



ARTÍCULOS

## La pobreza en los hogares del Gran Córdoba: aplicación del modelo de regresión logística

Silvia Ojeda , Liliana E. Pereyra y Sonia Gontero

Revista de Economía y Estadística, Cuarta Época, Vol. 43, No. 1 (2005), pp. 99-121.

<http://revistas.unc.edu.ar/index.php/REyE/article/view/3813>



La Revista de Economía y Estadística, se edita desde el año 1939. Es una publicación semestral del Instituto de Economía y Finanzas (IEF), Facultad de Ciencias Económicas, Universidad Nacional de Córdoba, Av. Valparaíso s/n, Ciudad Universitaria. X5000HRV, Córdoba, Argentina.

Teléfono: 00 - 54 - 351 - 4437300 interno 253.

Contacto: [rev\\_eco\\_estad@eco.unc.edu.ar](mailto:rev_eco_estad@eco.unc.edu.ar)

Dirección web <http://revistas.unc.edu.ar/index.php/REyE/index>

### Cómo citar este documento:

Ojeda, S., Pereyra L. y Gontero S. (2005). La pobreza en los hogares del Gran Córdoba: aplicación del modelo de regresión logística. *Revista de Economía y Estadística*, Cuarta Época, Vol. 43, No. 1, pp. 99-121.

Disponible en: <http://revistas.unc.edu.ar/index.php/REyE/article/view/3813>

El Portal de Revistas de la Universidad Nacional de Córdoba es un espacio destinado a la difusión de las investigaciones realizadas por los miembros de la Universidad y a los contenidos académicos y culturales desarrollados en las revistas electrónicas de la Universidad Nacional de Córdoba. Considerando que la Ciencia es un recurso público, es que la Universidad ofrece a toda la comunidad, el acceso libre de su producción científica, académica y cultural.

<http://revistas.unc.edu.ar/index.php/index>



## **La pobreza en los hogares del Gran Córdoba: aplicación del modelo de regresión logística\***

**SILVIA OJEDA**

Universidad Nacional de Córdoba. Instituto de Economía y Finanzas Facultad de Ciencias Económicas y Facultad de Matemática Astronomía y Física  
ojeda@famaf.unc.edu.ar

**LILIANA E. PEREYRA**

Universidad Nacional de Córdoba. Instituto de Economía y Finanzas Facultad de Ciencias Económicas.  
lpereyra@eco.unc.edu.ar

**SONIA GONTERO**

Universidad Nacional de Córdoba. Instituto de Economía y Finanzas Facultad de Ciencias Económicas.  
sonia.gontero@gmail.com

### **Resumen**

*Con el propósito de identificar perfiles de bienestar de la población cordobesa, este trabajo indaga sobre las variables que inciden en la probabilidad de un hogar de ser o no pobre, mediante un análisis de regresión logística. La medida utilizada para la identificación de los hogares pobres se realiza a partir del método de la línea de la pobreza absoluta y los datos utilizados corresponden a los proporcionados por la Encuesta Permanente de Hogares para octubre de 2001. Su contribución al estudio del problema de la pobreza radica en que proporciona herramientas de utilidad para el diseño de políticas destinadas a paliar este flagelo.*

\* Las autoras agradecen los comentarios del Dr. Alejandro Jacobo.

## **Abstract**

*In order to identify profiles of well-being of the population of Córdoba, Argentina, this paper studies the variables that affect the probability of a home of being or not being poor, using a logistic regression analysis. The measure used for the identification of the poor homes is the absolute poverty line and data used are provided by the Permanent Homes Survey for October of 2001. Its contribution to the study of the poverty problem is that it provides tools of utility for the design of policies destined to palliate this flagellum.*

Palabras Claves: Línea de la pobreza absoluta. Perfiles de bienestar. Regresión logística. Encuesta Permanente de Hogares (EPH).

Clasificación JEL: I32; C25

## **I. INTRODUCCIÓN**

En Argentina, el tema de la pobreza, es reconocido como uno de los problemas más acuciantes que enfrenta actualmente el país. Tanto a nivel nacional como provincial y municipal, la discusión se centra, por una parte en la identificación de indicadores que caracterizan la pobreza, por otra en la formulación de políticas que disminuyan las condiciones desfavorables de los grupos más vulnerables de la población.

En los últimos años la mayor parte de los trabajos empíricos referidos a la pobreza en Argentina se han ocupado principalmente del área del Gran Buenos Aires<sup>1</sup>. Sin embargo, resulta de especial interés el desarrollo de estudios a nivel regional, entre otras razones, porque las grandes diferencias sociales, económicas y culturales que presenta nuestro país, pueden determinar que las políticas sociales diseñadas para ese centro urbano, no resulten apropiadas para el resto del país<sup>2</sup>.

<sup>1</sup> Llach y Montoya (1999); Montoya y Olivero (1998); Albornoz y Petrecolla (1996); Mitnik y Montoya (1995); Montoya y Mitnik (1993).

<sup>2</sup> Algunos estudios que trabajan sobre el tema de la pobreza en aglomerados del interior del país son: Paz, y Piselli (1999) para distintas ciudades entre 1995 y 1999, Mondino (2001) para el Gran Mendoza, Mitnik (1996) para el Gran Córdoba y Felíz y Panigo (2000) para las regiones argentinas.

El Gran Córdoba es uno de los aglomerados con mayor población de Argentina. Su ubicación geográfica de privilegio, en el centro del país, su nivel de desarrollo humano y productivo y su riqueza cultural, entre otros aspectos, posicionan a esta ciudad y sus alrededores como uno de los centros con mayor potencial para la erradicación de la pobreza. No obstante, en octubre de 2001 se estimó que un 36,2% de sus habitantes eran pobres, y un 12,0% indigentes. Luego de la devaluación de la moneda registrada durante el año 2002, se verificaron incrementos en los precios de los bienes que componen la canasta básica utilizada para establecer el umbral de pobreza. Debido a que los salarios no registraron aumentos similares se produjo un importante deterioro del poder adquisitivo de las personas, lo desembocó en un incremento significativo en la cantidad de pobres, que un año después alcanzaba al 58,5% de las personas<sup>3</sup>. Por supuesto este incremento de la pobreza en Córdoba, también reconoce entre sus causas a las elevadas tasas de desempleo. En esta ciudad, la cantidad de personas que buscan empleo sin conseguirlo se incrementó del 16% en octubre de 2001 al 19,4% en octubre de 2002<sup>4</sup>.

Con el propósito de identificar perfiles de bienestar de la población cordobesa, este trabajo indaga sobre las variables que inciden en la probabilidad de un hogar de ser o no pobre, mediante un análisis de regresión logística. Este modelo, es una herramienta que proporciona la estadística multivariante para el análisis de datos y constituye una alternativa interesante para estudiar la relación entre la probabilidad de que un evento suceda y la probabilidad de que no ocurra, pues permite identificar las variables que, dentro de un conjunto previamente propuesto, participan o inciden de manera importante en la probabilidad de ocurrencia de dicho evento; en este caso "ser pobre"<sup>5</sup>.

En la próxima sección se desarrollan algunos aspectos teóricos relacionados con la medición y el concepto de pobreza, mientras que la tercera hace lo propio con el modelo de regresión logística. En el cuarto apartado se plantea el modelo a estimar y en la quinta se detallan los resultados obtenidos. Finalmente, se presentan las principales conclusiones del trabajo.

<sup>3</sup> Ver Informes de Prensa del INDEC.

<sup>4</sup> Seguramente este indicador sería más elevado si no se consideraran como ocupados a las personas que reciben subsidios del Plan Jefes de Hogar Desocupados.

<sup>5</sup> Un trabajo que puede ser considerado un antecedente de la presente investigación es el de Mitnik (1996).

## **II. MENSURABILIDAD DE LA POBREZA EN ARGENTINA**

Una de las mayores dificultades que se presenta en toda investigación sobre la pobreza es la referida a la metodología que debe utilizarse para su medición. Este hecho surge de la dificultad para definir el concepto mismo de pobreza, que puede considerarse simplemente como la falta de dinero o de medios para vivir dignamente. Pero esta definición también incluye aspectos valorativos: ¿qué significa vivir dignamente?, ¿debemos medir la pobreza en términos absolutos o en términos relativos?, ¿son suficientes los criterios de subsistencia o de necesidades básicas insatisfechas? o, ¿es necesaria la consideración de otras privaciones como derechos, oportunidades o elecciones básicas?

El debate surge por la diversidad de enfoques que se plantean para considerar este problema y que van desde una concepción biológica, que se basa en criterios de subsistencia o líneas de pobreza, hasta los enfoques de privación relativa que enfatizan las condiciones y sentimientos de las personas sobre ciertos atributos deseables y que trascienden el campo de lo económico. Mientras que el enfoque de las necesidades biológicas y nutricionales provee un concepto elemental de necesidades de subsistencia básicas, la noción de carencias relativas es un concepto más amplio, en el cual influyen los valores de cada sociedad (Vivas Pacheco, 1996).

No obstante, y a los fines de encontrar medidas que puedan ser utilizadas para comparaciones internacionales, que permitan realizar seguimientos en el tiempo y aplicar políticas para reducir la pobreza, se hace necesaria la identificación y cuantificación de algunos indicadores. Cada uno de ellos presentará ventajas e inconvenientes que el investigador deberá tener en cuenta, atendiendo especialmente a cuáles sean sus objetivos. El desarrollo de cada una de estas metodologías excede el objetivo de esta investigación<sup>6</sup>, pero como ejemplo podemos mencionar que en los países más ricos no se considera como un problema fundamental la pobreza extrema por lo que se utilizan enfoques relativos de pobreza; en cambio, en los países menos desarrollados, existe un elevado grado de insatisfacción de las necesidades básicas, por lo que resulta relevante trabajar con mediciones de pobreza absoluta.

<sup>6</sup>Una primera aproximación a cada uno de estos indicadores puede encontrarse en Sierra (1997). También puede consultarse Feres y Macero (1999).

En Argentina los métodos más utilizados son el de las Necesidades Básicas Insatisfechas y el de la Línea de Pobreza Absoluta. El primer método considera pobre a todo individuo cuyo hogar tiene ciertas características que reflejan un consumo permanente bajo, o que condicionan seriamente su capacidad de generar ingresos suficientes en el futuro<sup>7</sup> y se elabora básicamente a través de información proveniente de los Censos.

Por su parte, el método de la Línea de Pobreza permite considerar dos grupos: los hogares (o personas) bajo la línea de indigencia (LI) y los hogares (o personas) bajo la línea de pobreza (LP). Las personas indigentes serán todas aquellas cuyos ingresos no puedan cubrir una canasta de alimentos básicos, mientras que las personas pobres serán las que no puedan adquirir una canasta de alimentos y otros servicios considerados esenciales. Este método será el criterio utilizado en el presente trabajo por lo que se decidió presentar con mayor detalle las características del mismo.

## II.1. Línea de Pobreza

El enfoque de la Línea de Pobreza (LP) es el procedimiento utilizado por el Instituto Nacional de Estadísticas y Censos (INDEC) y parte de determinar una Canasta Básica de Alimentos (CBA) que permitiría la subsistencia de un varón adulto con actividad moderada (adulto equivalente) según los hábitos de consumo de la zona que habite<sup>8</sup>. Una vez establecida esa canasta, se valoran sus componentes según los precios mínimos vigentes en el mercado. El valor de esta canasta constituye la "línea de indigencia" para esa persona. Para calcular la "línea de pobreza" se amplía esta canasta incluyendo bienes y servicios no alimentarios (vestimenta, transporte, educación, salud, etc.) y se obtiene lo que se denomina Canasta Básica Total (CBT). Debido a limitaciones en los datos, su estimación se realiza a partir de su relación con la CBA y su forma de cálculo viene dada por:

$$CBT = CBA * \frac{1}{\text{Coef.E.}} \quad (1)$$

donde Coef. E es el "Coeficiente de Engel", definido como la relación entre los gastos alimentarios y los gastos totales observados en la población de

<sup>7</sup> Gasparini, Marquioni, Sosa Escudero (2001) Estas características incluyen hacinamiento, precariedad en la vivienda, no asistencia a la escuela de los niños y baja educación del jefe de familia combinada con baja relación de miembros activos/inactivos del hogar.

<sup>8</sup> Estos se estiman a partir de la Encuesta de Gastos e Ingresos de los Hogares

referencia, ambos se actualizan según la variación de los precios correspondientes al Índice de Precios del Consumidor (IPC).

Por último, el valor de la CBT de cada hogar<sup>10</sup> se compara con el ingreso total familiar de dicho hogar. Si el ingreso es inferior al valor de la CBT se considera que el hogar y los individuos que lo componen se hallan por debajo de la línea de pobreza, de lo contrario se encontrarán en el grupo de hogares y personas no pobres

A partir del año 2001, el INDEC trabaja en la conformación de canastas regionales, pero por el momento utiliza "una metodología de transición"<sup>11</sup>, según la cual se ajusta la canasta básica de alimentos utilizada para el Gran Buenos Aires, con los coeficientes elaborados en base a los precios registrados en las distintas provincias. Con esta información se calculan luego los valores regionales de CBA y CBT utilizadas para determinar la línea de pobreza para cada región.

El área del Gran Córdoba, que nos interesa particularmente en este estudio, forma parte de la Región Pampeana, y los valores de la CBA y de la CBT se establecieron en \$57,6 y \$135,9 respectivamente para septiembre de 2001<sup>12</sup>.

Este método adolece de limitaciones de diferente índole. Para comenzar se puede partir del hecho que las estimaciones se realizan en base a encuestas de hogares, en donde uno de los aspectos más controvertidos es la calidad de las mediciones de la variable "ingreso", indispensable para la correcta interpretación del método. También es posible encontrar algunas críticas a los supuestos con los cuales se elabora la canasta de alimentos (CBA) y más aún a los métodos utilizados para la estimación de la canasta de bienes totales (CBT). Además, se debe señalar que este indicador no

<sup>9</sup> Según el tamaño y composición del hogar se estima un valor de CBA para cada hogar.

<sup>10</sup> Ver Informes de Prensa del INDEC en [www.indec.gov.ar](http://www.indec.gov.ar)

Estos fueron los valores utilizados para las estimaciones de pobreza correspondientes a octubre de 2001.

<sup>11</sup> Sen (1992)

<sup>12</sup> Luego de esta fecha se verificó en Argentina un período de inflación que redujo los salarios reales e incrementó sustancialmente la cantidad de hogares que no cubren una canasta básica y que por lo tanto son considerados "nuevos pobres". En este caso en particular, nos interesa identificar las características de las personas pobres, sin tener en cuenta aquellas que se encuentran en esta situación sólo circunstancialmente. Por ello no se considera desactualizada la información de octubre de 2001.

considera aspectos distributivos y que no satisface ciertos requisitos que debería cumplir como buen indicador de pobreza<sup>13</sup>.

Sin embargo, el método de la línea de pobreza es el que se utiliza actualmente tanto en Argentina como en otros países de Latinoamérica y, sin descartar los aportes que pudieran realizar otras metodologías, es el que se sigue en este trabajo, tratando de obtener el mayor provecho posible de la información provista por las encuestas de hogares. En este sentido, uno de los recursos que brinda herramientas para hacer, consiste en el desarrollo y aplicación de metodologías estadísticas provenientes del análisis multivariado. En este caso en particular, se propone un modelo de Regresión Logística, con el cual se pretende identificar qué variables y en qué sentido influyen en la probabilidad de ser pobre. La fuente de información utilizada es la Encuesta Permanente de Hogares del Gran Córdoba, para octubre de 2001<sup>14</sup>.

### III. EL MODELO DE REGRESIÓN LOGÍSTICA

De acuerdo con el modelo de regresión logística, si  $Y$  es una variable aleatoria con respuesta binaria<sup>15</sup>, un conjunto de factores, recogidos en un vector  $X = (X_1, X_2, \dots, X_p, \dots, X_k)$ , es capaz de explicar la probabilidad de que  $Y$  valga 1. En términos matemáticos, la probabilidad condicional del evento ( $Y=1$ ), por ejemplo "hogar pobre", considerando la información proporcionada por el vector aleatorio  $X=(X_1, X_2, \dots, X_p, \dots, X_k)$ , está dada por:

$$P [Y = 1 / X_1, X_2, \dots, X_k] = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}} \quad (2)$$

donde  $\beta = (\beta_0, \beta_1, \dots, \beta_k)$  es el vector de parámetros a estimar.

<sup>13</sup> Esta variable vale 1 si se observa una característica específica, como por ejemplo "hogar pobre" y 0 en caso contrario.

<sup>14</sup> Este cociente es frecuentemente llamado "odd ratio" (cociente de riesgo), cuando el valor de la variable  $X_i$  aumenta en una unidad. Por su parte  $e^{\beta_0}$  es un factor de escala.

<sup>15</sup> En el modelo de regresión lineal  $\beta_1$  mide directamente el impacto que un cambio en  $X_1$  produce en la variable dependiente; en cambio en el modelo de regresión logística,  $\beta_1$  cuantifica el efecto que un cambio en  $X_1$  tiene en la probabilidad de ocurrencia del evento ( $Y=1 / X_1, X_2, \dots, X_p, \dots, X_k$ ) con respecto a la probabilidad de su no ocurrencia.

A partir de (2) y operando algebraicamente, se obtiene para el modelo la siguiente expresión:

$$\frac{P[Y = 1/X_1 X_2 \dots X_k]}{1 - P[Y = 1/X_1 X_2 \dots X_k]} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k} \tag{3}$$

y designando  $P_y(X)$  a  $P[Y = 1/X_1 X_2 \dots X_k]$  y  $Q_y(X) = 1 - P_y(X)$ , se tiene:

$$\frac{P_y(X)}{Q_y(X)} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k} \tag{4}$$

Para interpretar el impacto de cada uno de los parámetros sobre el cociente  $\frac{P_y(X)}{Q_y(X)}$  es conveniente expresar (4) como:

$$\frac{P_y(X)}{Q_y(X)} = e^{\beta_0} (e^{\beta_1})^{X_1} \dots (e^{\beta_k})^{X_k} \tag{5}$$

Suponiendo que las variables  $X_j$ , permanecen constantes para todo  $j \neq i$ , y que sólo varía  $X_i$ , (5) se expresa de la siguiente manera:

$$\frac{P_y(X)}{Q_y(X)} = c e^{\beta_i X_i} \tag{6}$$

donde  $c = \prod_{j=1, j \neq i}^k e^{\beta_j X_j}$ , es un número positivo y  $e^{\beta_i}$  es el factor de cambio en el

cociente  $\frac{P_y(X)}{Q_y(X)}$  <sup>(16)</sup>. El efecto de  $\beta_i$  sobre  $\frac{P_y(X)}{Q_y(X)}$  resulta evidente a partir de

(6); si  $\beta_i > 0$  y  $X_i$  aumenta en una unidad,  $\frac{P_y(X)}{Q_y(X)}$  también aumenta, indicando en qué medida  $P_y(X)$  supera a  $Q_y(X)$  cuando  $X_i$  aumenta una unidad. Por el contrario, si  $\beta_i < 0$  y  $X_i$  aumenta en una unidad,  $\frac{P_y(X)}{Q_y(X)}$  disminuye y su valor establece en qué medida  $P_y(X)$  está por debajo de  $Q_y(X)$  cuando  $X_i$  aumenta una unidad. Por ejemplo, pensando en el modelo de regresión logística

<sup>16</sup> Vivas Pacheco, H. (1996), Mitnik (1996), Dagum, C. (2002)

aplicado al estudio de la pobreza en Córdoba, si  $X_1$  indica número de personas en el hogar y  $\beta_1 > 0$ , al aumentar el número de integrantes del hogar, digamos de 4 a 5, aumentará, indicando en qué medida la probabilidad de ser pobre supera a la probabilidad de no serlo, cuando el número de integrantes del hogar varía de 4 a 5. Un análisis similar es válido en el caso en que  $\beta_1 < 0$  <sup>(17)</sup>.

La estimación de los parámetros del modelo se realiza por el método de Máxima Verosimilitud, asumiendo distribución de Bernoulli con parámetro  $P_Y(X) = P(Y=1 / X_1 = x_1, X_2 = x_2, \dots, X_k = x_k)$ , para la variable  $Y / (X_1 = x_1, X_2 = x_2, \dots, X_k = x_k)$ . Para una muestra de tamaño  $n$ , expresada como  $\{y^j \times x^j\}_{j=1}^n$ , siendo  $x^j = (x_1^j, x_2^j, \dots, x_k^j)$ ,  $j=1, \dots, n$  valores del vector  $(X_1, X_2, \dots, X_k)$ , e  $y^j$  igual a 0 ó 1, la función de verosimilitud de  $Y/X_1, X_2, \dots, X_k$  viene dada por:

$$L(\beta / (y^j, x^j)) = \prod_{j=1}^n \left[ P_Y(x^j) \right]^{y^j} \left[ 1 - P_Y(x^j) \right]^{1-y^j} \quad (7)$$

con  $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ , vector de parámetros a estimar. El método de Máxima verosimilitud, permite estimar  $\beta$ , maximizando la función (7), lo cual implica resolver mediante métodos iterativos el siguiente sistema de ecuaciones:

$$\left[ \frac{\partial \log L}{\partial \beta} \right] = \sum_{j=1}^n x^j (y^j - P_Y(x^j)) = 0 \quad (8)$$

Uno de los problemas que se presenta habitualmente en el modelo de regresión logística, es la colinealidad o casi colinealidad entre las variables regresoras. Ello determina modelos desprovistos de sentido en términos de explicar el comportamiento de  $Y$  a partir de las variables independientes  $X_1, \dots, X_k$ ; pues si bien el algoritmo de cálculo encuentra una solución para la estimación de los coeficientes  $\beta$ , ésta suele no ser precisa, resultando nocivamente dependiente de los datos, de tal manera que una pequeña variación en ellos determina un cambio considerable en los valores de los

<sup>17</sup> En el modelo de regresión lineal  $\beta_1$  mide directamente el impacto que un cambio en  $X_1$  produce en la variable dependiente; en cambio en el modelo de regresión logística,  $b_1$  cuantifica el efecto que un cambio en  $X_1$  tiene en la probabilidad de ocurrencia del evento  $(Y=1 / X_1, X_2, \dots, X_{j-1}, X_{j+1}, \dots, X_k)$  con respecto a la probabilidad de su no ocurrencia.

coeficientes de la ecuación. De allí que resulte importante descartar aquellos modelos en los cuales se detecte colinealidad entre dos o más variables independientes.

Finalmente, para controlar la bondad del ajuste global del modelo se analiza el estadístico de bondad de ajuste de Hosmer-Lemeshow dado por:

$$HL = \sum_{i=1}^n \left[ \frac{(y_i - \hat{\pi}_i)^2}{\hat{\pi}_i(1 - \hat{\pi}_i)} \right] \tag{9}$$

donde

$$\hat{\pi}_i = \left[ \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_k x_{ik}}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_k x_{ik}}} \right]$$

Bajo la hipótesis de que el modelo es adecuado, el estadístico sigue una distribución .

#### **IV. PLANTEO DEL PROBLEMA**

Para el estudio de la pobreza en el Gran Córdoba se planteó un modelo de regresión logística que expresa el cociente entre la probabilidad de ser pobre de un hogar y la probabilidad de no serlo, como combinación lineal de ciertas variables independientes seleccionadas a priori; tal probabilidad estará dada por la fórmula (3) donde Y es la variable POBREZA, la cual se define como:

$$Y(i) = \begin{cases} 1 & \text{si el hogar } i \text{ es pobre} \\ 0 & \text{en caso contrario} \end{cases}$$

El criterio de selección de las variables independientes del modelo (ver Cuadro 1) presenta algún grado de subjetividad y está basado en la hipótesis de influencia de ciertos factores socioeconómicos y demográficos de los hogares sobre su condición de pobreza; sin embargo, este procedimiento de selección ha sido utilizado en numerosos estudios referidos a este

<sup>17</sup> Vivas Pacheco, H. (1996), Mitnik (1996), Dagum, C. (2002).

tema<sup>17</sup>.

**Cuadro 1**  
**Definición de las variables utilizadas**

Notación	Variable	Definición
$X_1$	Tamaño del hogar	Cantidad de personas que habitan en el hogar
$X_2$	Tasa de participación esperada	La razón entre las personas en edad de trabajar y la cantidad total de miembros del hogar
$X_3$	Edad del jefe del hogar	Años de edad del jefe de hogar
$X_4$	Años de educación del jefe de hogar	Años de educación del jefe de hogar
$X_5$	Estado civil del jefe de hogar	Estado Civil del jefe del hogar: soltero, unido, casado, separado o viudo
$X_6$	Años de experiencia del jefe de hogar	Años de experiencia del jefe de hogar en la ocupación principal
$X_7$	Años de estudio de la persona que más estudió en el hogar	Años de educación de la persona que más estudió en el hogar
$X_8$	Promedio de horas trabajadas en el hogar	Media de horas semanales trabajadas en el hogar
$X_9$	Años de educación promedio de las personas del hogar en edad de trabajar	Media de años de educación de las personas en edad de trabajar del hogar
$X_{10}$	Ingresos por hora del jefe de hogar	Ingreso horario de la ocupación principal del jefe de hogar (variable construida por el INDEC)
Y	Condición de pobreza	La condición de pobreza indica si el hogar es pobre o no pobre

En primera instancia se estudió la posible relación de colinealidad entre las variables independientes, detectándose entre  $X_3$  y  $X_6$  y entre  $X_7$  y  $X_9$ ; este hecho, que puede apreciarse en los Gráficos 1 y 2, (scatterplot de las variables mencionadas) determinó la exclusión del análisis de las variables  $X_3$  y  $X_7$ . El criterio de exclusión se basó en la comparación de los de

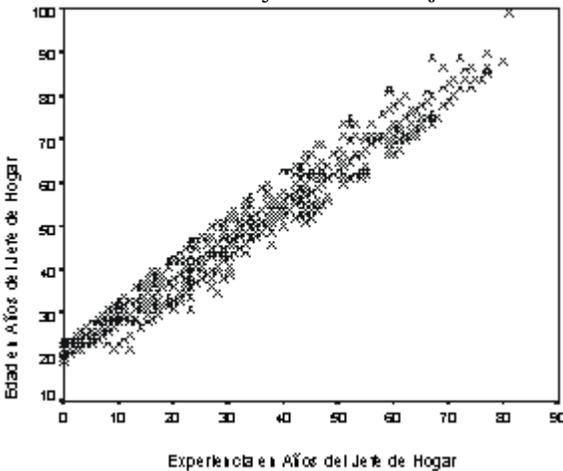
resultados obtenidos a partir de la aplicación del modelo de regresión logística, considerando inicialmente como única variable regresora una de las variables de cada par y luego reemplazando ésta por la otra. Para el caso del par  $(X_3, X_6)$ , los modelos comparados fueron:

$$P[Y = 1 / X_3] = \frac{e^{\sigma_6 + \delta_1 X_3}}{1 + e^{\sigma_6 + \delta_1 X_3}} \tag{10}$$

$$P[Y = 1 / X_6] = \frac{e^{\delta_3 + \delta_1 X_6}}{1 + e^{\delta_3 + \delta_1 X_6}} \tag{11}$$

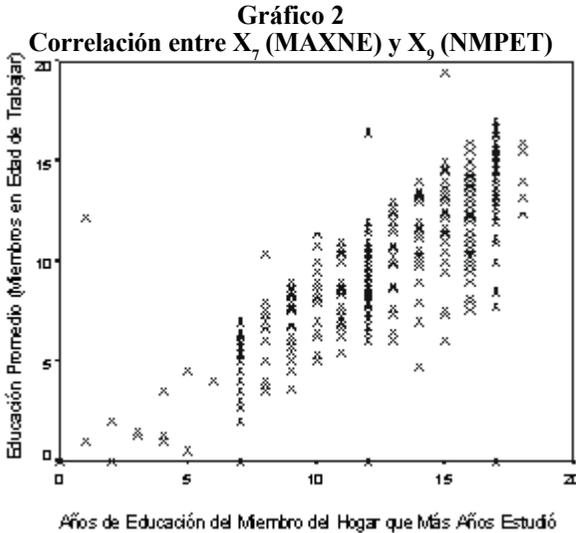
El impacto de  $e^{\alpha_1}$  sobre el cociente de riesgo de (10) resultó mayor que el impacto de  $e^{\delta_1}$  sobre el cociente de riesgo de (11), por lo que se eliminó  $X_6$ . Un análisis similar se realizó para seleccionar  $X_3$  en lugar de  $X_7$ .

**Gráfico 1**  
Correlación entre  $X_3$  (EDADJ) y  $X_6$  (EXPERJ)



La unidad de observación considerada fue el hogar. Del total de hogares de la muestra, se separó una submuestra aleatoria del 80% con el propósito de evaluar luego, a partir del 20% restante, la capacidad predictiva del modelo. Además, en el caso en que se registraran valores ausentes de alguna de las variables explicativas, se utilizó el criterio de eliminar ese hogar

la muestra por lo que el total de hogares efectivamente procesados en el estudio de regresión logística se redujo a 538 casos.



La estimación se realizó mediante el software estadístico SPSS; se aplicaron técnicas forward stepwise y backward stepwise obteniéndose diferentes modelos que fueron comparados bajo distintos criterios, a saber: estadísticos de bondad de ajuste (-2loglikelihood, Cox & Snell, R Square, Nagelkake R Square, Hosmer and Lemeshow), capacidad de predicción y el criterio de parsimonia.

## V. RESULTADOS OBTENIDOS

Los resultados de la regresión se presentan en el Cuadro 2, en el cual se observa la estimación de máxima verosimilitud de los parámetros del modelo para las variables que resultaron estadísticamente significativas. Con un nivel de significación del 5%, los coeficientes beta de dichas variables son estadísticamente diferentes de cero (test de Wald). Ello indica que las mismas explican el cociente de riesgo del modelo, o dicho en otras palabras, tienen capacidad para describir la probabilidad de que un hogar sea pobre con respecto a la probabilidad de no serlo. En los casos en los cuales la estimación de beta sea un número negativo, implicará que un aumento en las respectivas variables producirá una disminución en la probabilidad de ser

pobre, y viceversa, en el caso que la estimación del coeficiente beta sea positivo. Los signos de las estimaciones de los parámetros corresponden a los resultados esperados.

En la última columna del Cuadro 2 se observa el valor de  $e^\beta$  para cada variable estadísticamente significativa del modelo. Este valor cuantifica los cambios en el cociente de riesgo si la variable correspondiente aumenta en una unidad. En este caso, un incremento unitario en las variables, tasa de participación esperada, edad del jefe, años de educación promedio de las personas en edad de trabajar y en el ingreso por hora del jefe de hogar, disminuye el factor de riesgo en un 83%, 2,2%, 22,5% y 30,5% respectivamente. De ello se deduce que aumentos en dichas variables amortiguan la probabilidad de ser pobre del hogar.

En el caso de la variable tamaño del hogar, el valor de  $e^\beta$  es 1,473; en consecuencia un aumento unitario en esta variable determina un incremento del 47,3 % en el cociente de riesgo; así el aumento en el número de integrantes del hogar determina una mayor probabilidad de pobreza del mismo.

**Cuadro 2**  
**Estimación de los coeficientes en el modelo**

Variables en la ecuación	$\beta$	SE	p-valor (estadístico de Wald)	$e^\beta$
$X_1$ : Tamaño del hogar	0,388	0,073	0,000	1,473
$X_2$ : Tasa de participación esperada	-1,771	0,549	0,001	0,17
$X_3$ : Edad del jefe del hogar	-0,022	0,009	0,009	0,978
$X_9$ : Años de educación promedio de las personas del hogar en edad de trabajar	-0,254	0,037	0,000	0,775
$X_{10}$ : Ingresos por hora del jefe de hogar	-0,364	0,077	0,000	0,695
CONSTANTE	2,945	2,945	0,001	19,010

-2L log likelihood= 441,625

Hosmer and Lemeshow test: 10,009 (p-valor 0,264)

La estimación de máxima verosimilitud de los parámetros del modelo proporciona como estimación de  $P(Y=1/ X_1, X_2, \dots, X_k)$  a:

$$(Y=1/ X_1, X_2, \dots, X_k) = \left[ \frac{e^{2,945+0,388X_1-1,771 X_2 -0,022 X_3 -0,254X_9 -0,364X_{10}}}{1+e^{2,945+0,388X_1-1,771 X_2 -0,022 X_3 -0,254X_9 -0,364X_{10}}} \right] \tag{12}$$

La expresión (12), permite estimar para cada hogar la probabilidad de ser pobre, conociendo los valores de las variables  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_9$  y  $X_{10}$  para dicho hogar.

Por otro lado resulta de interés analizar la capacidad predictiva del modelo, lo cual se sintetiza en el Cuadro 3.

**Cuadro 3**  
**Capacidad predictiva del modelo**

Condición de pobreza observada	Condición de pobreza predicha					
	Casos seleccionados			Casos no seleccionados		
	No pobre	Pobre	% correcto	No pobre	Pobre	% correcto
No pobre	295	97	75,3%	78	34	69,6%
Pobre	34	112	76,7%	7	33	82,5%
Total	329	209	75,7%	85	67	73%

En el caso de los hogares seleccionados en la muestra original, se observa que el porcentaje bien clasificado es del 75,7% y no se distinguen diferencias importantes entre hogares pobres (76,7 %) y no pobres (75,3 %).

En el caso de hogares seleccionados en la muestra de validación, los bien clasificados fueron en promedio el 73%; sin embargo, en este caso se observa que el modelo clasifica mejor a los hogares pobres que a los no pobres pues ubica correctamente al 82,5% de los primeros y al 69,6% de los no pobres.

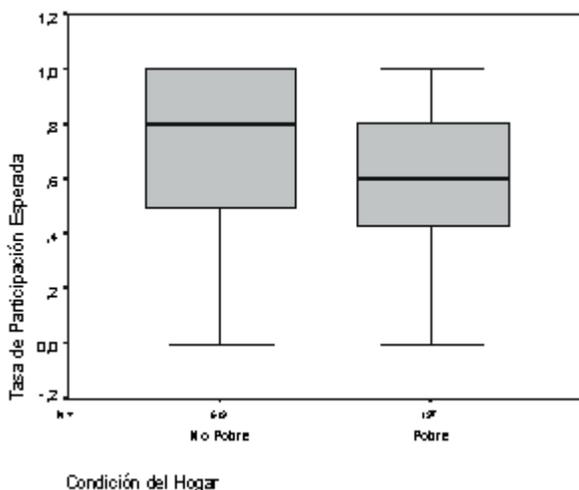
Por último, y con el propósito de visualizar el comportamiento de las variables estadísticamente significativas del modelo, se presenta a continuación un análisis más detallado de las características de las mismas.

**V.1. Pobreza y población en edad de trabajar**

La tasa de participación esperada, se ha definido como el cociente entre las personas en edad de trabajar y el total de miembros de un hogar,

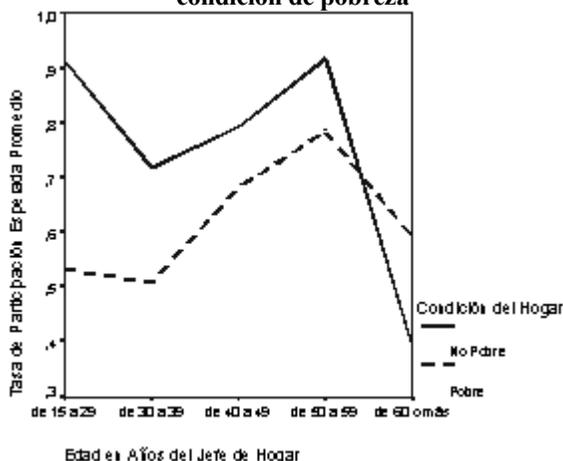
e indica la posibilidad de un hogar de generar ingresos. Los Gráficos 3 y 4 indagan sobre distintos aspectos de esta variable. El primero presenta la distribución de esta variable de acuerdo con la condición de pobreza; en tanto el Gráfico 4 estudia su patrón de comportamiento como función de la edad del jefe de hogar, discriminando entre pobres y no pobres. En el Gráfico 3 se observa que los hogares pobres tienen una tasa de participación esperada que oscila entre 0,4 y 0,8; mientras que para los hogares no pobres varía entre 0,5 y 1,0. Esto implica que, en general, los hogares pobres tienen mayor cantidad de miembros dependientes (personas que no están en edad de trabajar) que los hogares no pobres<sup>18</sup>. Por su parte, el gráfico 4 muestra que los hogares pobres cuyos jefes de hogar son jóvenes tienen menores tasas de participación esperada que los hogares no pobres con jefes dentro del mismo grupo etario. Además se puede destacar que para hogares con jefes cuya edad es de alrededor de 50 años o más, la tasa de participación esperada disminuye drásticamente en el grupo de los no pobres, posiblemente por el retiro de la actividad productiva al acceder a una jubilación. Sin embargo, en el caso de los hogares pobres esta disminución es más moderada; este hecho podría deberse a que por falta de recursos para establecerse como hogar independiente, cohabitan en una misma vivienda los hijos del hogar con sus parejas, incorporando al núcleo integrantes en edad de trabajar

**Gráfico 3**  
**Tasa de participación esperada según condición de pobreza**



<sup>18</sup> Este gráfico se denomina Boxplot, en ellos, la línea interna a cada caja representa la mediana de la distribución de la tasa de participación esperada, y las barbas los puntos máximos y mínimos asumidos en la muestra por esta variable.

**Gráfico 4**  
**Tasa de participación esperada según edad del jefe de hogar y**  
**condición de pobreza**



Fuente: Elaboración propia en base datos EPH-INDEC

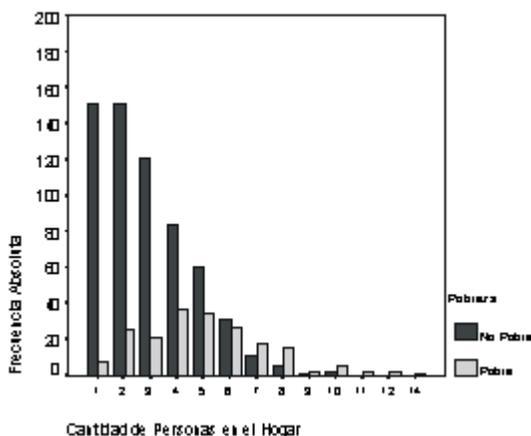
## V.2. Pobreza y tamaño de los hogares

Según el estudio previo, una de las variables con mayor incidencia en la condición de pobreza es el tamaño del hogar. El Gráfico 5 presenta la distribución de frecuencias de esta variable según la condición de pobreza; se observa que en el caso de los hogares pobres la distribución es aproximadamente simétrica, mientras que para los hogares no pobres las frecuencias más altas corresponden a hogares con menor número de integrantes. Así, mientras la cantidad promedio de personas del hogar es de 3 en el caso de los hogares no pobres, esta medida se incrementa a 5 en el caso de los hogares pobres<sup>19</sup>. El Gráfico 6 permite complementar este análisis al relacionar el tamaño del hogar con la edad de los jefes según su condición de pobreza. Se pueden señalar los siguientes aspectos: 1) independientemente de la edad del jefe de hogar el tamaño del hogar es superior en hogares pobres 2) dentro de los hogares pobres tienen mayor tamaño aquellos cuyo jefe de hogar varía entre 30 y 39 años; mientras que dentro de los hogares no pobres, los de mayor tamaño, corresponden a hogares cuyo jefe tiene

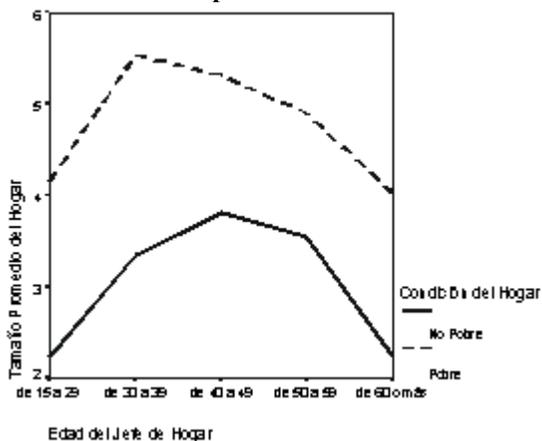
<sup>19</sup> Con una alta concentración en hogares de 1 y 2 miembros.

entre 40 y 49 años<sup>20</sup>. Este hecho podría ser el resultado de una mejor planificación familiar por parte de los hogares no pobres, en el sentido de postergar en el tiempo la llegada de los hijos al hogar.

**Gráfico 5**  
**Tamaño del hogar por condición de pobreza**



**Gráfico 6**  
**Tamaño promedio del hogar por grupo de edad del jefe y condición de pobreza**



Fuente: *Elaboración propia en base datos EPH-INDEC*

<sup>20</sup>Cabe aclarar que este máximo no alcanza el tamaño mínimo de los hogares pobres, ya sea con jefes muy jóvenes (de 15 a 29 años) o mayores (de 60 años o más).

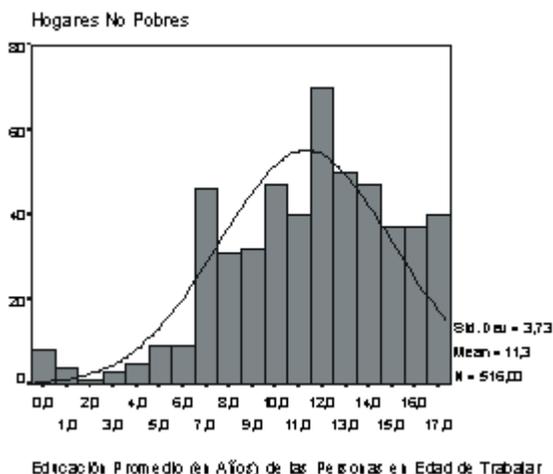
### V.3. Pobreza y educación

Por último, los Gráficos 7 y 8 relacionan la condición de pobreza de los hogares con el nivel educativo de sus integrantes. Esta característica se analiza a través de la variable "Años de Educación Promedio de las personas en edad de trabajar".

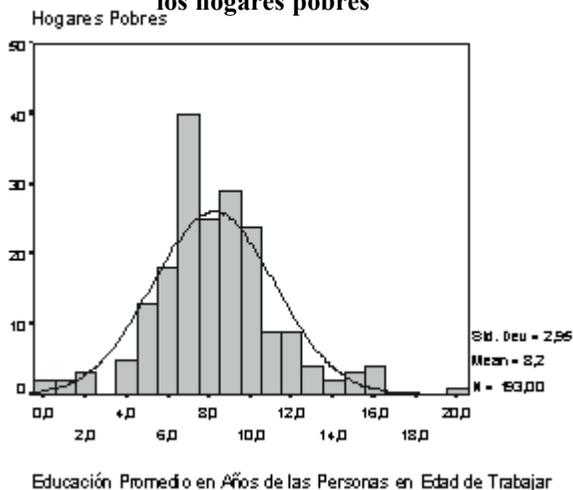
Se puede advertir que las personas en edad de trabajar en hogares no pobres se educan, en promedio, 11,3 años, mientras que en los hogares pobres esta media es de 8,2 años. Para este último grupo, el histograma presenta un comportamiento más simétrico que en el caso de los hogares no pobres, lo que determina que su media sea más representativa de la distribución de la educación. En cambio, en el caso de los hogares no pobres, la cantidad de años de educación se encuentra mejor representada por la mediana, la cual está desplazada hacia la derecha del valor 11.

En síntesis, la diferencia en capital humano parece ser una característica que diferencia a los hogares pobres de los que no lo son. Mientras que en estos últimos, sus integrantes se encuentran culminando los estudios de nivel medio, en los hogares pobres se interrumpe la educación formal aproximadamente en el primer año del nivel medio.

**Gráfico 7**  
**Años de educación promedio de las personas en edad de trabajar para los hogares no pobres**



**Gráfico 8**  
**Años de educación promedio de las personas en edad de trabajar para los hogares pobres**



*Fuente: Elaboración propia en base datos EPH-INDEC*

## VI. CONCLUSIONES

En los últimos años se ha incrementado el debate sobre los niveles de pobreza alcanzados en Argentina. Esto se debe no sólo a que es posible intuir mayores carencias en la sociedad, sino también a que INDEC ha comenzado con la elaboración de canastas valuadas a precios regionales que permiten realizar estimaciones de pobreza en varias ciudades del país, y no limitar su estudio solamente al área del Gran Buenos Aires.

Como un aporte al estudio de esta problemática, se decidió analizar las características de la población de la ciudad de Córdoba que constituye uno de los centros más poblados del país. El trabajo se basó en la Encuesta de Hogares realizada en octubre de 2001, coincidente con el momento histórico de finalización del programa de convertibilidad implantado en Argentina desde 1991. En este sentido, marca una referencia importante sobre las variables de peso en la condición de pobreza para el cierre de este período. A partir de esta fecha, un proceso de alta inflación redujo los salarios reales

y por lo tanto se incrementó fuertemente la cantidad de personas pobres que son captados por el método de la línea de pobreza.

El modelo propuesto para estudiar el comportamiento de la probabilidad de que un hogar sea clasificado como pobre fue el de regresión logística, ya que el mismo constituye una herramienta útil para determinar a partir de un conjunto de variables propuestas a priori, cuáles de ellas tienen mayor influencia sobre una variable dicotómica dada.

Las estimaciones realizadas muestran que las variables de mayor influencia en la probabilidad de ser pobre de los hogares cordobeses son: el tamaño del hogar, la tasa de participación esperada, la edad del jefe de hogar, los años de educación promedio de las personas en edad de trabajar del hogar y los ingresos por hora del jefe de hogar. Esto implica que un hogar conformado por varios menores, con jefes de hogar muy jóvenes y baja acumulación de capital humano constituye un grupo con alta probabilidad de pertenecer al conjunto de los hogares pobres. Variables tales como: "empleo informal", "participación de la mujer en el mercado laboral", relacionadas con la ocupación de los miembros del hogar, no fueron incluidas en este estudio por encontrarse altamente correlacionadas con otras variables con mayor significación estadística en la regresión logística. Con respecto al sexo del jefe de hogar, no resultó una variable significativa de la condición de pobreza.

Los resultados encontrados permitieron dar cuenta de la naturaleza multidimensional de la pobreza identificando aquellos factores que tienen mayor peso en la determinación de la probabilidad de que un hogar sea pobre. El análisis realizado permite señalar algunas pautas hacia las cuales deberían orientarse las políticas sociales destinadas a reducir la pobreza en esta ciudad. Estas deberían encaminarse a realizar esfuerzos en el área de la salud reproductiva, de la educación y del trabajo; pero dada la naturaleza multifacética que evidencia la pobreza a través del estudio efectuado, es de fundamental importancia que estas políticas se lleven a cabo en forma conjunta y a través de esfuerzos coordinados.

En términos generales, el mensaje es claro con respecto a la asistencia en materia de salud reproductiva; es necesario lograr una mayor educación, conciencia y responsabilidad en la planificación del tamaño del hogar, principalmente entre los jóvenes de grupos vulnerables .

Por otro lado, y confirmando la amplia evidencia en este sentido, se corrobora la importancia de la inversión en educación. En este caso en particular, se advierte la necesidad de políticas educativas destinadas a

lograr una mayor retención escolar, aunque seguramente la mejora en la calidad de la enseñanza es un hecho que tampoco puede dejar de considerarse.

Finalmente, resulta relevante destacar que el hecho de que el método de la línea de pobreza presente algunas limitaciones hace necesaria la revisión de las metodologías alternativas de medición de pobreza; además como se considera que el ingreso es la variable relevante para la determinación de la pobreza, sería importante avanzar en la aplicación de métodos de corrección de ingresos.

## VII. REFERENCIAS

- Albornoz, F. y Petrecolla D. (1996). "Medidas alternativas de la pobreza por ingresos para el Gran Buenos Aires 1980-1995" *Económica*, La Plata, Vol XLII N<sup>o</sup> 1-2.
- Bartels, C. (1982). *Economía del bienestar, distribución del ingreso y desempleo*. México. FCE.
- Beccaria, L.; Feres, J. C. y Sainz, P. (1999). "Medición de la Pobreza: situación actual de los conceptos y métodos" 4to. Taller Regional del MECOVI. <http://www.eclac.cl/deype/noticias/proyectos/7/5497/mecovi/ACTIVIDADREGIONAL/TALLERES/4TOTALLER/4TOTALLER.html>
- Banco Interamericano de Desarrollo, BID (1999). *América Latina frente a la desigualdad*.
- Dagum, C. (2002). "Analysis and measurement of poverty and social exclusion using fuzzy set theory Applications and policy implications". *Sociedad Argentina de Estadística*.
- Feres, J. C. (1999). "Notas sobre la medición de la pobreza según el método del ingreso" 4to. Taller Regional del MECOVI. <http://www.eclac.cl/deype/noticias/proyectos/7/5497/mecovi/ACTIVIDADREGIONAL/TALLERES/4TOTALLER/4TOTALLER.html>
- Feres, J. C. y Mancero, X. (1999). "Enfoques para la medición de la pobreza. Breve revisión de la Literatura" 4to. Taller Regional del MECOVI.
- Feliz, M. y Panigo, D. (2000). "Desigualdad, Pobreza y Bienestar en las regiones argentinas" Universidad Torcuato Di Tella en <http://www.utdt.edu/~redpobreza/>
- Ferran Aranaz, M. (1996). *SPSS para Windows. Programación y Análisis Estadístico*. Mc. Graw-Hill.

- Gasparini, L; Marchioni, M; Sosa Escudero, W (2001) *Distribución del ingreso en la Argentina : perspectivas y efectos sobre el bienestar*. Fundación ARCOR.
- Hair, J, Anderson, R. Tatham, R y Black, W. (1999). *Análisis Multivariante*. Prentice Hall.
- INDEC (2000) "Las Necesidades Básicas Insatisfechas: sus deficiencias técnicas y su impacto en la definición de políticas sociales" 5to. *TALLER REGIONAL DEL MECOVI*.
- LLach, J. J. y Montoya, S. (1999). *En pos de la equidad. La pobreza y la distribución del ingreso en el Area Metropolitana de Gran Buenos Aires: diagnóstico y alternativas de políticas*. IERAL, Buenos Aires.
- Mitnik, O. (1996) "Educación y pobreza. Un estudio sobre el Gran Córdoba" *Estudios*, Año XIX, N° 79, octubre- diciembre.
- Mitnik, O y Montoya, S. (1995) "Pobreza y Distribución del Ingreso: dinámica y características Gran Buenos Aires, 1974 - 1994". *Revista Estudios* Julio / Septiembre.
- Mondino, G. (2001) "Pobreza y Distribución del Ingreso en el Gran Mendoza" INDEC, *Programa MECOVI-ARGENTINA*, BID-BM-CEPAL.
- Montoya S. y Olivero M. (1998) "Claves para reducir la pobreza agregada disminuyendo las disparidades regionales. Argentina 1990/1997" *Revista Estudios*, Fundación Mediterránea Año XX, N°85, Abril/Junio 1998.
- Montoya, S. Y Mitnik, O (1993) "La pobreza urbana en Argentina. El caso del Gran Buenos Aires" *Estudios*, Abril / Junio 1993.
- Paz, J. y Pisselli C. (2000) "Desigualdad de ingresos y pobreza en Argentina" *Anales de la Asociación Argentina de Economía Política*, Córdoba.
- Sen, A (1992) "Sobre conceptos y medidas de pobreza" *Revista Comercio Exterior*, Vol.42, Nro. 4, México, Abril.
- Sierra, L. (1997a) "Determinación de Umbrales de Pobreza" *Estudios Empresariales* N° 94 1997/2.
- (1997b) "Determinación de Umbrales de Pobreza II" *Estudios Empresariales* N° 95 1997/3.
- Vivas Pacheco, H. (1996). "Perfiles de Bienestar de los Grupos Vulnerables en Cali-Yumbo". *Centro de Investigaciones y Documentación Socioeconómica CIDSE*, Universidad del Valle, Colombia. Documento de Trabajo No. 26, Septiembre.
- Ray, D (1998) *Economic Development*. Princeton University. Press.