

**Universidad de Puerto Rico  
Recinto de Río Piedras  
Departamento de Economía  
Unidad de Investigaciones Económicas**

**Serie de Ensayos y Monografías: Núm. 8**

***Estudio sobre factores que determinan  
la pobreza en Puerto Rico***

**Pedro A. Cantero  
José Rodríguez  
Mary H. Alverio  
Revisado: Angel L. Ruiz  
Septiembre 1979**

## INDICE

I.	<u>Introducción.....</u>	Pág. 1
	a) Problema de la definición de pobreza; dependiendo del campo de estudio desde donde se analice.	
	b) Definición de pobreza en su contexto social.	
	c) Por qué seleccionamos esta área de estudio	
II.	<u>Formulación del Problema.....</u>	Pág. 5
	a) Formulación del problema en específico.	
	b) Factores que afectan nuestra área de estudio.	
	c) Implicaciones de nuestro problema.	
	d) Importancia del estudio.	
III.	<u>Desarrollo de las Hipótesis del Trabajo.....</u>	Pág. 13
	a) Definición de la variable dependiente	
	b) Motivos para la inclusión de las variables independientes en el estudio.	
	c) Hipótesis para cada variable independiente	
	d) Hipótesis general del estudio.	
IV.	<u>Planificación del Estudio.....</u>	Pág. 19
	a) Métodos utilizados para recolectar la data de las variables.	
	b) Fuentes de información de donde fue recopilada la data.	
	c) Problemas encontrados al recolectar la data.	
	d) Soluciones seguidas para resolver los problemas.	
	e) Técnicas estadísticas a utilizarse para analizar la data.	

V. Presentación y análisis de los resultados.....Pág. 25

- a) Definición del análisis de regresión múltiple y para qué se utiliza.
- b) Interpretación de los resultados generales de nuestro análisis
- c) Aceptación o rechazo de nuestras hipótesis de trabajo.
- d) Aceptación o rechazo del modelo general
- e) Importancia relativa de cada una de nuestras variables independientes sobre nuestra variable dependiente.

VI. Conclusiones y Recomendaciones.....Pág. 40

Bibliografía.....Pág. 43

Apendice.....Pág. 45

NOTA

Este trabajo surge como requisito parcial del curso Métodos de Investigación Económica (Econ. 493). El mismo ha sido un esfuerzo conjunto de los estudiantes Pedro A. Cantero Frau (líder del grupo), José Rodríguez García y Mary H. Alverio Rodríguez, bajo la supervisión del Profesor Alfredo González Martínez, Catedrático Asociado del Departamento de Economía del Recinto Universitario de Mayaguez.



## I. Introducción

La pobreza es un problema que pasa muchas veces por desapercibido para una parte de la población puertorriqueña. Día a día nuestros medios de comunicación hacen grandes alardes de los méritos y los triunfos de ese grupo del país que mediante su gestión productiva en la economía, logran asegurarse unos medios que les permiten vivir "bien o bastante bien". Mientras la atención pública se dirige, principalmente, hacia estos grupos, dentro de la economía, existe otro grupo ignorado y enajenado: los pobres. Es hacia este grupo y su problemática que va dirigido este estudio.

Definir qué es pobreza no es un asunto fácil, debido a la relatividad y a la subjetividad que estaría envuelta en cualquier definición que logremos formular. La gran mayoría de las definiciones que se lograron obtener de los libros técnicos y de otras publicaciones, tienden a mirar el problema desde un punto de vista en específico, dependiendo del campo de estudio desde donde se analice.

Por ejemplo: Campbell R. McConnell da una definición de pobreza relacionada a su campo de estudio, la economía:

"...a family lives in poverty when its basic needs exceed its available means of satisfying them"<sup>1/</sup>

---

<sup>1/</sup> McConnell, Campbell R., Economics Principles, Problems and Policies, sixth edition, New York, McGraw Hill Book Co., Inc., 1975, p. 736.

Rafael Ramírez, sociólogo, por su parte da su versión de qué es pobreza:

"...pobreza es el resultado de una condición de desigualdad que establece diferencias en la accesibilidad a unos bienes materiales que se consideran necesarios y deseables. La desigualdad es producto de una estructura social y económica que limita las oportunidades de un sector de la población.. "2/

Oscar Lewis, conocido antropólogo, habla más bien del medio ambiente en que viven los pobres:

"The culture of poverty is both an adaptation and a reaction of the poor to their marginal position in classstratified, highly individualistic, capitalistic society"3/

Para entrar en el objeto de este estudio, en específico conviene dividir la diversidad de definiciones que se puedan dar sobre la pobreza, en tres conceptos principales: pobreza moral, pauperismo y pobreza social.<sup>4/</sup>

La pobreza moral se refiere más bien a la posición en la escala de valores de una sociedad, de la pobreza define si la pobreza es moralmente aceptable y qué posición social se confiere o de cuál se le priva al pobre.

---

<sup>2/</sup> Ramírez, Rafael, Pobreza en Puerto Rico, Teoría y Praxis, San Juan, Junta de Planificación, 1973, p. 2.

<sup>3/</sup> Lewis, Oscar, A Study of Slum Culture, New York: Random House, p. 5

<sup>4/</sup> Enciclopedia Internacional de las Ciencias Sociales, Madrid, Aguilar, 1976, Vol. 8 p. 289-293.

El pauperismo por su parte, describe a una categoría de personas incapaces de mantenerse a sí mismas en absoluto, o de mantenerse al nivel convencionalmente considerado mínimo, sin asistencia exterior.

El concepto de pobreza social, además de inferir una situación de desigualdad social, es decir, una relación de inferioridad, dependencia o explotación, implica una situación de desigualdad económica (desigualdad en la posesión de propiedades, en los ingresos, etc.).

El presente estudio bregará con el concepto de pobreza social, en el sentido de desigualdad económica. En otras palabras, qué características económicas hacen que una persona sea pobre.

La selección de este tema se debió, principalmente, a la preocupación por desigualdad existente en Puerto Rico en la distribución del ingreso.

Este trabajo estará constituido de cinco partes principales. La primera parte bregará con la formulación del problema. En esta parte se tratará de delimitar al área de estudio, la distribución de los ingresos, hasta llegar a la formulación del problema que será objeto de este trabajo. Así se presentarán, también, los factores que posiblemente influirán sobre el objeto de estudio, procediendo luego a explicar en qué consiste cada uno de estos factores. Se hablará, también, de las implicaciones que tiene este problema y de la importancia que tiene el estudio.

En la segunda parte principal, se desarrollarán las hipótesis del trabajo. En esta parte se definirá la variable dependiente y la relación que teóricamente guarda ésta, con cada una de las variables independientes, o factores tomados en cuenta en este trabajo. También, se formulará la hipótesis general que se considera en este estudio.

La tercera parte será la planificación de este estudio. Aquí se presentarán los métodos utilizados para recolectar la data de cada una de las variables, exponiendo también, las fuentes de información utilizadas para recolectar la misma. Se analizará, a su vez, los problemas o limitaciones que se encontraron en este proceso de recopilación, pasando luego a buscar soluciones a los mismos. Por último, se presentará en esta parte, la técnica estadística a utilizarse para analizar la data. Para fines de este estudio se hará un análisis de corte seccional de la pobreza en los municipios de Puerto Rico para el año 1969, utilizando la técnica de regresión múltiple para obtener los resultados.

La cuarta parte principal de este trabajo constará de la presentación y análisis de los resultados. En primer lugar se explicará qué es un análisis de regresión múltiple, y para qué se utiliza. Luego se pasará a explicar los resultados generales de este estudio, procediendo después a verificar la validez empírica de las hipótesis de trabajo formulados anteriormente.

Por último, la quinta parte del trabajo estará constituida por las conclusiones y recomendaciones que se derivaron de los resultados obtenidos en este estudio.

## II. Formulación del Problema

Habiendo definido ya, como objeto de estudio, la pobreza social, en el sentido de desigualdad económica, se pasará ahora a ver cuál va a ser el problema que se intentó investigar en el transcurso del estudio.

El problema que ocupa a los autores, en este trabajo, surge como consecuencia de una preocupación por una posible desigualdad en la distribución de los ingresos existentes en Puerto Rico. La posible existencia de una desigualdad en la distribución de los ingresos, puede traer como resultado el que una considerable porción de las familias puertorriqueñas, caigan por debajo de un nivel de ingreso tan insuficiente que no les permita el obtener unos bienes y servicios básicos necesarios para la satisfacción de sus necesidades. Es esta deficiente estructuración en la distribución de los ingresos en la economía, la que puede enajenar a una parte bastante considerable de la población puertorriqueña de sus funciones sociales y económicas.

Antes de proseguir, sería bueno dejar establecido, que para objeto de este estudio se va a considerar a una persona o familia como pobre, cuando su nivel de ingresos esté por debajo de un nivel mínimo establecido, sin tomar en consideración los aspectos psicológicos que puedan determinar o influir en el que una persona sea pobre.

Cuando decimos que para el 1970 el 59.6 por ciento de las familias puertorriqueñas eran pobres (según el Censo Poblacional de 1970)<sup>5/</sup>, a lo que en realidad nos referimos es que para esta fecha, el 59.6 por ciento de las familias puertorriqueñas, estaban por debajo de un nivel mínimo de ingresos establecidos por el mismo Censo Poblacional. Esta medida puede resultar arbitraria debido a que, para la formulación de la misma, el Censo tuvo que hacer una serie de clasificaciones, como se verá más adelante, que no necesariamente demuestran la magnitud real de la pobreza.

Hecha esta aclaración, se pasará ahora a hacer un inventario de los factores que a nuestro juicio influyen sobre nuestro problema, o dicho en otras palabras, los factores que determinan el que una familia posea un nivel de ingresos tan bajo, como para ser clasificada pobre.

Entre estos factores se encuentran los siguientes: la proporción del total de familias que viven en áreas rurales, la proporción del total de la fuerza laboral civil que está desempleada y el nivel de escolaridad que poseen los individuos.

El primer factor mencionado lo componen aquellas familias que viven en áreas rurales, expresadas como un porcentaje del total de familias. En general, se puede definir un área rural como aquella que está apartada de una área densamente poblada. Pero, para fines del estudio, se seguirá la definición más específica de lo que es una área rural, que da el Censo Poblacional de 1970<sup>6/</sup>. Según el Censo, se define como población rural, a

---

<sup>5/</sup> Censo Poblacional de 1970, Publicación para Puerto Rico del Departamento de Comercio de Estados Unidos, 1970, p. 53-185.

<sup>6/</sup> El Censo Poblacional de 1970 lo que no da es la definición de lo que constituye un área urbana, pasando luego a decir que todo lo que no caiga bajo la definición de área urbana será área rural, Censo Poblacional, p. APP-3.

a todas aquellas personas que viven en áreas no urbanizadas y en lugares de menos de 2,500 habitantes. A la vez, la población rural está dividida en residentes rurales que viven en fincas y en residentes rurales que no viven en fincas. La población rural que vive en fincas, consiste de todas aquellas personas que habitan en lugares de tres cuerdas o más, en las cuales había algún ganado o quince (15) o más aves; también, se consideró finca, aquellos lugares de menos de tres cuerdas, si las ventas brutas de productos agrícolas durante el año anterior, o ventas brutas anticipadas para un próximo año eran de \$100 o más. Los que no llenaban este requisito eran considerados como la población rural en no fincas.

Para propósito del estudio no se tomará en consideración la división que hace el Censo, en cuanto a residentes rurales y residentes rurales que no habitaban en fincas. Se utilizará un solo índice de ruralidad, en el cual estarán comprendidas estas dos subdivisiones.

En opinión de los autores, el lugar de residencia de las familias, es un factor muy importante, que afectará el nivel de ingresos que esas familias pueden percibir. Por lo general, en las áreas urbanas es que están concentrados los empleos con mejores remuneraciones, mientras que en las áreas rurales se encuentran los empleos menos remunerados.

Un segundo factor que, a nuestro juicio, influye en el que una persona reciba un nivel de ingresos tal que no le permita la satisfacción de todas sus necesidades básicas, es el desempleo. Al estar las personas desempleadas, no tendrán una fuente de ingresos fijos y tendrán que estar dependiendo, casi exclusivamente, de ayudas gubernamentales para poder satisfacer algunas de sus necesidades básicas.

Este factor, igual que el anterior, no es uno cuya definición sea igual para todos los casos o situaciones. Para fines de este trabajo se utilizará la definición de desempleo que da el Departamento del Trabajo de Puerto Rico<sup>7</sup>.

Según el Departamento del Trabajo, una persona se clasifica como desempleada si hizo gestiones para conseguir trabajo durante la semana de referencia y además, no trabajó durante esa semana. El que una persona esté buscando trabajo puede consistir en actividades tales como: 1) registrarse en agencias de empleos, públicos o privadas, 2) visitar patronos, 3) poner o contestar anuncios, 4) escribir cartas de solicitud y 5) esperar por el resultado de haberse registrado en agencias de empleos dentro de los últimos 60 días. Algunos grupos pequeños se cuentan como desempleados a pesar de no estar buscando trabajo activamente, esto es así porque se presume que en circunstancias ordinarias estarían buscando trabajo. Estos grupos incluyen, entre otros, a personas que estaban enfermas temporeramente, a personas que estaban esperando que los llamaran otra vez a un empleo del cual habían sido suspendidos por un período de 30 días o más a partir de la fecha de la enumeración.

Cuando se va a determinar si una persona está empleada, desempleada, o fuera del grupo trabajador, a base de las actividades de esa persona durante una semana determinada, surge el problema de que muchos individuos han realizado más de una actividad durante la semana, y por lo tanto, hay que decidir a base de qué actividad se ha de clasificar la persona. Para esto, se ha establecido un sistema de prioridades:

---

<sup>7</sup>/Vilches, Rubén A., La Medición del Desempleo en Puerto Rico, Revisión de un artículo publicado en la Revista del Trabajo, Departamento del Trabajo de Puerto Rico, Negociado de Estadísticas, San Juan, 1975, pp. 14-20.

1) las actividades dentro del grupo trabajador tienen prioridad sobre las actividades fuera del grupo trabajador. 2) trabajar tiene prioridad sobre buscar trabajo. 3) buscar trabajo tiene prioridad sobre estar fuera de un empleo por vacaciones, disputas obreras, etc.

Un tercer factor, que posiblemente afecte el objetivo de este estudio, lo es el nivel de escolaridad de los individuos. Creemos, que el total de años de educación completados por un individuo puede afectar el nivel de ingresos que ese individuo pueda generar en un futuro.

Existen varias maneras mediante las cuales se puede definir el nivel de escolaridad de los individuos en la sociedad. Entre éstos se ha tomado la definición proporcionada por el Censo Poblacional de 1970, por conveniencias metodológicas.<sup>8/</sup> El censo toma, como medida de escolaridad la mediana de años de escuela completados. Esta se define, como el valor que divide el grupo poblacional en dos partes iguales, una con mayor años de educación que la mediana. Esta mediana se calcula convirtiendo a una serie continua de números, los años de escolaridad completados por un individuo (ejemplo: el primer año de universidad es tomado como el 13er año y así sucesivamente).

Estos datos recopilados del Censo, tienden a subestimar los años de escuela completados por las personas más jóvenes, ya que al tomarse en consideración el año más alto completado, se elimina aquella parte del año que está en proceso. Por ejemplo, una persona que estuviera cursando el segundo semestre de su cuarto año de universidad, para efectos del Censo, tendría un nivel de educación equivalente a su tercer año de universidad. Debido a esto, esta data no sería muy confiable en cuanto a una medida absoluta de instrucción alcanzada, aunque si puede ser utilizada, acertadamente, al medir relaciones entre grupos o de un mismo grupo en diferentes fechas

---

<sup>8/</sup>

Censo Poblacional de 1970, op. cit., p. App 15-17

Debido a la naturaleza del curso y a las limitaciones en cuanto al tiempo disponible, nos vemos precisados a limitar nuestra investigación a la relación que guardan estos tres factores mencionados, con nuestro objeto de estudio. De ninguna manera esto significa que no existan otros factores que puedan influir sobre el problema que intentamos investigar. Como en la generalidad de las investigaciones económicas, nuestro objeto de estudio se ve afectado, de una u otra forma, por un número considerable de factores. Por la complejidad que presenta este problema y por las limitaciones dichas anteriormente, nos es imposible el intentar hacer una lista exhaustiva de todos los factores, aunque sí, debieramos mencionar algunos otros.

Uno de estos factores sería la composición demográfica de la familia. El que una familia esté encabezada por una mujer o por un adulto mayor de 65 años de edad, puede que influya en el nivel de ingresos que esta pueda obtener. Para el año 1969, el 71.1 por ciento de las familias encabezadas por personas de 65 años o más y el 73.7 por ciento de las familias encabezadas por mujeres, caían por debajo del nivel de pobreza<sup>9/</sup>. Estos datos, aparentemente, apuntan a una relación entre las características demográficas de las familias y la incidencia de la pobreza.

Otro factor que también puede influenciar nuestro objeto de estudio lo es, la influencia que ejerce la condición de pobreza de una familia sobre sus descendientes. Es muy probable que el hecho de que una familia sea pobre, afecte directamente a sus descendientes, de manera que estos adquieran las mismas características que tenían sus antepasados.

---

<sup>9/</sup> Ibid, 53-179

A primera vista este factor parece ser uno meramente psicológico, pero si nos deteneros y lo analizamos a fondo veremos que la condición de pobreza de una familia no le permite, generalmente, educar a sus hijos, ni social, ni académicamente, de manera que estos puedan desarro'lar sus capacidades innatas y logren salir de la condición en que se encontraban sus padres.<sup>10/</sup>

El problema de la pobreza siempre ha tenido, tiene y tendrá unas implicaciones indeseables para la sociedad en general. Ese 59.6 por ciento de las familias puertorriqueñas que vivían por debajo del nivel de pobreza para 1969, habitaban, en muchos casos, en un ambiente cuyas características principales, entre otras, lo eran la privación material, el crimen, el vicio, la codicia y el deseo de salir de ese medio ambiente a como diera lugar. Es a este medio ambiente, al que Oscar Lewis describe como la cultura de la pobreza, ya que (generalizando) los pobres tienden a compartir un mismo sistema de valores que los lleva a vivir la vida de una forma muy parecida.

Estas características existentes en el medio ambiente en que conviven los pobres se debe en gran medida a la falta de unos ingresos adecuados que les permitan a estos, el obtener unos bienes y servicios, que no tan sólo les satisfagan sus necesidades básicas, sino también les permita satisfacer algunos de sus deseos, no básicos, para la supervivencia.

El problema surge cuando estas personas se hacen conscientes de la disparidad existente entre su nivel de vida y el resto de la sociedad.

---

<sup>10/</sup> Tanto Oscar Lewis como Herman Miller reconocen la importancia de este factor.

Viendo esto desde este punto de vista, es de suponer que el pobre busque, de alguna manera, llegar a obtener eso que ve que otros tienen en abundancia. En ciertos casos, si no lo pueden obtener con los ingresos que reciben, se idean métodos ilegales para poder obtenerlos.

De aquí es, entonces, que posiblemente surge la mayoría de los problemas sociales que tanto han aquejado al pueblo puertorriqueño desde hace varios años. Es el propósito de este estudio identificar los factores, que a nuestro juicio, determinan el que una familia caiga en la posición desventajosa que se ha explicado anteriormente. Una vez determinada la importancia relativa de estos factores sobre el objetivo de este estudio, se pasará a sugerir las áreas en las que se podrían trabajar para tratar de erradicar, al menos parcialmente, el problema de la pobreza.

### III. Desarrollo de la hipótesis del trabajo

Habiendo definido ya el objetivo de este estudio y los factores que, a nuestro juicio, lo afectan más significativamente, conviene ahora, para fines analíticos, convertir este objetivo y estos factores en variables. El índice de pobreza será la variable dependiente y los factores que actúan sobre este, serán las variables independientes.

Antes de proceder a la formulación de las relaciones entre la variable dependiente y las independientes, conviene aclarar la definición específica que se utiliza para la variable dependiente.

Lo que presenta esta variable es el porcentaje de total de familias con ingresos por debajo del nivel de pobreza. Para establecer el nivel de pobreza se provee una serie de límites de ingresos de

pobreza ajustados por factores tales como tamaño de la familia, sexo del jefe de familia, número de hijos menores de 18 años de edad y residencia en fincas y en no fincas. También está ajustado para tomar en consideración un plan de alimentación nutricionalmente adecuado, diseñado por el Departamento de Agricultura, y para tomar en consideración las diferencias en el costo de la vida entre las familias que viven en fincas y las que no viven en fincas.

Para el 1969, los umbrales de pobreza fluctuaban entre \$1,487 para una mujer no relacionada de 65 años o más que vivía en una finca y \$6,116 para una familia con siete o más personas que no vivía en finca y que tenía un jefe varón. El umbral de pobreza promedio de una familia de cuatro que no vivía en finca y cuyo jefe era varón, era de \$3,745<sup>11/</sup>.

Ya teniendo en claro las cuatro variables que se habrán de utilizar en el estudio, se procederá ahora a la formulación de las hipótesis. Las hipótesis que se expondrán son tan sólo suposiciones tentativas, basadas en nuestras lecturas sobre el tema, principalmente sobre el tipo de relación existente entre la variable dependiente y las variables independientes. Esta relación puede ser: directa, si un cambio en la variable independiente ocasiona un cambio en igual dirección de la variable dependiente; inversa, si el cambio ocasionado por la variable independiente sobre la variable dependiente, es en dirección contraria; o neutral si un cambio en la variable independiente no ocasiona ningún tipo de cambio en la variable dependiente.

La primera hipótesis que se establecerá para tratar de aclarar los factores que afectan el nivel de ingresos recibidos por una persona o familia, es la de la relación existente entre el nivel de educación que posea una persona y el nivel de ingresos que esta educación le permita obtener a la misma.

En general en la sociedad moderna, tiende a haber un alto grado de correlación positiva entre el nivel de educación de un individuo y el nivel de ingresos que éste recibe. En el caso de Puerto Rico esto puede ser así debido a la peculiaridad de la transformación económica que ha ocurrido en Puerto Rico en las últimas décadas. Las nuevas industrias que se establecen ahora en Puerto Rico son generalmente, de uso intensivo de tecnología avanzada, y por lo tanto, los requerimientos de empleos de estas nuevas industrias son por trabajadores diestros con cierto grado de escolaridad. Por lo tanto es de esperarse que mientras menor sea el nivel de escolaridad de una región, menor va a ser el nivel de ingresos de los individuos en esa región y por ende mayor va a ser el número de familias que caerán por debajo del nivel de pobreza.

Es sumamente interesante el notar que para 1970, sólo entre el 21 y 27 por ciento de la población universitaria en Puerto Rico provenían de familias con ingresos menores de \$3,000, no empuje a que éstos formaban el 49 por ciento de la población total<sup>12/</sup>. Estos datos tienden a indicar la existencia de algún tipo de relación directa entre el nivel de escolaridad y el nivel de ingreso de una persona o familia.

Todo esto sugiere la existencia de una relación inversa entre el nivel de escolaridad y la variable dependiente, lo que traducido a términos matemáticos se define como:

1)  $P_b = f(X_1)$ : y la relación es de tal forma que:

2)  $P_b = a - b_1(X_1)$

donde:

$P_b$  = es el índice de pobreza.

$X_1$  = es el índice de escolaridad.

$a$  = es el intercepto de la función con el eje de  $P_b$ .

$b_1$  = es el coeficiente de regresión entre la variable  $P_b$  y  $X_1$ .

---

<sup>12/</sup> La Alternativa Socialista, Tesis Política del Partido Socialista Puertorriqueño, San Juan, 1974, p. 29.

La primera ecuación lo que hace es indicarnos la existencia de una relación entre la variable dependiente,  $P_b$ , y la variable independiente  $X_1$ . La segunda ecuación hace más específica esta relación, estableciéndola de forma lineal y demostrando la suposición de la existencia de una relación inversa entre el nivel de escolaridad y la variable dependiente.

La segunda hipótesis que se formulará es sobre la relación existente entre el índice de desempleo y el índice de pobreza. Es de esperarse que si el nivel de desempleo existente en un lugar determinado es alto, la proporción de personas, en ese lugar, que caerá por debajo del nivel de pobreza también va a ser alto. Al una persona estar desempleada, por lo general, no tendrá otras fuentes de ingresos que no sean aquellas proporcionadas por el gobierno en calidad de asistencia pública, las cuales, en la mayoría de los casos, no son suficientes como para lograr que una persona desempleada sobrepase el límite establecido de la pobreza.

Esto apunta la existencia de una relación directa entre el índice de desempleo y el índice de pobreza que utilizaremos en este estudio.

Transformando esta relación a términos matemáticos tenemos que:

1)  $P_b = f(X_2)$ : y la relación es de tal forma que:

2)  $P_b = a + b_2(X_2)$

donde:

$P_b$  = es el índice de pobreza

$X_2$  = es el índice de desempleo

$a$  = es el intercepto de la función con el eje de  $P_b$ .

$b_2$  = es el coeficiente de regresión entre la variable  $P_b$  y  $X_2$ .

La primera ecuación : indica la existencia de una relación entre el índice de desempleo,  $X_2$  y el índice de pobreza  $P_b$ . La segunda ecuación hace más específica esta relación, estableciéndola de forma lineal y demostrando la supuesta existencia de una relación directa entre el nivel de desempleo y nuestra variable dependiente.

La tercera hipótesis que se expondrá es sobre la relación existente entre la tercera variable independiente, el índice de ruralidad y la variable dependiente, el índice de pobreza.

En el caso de Puerto Rico es de esperarse que mientras mayor sea la proporción de habitantes que vivan en áreas rurales, en un lugar determinado, mayor va a ser la cantidad de personas que van a estar por debajo del nivel de pobreza en ese lugar. Esta hipótesis se basa en la suposición, bastante acertada, de que cuando una persona vive en una área rural, no tiene a su alcance las mismas oportunidades de trabajo, ni los mismos tipos de empleos bien remunerados que están al alcance de las personas que viven en áreas urbanas.

Esto, principalmente, se debe a lo desbalanceado que ha sido el desarrollo económico sufrido por Puerto Rico en las últimas décadas. Las nuevas industrias que han venido y que vienen a la Isla por gestiones de la Administración de Fomento Económico, han preferido en la mayoría de los casos, establecerse en áreas urbanas por la conveniencia que esto les representa, causando esto una situación de inferioridad económica en las áreas rurales <sup>13/</sup>.

Otra posible explicación para la situación desventajosa en que se encuentran las áreas rurales es que tradicionalmente los empleos existentes en las áreas rurales han sido provistos por la agricultura, cuya capacidad de generar ingresos y empleos ha ido decayendo en las últimas décadas en Puerto Rico <sup>14/</sup>.

Todo esto apunta hacia la existencia de una relación directa entre el índice de ruralidad y el índice de pobreza utilizados en este estudio.

---

<sup>13/</sup> González, Antonio J., Economía Política de Puerto Rico, San Juan, Editorial Cordillera, Inc., 1971, p. 148.

<sup>14/</sup> Censo Poblacional de 1970, op.cit., p. 53-203. Para 1969, alrededor de 20% de las familias rurales trabajan en la agricultura.

Transformando esta relación a términos matemáticos tenemos que:

1)  $P_b = f(X_3)$ : y la relación es de tal forma que:

2)  $P_b = a + b_3 (X_3)$

donde:

$P_b$  = es el índice de pobreza

$X_3$  = es el índice de ruralidad.

$a$  = es el intercepto de la función con el eje de  $P_b$ .

$b_3$  = es el coeficiente de regresión entre la variable  $P_b$  y  $X_3$ .

La primera ecuación indica la existencia de una relación entre la variable dependiente,  $P_b$ , y la variable independiente,  $X_3$ . La segunda ecuación hace más específica la relación, estableciéndola de forma lineal y demostrando la suposición de la existencia de una relación directa entre las dos variables.

Las tres hipótesis de trabajo formuladas anteriormente tienden a señalar la existencia de una relación lineal entre nuestras tres variables independientes y la variable dependiente. Transformando estas relaciones de forma que se establezca un sólo modelo general del problema, se obtendrá la siguiente ecuación:

$$P_b = a - b_1 (X_1) + b_2 (X_2) + b_3 (X_3)$$

donde:

$P_b$  = índice de pobreza

$a$  = el punto donde la función corta el eje de  $P_b$ .

$b_1, b_2, b_3$  = son los parámetros de las variables  $X_1, X_2, X_3$  respectivamente.

$X_1$  = es el índice de escolaridad

$X_2$  = es el índice de desempleo

$X_3$  = es el índice de ruralidad.

Esta ecuación señala a la existencia de una relación de tipo lineal entre las variables independientes y la variable dependiente. Para fines de estudio se supone que sólo estas tres variables independientes van a determinar el comportamiento de la variable dependiente.

#### IV. Planificación del Estudio

Habiéndose planteado el problema que se intenta investigar y teniendo ya las hipótesis del trabajo formuladas, los próximos pasos serán el describir el método utilizado para las fuentes de información de donde se recopiló la misma, el procedimiento utilizado para su análisis y los problemas encontrados en el transcurso de estos pasos.

Al principio del estudio, cuando se comenzó a revisar la literatura<sup>15/</sup> sobre el tema que aquí se plantea, se encontró con una gran variedad de factores que posiblemente pudieran influir sobre el problema bajo estudio. Luego de consultar con el profesor que dicta el curso, se decidió a escoger los tres factores ya mencionados.

Después de tener en claro cuál iba a ser el objetivo del estudio y los tres factores, que se opinó lo iban a influenciar más, el próximo paso fue el de buscar una variable que tomara el lugar de la variable dependiente. En otras palabras, se tenía que buscar una variable que definiera la pobreza. Esto resultó ser el primer problema ya que no se sabía de qué forma se podía obtener una variable que indicara la situación de pobreza. A sugerencia del profesor, los autores de este estudio se refirieron al Censo Poblacional de 1970, donde se encontró una variable que muy bien podía ser utilizada como la variable dependiente. Esta indicaba la proporción de personas que caían por debajo de un nivel de pobreza establecido por el mismo Censo Poblacional de 1970. De esta misma publicación fue de donde se obtuvieron las series de estadísticas necesarias para esta variable.

---

<sup>15/</sup> Ver Bibliografía.

Luego de esto, se procedió a obtener la data necesaria para la formulación de las variables independientes. Primero, se pasó a buscar series numéricas que pudieran convertir el primer factor, la escolaridad, en una variable independiente. En el mismo Censo Poblacional de 1970 se encontraron las series numéricas del nivel de escolaridad de la población que hacían falta.

Se prosiguió luego a la obtención de una serie numérica para el segundo factor, el desempleo. En este caso los datos del Censo no son confiables debido a la definición de desempleo que utiliza el mismo. Por ejemplo, una persona que estuviera de vacaciones al momento de llevarse a cabo el Censo, era considerada por éstos como un desempleado. Por ésta y otras irregularidades se descartó la definición del desempleo ofrecida por el Censo. El próximo paso fue buscar la definición de desempleo que utiliza el Departamento del Trabajo de Puerto Rico, encontrando que esta definición era mucho más completa y razonable<sup>16/</sup>.

Una vez solucionado este problema, se procedió a la búsqueda de series numéricas para el tercer factor, la ruralidad. Esta data fue fácil de obtener, ya que el mismo Censo Poblacional de 1970 la proveía. El único inconveniente fue que el Censo dividía la población rural entre aquellos que vivían en fincas y aquellos que no vivían en fincas, pero este problema fue rápidamente resuelto sumando ambas cantidades para obtener un sólo índice de ruralidad.

Luego de haber obtenido los datos para cada una de las variables, se procedió a tabular la información donde se presentan estas cuatro variables para cada uno de los municipios de Puerto Rico, para el año 1969. La razón para haber tomado este año en particular, es que para hacer un análisis de corte seccional, por municipios, sobre la pobreza, la data más reciente que existía era la data proporcionada por el Censo Poblacional de 1970.

<sup>16/</sup> Las series numéricas para esta variable fueron conseguidas fácilmente, gracias a la ayuda del profesor.

## Variables Independientes

Municipios	Variable Dependiente	Indice de	Indice de	Indice de
	Indice de Pobreza (Ph)	Escolaridad (X <sub>1</sub> )	Desempleo (X <sub>2</sub> )	Ruralidad (X <sub>3</sub> )
Adjuntas	.807	4.4	.233	.715
Aguada	.784	3.8	.240	.621
Aguadilla	.605	7.5	.240	.442
Aguas Buenas	.794	4.5	.185	.916
Albionito	.711	5.2	.187	.622
Arasco	.712	4.9	.240	.773
Arecibo	.670	6.1	.185	.517
Arovo	.792	5.9	.201	.583
Barceloneta	.743	4.6	.233	.763
Barranquitas	.828	4.5	.192	.776
Bayamón	.398	9.5	.047	.056
Cabo Rojo	.476	5.2	.210	.725
Cañas	.513	9.1	.185	.329
Camuy	.751	4.7	.185	.805
Carolina	.362	11.1	.047	.100
Cataño	.695	6.2	.047	.000
Cavey	.685	5.1	.187	.439
Ceiba	.489	8.3	.168	.721
Ciales	.740	4.4	.233	.740
Cidra	.742	4.4	.185	.736
Coamo	.379	5.5	.192	.544
Comerio	.830	3.6	.233	.665
Corozal	.796	4.8	.192	.798
Culebra	.561	7.0	.239	.100
Dorado	.659	6.2	.098	.753
Fajardo	.500	7.5	.168	.207
Guánica	.715	6.2	.196	.397
Guayama	.649	6.6	.201	.364
Guayanilla	.669	5.8	.196	.714
Guaynabo	.408	10.4	.047	.175
Guayama	.556	5.6	.185	.591
Hatillo	.759	4.4	.185	.874

Municipios	Variables Independientes			
	Variable Dependiente Indice de Pobreza (Pb)	Indice de Escolaridad (X <sub>1</sub> )	Indice de Desempleo (X <sub>2</sub> )	Indice de Ruralidad (X <sub>3</sub> )
Hormigueros	.523	7.3	.210	.397
Humacao	.661	5.9	.239	.655
Isabela	.762	4.8	.240	.667
Jayuya	.838	4.0	.233	.718
Juana Díaz	.723	5.7	.192	.636
Juncos	.649	5.3	.185	.633
Lajas	.668	5.9	.210	.795
Lares	.835	4.5	.233	.620
Las Marias	.836	4.0	.240	1.000
Las Piedras	.684	4.8	.239	.744
Lotza	.663	6.1	.086	.713
Luquillo	.632	6.0	.166	1.000
Manatí	.696	5.1	.233	.559
Martícano	.864	4.6	.240	1.000
Maurabo	.820	4.5	.279	1.000
Mayaguez	.579	7.0	.128	.190
Roca	.836	3.8	.240	1.000
Rorovis	.813	3.9	.233	.846
Maguabo	.706	4.4	.241	.745
Naranjito	.757	5.3	.192	.836
Orocovis	.878	4.0	.233	.818
Patillas	.766	4.0	.201	.857
Peñuelas	.768	4.9	.196	.802
Ponce	.573	7.7	.152	.171
Quebradillas	.736	5.6	.229	.816
Rincón	.789	4.2	.240	1.000
Río Grande	.661	6.1	.120	.632
Sabana Grande	.712	4.9	.210	.660
Salinas	.733	5.6	.250	.540
San Germán	.683	5.8	.210	.585
San Juan	.432	9.6	.047	.023
San Lorenzo	.735	4.2	.185	.723

## Variables Independientes

Municipios	Variable Dependiente	Indice de	Indice de	Indice de
	Indice de Pobreza (Pb)	Escolaridad (X <sub>1</sub> )	Desempleo (X <sub>2</sub> )	Ruralidad (X <sub>3</sub> )
San Sebastián	.794	4.8	.240	.763
Santa Isabel	.698	5.5	.192	.720
Toa Alta	.730	5.0	.253	.832
Toa Baja	.514	3.7	.111	.217
Trujillo Alto	.558	8.1	.110	.397
Ituado	.850	4.4	.233	.674
Vega Alta	.676	5.7	.290	.619
Vega Baja	.677	5.9	.237	.516
Vieques	.694	6.0	.239	1.000
Villalba	.839	4.3	.192	.779
Yabucoa	.814	4.9	.239	.732
Yauco	.746	5.6	.196	.632

## V. Presentación y Análisis de los Resultados

Una vez procesado el programa en la computadora del R.U.M., obtuvimos los resultados principales.

La técnica de regresión múltiple, utilizada en este estudio, lo que hace es estimar, cómo una variable dependiente está relacionada con dos o más variables independientes. Esta técnica de análisis, no solamente permite el detectar la existencia de una relación entre la variable dependiente y las variables independientes, sino que también permite ver la magnitud y la dirección de la dependencia entre estos dos tipos de variables.

Una vez obtenidos los resultados mediante este análisis, se procedió a substituir estos resultados en la ecuación general formulada anteriormente como se muestra a continuación<sup>17/</sup>.

$$P_b = 1.0451 - 0.0649 (X_1) + 0.0150 (X_2) + 0.0240 (X_3)$$

$$(0.0551) \quad (0.0053) \quad (0.1193) \quad (0.0298)$$

$$\text{Coeficiente de Determinación (CD)} = 0.8810$$

$$\text{Coeficiente de Correlación Múltiple (CCM)} = 0.9386$$

$$\text{Error estándar de regresión (EER)} = 0.0404$$

Como se puede observar en la ecuación anterior, los parámetros utilizados anteriormente ( $a_1, b_1, b_2, b_3$ ) han sido substituidos, por los valores numéricos dados en el análisis de regresión múltiple. Estos parámetros permanecerán constantes en el modelo. Los valores que aparecen debajo de los parámetros de las variables en la ecuación, representan los errores estándar de los coeficientes de regresión.

---

<sup>17/</sup> Ver apéndice para otros resultados del análisis de este modelo.

El primer parámetro, que es el intercepto de la función con el eje de Pb, lo que hace es recoger la variación de la variable dependiente, que no está explicada por las tres variables independientes formuladas.

Un alto valor relativo de este parámetro significa que existen otros factores que afectan la variable dependiente, que no se han considerado en el modelo.

Los otros tres parámetros lo que hacen es describir el cambio marginal que ocurría en la variable dependiente a consecuencia de un cambio en una de las variables independientes, manteniendo las otras dos constantes.

El signo que precede a estos parámetros muestra el tipo de relación existente entre la variable independiente en cuestión y la variable dependiente. Si el signo es negativo la relación será inversa y si es positivo la relación será directa. En este caso, la relación existente entre la variable dependiente y el índice de escolaridad resultó ser una relación inversa, indicando que a medida que aumenta el índice de escolaridad, la proporción de personas que caerán por debajo del nivel de pobreza disminuirá, tal y como se había supuesto en la formulación de las hipótesis. En el caso de los índices de desempleo y ruralidad, la relación de estas con la variable dependiente, resultó ser una de tipo directa. En otras palabras, un aumento en los índices de ruralidad o de desempleo causarán un aumento en el número de personas que caerán por debajo del nivel de pobreza, como también se supuso en el desarrollo de las hipótesis de trabajo.

Por otra parte, el coeficiente de determinación (CD) expuesto anteriormente, indica el porcentaje de la variación en nuestra variable dependiente que está determinada por la acción conjunta y simultánea de las tres variables independientes. Para nuestro trabajo, este coeficiente, CD, resultó ser de 88.10%, lo que nos indica que el 88.10% de las variaciones en la variable dependiente se debieron a la acción conjunta y simultánea de las tres variables independientes: el índice de escolaridad,

el índice de desempleo y el índice de ruralidad. También, se puede concluir en base de este coeficiente que el 11.90% de la variación en la variable dependiente se debió a la acción de otros factores no incluidos en el modelo general.

El otro coeficiente expuesto, es el coeficiente de correlación múltiple (CCM). Este coeficiente lo que hace es medir, en términos porcentuales, el grado de asociación o correlación existente entre las variables independientes y la variable dependiente. En el presente caso, este coeficiente, al igual que el anterior, resultó ser relativamente alto 93.86%. Esto indica que en el modelo, existe un alto grado de correlación o asociación entre las variables independientes y la variable dependiente.

El último de los coeficientes expuestos anteriormente es el error estándar de regresión (EER). Este coeficiente lo que hace es medir el error promedio de estimación al utilizar la ecuación ajustada. Este error resultó ser de .0404, lo que indica que cualquier valor del índice de pobreza que se utilice en la ecuación, va a tener un error promedio de estimación de .0404 unidades.

Si se divide este error entre el promedio del índice de pobreza dará el grado de desconfiabilidad general del modelo. En este caso esto sería igual a:

$$\frac{\text{EER}}{\bar{\text{Pb}}} = \frac{.0404}{.6956} = .0581$$

donde:

EER = error estándar de regresión

$\bar{\text{Pb}}$  = promedio del índice de pobreza

A base de esto podemos concluir que nuestro modelo tiene un grado de confiabilidad promedio del 94.2%.

Luego de haber visto los resultados generales del modelo general, se procederá ahora a la comprobación de las hipótesis formuladas anteriormente,

entre las variables independientes y la variable dependiente.

El propósito de esta sección será la de determinar si las suposiciones teóricas que se formularon son comprobadas por el análisis empírico que se llevó a cabo.

La primera de estas hipótesis formuladas fue sobre la existencia de una relación inversa entre el índice de pobreza y el índice de escolaridad. Lo que se quiere averiguar es, si la variable independiente, el índice de escolaridad, es una variable significativa en el modelo. Para esto se utilizará la prueba estadística "t-student", la cual permite determinar si la variable independiente en cuestión está relacionada significativamente con la variable dependiente o no.

Para calcular la estadística de "t-student" se hace uso de la siguiente fórmula:

$$t_c = \frac{b_i}{S_{b_i}}$$

donde:

$t_c$  = es el valor calculado de t

$b_i$  = el valor del parámetro al cual se quiere aplicar la prueba de t.

$S_{b_i}$  = es el error estimado de  $b_i$ .

Para saber si el coeficiente de regresión de la primera variable independiente,  $X_1$ , (índice de escolaridad) es significativo o no, se formularán dos hipótesis generales que serán aceptadas o rechazadas dependiendo del valor de t para el parámetro en cuestión.

Hipótesis nula:  $H_0 : b_1 = 0$ ; esto ocurrirá cuando el parámetro  $b_1$  no sea, significativamente diferente de 0. Si esto ocurre se puede concluir que esta variable  $X_1$  no es significativa en el modelo y por lo tanto no existe una relación lineal entre la variable dependiente y la independiente.

Hipótesis alternativa:  $H_a : b_1 \neq 0$ ; esto ocurrirá cuando el parámetro  $b_1$  sea significativamente diferente de 0. De ocurrir esto, se puede concluir que la variable  $X_1$  es significativa en el modelo y que va a tener una relación lineal con la variable dependiente.

Para el parámetro del índice de escolaridad el valor de  $t$  calculado fue:

$$t_c = \frac{-0.0649}{0.0053} = -12.1939$$

Este valor,  $t_c$ , hay que compararlo con un valor de  $t$  tabular  $(t_t)^{18/}$ , utilizando un nivel de confiabilidad del 95% con 72 grados de libertad. Si el valor de  $t_c$  es mayor que el valor  $t_t$ , procederíamos inmediatamente a rechazar la hipótesis nula. De ocurrir lo contrario, que  $t_c$  sea mayor que  $t_t$ , se tendría entonces, que aceptar la hipótesis nula y rechazar la hipótesis alternativa.

Al buscar en la tabla apropiada se encontró que el valor de  $t$  tabular, en nuestro caso, es igual a:

$$t_t = -1.9935$$

Por lo tanto, se puede apreciar que el valor de  $t_c$  para  $b_1$  es mayor que el valor  $t_t$ , lo que permite rechazar la hipótesis nula y aceptar la hipótesis alternativa. El parámetro  $b_1$  es significativamente diferente de 0, por lo cual se puede concluir que el índice de escolaridad es una variable significativa en el modelo general y que va a tener una relación de tipo lineal con el índice de pobreza.

Se procederá ahora a comprobar si la segunda variable independiente formulada,  $X_2$  (índice de desempleo) es significativa o no, en nuestro modelo. Nuevamente se establecerán dos hipótesis generales:

---

<sup>18/</sup> Sahai, Hardeo, A Vade Mecum of Commonly Used Statistical Tables with Explanation of Their Uses, Mayaguez, 1976, p. 42, Table 10.

Hipótesis nula:  $H_0 : b_2 = 0$ ; esto ocurrirá si el parámetro

$b_2$  no es significativamente diferente de 0;  $X_2$  (índice de desempleo) no será una variable significativa en el modelo, por lo tanto no existirá relación lineal entre  $X_2$  y  $P_b$ .

Hipótesis alternativa;  $H_a : b_2 \neq 0$ ; esto ocurrirá si el parámetro

$b_2$  es significativamente diferente de 0;  $X_2$  (índice de desempleo) será una variable significativa en el modelo, por lo tanto existirá relación lineal entre  $X_2$  y  $P_b$ .

El valor calculado de  $t$  para el parámetro de la variable  $X_2$

fue:

$$t_c = \frac{0.0150}{0.1193} = 0.1257$$

y el valor de  $t_c = 1.9935$

Como se puede ver el valor de  $t_c$ , 1.9935, es mayor que el valor de  $t_c$ , 0.1257, por lo tanto se procede a aceptar la hipótesis nula, que indica que el parámetro  $b_2$  no es significativamente diferente de 0 y que por lo tanto no existe una relación de tipo lineal entre el índice de desempleo y la variable dependiente, el índice de pobreza.

Al analizar los resultados del modelo general, se comprobó de que posiblemente existía el problema de multicolinealidad, que surge cuando existe una alta correlación entre las variables independientes. Inmediatamente, a la matriz de correlaciones parciales y se comprobó que esto era cierto. El coeficiente de correlación entre  $X_2$  y  $X_3$  era de .618. Por otra parte, el coeficiente de correlación entre  $X_1$  y  $X_3$  era de .781. Estos resultados demuestran una alta correlación entre las variables independientes  $X_1$ ,  $X_2$  y  $X_3$ . Este problema se analizará más a fondo luego de que se verifiquen la hipótesis para la última variable independiente,  $X_3$  y con la comprobación de la significancia de la constante de la ecuación.

Haciendo la misma prueba de la significación que se hizo con las otras dos variables independientes, se verificará la importancia de la tercera variable independiente ( $X_3$ ) (índice de ruralidad). Nuevamente se formulan las dos hipótesis generales:

Hipótesis nula;  $H_0$  ;  $b_3 = 0$ ; esto ocurrirá si el parámetro

$b_3$  no es significativamente diferente de 0;  $X_3$  (índice de ruralidad) no será una variable significativa en el modelo, por lo tanto no existirá relación lineal entre  $X_3$  y  $P_b$ .

Hipótesis alternativa;  $H_a$  :  $b_3 \neq 0$  esto ocurrirá si el

parámetro  $b_3$  es significativa en el modelo, por lo tanto existirá relación lineal entre  $X_3$  y  $P_b$ .

El valor calculado de  $t$ , para el parámetro de la variable  $X_3$  es:

$$t_c = \frac{.0240}{.0298} = .8046$$

y el valor de  $t_t$  es de:

$$t_t = 1.9935$$

Se observó que el valor de  $t_t$  (1.9935) es mayor que el valor  $t_c$  (.8046) por lo tanto se procede a aceptar la hipótesis nula que indica que el parámetro  $b_3$  no es significativamente diferente de 0 y por lo tanto no existe una relación de tipo lineal entre la variable dependiente, el índice de pobreza, y la variable independiente,  $X_3$  (índice de ruralidad).

Se observa el mismo problema que surgió con la variable  $X_2$ , en donde el análisis indica que la variable  $X_3$  no es significativa en el modelo general, rechazando la hipótesis que se formuló anteriormente.

Posiblemente esto se debe al problema de la multicolinealidad existente entre las variables independientes, como ya se ha señalado anteriormente.

Este problema de la multicolinealidad puede llevar a conclusiones erróneas sobre la significancia de las variables independientes en el modelo general, por lo tanto es necesario hacer otros tipos de análisis que arrojen más luz sobre este tipo de problema. Antes de pasar a estos análisis conviene comprobar la significancia de la constante de la ecuación,  $a$ :

Se formulan de nuevo las dos hipótesis generales:

Hipótesis nula:  $H_0 : a = 0$ ; esto sucederá si el intercepto de la función con el eje de  $P_b$  no es significativamente diferente de 0.

Hipótesis alternativa:  $H_a : a \neq 0$ ; esto sucederá si el intercepto de la función con el eje de  $P_b$  es significativamente diferente de 0.

El valor calculado de  $t$ , para el intercepto de la función,  $a$ , es:

$$t_c = \frac{1.0451}{0.0551} = 18.9559$$

y el valor de la  $t$  tabular es:

$$t_t = 1.9935$$

Por lo tanto, se rechaza la hipótesis nula y se acepta la hipótesis alternativa, ya que el valor de  $t_c$  (18.9559) es mayor que el valor de  $t_t$  (1.9935). Esto quiere decir que el valor de intercepto es significativamente diferente de 0, y que por lo tanto deben existir otros factores, no incluidos en este trabajo, que afectan la variable dependiente.

Como se dijo anteriormente, se pasará ahora a bregar con el problema de la multicolinealidad entre las variables independientes. Cuando se observó en los resultados que dos de las tres variables independientes,

$X_2$  y  $X_3$ , no eran supuestamente significativas y confiándose en que el razonamiento teórico para la inclusión de estas dos variables era suficientemente fuerte, se consultó con el Profesor que dicta el curso quien sugirió varias maneras mediante las cuales se podría resolver este problema.

De entre los métodos recomendados, los dos que se analizarán a fondo serán: 1) hacer regresiones simples entre las variables independientes  $X_2$  y  $X_3$ , por separado con la variable dependiente,  $P_b$ , y 2) hacer una regresión múltiple entre  $X_2$  y  $X_3$  y el índice de pobreza.

El primer método sugerido fue el de hacer regresiones simples entre estas dos variables, por separado, con la variable dependiente, para poder determinar si en verdad estas variables eran insignificantes o si lo que ocurría era el problema de la multicolinealidad.

La regresión simple entre la variable independiente  $X_2$  (índice de desempleo) y la variable dependiente (el índice de pobreza), arrojó los siguientes resultados: el coeficiente de determinación fue de 0.6691 y el error estándar de regresión fue de 0.0853. El valor de  $t$  calculada ( $t_c$ ) para el parámetro del índice de desempleo, en esta ocasión fue de 7.7459, lo cual comparado con el valor de  $t$  tabular  $\frac{19}{74}$ ,  $t_t$ , (con un grado de 95% de confiabilidad y 74 grados de libertad que igual a 1.9925, nos da el resultado de que  $b_2$  sí es significativamente diferente de 0 y que por lo tanto existe una relación lineal entre  $X_2$  (índice de desempleo) y la variable dependiente.

Por otra parte, la regresión simple llevada a cabo entre la variable independiente  $X_3$  (índice de ruralidad) y la variable dependiente,  $P_b$ ; nos dió los siguientes resultados: el coeficiente de determinación fue 0.5683, el coeficiente de correlación múltiple fue de 0.7539 y el error

---

<sup>19/</sup> Ibid, p. 42, Tabla 10.

estándar de regresión fue de 0.0758. El valor de  $t$  calculada para el parámetro  $b_3$ , de la variable  $X_3$ , fue de 9.8700, lo cual comparado con el valor de  $t$  tabular,  $t_t$ , (con un grado de 95% de confiabilidad y 74 grados de libertad) que es igual a 1.9925, nos dá el resultado de que  $b_3$  sí es significativamente diferente de 0 y que por lo tanto existe una relación lineal entre la variable independiente  $X_3$  (índice de ruralidad) y la variable dependiente,  $P_b$ .

Además, en ambas regresiones, se notó que los coeficientes de correlación y de determinación eran relativamente altos, si se toma en cuenta que eran regresiones simples y que se estaba efectuando un análisis de corte seccional, lo que favorece el argumento de que las variables independientes  $X_2$  y  $X_3$ , sí son significativas en el modelo general.

Otro método que se utilizó para dilucidar el problema, fue el de una regresión múltiple entre las variables independientes,  $X_2$  y  $X_3$ , y la variable dependiente,  $P_b$ . Se cree que eliminando la variable  $X_1$ , que resulta ser la más significativa en el modelo y que, a la vez, parece absorber el efecto de las otras dos variables sobre la variable dependiente, se puede ver mejor el efecto de las variables  $X_2$  y  $X_3$  sobre la variable dependiente.

Este modelo arrojó los siguientes resultados: el coeficiente de determinación fue de 0.6352, el coeficiente de correlación múltiple fue de 0.7970 y el error estándar de la regresión fue de 0.0702. El valor de  $t$  calculada,  $t_c$ , para  $b_2$  fue de 3.6586 y el de  $b_3$  fue de 6.1243, lo cual comparado con el valor de  $t$  tabular,  $t_t$ , (con un grado de 95% de confiabilidad y 73 grados de libertad), que es igual a 1.9930 <sup>20/</sup>, dá el resultado de que  $b_2$  y  $b_3$  sí son significativamente diferentes de 0 y que por lo tanto, existe una relación lineal entre las variables independientes,  $X_2$  y  $X_3$  y la variable dependiente  $P_b$ .

Como se puede observar, al eliminar el efecto de la variable independiente  $X_1$ , sobre la variable dependiente y sobre las otras dos variables independientes, se encontró que las otras dos variables,  $X_2$  y  $X_3$ , son altamente significativas, ya que existe un grado de correlación entre estas dos variables y la variable dependiente de un 9.0%, lo cual resulta ser bastante alto.

Se hicieron otros análisis para tratar de esclarecer el problema que surgió, y todos estos dieron los mismos resultados: las variables  $X_2$  y  $X_3$ , si son significativas en el modelo general.

Debido a las limitaciones del tiempo sería prácticamente imposible analizar en detalle los diferentes métodos utilizados para atacar el problema, aunque sí conviene mencionar algunos de estos:

- 1) Se seleccionó una muestra de 15 municipios los cuales tuvieran un índice de escolaridad relativamente constante, entre 4.2 y 4.6 y luego se procedió a hacer un análisis de regresión múltiple entre las tres variables independientes y la variable dependiente.
- 2) Se hizo análisis de regresión múltiple entre:
  - a) la variable dependiente y las variables independientes  $X_1$  y  $X_3$ .
  - b) la variable dependiente y las variables independientes  $X_1$  y  $X_2$ .
- 3) Se hizo análisis de regresión simple entre:
  - a) la variable  $X_1$ , utilizada como variable dependiente, y la variable independiente  $X_2$ .
  - b) la variable  $X_1$ , utilizada como variable dependiente, y la variable independiente  $X_3$ .
  - c) la variable  $X_2$ , utilizada como variable dependiente, y la variable independiente  $X_3$ .

El análisis de todas estas regresiones, como se dijo anteriormente, apuntan hacia dos resultados básicos:

- 1) existe una alta correlación entre las variables independientes, lo que crea el problema de la multicolinealidad. Esta alta correlación entre las variables independientes, fue la causante de que las variables  $X_2$  y  $X_3$  resultaran no significativas en el modelo general, en los análisis preliminares.
- 2) las variables independientes,  $X_2$  y  $X_3$ , resultan ser significativas en el modelo general, no empuje a la existencia de la multicolinealidad, que hace creer lo contrario.

Analizadas las pruebas de significancia para los coeficientes de regresión y para la constante de regresión, se procederá ahora a hacerle una prueba de significancia al coeficiente de determinación (CD) del modelo general. Lo que se quiere averiguar es si el coeficiente de determinación es significativamente diferente de 0. Si CD resulta ser significativamente diferente de 0, entonces se podría concluir que las tres variables independientes, conjunta y simultáneamente explican el comportamiento de la variable dependiente en la magnitud del valor de CD.

Para esta prueba de significancia se utilizará la estadística F, la cual se comparará con un valor tabular de F ( $F_t$ ), siguiendo un criterio muy parecido al de la prueba de "t-student".

Se establecerá aquí también, las dos hipótesis generales:

Hipótesis nula;  $H_0$  :  $CD = 0$ ; esto significaría que la acción conjunta y simultánea de las tres variables independientes no explican el comportamiento de la variable dependiente.

Hipótesis alternativa;  $H_a$  :  $CD \neq 0$ ; esto significaría que la acción conjunta y simultánea de las tres variables independientes sí explican el comportamiento de la variable dependiente.

El valor de  $F$  computado resultó ser igual a:

$$F_c = 177.6502$$

Y el valor tabular de  $F^{21}$  a un grado de confianza de 95%, con 72 grados de libertad, resultó ser igual a:

$$F_t = 2.76$$

Como se puede ver,  $F_c$  es mayor que  $F_t$ , lo que significa que se debe rechazar la hipótesis nula y aceptar la hipótesis alternativa que indica que  $CD$  vá a ser significativamente diferente de 0 y que por lo tanto la acción conjunta y simultánea de las tres variables independientes, sí van a explicar el comportamiento de la variable dependiente. Las variables independientes, conjunta y simultáneamente, van a estar relacionadas sistemáticamente con la variable dependiente.

Habiendo ya comprobado la significancia, de los primeros parámetros de las variables independientes y del coeficiente de la constante de regresión y habiendo, también comprobado que la acción conjunta y simultánea de las tres variables independientes explican el comportamiento de la variable dependiente, lo único que quedaría sería, el analizar cuál de las variables independientes tuvo un impacto mayor sobre el comportamiento de la variable dependiente.

Esto se puede lograr si utilizamos el coeficiente de  $B$  (beta). El coeficiente de  $B$  lo que hace es ordenar de mayor a menor grado de importancia las variables independientes que resultaron ser significantes en el modelo. Mientras más alto sea el valor de  $B$  para una variable independiente, mayor será la influencia que tendrá ésta sobre la variable dependiente.

El coeficiente de  $B$  para la primera variable independiente,  $X_1$  (índice de escolaridad) fue de:

---

21/ Véase nota al calce 16, p. 48, Tabla 12.

$$B_1 = 0.8914$$

El valor del coeficiente de B para la variable  $X_2$  (índice de desempleo) fue de:

$$B_2 = 0.0073$$

Y por último, el valor del coeficiente de B para la variable  $X_3$  (índice de ruralidad) fue de:

$$B_3 = 0.530$$

En base a estos resultados se puede concluir que en el modelo general la variable independiente de mayor importancia, o de mayor impacto, sobre la variable dependiente lo fue el índice de escolaridad. Luego de ésta, le siguió en importancia el índice de ruralidad; y como la variable independiente, de menor importancia o de menor efecto sobre nuestra variable dependiente estuvo el índice de desempleo.

## VI. Conclusiones y Recomendaciones

Luego de haber analizado todos los resultados obtenidos de los análisis anteriormente descritos, podemos llegar a las siguientes conclusiones:

- 1) Con respecto a la primera variable independiente, el índice de escolaridad, se encontró que ésta está altamente correlacionada, de forma lineal e inversa, con el índice de pobreza, la variable dependiente.
- 2) Cuando se analizó la variable independiente correspondiente al índice de desempleo en el modelo general formulado, se encontró que esta variable no era significativa. Se previó aquí, que posiblemente ocurría el problema de la multicolinealidad y se hicieron análisis posteriores. Estos análisis indicaron que sí existía una relación significativa, de tipo lineal y directa entre el índice de desempleo y el índice de pobreza.

- 3) Con respecto a la tercera variable independiente, el índice de ruralidad, se encontró que en el modelo general formulado, la misma resultó ser no significativa. Se previó aquí también, el problema de la multicolinealidad y al igual que en el caso anterior, se procedió a hacer análisis posteriores. Estos análisis indicaron que sí existía una relación significativa, de tipo lineal y directa, entre el índice de ruralidad y el índice de pobreza.
- 4) El problema de la multicolinealidad se debió a los altos grados de correlación parcial, existente entre las variables independientes.
- 5) Se encontró que la variable independiente más importante o con un mayor efecto sobre el nivel de pobreza fue el índice de escolaridad. Luego le siguió en importancia el índice de ruralidad y por último, como variable menos importante, el índice de desempleo.
- 6) No empece al problema de la multicolinealidad, se pudo aceptar la hipótesis general que indicaba que las tres variables independientes estaban relacionadas de forma lineal a la variable dependiente.

Antes de terminar, se formularán dos recomendaciones, que a juicio de los autores, servirían para mejorar el presente estudio.

En primer lugar, se deberían buscar otros factores que ayudaran a explicar mejor el problema de la pobreza en Puerto Rico. Se observa todavía que el coeficiente de la constante de regresión es relativamente alto, lo que indica la ausencia en el modelo de otros factores importantes en la determinación de la condición de pobreza.

-Bibliografía-

1. Davis, Kenneth S., The Paradox of Poverty in America, New York, The H. W. Wilson Company, 1969.
2. Allen, Vernon L., Psychological Factors in Poverty, Chicago, Mark Ham Publishing Company, 1971.
3. Lewis, Oscar, A Study of Slum Culture, Backgrounds for La Vida, New York, Random House, 1968.
4. Lewis, Oscar, La Vida: A Puerto Rican Family in the Culture of Poverty, San Juan and New York, New York, New York, New York Vintage Books, 1965.
5. Miller, Herman, Poverty in Puerto Rico, San Juan, Junta de Planificación de Puerto Rico, 1964.
6. González Antonio., Economía Política de Puerto Rico, San Juan, Editorial Cordillera, Inc., 1971.
7. La Alternativa Socialista, Tesis Política del Partido Socialista Puertorriqueño, Río Piedras, Ediciones Puerto Rico, 1974.
8. McConnell, Campbell R., Economics, New York, McGraw-Hill Book Company, 1975.
9. Wallis, Kenneth E., Introducción a la Econometría, Madrid, Alianza Editorial, 1976.
10. Dutta, H., Econometrics Methods, Cincinnati, South-Western Publishing Company, 1975.
11. Wonnacott and Wonnacott, Econometrics, New York, John Willey and Sons, 1970.
12. Dagon y Dagon, Introducción a la Econometría, Argentina Editorial Siglo XXI, S.A., 1971.
13. Sahai, Hardeo, A Vade Mecum of Commonly Used Statistical Tables with Explanations of their Uses, Mayaguez, Universidad de Puerto Rico, 1976.
14. Ferber and Verdoorn, Research Methods in Economics and Business, 6th Edition; New York, McMillan Company, 1970.
15. Enciclopedia Internacional de las Ciencias Sociales, Madrid, Aguilár, 1976. Vol. 3.

15. Palgrave, Robert I., Palgrave's Dictionary of Political Economy, reprinted revised edition, New York, Augustus H. Kelly Bookseller, 1963.
17. McGraw Hill Economic Dictionary, McGraw Hill Book Company, New York, 1955, pp. 387-388.
18. Horton, Bryne Joseph, Dictionary of Modern Economics, Washington D.C. American Council of Public Affairs, pp. 250-251.
19. Censo Poblacional de 1970, Publicación para Puerto Rico del Departamento de Comercio de Estados Unidos. 1970.

PUBLICACIONES DE LA  
UNIDAD DE INVESTIGACIONES ECONOMICAS

TEMAS SOBRE ECONOMIA DE PUERTO RICO

Serie de Conferencias y Foros

- Núm. 1: José A. Herrero, La economía de Puerto Rico: El presente crítico, noviembre 1975.
- Núm. 2: Jaime A. Santiago, Impacto de la recesión en Puerto Rico, febrero 1976.
- Núm. 3: Jaime A. Santiago, Objectives and Strategy for Economic Development: A Challenge, julio 1976.
- Núm. 4: Jaime A. Santiago, Conferencia sobre economía y población, septiembre 1976.
- Núm. 5: Rubén A. Vilches, Current Labor Force Perspectives & Problems in Puerto Rico, octubre 1976.
- Núm. 6: Jaime A. Santiago, La situación presupuestaria de 1976 y 1977 y sus efectos sobre la economía, diciembre 1976.
- Núm. 7: Elías R. Asón, Comentarios sobre las recomendaciones del uso de los fondos de las corporaciones 936, mayo 1977.
- Núm. 8: Irma G. Tirado de Alonso, Equity and Efficiency Considerations Concerning Higher Education in Puerto Rico, diciembre 1977.

Serie de Ensayos y Monografías

- Núm. 1: Alejandro Asmar, Análisis breve del Informe Tobin, junio 1976.
- Núm. 2: Samuel Torres Román, La problemática económica de Puerto Rico, diciembre 1976.
- Núm. 3: Angel L. Ruiz, Enfoque de insumo-producto para determinar una política de sustitución de importaciones en Puerto Rico, mayo 1977.
- Núm. 4: Angel A. Rivera, Modelo de costo-beneficio para la evaluación de proyectos que sustituyan importaciones, junio 1977.
- Núm. 5: Angel L. Ruiz, Un análisis de la productividad en la industria de centrales y refinerías en Puerto Rico, diciembre 1977.

(sigue)

## SERIE DE ENSAYOS Y MONOGRAFÍAS (Cont'n)

- Núm. 6: Angel L. Rufz, Análisis estructural de la industria de la construcción, Años 1962-63 a 1971-72, septiembre 1978.
- Núm. 7: Angel L. Rufz y Parimal Choudhury, The Impact of Food Stamp Program on the Puerto Rican Economy: An Input-Output Approach, septiembre 1978.
- Núm. 8: Pedro A. Cantero Frau, José Rodríguez García y Mary H. Alverio Rodríguez, Estudio sobre Factores que determinan la pobreza en Puerto Rico, diciembre 1978.

## SERIE DE TEMAS RELACIONADOS

- Núm. 1: Walter Bruckman, Diseño y método de investigación en la ciencia económica, mayo 1977.
- Núm. 2: Stephen J. Hienstra, Perspectives on Food Program Evaluation and Research, mayo 1977.
- Núm. 3: Angel L. Rufz, Las conferencias cimeras de Rambouillet y Puerto Rico, mayo 1977.
- Núm. 4: Jaime Santiago Meléndez, El presupuesto ejecutivo: Reto al poder ejecutivo, octubre 1977.
- Núm. 5: Miguel A. Ramírez Pérez y Migdalia Rivera Avilés, Recesión e inflación: Inaplicabilidad de la Curva Phillips a la economía de Puerto Rico, diciembre 1977.

## TEMAS DIVERSOS DE ECONOMIA

- Núm. 1: Fuat y Suphan Andic, Public Finance, Development, and the Third World, diciembre 1976.
- Núm. 2: Paul W. Strassman, Housing Priorities in Developing Countries: A Planning Model, enero 1977.
- Núm. 3: William J. Baumol, Quasi Optimality, The Price we must pay for a Price System, marzo 1977.
- Núm. 4: Angel L. Rufz, La recuperación económica norteamericana y perspectivas para 1977, junio 1977.
- Núm. 5: J. Vanek, Self-management, Workers' Management and Labor Management in Theory and Practice: A Comparative Study, julio 1977.
- Núm. 6: J. Vanek, The Absurdity of the Rich Men's Trade Doctrine and Institutions for the Present Day World Economy and an Attempt to Reformulate, agosto 1977.
- Núm. 7: Karel Holbik, The International Monetary System and the New International Order, noviembre 1977.
- Núm. 8: J.R. Dosworth y T.V. Somersall, The Analysis of Taxable Capacity of LDC's, noviembre 1977.

- Núm. 9: Ramón J. Cao-García, Equilibrio y análisis económico, noviembre 1977.
- Núm. 10: Suphan Andic, Industrialization and Economic Integration: An Evaluation of Latin American Performance, diciembre 1977.
- Núm. 11: Víctor A. Sepúlveda-Rodríguez, Nunca veremos la mano invisible: El Proceso de Difusión de Capital, diciembre 1977.
- Núm. 12: Evelyn Ramos de Jesús, Apuntes sobre desarrollo económico: I, diciembre 1977.
- Núm. 13: Evelyn Ramos de Jesús, Apuntes sobre desarrollo económico: II, diciembre 1977.
- Núm. 14: Gilberto Gómez Arango, El papel del sector informal en la economía capitalista dependiente, junio 1978.
- Núm. 15: Daniel M. Schdlowsky, Capital Utilization, Growth, Employment, Balance of Payments and Price Stabilization, junio 1978.
- Núm. 16: Ramón J. Cao, Zona franca industrial y comercial de Cartagena: Un análisis de costo y beneficio social, noviembre 1978.
- Núm. 17: Joachim Singelmann, Population, Income, and Employment Growth in Services, 1960-1970; May 1978.

## SEPARATAS

- Núm. 1: Richard A. Posner, The Economic Approach to Law, 1976.
- Núm. 2: Richard A. Musgrave, ET., OT., and S.B.T., 1976
- Núm. 3: Fuat M. Andic y Arthur Mann, Secular Tendencies in the Inequality of Earnings in Puerto Rico, febrero 1977.
- Núm. 4: Arthur J. Mann, Public Expenditure Patterns in the Dominican Republic and Puerto Rico, 1930-1970, septiembre 1977.
- Núm. 5: Arthur J. Mann, The Fiscal System and Income Distribution: The Case of Puerto Rico, septiembre 1977.
- Núm. 6: Arthur J. Mann y Donald Perlis, Utility Functions, Public Goods, and Income Distribution, septiembre 1977.
- Núm. 7: Suphan Andic y Parimal Choudhury, Direct Subsidies and Income Redistribution: The Food Stamp Program in Puerto Rico, octubre 1977.
- Núm. 8: Fuat M. Andic, Poverty and Tax Incidence in West Malaysia, diciembre 1977.

## RESEÑAS

- Núm. 1: Elena M. Díaz, Economic Theory and the Developing Countries - por Ajit K. Dasgupta, diciembre 1977.
- Núm. 2: Miriam Borrero, Economic Theory and the Developing Countries - Ajit K. Dasgupta, diciembre 1977.