En el caso de una distribución de Pareto, resulta

$$F(x'') - F(x') = \left(\frac{k}{x''}\right)^{\alpha} - \left(\frac{k}{x'}\right)^{\alpha} \quad (b)$$

Por otra parte,

$$\int_{x'}^{x''} x f(x) dx = \int_{x'}^{x''} x \frac{\alpha k^{\alpha}}{x^{\alpha+1}} dx = \frac{\alpha k^{\alpha}}{\alpha-1} (x''^{1-\alpha} - x'^{1-\alpha}) \quad (c)$$

Dividiendo (c) por (b) :

$$\bar{x} = \frac{\frac{\alpha k^{\alpha}}{\alpha-1} (x''^{1-\alpha} - x'^{1-\alpha})}{k^{\alpha} (x''^{-\alpha} - x'^{-\alpha})} = \frac{\alpha}{\alpha-1} \cdot \frac{x''^{1-\alpha} - x'^{1-\alpha}}{x''^{-\alpha} - x'^{-\alpha}} \quad (d)$$

En particular, para el último intervalo hay que hacer $x'' \longrightarrow +\infty$, lo que entrega el siguiente resultado

$$\bar{x}_{n+1} = \frac{\alpha}{\alpha-1} x_n \quad (e)$$

Un estimador para estas "marcas de clase", que gozará de la propiedad de consistencia, se obtiene insertando en (d) o (e) un estimador de α .

Los párrafos que vienen a continuación exponen los métodos de estimación de α que fueron ensayados y cuyos resultados figuran en la sección 3.4.

2.4.3 Estimación de parámetros mediante cuantiles

Se eligen dos niveles de probabilidad (en nuestro caso, las probabilidades acumuladas en los dos últimos intervalos conocidos) P_1 y P_2 y se determinan las correspondientes cuantiles x_{P_1} y x_{P_2} mediante

$$P_1 = 1 - \left(\frac{k}{x_{P_1}} \right)^\alpha \quad (a)$$

$$P_2 = 1 - \left(\frac{k}{x_{P_2}} \right)^\alpha \quad (b)$$

Tomando logaritmos,

$$\alpha \ln k - \alpha \ln x_{P_1} = \ln(1 - P_1)$$

$$\alpha \ln k - \alpha \ln x_{P_2} = \ln(1 - P_2)$$

$$\text{de donde } \hat{\alpha} = \frac{\ln \frac{1 - P_1}{1 - P_2}}{\ln \frac{x_{P_2}}{x_{P_1}}} \quad (c)$$

Dado que las cuantiles muestrales son estimadores consistentes de las cuantiles poblacionales, se sigue que $\hat{\alpha}$ es también consistente.

El valor de \hat{k} se obtiene sustituyendo $\hat{\alpha}$ en (a) o (b).

2.4.4 Estimación de parámetro por máxima verosimilitud

La función de verosimilitud de una muestra (x_1, \dots, x_n) de una población paretiana es

$$G_i = \frac{\alpha^n k^{\alpha n}}{\left(\prod_i x_i \right)^{\alpha + 1}}$$

Tomando logaritmos,

$$L = \ln G = n \ln \alpha + \alpha n \ln k - (\alpha + 1) \sum_i \ln x_i$$

Por lo tanto,

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} + n \ln k - \sum_i \ln x_i = 0$$

y obtenemos el estimador

$$\hat{\alpha} = \frac{n}{\sum_i \ln \left(\frac{x_i}{\hat{k}} \right)} \quad (2)$$

No es posible obtener \hat{k} por la misma técnica puesto que $\partial L / \partial k = \alpha n / k$ carece de máximo. Pero dado que k es el mínimo de la variable aleatoria x , L puede ser maximizada si

$$\hat{k} \leq \min_i x_i \quad (b)$$

Es evidente que esto se verifica cuando

$$\hat{k} = \min_i x_i = x_0 \quad (d)$$

También es posible demostrar la consistencia de estos estimadores máximo-verosímiles.

2.4.5 Estimación de parámetros por mínimos cuadrados

Para los efectos de este método, conviene "normalizar" los valores de x , dividiéndolos por su mínimo x_0 . De este modo, los intervalos quedan definidos en la forma

$$1 = x_0 < x_1 < \dots < x_n < x_{n+1} = +\infty.$$

En este caso, la función de distribución de Pareto depende sólo del parámetro α y su forma es

$$F(x) = 1 - x^{-\alpha}, \quad \alpha > 0, \quad x \geq 1.$$

Pero entonces, la variable $y = 1 - F(x) = x^{-\alpha}$ representa la distribución acumulada "descendente" ($y = \Pr(X \geq x)$). Tomando logaritmos,

$$\ln y = -\alpha \ln x \quad (a)$$

y esta es una relación lineal a la cual se puede aplicar la mecánica de los mínimos cuadrados. El estimador resultante es

$$\hat{\alpha} = - \frac{\sum \ln x \cdot \ln y}{\sum (\ln x)^2}$$

estimador que también es consistente.

2.4.6 Estimación por el método de los momentos

Consiste en igualar la media muestral con el valor esperado de la distribución (ver párrafo 2.4.1):

$$\bar{x} = \frac{\hat{\alpha} \hat{k}}{\hat{\alpha} - 1}$$

o sea

$$\hat{\alpha} = \frac{\bar{x}}{\bar{x} - \hat{k}} \quad (a)$$

La estimación de k se realiza del modo siguiente:

$$\Pr(X \geq x) = 1 - F(x) = \left(\frac{k}{x}\right)^{\alpha}$$

Por lo tanto la probabilidad de que los n valores muestrales sean mayores que x será $(k/x)^{\alpha n}$. Esta es también la probabilidad de que el menor valor observado sea mayor que x .

En consecuencia la función de distribución del mínimo muestral es

$$G(x) = 1 - \left(\frac{k}{x}\right)^{\alpha n}, \quad (d)$$

su densidad es $g(x) = \frac{\alpha n k^{\alpha n}}{x^{\alpha n + 1}}$

y su valor esperado $\int_k^{\infty} x \cdot \frac{\alpha n k^{\alpha n}}{x^{\alpha n + 1}} = \frac{\alpha n k}{\alpha n - 1}$ (e)

Igualando esta esperanza con el mínimo observado x_0 , se tendrá:

$$x_0 = \frac{\hat{\alpha} n k}{\hat{\alpha} n - 1}$$

o sea, $\hat{k} = \frac{(\hat{\alpha} n - 1) x_0}{\hat{\alpha} n}$ (f)

Por lo tanto, $\hat{\alpha} = \frac{\bar{x}}{\bar{x} - \frac{(\hat{\alpha} n - 1) x_0}{\hat{\alpha} n}}$

de donde $\hat{\alpha} = \frac{n\bar{x} - x_0}{n(\bar{x} - x_0)}$ (g)

Las ecuaciones (f) y (g) proporcionan los estimadores por el método de los momentos. Es posible demostrar que poseen la propiedad de consistencia.

3. RESULTADOS DE LOS ENSAYOS

3.1 LA BASE DE DATOS UTILIZADA

Para los efectos de ensayar los métodos descritos en el capítulo anterior relativo a la estimación de la marca de clase en una distribución de frecuencias de ingresos con datos agrupados, se seleccionó información de 11 encuestas de hogares pertenecientes a 6 países latinoamericanos. En todos esos casos las distribuciones disponibles registraban el número de unidades receptoras y el ingreso agregado de cada intervalo, posibilitando de este modo el cálculo del ingreso medio del tramo. De allí, entonces, la factibilidad de evaluar la discrepancia entre el "valor observado" de los ingresos medios y su correspondiente "valor estimado" obtenido mediante la aplicación de los procedimientos de ajuste.

En el cuadro 1 se sintetizan algunas características de las encuestas consideradas. Con excepción de los casos de Chile y México, la mayoría de esas encuestas forma parte de programas permanentes establecidos por los países y son de diversa naturaleza en cuanto al propósito central de la investigación: empleo y desempleo, ingresos y presupuestos familiares. Se hace notar este rasgo por cuanto es importante de tener en cuenta, al igual que el concepto de ingreso investigado, para la determinación de su probable incidencia en la confiabilidad de los datos de ingreso, así como en el tipo y magnitud de los sesgos y errores de medición, lo que puede afectar la bondad de los ajustes propuestos.

Los relevamientos señalados están dentro del período 1963-1978 y son de cobertura nacional, salvo Argentina 1975 que cubre sólo la Capital Federal y Gran Buenos Aires. Por otra parte, las distribuciones que sirven de base a este ejercicio corresponden a los hogares ordenados según clase y cuantía del ingreso familiar, y en el caso de Venezuela 1971 incluye además una distribución de los receptores individuales.

3.2 EL AJUSTE PARA LOS INTERVALOS INTERMEDIOS

Tal como se anticipara en la sección 2.2, los resultados empíricos sugieren en estos casos que la marca de clase del intervalo se comporta como un estimador satisfactorio del ingreso medio del tramo, lo que se verifica en prácticamente todas las distribuciones analizadas (ver cuadro 2). La magnitud de los desvíos es en general inferior al $\pm 5\%$ y en aquellas situaciones en que supera esta cifra -por ejemplo México 1968- suele corresponderse con tramos que concentran una alta frecuencia relativa de casos (igual o mayor al 10%), o bien se ubican muy cerca del extremo superior de la distribución. Cabe observar sí, respecto de esto último, que aún a nivel del percentil 90-95% el ajuste lineal para la determinación del ingreso medio es todavía razonablemente bueno. Por otra parte, el signo de las desviaciones no es sistemático y su significado pierde relevancia debido a la reducida dimensión de los sesgos.

Alternativamente se podría haber estimado los ingresos medios de estos intervalos por la vía del ajuste de una función de Pareto a todos los puntos ubicados a la derecha del punto modal. Sin embargo, éste método no fue ensayado, en parte por los excelentes resultados obtenidos mediante la interpolación lineal dentro de cada tramo de ingreso.

3.3 ESTIMACION DEL INGRESO MEDIO DEL PRIMER INTERVALO

Para la estimación del ingreso medio del primer intervalo de una distribución con datos agrupados se testearon dos hipótesis: la del punto medio entre un valor inicial igual a cero y el límite derecho del intervalo, y el ajuste de una parábola cúbica, basada en el supuesto de que las unidades tienden sistemáticamente a agruparse más a medida que se alejan

del valor mínimo. En ambos casos se intentó medir, además, una cierta sensibilidad de la hipótesis de estimación con respecto a la longitud del primer intervalo, para lo cual, en aquellas distribuciones que resultaba factible, se agruparon sucesivamente los dos o tres primeros tramos de modo de aumentar la frecuencia acumulada de observaciones hasta un nivel cercano al 15%.

Los resultados de este ejercicio están contenidos en el cuadro 3. En él se aprecia, primeramente, que la opción del punto medio, a diferencia de los casos tratados en el acápite 3.2, proporciona estimaciones fuertemente sesgadas, cualquiera sea el nivel de frecuencia del primer intervalo. Esto último es válido incluso para los intervalos pequeños (inferiores al 5% de casos). En concreto se observa una subestimación sistemática que oscila entre el 17% y el 32%, donde además la mayoría de los casos presenta una subestimación del ingreso medio del tramo que supera el 20%.

Por su parte, en lo relativo al ajuste de una parábola cúbica, los resultados son bastante más satisfactorios. Por la forma de especificación de la función descrita en el apartado 2.3.2 el estimador es siempre mayor que el punto medio, lo que tiende a contrarrestar el sesgo implícito en este último. No obstante, para 15 de las 19 observaciones este factor de "desplazamiento" del ingreso medio estimado resultó insuficiente, manteniéndose siempre algún grado (menor) de subestimación. Las discrepancias obtenidas fueron en 8 casos inferiores al -10%, de los cuales la mitad se ubica bajo el -5%. En general, a medida que el intervalo es de menor frecuencia relativa el ajuste tiende a mejorar.

3.4 ESTIMACION DEL INGRESO MEDIO DEL INTERVALO SUPERIOR

3.4.1 AJUSTE DE UNA FUNCION DE PARETO POR EL METODO DE CUANTILAS

El primer procedimiento ensayado para la estimación de los parámetros de una función de Pareto fue el descrito en la sección 2.4.3 y conocido como método de las cuantilas. Los resultados obtenidos de sus diversas aplicaciones se recogen en el cuadro 4. Allí resalta, primero que nada, una clara tendencia a la sobre-estimación de los ingresos medios del intervalo final de la distribución, el que generalmente se presenta abierto en su extremo derecho. Esto es equivalente a señalar de que el α estimado sería sistemáticamente menor que el α "implícito" en los datos observados, cuestión que a su vez no parece depender de manera estrecha -según esta evidencia- de la probabilidad acumulada hasta el penúltimo y último tramo, dentro de los límites de lo que podría considerarse como el segmento paretiano de la distribución (10 a 15% de las observaciones de más altos ingresos).

La mayoría de los ajustes efectuados mediante este método arrojaron sobreestimaciones del ingreso medio del intervalo superior abierto que oscila en torno a un 20%, o algo por debajo de este nivel. Aquellos casos en que no fue así y la discrepancia excede dicho porcentaje, expresan una mayor disparidad entre el valor estimado para el parámetro α y su valor observado, o bien el hecho que para el primero de ellos su valor absoluto resultó excepcionalmente bajo (inferior a 2.00).^{1/} En particular para este último evento se propuso entonces la fijación de α , asignándole exógenamente un valor igual a 2, procedimiento que, como es obvio, permitió reducir el margen de sobreestimación de los ingresos medios. Estos resultados también se incluyen en el cuadro 4 donde además se puede percibir

^{1/} Ya en la sección 2.4.1 se hizo referencia a que situaciones como éstas dificultan los trabajos de estimación.

que para un tercio de los casos la citada reducción llega a extremos en que el error cambia de signo pero siempre estableciéndose una disminución en el módulo de la discrepancia.

3.4.2 Ajuste de una función de Pareto por el método de máxima verosimilitud

Otro método utilizado para estimar los parámetros de una función de Pareto fue el de máxima verosimilitud (sección 2.4.4). En este caso los ensayos se efectuaron aplicando el ajuste a los tramos finales de la distribución, donde se hizo variar el largo de dicho segmento de modo de incorporar cada vez una mayor cantidad de puntos. Los resultados (ver cuadro 5) se organizaron en torno al 5%, 10% y 20% de las unidades receptoras, de más altos ingresos, pero siempre la discrepancia en el ingreso medio estimado corresponde al último intervalo tal como éste viene definido en la publicación de los datos originales.

A medida que crece la longitud de "cola" considerada se observa la tendencia a un aumento de la sobreestimación del ingreso medio del intervalo final, o a que disminuya la subestimación. Es así como al nivel del 5% en la mayoría de los casos se subestima, aunque en porcentajes muy variables: de un máximo de 42% (Chile 1967-68) a un mínimo de 3% (Venezuela 1971 H). Al pasar al nivel del 10% la situación se hace menos nítida, dado que persisten casos con marcado sesgo negativo (Chile 1967-68 : 38%) frente a otros donde se acentúa a límites extremos la sobreestimación (Colombia 1971 : 64%). Esta tendencia se refuerza luego aún más al ampliar el segmento al cual se aplica la función de Pareto, cuestión que se aprecia al llevar el ajuste a los niveles del 20% de las observaciones de mayores ingresos.

Por otro lado, en cuanto al número de puntos comprendidos en cada segmento y que participan en el ajuste, no parece posible desprender una incidencia clara sobre los resultados del mismo.

3.4.3 Ajuste de una función de Pareto por el método de mínimos cuadrados

Finalmente, se ensayó también el método de los mínimos cuadrados (sección 2.4.5). Su aplicación, así como la presentación de resultados empíricos, responde a un esquema similar al empleado en el procedimiento anterior. Aunque el cuadro 6 recoge información un poco más incompleta que la exhibida en el análisis previo, es igualmente clara la tendencia que mientras más grande el segmento de la distribución seleccionado para el ajuste, mayor es el error implícito en la estimación del ingreso medio del último intervalo. Tanto más cuanto que los sesgos en este caso son prácticamente todos positivos. Las únicas desviaciones de la tendencia señalada se observan para Argentina 1975 y Chile 1967-68.

Ahora bien, consecuente con esto sólo cabe agregar que es en el ajuste a nivel del menor segmento estudiado (alrededor del 5%) donde se da un mayor número de estimaciones con un coeficiente de discrepancias razonablemente bueno (inferior a 13%). Para el resto de los niveles (10% y 20%) la calidad de los ajustes empieza a sufrir un evidente deterioro. Tampoco se visualiza aquí un vínculo claro entre la bondad de la estimación y el número de puntos incluidos en cada segmento.

4. EL PROCEDIMIENTO ELEGIDO

En base a los atributos teóricos y propiedades estadísticas de las funciones descritas, así como de los resultados de las aplicaciones empíricas, se adoptó un método estandar de estimación de la distribución de los ingresos para todos aquellos casos en que, conociéndose un perfil de datos agrupados, éste no incluya los ingresos medios ni los ingresos agregados de cada intervalo.

En ausencia de una forma funcional que represente adecuadamente la distribución de ingreso en toda su longitud, dicho procedimiento parte por dividir la distribución total en tres segmentos distintos. En primer lugar, aquel tramo que reúne las unidades receptoras de más bajos ingresos, o intervalo "inferior" de la distribución. En este caso y por los motivos expuestos en la sección 3.3, se asume el ajuste de una función de tercer grado, la cual comparada con la hipótesis del punto medio del tramo, desplaza hacia la derecha el valor estimado de los ingresos medios. Los sesgos de subestimación implícitos en esta alternativa se ubican dentro de márgenes aceptables en la mayoría de los ajustes realizados, sin perjuicio de que la escasa masa de ingresos captada corrientemente por las observaciones de este intervalo concede una mayor tolerancia de los errores respecto de su incidencia en la estructura global de la distribución.

La fórmula de cálculo de los ingresos medios del primer intervalo está dada por:

$$\bar{x}_1 = \frac{x_1}{2} + H \frac{x_1^2}{12y_1}$$

$$\text{donde } H = \frac{1}{2} \left[\frac{y_1}{x_1} + \frac{y_2 - y_1}{x_2 - x_1} \right]$$

x_i ($i = 1, 2$) : valor del ingreso del límite derecho del intervalo i ^{2/}

y_i ($i = 1, 2$) : probabilidad acumulada de casos hasta el intervalo i

Luego se define el segmento de los intervalos intermedios o no extremos, para los cuales la evidencia presentada avala fuertemente la elección del punto medio del tramo (o ajuste lineal) como estimador de los ingresos medios.

Vale decir,

$$\bar{x}_i = \frac{x_i - x_{(i-1)}}{2}$$

x_i ($i = 2, \dots, n$) : valor del ingreso del límite derecho del intervalo i

Finalmente se especifica el intervalo superior o "cola" de la distribución. Por lo general este es un tramo que se presenta no acotado en su extremo derecho y tiene gran importancia debido a que, pese a acumular una reducida frecuencia de casos, suele en cambio captar una

^{2/} Recordemos que el ajuste de esta parábola cúbica es una interpolación que supone el paso de la curva por el punto (0, 0). (Ver sección 2.3.2).

alta proporción del ingreso global. Este hecho hace que pequeños errores en la estimación de los ingresos del intervalo afecten sensiblemente el perfil distributivo y su grado de concentración.

En este último segmento, que reúne a las unidades receptoras de más altos ingresos, se aplicó el ajuste de una función de Pareto. A su vez, de los diversos métodos de estimación de los parámetros de una función de este tipo reseñados en el acápite 2.4, se viene aplicando preferentemente el método de las cuantilas, tanto por la reducida diferencia que exhiben los resultados comparados con aquellos obtenidos de la aplicación de los otros procedimientos, como por su mayor simplicidad operacional.

En definitiva, entonces, la fórmula utilizada para la estimación del ingreso medio del intervalo superior abierto es:

$$\bar{x}_{n+1} = \frac{\alpha}{\alpha - 1} x_n$$

$$\text{donde } \alpha = \frac{\log\left(\frac{1 - y_n}{1 - y_{n-1}}\right)}{\log\left(\frac{x_{n-1}}{x_n}\right)}$$

- x_n : valor del ingreso del límite derecho del penúltimo intervalo
- x_{n-1} : valor del ingreso del límite derecho del antepenúltimo intervalo
- y_n : probabilidad acumulada de casos hasta el penúltimo intervalo
- y_{n-1} : probabilidad acumulada de casos hasta el antepenúltimo intervalo.

CUADROS ESTADÍSTICOS

CARACTERÍSTICAS DE LAS ENCUESTAS UTILIZADAS

| PAIS | Nombre de la Encuesta | Año | Tipo de encuesta a/ | Cober- tura geográ- fica b/ | Ingreso investigado | |
|-----------|---|---------|------------------------|---|-------------------------|------------------------|
| | | | | | Recep- tores | Tipos de ingreso |
| ARGENTINA | Encuesta Permanente de Hogares | 1975 | E | A.M. | Todos | Todos |
| BRASIL | Pesquisa Nacional por Amostra de Domicilios (PNAD) | 1972 | I | N | Todos | Todos |
| | id | 1976 | E | N | Todos | Todos |
| COLOMBIA | Encuesta Nacional de Hogares | 1971 | P.F. | N | Todos | Todos |
| | id | 1972 | P.F. | N | Todos | Todos |
| CHILE | Encuesta Continua de Mano de Obra | 1967-68 | E | N | Ocupados no agrícola | Ocupación principal |
| | id | 1968 | I | N | Todos | Todos |
| México | Encuesta sobre Ingresos y Gastos Familiares | 1963 | P.F. | N | Todos | Todos |
| | Estudio de Ingresos y Gastos de las Familias | 1968 | P.F. | N | Todos | Todos |
| | Encuesta Nacional de Ingresos y Gastos | 1977 | P.F. | N | Todos | Todos |

(Continúa)

Cuadro 1 (conclusión)

| PAIS | Nombre de la Encuesta | Año | Tipo de encuesta <u>a/</u> | Cober- tura geográ- fica <u>b/</u> | Ingreso investigado | |
|--------------|-----------------------|------|-------------------------------|--|---------------------|------------------------|
| | | | | | Recep- tores | Tipos de ingreso |
| VENEZUELA | Encuesta de Hogares | 1971 | E | N | Activos no agrícola | Ingresos en efectivo |
| (Receptores) | id | 1971 | E | N | Activos no agrícola | Ingresos en efectivo |

a/ E: Empleo y Desempleo

I: Ingresos

P.F.: Presupuestos Familiares

b/ A.M.: Area Metropolitana

N: Nacional

Cuadro 2 (continuación)

| Colombia 1972 | | | Chile 1967-68 | | | Chile 1968 | | | México 1963 | | |
|------------------------|------------------------------|---|------------------------|------------------------------|---|------------------------|------------------------------|---|------------------------|------------------------------|---|
| Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) |
| 10.3-35.6 | 25.3 | -3.75 | 11.6-22.0 | 10.4 | -0.93 | 28.4-60.6 | 32.2 | 3.78 | 18.5-43.5 | 25.0 | 2.90 |
| 35.7-58.3 | 22.6 | 0.24 | 22.1-32.9 | 10.8 | -0.75 | 60.7-78.6 | 17.9 | 2.66 | 43.6-65.1 | 21.5 | 2.97 |
| 58.4-72.1 | 13.7 | -0.38 | 33.0-47.8 | 14.8 | -2.19 | 78.7-86.1 | 7.4 | 2.41 | 65.2-76.0 | 10.8 | 0.76 |
| 72.2-78.7 | 6.5 | -1.68 | 47.9-60.5 | 12.6 | -0.08 | 86.2-90.7 | 4.5 | 1.93 | 76.1-91.0 | 14.9 | 7.12 |
| 78.8-85.0 | 6.2 | -0.38 | 60.6-69.4 | 8.8 | -0.92 | 90.8-93.7 | 2.9 | 1.36 | 91.1-95.6 | 4.5 | 2.72 |
| 85.1-90.3 | 5.2 | 0.37 | 69.5-79.0 | 9.5 | -0.29 | 93.8-96.4 | 2.6 | 3.10 | 95.7-97.5 | 1.8 | 0.92 |
| 90.4-95.2 | 4.8 | 3.62 | 79.1-85.4 | 6.3 | -0.64 | 96.5-98.0 | 1.5 | 2.62 | 97.6-99.1 | 1.5 | 4.40 |
| 95.3-97.9 | 2.6 | 4.29 | 85.5-90.1 | 4.6 | -0.74 | | | | | | |
| | | | 90.2-93.1 | 2.9 | -0.87 | | | | | | |
| | | | 93.2-95.8 | 2.6 | -0.77 | | | | | | |
| | | | 95.9-97.0 | 1.1 | -0.73 | | | | | | |

Cuadro 2 (conclusión)

| México 1968 | | | México 1977 | | | Venezuela 1971 (H) | | | Venezuela 1971 (R) | | |
|------------------------|------------------------------|---|------------------------|------------------------------|---|------------------------|------------------------------|---|------------------------|------------------------------|---|
| Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | Percentiles de Hogares | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) |
| 0.7- 3.9 | 3.2 | 7.34 | 10.3-16.5 | 6.2 | -2.60 | 6.0-17.3 | 11.3 | -3.02 | 11.0-26.8 | 15.8 | -2.60 |
| 4.0-11.6 | 7.6 | 4.48 | 16.6-23.2 | 6.6 | -1.22 | 17.4-36.5 | 19.1 | -1.75 | 26.9-46.2 | 19.3 | -1.96 |
| 11.7-45.2 | 33.5 | 22.59 | 23.3-31.8 | 8.5 | -0.16 | 36.6-46.8 | 10.2 | -1.82 | 46.3-64.4 | 18.1 | -1.80 |
| 45.3-70.6 | 25.3 | 16.38 | 31.9-41.3 | 9.4 | -0.64 | 46.9-55.1 | 8.2 | -1.96 | 64.5-80.7 | 16.2 | -1.62 |
| 70.7-82.8 | 12.1 | 11.24 | 41.4-51.5 | 10.1 | -0.62 | 55.2-69.1 | 13.9 | -0.7 | 80.8-91.0 | 10.2 | -0.95 |
| | | | 51.6-64.1 | 12.5 | 1.50 | 69.2-79.1 | 9.9 | -1.07 | 91.1-95.2 | 4.1 | -3.69 |
| | | | 64.2-74.5 | 10.3 | 1.52 | 79.2-85.1 | 5.9 | -1.40 | 95.3-96.8 | 1.5 | -4.38 |
| | | | 74.6-83.3 | 8.7 | 1.49 | 85.2-89.8 | 4.6 | -1.65 | 96.9-98.3 | 1.4 | -6.11 |
| | | | 83.4-90.3 | 6.9 | 0.13 | 89.9-95.9 | 6.0 | 0.94 | | | |
| | | | 90.4-94.5 | 4.1 | 0.81 | 96.0-98.2 | 2.2 | -1.20 | | | |
| | | | 94.6-97.4 | 2.8 | 1.73 | 98.3-99.2 | 0.9 | 1.57 | | | |

Cuadro 3

ESTIMACIONES DEL INGRESO MEDIO DEL PRIMER INTERVALO

| | Frecuencia relativa de casos | Discrepancia del ingreso medio estimado (%) | |
|--------------------|------------------------------------|---|-------------------|
| | | Punto medio | Función cúbica |
| ARGENTINA 1975 | 4.8 | - 25.00 | 5.49 |
| | 10.4 | - 29.85 | - 10.87 |
| | 15.7 | - 31.97 | - 19.45 |
| BRASIL 1972 | 6.5 | - 22.09 | - 4.21 |
| | 17.9 | - 17.49 | - 2.83 |
| BRASIL 1976 | 17.4 | - 21.84 | - 6.12 |
| COLOMBIA 1971 | 19.6 | - 30.43 | - 15.96 |
| COLOMBIA 1972 | 10.2 | - 31.79 | - 11.95 |
| CHILE 1967-68 | 28.3 | - 20.31 | - 6.11 |
| MEXICO 1963 | 18.4 | - 30.43 | - 16.72 |
| MEXICO 1968 | 0.6 | - 30.59 | 7.01 |
| | 3.9 | - 22.68 | 2.84 |
| | 11.6 | - 21.72 | - 5.74 |
| MEXICO 1977 | 10.2 | - 18.60 | - 2.05 |
| | 16.5 | - 16.53 | - 1.50 |
| VENEZUELA 1971 (H) | 5.9 | - 27.88 | - 10.26 |
| | 17.3 | - 20.34 | - 6.33 |
| VENEZUELA 1971 (R) | 10.9 | - 26.47 | - 11.40 |

ESTIMACIONES DEL INGRESO MEDIO DEL ULTIMO INTERVALO POR AJUSTE DE
UNA FUNCIÓN DE PARETO: METODO DE CUANTILAS

| Encuesta | Frecuencia relativa de casos | | α "Observed" | α "Estimated" | Discrepancia del ingreso medio estimado(%) | |
|----------------|------------------------------|---------------------|------------------------|-------------------------|--|------------------|
| | Ultimo intervalo | Penúltimo intervalo | | | Con $\hat{\alpha}$ | Con $\alpha = 2$ |
| ARGENTINA 1975 | 5.1 | 4.7 | 3.76 | 2.80 | 14.18 | |
| | 5.1 | 9.8 | 3.76 | 2.67 | 17.35 | |
| | 9.8 | 5.1 | 3.31 | 2.54 | 15.09 | |
| | 9.8 | 10.4 | 3.31 | 2.76 | 9.37 | |
| BRASIL 1972 | 0.2 | 0.3 | 1.92 | 2.26 | - 14.27 | |
| | 0.2 | 0.6 | 1.92 | 2.00 | - 4.41 | |
| | 0.5 | 0.3 | 2.04 | 1.63 | 31.52 | 2.04 |
| | 0.8 | 2.0 | 1.94 | 1.81 | 8.29 | - 3.25 |
| | 2.8 | 2.8 | 1.86 | 1.71 | 11.32 | - 7.60 |
| | 5.6 | 3.8 | 1.81 | 1.45 | 43.66 | -10.42 |
| | 9.4 | 4.8 | 1.73 | 1.23 | 129.04 | -15.77 |
| BRASIL 1976 | 1.4 | 2.8 | 2.13 | 1.58 | 44.76 | 6.28 |
| | 1.4 | 10.0 | 2.13 | 1.51 | 57.33 | |
| | 4.2 | 7.2 | 1.84 | 1.44 | 49.09 | - 8.89 |
| | 11.4 | 9.7 | 1.67 | 1.21 | 132.17 | -19.41 |

(Continúa)

Cuadro 4 (continuación)

| Encuesta | Frecuencia relativa de casos | | α "Observado" | α "Estimado" | Discrepancia del ingreso medio estimado (%) | |
|---------------|------------------------------|---------------------|-------------------------|------------------------|---|------------------|
| | Ultimo intervalo | Penúltimo intervalo | | | Con $\hat{\alpha}$ | Con $\alpha = 2$ |
| COLOMBIA 1971 | 1.1 | 2.1 | 6.13 | 2.09 | 60.45 | |
| | 1.1 | 5.5 | 6.13 | 1.96 | 71.29 | 67.39 |
| | 3.2 | 3.4 | 3.09 | 1.79 | 53.80 | 35.32 |
| | 6.6 | 3.3 | 2.44 | 1.41 | 103.18 | 18.05 |
| COLOMBIA 1972 | 2.1 | 2.7 | 2.56 | 1.62 | 59.54 | 21.91 |
| | 2.1 | 7.6 | 2.56 | 1.67 | 51.94 | |
| | 4.8 | 4.9 | 2.09 | 1.74 | 23.24 | 4.42 |
| | 4.8 | 10.2 | 2.09 | 1.64 | 33.79 | |
| | 9.7 | 5.3 | 1.97 | 1.52 | 43.92 | - 1.53 |
| CHILE 1967-68 | 3.0 | 1.2 | 1.81 | 1.51 | 32.50 | - 10.51 |
| | 3.0 | 3.9 | 1.81 | 1.63 | 15.77 | |
| | 4.2 | 2.7 | 1.76 | 1.72 | 3.00 | - 13.77 |
| | 6.9 | 3.0 | 1.74 | 1.98 | - 13.78 | - 14.64 |
| | 9.9 | 4.7 | 1.77 | 1.94 | - 10.11 | - 12.90 |
| CHILE 1968 | 2.0 | 1.6 | 3.43 | 2.63 | 14.17 | |
| | 2.0 | 4.3 | 3.43 | 2.25 | 27.66 | |
| | 3.6 | 2.7 | 3.14 | 1.95 | 40.21 | 36.27 |
| | 3.6 | 5.7 | 3.14 | 2.02 | 34.96 | |
| | 6.3 | 7.6 | 2.66 | 1.95 | 28.04 | 24.75 |

(Continúa)

Cuadro 4 (continuación)

| Encuesta | Frecuencia relativa de casos | | α "Observado" | α "Estimado" | Discrepancia del ingreso medio estimado(%) | |
|--------------------|------------------------------|---------------------|-------------------------|------------------------|--|------------------|
| | Ultimo intervalo | Penúltimo intervalo | | | Con $\hat{\alpha}$ | Con $\alpha = 2$ |
| MEXICO 1963 | 0.9 | 1.6 | 3.83 | 2.00 | 47.73 | |
| | 0.9 | 3.5 | 3.83 | 1.66 | 85.78 | 47.73 |
| | 2.5 | 1.9 | 2.59 | 1.97 | 24.95 | 22.73 |
| | 2.5 | 6.5 | 2.59 | 1.85 | 33.56 | |
| | 4.4 | 4.6 | 2.36 | 1.76 | 33.08 | 15.35 |
| MEXICO 1968 | 17.2 | 12.2 | 2.52 | 1.05 | 1180.24 | 20.65 |
| MEXICO 1977 | 2.6 | 2.9 | 2.80 | 2.54 | 6.06 | |
| | 2.6 | 7.1 | 2.80 | 2.30 | 13.77 | |
| | 5.5 | 4.2 | 2.68 | 2.04 | 23.00 | |
| | 5.5 | 11.2 | 2.68 | 1.91 | 31.61 | 25.41 |
| | 9.7 | 7.0 | 2.45 | 1.80 | 32.98 | 18.20 |
| VENEZUELA 1971 (4) | 0.8 | 1.0 | 3.75 | 2.82 | 13.63 | |
| | 0.8 | 3.3 | 3.75 | 2.84 | 13.19 | |
| | 1.8 | 2.3 | 3.24 | 2.86 | 6.32 | |
| | 4.1 | 6.1 | 2.96 | 2.25 | 19.28 | |
| | 4.1 | 10.8 | 2.96 | 2.20 | 21.49 | |
| | 10.2 | 4.7 | 2.54 | 2.08 | 16.70 | |

(Continúa)

Cuadro 4 (conclusión)

| Encuesta | Frecuencia relativa de casos | | α "Observed" | α "Estimated" | Discrepancia del ingreso medio estimado % | |
|--------------------|---------------------------------|------------------------|------------------------|-------------------------|---|------------------|
| | Ultimo intervalo | Penúltimo intervalo | | | Con $\hat{\alpha}$ | Con $\alpha = 2$ |
| VENEZUELA 1971 (R) | 1.7 | 1.5 | 2.85 | 3.47 | -8.85 | |
| | 1.7 | 3.1 | 2.85 | 2.56 | 6.47 | |
| | 3.2 | 1.6 | 2.88 | 1.82 | 44.90 | 30.57 |
| | 3.2 | 5.8 | 2.88 | 2.02 | 29.29 | |
| | 4.8 | 4.2 | 2.50 | 2.19 | 10.30 | |

Cuadro 5

ESTIMACIONES DEL INGRESO MEDIO DEL ULTIMO INTERVALO POR AJUSTE DE UNA FUNCION DE PARETO:
METODO DE MAXIMA VEROSIMILITUD

| Encuesta | Ca. 5% | | | Ca. 10% | | | Ca. 20% | | |
|----------------|--|------------------------|-------------------|--|------------------------|-------------------|--|------------------------|-------------------|
| | Discrepancia del ingreso medio estima- do (%) | Número de puntos | μ Estimado | Discrepancia del ingreso medio estima- do (%) | Número de puntos | μ Estimado | Discrepancia del ingreso medio estima- do (%) | Número de puntos | μ Estimado |
| Argentina 1975 | | | | -16.08 (9.8) | 2 | 7.98 | -10.08 (14.9) | 3 | 5.44 |
| Brasil 1972 | 5.88 (2.8) | 4 | 1.82 | 28.11 (9.4) | 6 | 1.60 | 57.11 (14.2) | 7 | 1.44 |
| Brasil 1976 | -10.62 (4.2) | 2 | 2.47 | 32.69 (11.4) | 3 | 1.67 | 80.31 (21.1) | 4 | 1.42 |
| Colombia 1971 | 17.49 (3.2) | 2 | 3.48 | 63.86 (9.9) | 4 | 2.04 | 81.98 (15.3) | 5 | 1.85 |
| Colombia 1972 | | | | -0.05 (9.7) | 3 | 2.56 | 13.44 (15.0) | 4 | 2.16 |
| Chile 1967-68 | -41.76 (6.9) | 3 | 4.32 | -38.23 (9.9) | 4 | 3.63 | -33.68 (14.6) | 5 | 3.07 |
| Chile 1968 | -19.72 (3.6) | 2 | 8.49 | 0.10 (9.3) | 4 | 3.42 | 10.98 (13.9) | 5 | 2.76 |

Cuadro 5 (conclusión)

| Encuesta | Ca. 5% | | | Ca. 10% | | | Ca. 20% | | |
|--------------------|---|------------------|-------------------|---|------------------|-------------------|---|------------------|-------------------|
| | Discrepancia del ingreso medio estimado (%) | Número de puntos | α Estimado | Discrepancia del ingreso medio estimado (%) | Número de puntos | α Estimado | Discrepancia del ingreso medio estimado (%) | Número de puntos | α Estimado |
| México 1963 | 13.06 (4.4) | 3 | 2.88 | 34.72 (9.0) | 4 | 2.21 | 120.32 (24.0) | 5 | 1.50 |
| México 1968 | | | | | | | -15.31 (29.4) | 2 | 3.48 |
| México 1977 | -23.59 (5.5) | 2 | 6.31 | -12.39 (9.7) | 3 | 3.76 | 1.60 (16.7) | 4 | 2.72 |
| Venezuela 1971 (H) | -3.34 (4.1) | 3 | 4.14 | 16.75 (10.2) | 4 | 2.69 | 39.19 (20.9) | 6 | 2.11 |
| Venezuela 1971 (R) | -17.36 (4.8) | 3 | 4.65 | -7.04 (9.0) | 4 | 3.31 | 12.65 (19.3) | 5 | 2.36 |

Nota: Los valores entre paréntesis expresan la frecuencia de casos efectivamente considerada en cada segmento.

Cuadro 6

ESTIMACIONES DEL INGRESO MEDIO DEL ULTIMO INTERVALO POR AJUSTE DE UNA FUNCION DE PARETO:
METODO DE MINIMOS CUADRADOS

| Encuesta | Ca. 5% | | | Ca. 10% | | | Ca. 20% | | |
|----------------|--|------------------------|---------------|--|------------------------|---------------|--|------------------------|---------------|
| | Discrepancia del ingreso medio estima- do (%) | Número de puntos | σ Estimado | Discrepancia del ingreso medio estima- do (%) | Número de puntos | σ Estimado | Discrepancia del ingreso medio estima- do (%) | Número de puntos | σ Estimado |
| Argentina 1975 | | | | 17.80 (14.9) | 3 | 2.65 | 14.82 (20.2) | 4 | 2.77 |
| Brasil 1972 | 7.18 (5.6) | 5 | 1.80 | | | | | | |
| Brasil 1976 | | | | 59.77 (11.4) | 3 | 1.50 | 87.82 (21.1) | 4 | 1.39 |
| Colombia 1971 | 73.92 (6.6) | 3 | 1.93 | | | | | | |
| Colombia 1972 | | | | 50.52 (9.7) | 3 | 1.68 | | | |
| Chile 1967-68 | 13.35 (6.9) | 3 | 1.65 | 3.36 (9.9) | 4 | 1.76 | -2.01 (21.0) | 6 | 1.84 |
| Chile 1968 | 31.17 (6.3) | 3 | 2.17 | | | | | | |

Cuadro 6 (conclusión)

| Encuesta | Ca. 5% | | | Ca. 10% | | | Ca. 20% | | |
|--------------------|---|------------------|---------------|---|------------------|----------|---|------------------|----------|
| | Discrepancia del ingreso medio estimado (%) | Número de puntos | Estimado | Discrepancia del ingreso medio estimado (%) | Número de puntos | Estimado | Discrepancia del ingreso medio estimado (%) | Número de puntos | Estimado |
| México 1963 | 48.46 (4.4) | 3 | 1.98 | | | | | | |
| México 1968 | | | | | | | | | |
| México 1977 | 6.10 (5.5) | 2 | 2.54 (9.7) | 15.77 | 3 | 2.25 | 26.15 (16.7) | 4 | 2.04 |
| Venezuela 1971 (H) | 13.09 (4.1) | 3 | 2.84 | 21.20 (10.2) | 4 | 2.53 | 24.48 (14.9) | 5 | 2.43 |
| Venezuela 1971 (R) | 11.65 (4.8) | 3 | 2.39 | 16.29 (9.0) | 4 | 2.26 | 25.13 (19.3) | 5 | 2.08 |

Nota: Los valores entre paréntesis expresan la frecuencia de casos efectivamente considerada en cada segmento.

