

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

INSTITUTO DE CIENCIAS MATEMÁTICAS

Ingeniería en Estadística Informática

**“CONSTRUCCIÓN DE INDICADORES SOCIALES A PARTIR
DE LOS DATOS DE UN CENSO POBLACIONAL Y VIVIENDA”**

TESIS DE GRADO

Previa la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentada por:

VERÓNICA LORENA IBARRA GARCÍA



GUAYAQUIL – ECUADOR

AÑO

1.999

AGRADECIMIENTO

A DIOS, por su apoyo y ayuda en los momentos más difíciles.

DEDICATORIA

A DIOS
A MI FAMILIA
A MIS VERDADEROS AMIGOS

TRIBUNAL DE GRADUACIÓN

Ing. Félix Ramírez
DIRECTOR DEL ICM

Mat. John Ramírez
DIRECTOR DE TESIS

Ph. D. José Alfonso Roche
VOCAL

Ing. Luis Rodríguez O.
VOCAL

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de esta Tesis de Grado, me corresponden exclusivamente y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”

Verónica Lorena Ibarra García

RESUMEN

El presente proyecto establece el Análisis de Componentes Principales como método para la construcción de indicadores sociales, el cual se basan en información obtenida del Censo de Población y Vivienda realizado en el Ecuador en el año 1990. La interpretación hecha por especialistas en la materia, ayudará a la toma de decisiones sobre aspectos de total trascendencia en la vida social del país.

En la primera fase se da un enfoque teórico de lo que son los indicadores sociales, así como también el tipo de indicadores que se pueden realizar, para luego centrarse en los indicadores sintéticos que se utilizarán en la construcción de un indicador de pobreza.

Luego se estudia la técnica multivariante de Análisis de Componentes Principales, que posteriormente se aplica en la elaboración del indicador sintético.

La segunda fase –básicamente el capítulo IV- trata sobre el uso de Escalas de Likert como método para la construcción de indicadores parciales que forman parte del indicador sintético, así como el análisis estadístico de diferentes variables censales y su influencia en el indicador sintético, para llegar a las conclusiones debidas.

ÍNDICE GENERAL

	Pág.
RESUMEN	VI
ÍNDICE GENERAL	VII
ÍNDICE DE FIGURAS	X
ÍNDICE DE TABLAS	XII
ABREVIATURAS	XIV
INTRODUCCIÓN	15
I. ASPECTOS GENERALES DE LA TESIS	18
1.1 Definición del Tema de Tesis	18
1.2 Antecedentes de la Pobreza en el Ecuador	19
1.3 Finalidad de la Tesis	20
1.4 Objetivos Específicos de la Tesis	21
II. ASPECTOS CONCEPTUALES Y METODOLÓGICOS	23
2.1 Indicadores Sociales y Análisis Cuantitativo de Pobreza	23
2.2 Criterios para el Establecimiento de Indicadores Sociales	24
2.3 Método para el Establecimiento de Indicadores Sociales y para la Evaluación de la Pobreza	26

2.4	Pasos para el Establecimiento de Indicadores Sintéticos	27
2.5	Propiedades de los Indicadores Sintéticos	30
2.6	Fuentes de Datos que pueden usarse para la Evaluación de la Pobreza	32
2.7	Selección de Indicadores de Pobreza	36
III.	TÉCNICAS MULTIVARIANTES PARA LA CONSTRUCCIÓN DE INDICADORES SINTÉTICOS	39
3.1	Empleo de Técnicas Estadísticas Multivariantes para el Establecimiento de Indicadores Sintéticos de Pobreza	39
3.2	Análisis de Componentes Principales	41
3.2.1	Objetivos del Análisis de Componentes Principales	41
3.2.2	Características deseables de las Componentes Principales	42
3.2.3	Generación de las Componentes Principales	43
3.2.4	Selección del número de Componentes Principales	55
3.2.5	Correlación entre variables originales y componentes principales	58
IV.	APLICACIÓN DE COMPONENTES PRINCIPALES PARA LA CONSTRUCCIÓN DE INDICADORES SINTÉTICOS	61
4.1	Definición de observaciones	61
4.2	Definición de variables	63
4.3	Escala de Likert	64
4.4	Estadística descriptiva de las variables	68

4.5	Análisis de Componentes Principales	85
4.5.1	Interpretación de las componentes principales en función de su correlación con las variables originales	93
4.5.2	Interpretación de la nube de variables y gráficos de correlación	96
4.5.3	Interpretación de la nube de individuos	104
4.6	Construcción del Indicador Sintético de Pobreza	109
V. CONCLUSIONES Y RECOMENDACIONES		114
APÉNDICES		118
BIBLIOGRAFÍA		144

ÍNDICE DE FIGURAS

		Pág.
Figura 4.1	X1 Proporción de viviendas por provincia	69
Figura 4.2	X2 Indicador porcentual de tipo de vivienda	70
Figura 4.3	X3 Indicador porcentual de abastecimiento de agua	71
Figura 4.4	X4 Indicador porcentual del sistema de eliminación de aguas servidas	72
Figura 4.5	X5 Proporción de las viviendas que disponen de servicio eléctrico del total de viviendas provinciales	73
Figura 4.6	X6 Proporción de las viviendas que disponen de servicio telefónico del total de viviendas provinciales	74
Figura 4.7	X7 Indicador porcentual del sistema de eliminación de basura	75
Figura 4.8	X8 Indicador porcentual de disponibilidad de servicio higiénico	76
Figura 4.9	X9 Indicador porcentual de disponibilidad de ducha	77
Figura 4.10	X10 Indicador porcentual de tenencia de vivienda	78
Figura 4.11	X11 Proporción de la población por provincias	79
Figura 4.12	X12 Proporción de la población rural por provincias	80
Figura 4.13	X13 Indicador porcentual nivel de instrucción	81
Figura 4.14	X14 Proporción de la población analfabeta de la población total provincial	82

Figura 4.15	X15 Proporción de la población económicamente activa de la población total provincial	83
Figura 4.16	Varianza Explicada por cada componente	89
Figura 4.17	Gráficos de las variables sobre el plano formado por la Componente I y la Componente II y su correlación	97
Figura 4.18	Gráficos de las variables sobre el plano formado por la Componente I y la Componente III y su correlación	100
Figura 4.19	Gráficos de las variables sobre el plano formado por la Componente I y la Componente IV y su correlación	102
Figura 4.20	Gráficos de las provincias sobre el plano formado por la Componente I y la Componente II	106
Figura 4.21	Gráficos de las provincias sobre el plano formado por la Componente I y la Componente II	107
Figura 4.22	Gráficos de las provincias sobre el plano formado por la Componente I y la Componente IV	108

ÍNDICE DE TABLAS

	Pág.	
Tabla IV.I	Tabla de valores propios obtenidos de la matriz R y porcentaje de la varianza total explicada	88
Tabla IV.II	Tabla de valores propios con sus respectivos vectores propios, basados en el número de factores	81
Tabla IV.III	Tabla de Componentes Principales o Factores	92
Tabla IV.IV	Correlaciones de las variables con las Componentes Principales o factores	95
Tabla IV.V	Proyección de las provincias sobre las Componentes Principales	105
Tabla IV.VI	Indicador Sintético de Pobreza	113
Tabla B.I	X ₁ Proporción de viviendas por provincia	
Tabla B.II	X ₂ Indicador porcentual de tipo de vivienda	
Tabla B.III	X ₃ Indicador porcentual de abastecimiento de agua	
Tabla B.IV	X ₄ Indicador porcentual del sistema de eliminación de aguas servidas	
Tabla B.V	X ₅ Proporción de las viviendas que disponen de servicio eléctrico del total de viviendas provinciales	
Tabla B.VI	X ₆ Proporción de las viviendas que disponen de servicio telefónico del total de viviendas provinciales	
Tabla B.VII	X ₇ Indicador porcentual del sistema de eliminación de basura	

Tabla B.VIII	X ₈ Indicador porcentual de disponibilidad de servicio higiénico
Tabla B.IX	X ₉ Indicador porcentual de disponibilidad de ducha
Tabla B.X	X ₁₀ Indicador porcentual de tenencia de vivienda
Tabla B.XI	X ₁₁ Proporción de la población por provincias
Tabla B.XII	X ₁₂ Proporción de la población rural por provincias
Tabla B.XIII	X ₁₃ Indicador porcentual nivel de instrucción
Tabla B.XIV	X ₁₄ Proporción de la población analfabeta de la población total provincial
Tabla C.I	Tabla General de datos de variables

ABREVIATURAS

ACP Análisis de Componentes Principales

COMP Componente

INEC Instituto de Estadísticas y Censos

%Prov Porcentaje provincial

INTRODUCCIÓN

El presente trabajo es una propuesta para la “Construcción de Indicadores Sociales a Partir de los Datos de un Censo de Población y Vivienda”, los cuales reflejan la situación latente en la vida socio-económica del país. Se tomó como referencia para el cálculo los datos que se obtuvieron en el V Censo de Población y IV de Vivienda realizado en el año 1.990.

En el Ecuador se han realizado diferentes estudios que evalúan la pobreza, un fenómeno de difícil conceptualización y cuantificación, pues no sólo depende de variables económicas sino también de variables culturales y geográficas, por lo que en el presente estudio vamos a abordar a la pobreza directamente desde su naturaleza multivariante, obteniendo un indicador que a la vez que refleje esta complejidad multidimensional, sea de fácil interpretación y uso en los estudios de tipo social y económico.

El método central utilizado en la elaboración del indicador social de pobreza, fue la técnica multivariante del análisis de componentes principales (ACP), la cual nos provee de un indicador sintético que no es más que la combinación lineal de las variables utilizadas en el censo, que maximizan la varianza, y a su vez, por medio de

-

la correlación entre el indicador y las variables, permite la interpretación su incidencia dentro del mismo.

En el Ecuador se han realizado diferentes estudios que evalúan la pobreza, pero este indicador nos dará la posición de las provincias frente a otras, de una manera más sencilla

I. ASPECTOS GENERALES DE LA TESIS

1.1. Definición del tema de tesis

El tema ha sido propuesto con el título “Construcción de Indicadores Sociales a Partir de los Datos de un Censo de Población y Vivienda”.

Lo que se quiere lograr a lo largo de la misma es construir indicadores sociales, estadístico que sintetiza información relevante sobre aspectos importantes en la sociedad, mediante el uso particular de indicadores sintéticos, en el cual se contemplarán varias variables que deben ser analizadas y que sean determinantes en la problemática a tratar; así, tenemos que si el indicador social que se construirá es de pobreza, las variables pueden contemplar aspectos como: vivienda, salud, vestimenta, educación, etc.

Las condiciones de pobreza pueden ser identificables en los datos estadísticos de un censo de población y vivienda. Por la información en diversos factores que inciden en la pobreza y por ser parte de las variables del censo es que se decide trabajar con indicadores sintéticos, en los cuales podemos hacer una combinación de estas variables mediante los indicadores parciales que las involucren.

1.2. Antecedentes de la pobreza en el Ecuador

Históricamente el Ecuador ha sido catalogado como uno de los países más pobres en Latinoamérica, lo cual se refleja en los índices económicos y sociales, como por ejemplo el PIB, que es uno de los más bajos a escala mundial, y cuyo valor representa el grado de crecimiento que poseemos.

El concepto de pobreza es subjetivo, porque depende del criterio de los pobladores al definir sus necesidades insatisfechas; por lo cual siempre existirá y no puede llegar a convertirse en cero, pero se vuelve indispensable su tratamiento cuando hablamos de necesidades básicas que influyen en la calidad de vida de los ciudadanos.

Así, en otros países existe la debida preocupación por realizar este tipo de estudios que brinden a sus gobiernos la debida información con respecto a las carencias de servicios básicos de los habitantes, que sirvan de guía para la búsqueda de soluciones que ayuden a aliviarlas.

Existen ciertos estudios que sirven como apoyo para esta tesis, pero que tal vez no han recibido atención o que simplemente han quedado archivados para ser consultados por estudiantes y entendidos en la materia, pero que no tienen el poder para cambiar la situación.

Realizar este trabajo es una motivación para querer mejorar y concientizar a los pobladores que los recursos que poseemos, así como, el esfuerzo que podemos realizar para salir adelante, nos pueden llevar a ser un día un país menos pobre.

1.3. Finalidad de la tesis

Se llegará a construir indicadores sociales, básicamente el referente a pobreza, lo que se hará basado en la técnica multivariante de componentes principales lineales, tomando variables que son una representación de los datos del Censo de Población y Vivienda del Ecuador del año de 1.990, cuyo fin es determinar la posición de las provincias dentro de una escala de pobreza.

Luego se efectuará el debido análisis de los resultados obtenidos para continuar con las conclusiones y recomendaciones pertinentes.

1.4. Objetivos específicos de la tesis

- ❖ Dar una explicación general de la teoría de indicadores sociales, con los cuales podemos llegar a resultados que indiquen la situación del país frente a la pobreza. Es importante conocer sobre indicadores sociales para llegar a comprender lo que se puede obtener mediante el uso de los mismos.

- ❖ Centrar al lector en los indicadores de pobreza como una representación de los indicadores sociales y a su vez el uso del método de indicadores sintéticos como medio para tratar de abarcar la mayor información que se presenta en el censo.

- ❖ Definir variables que ayuden al estudio de la pobreza en el Ecuador, lo que está relacionado con aspectos de alimentación, salud, vestimenta, vivienda y educación.

- ❖ Describir la utilización del método de componentes principales para la elaboración de indicadores sintéticos.

- ❖ Hacer un análisis estadístico de los resultados obtenidos y dar las debidas conclusiones y recomendaciones que sean necesarias.

II. ASPECTOS CONCEPTUALES Y METODOLÓGICOS

2.1. Indicadores sociales y análisis cuantitativo de pobreza

Se define como estadístico una función de las variables aleatorias que se pueden observar en una muestra y de las constantes conocidas. Los estadísticos se utilizan para hacer inferencias (estimaciones o decisiones) con respecto a parámetros poblacionales desconocidos.

Los indicadores sociales son estadísticos, basados en observaciones que dicen algo acerca de un aspecto de la vida social en el que se está interesado. Dichas observaciones son normalmente cuantitativas. Cuando la información intenta mostrar cual es la posición o como está ella cambiando, se puede llamar

objetiva; mientras, cuando intenta mostrar como es considerada la posición objetiva o sus cambios por la comunidad en general, se denomina *subjetiva*.

En los tiempos actuales se necesita construir indicadores sociales que muestren los resultados del crecimiento, las causas que llevan a una situación actual y la manera de proyectarse al futuro, a la vez la manera en que se distribuyen, en función de bienestar y prosperidad de la población. Es por ello que los gobiernos son los mayores demandantes de estadísticas económicas, sociales y demográficas.

Se han desarrollado varios tipos de indicadores a partir de datos económicos o sociales, en este caso utilizaremos información social.

2.2. Criterios para el establecimiento de indicadores sociales

Existen seis criterios que deben ser tomados en cuenta para el establecimiento de indicadores sociales, estos son adecuación, capacidad de compendiación, coordinación, exactitud, oportunidad y frecuencia, y viabilidad.

1. Adecuación: Los indicadores sociales deben proporcionar medidas de niveles o tendencias absolutas que se ajusten al aspecto de las condiciones de

vida o al servicio social de que se trata; ésto, para que sean apropiados a la preocupación social relevante.

2. Capacidad de compendición: Para que los indicadores sociales se reduzcan al mínimo indispensable, se deben diseñar de modo que incorporen tanta información sobre una preocupación social dada como sea factible, sin ocultar información sobre las facetas principales de la preocupación, que se necesita para su evaluación.

3. Coordinación: Los indicadores sociales deben estar basados en un conjunto de datos coordinados para que puedan usarse en combinación, para observar y evaluar la situación socioeconómica.

4. Exactitud: El grado de exactitud y de capacidad de comparación de los indicadores sociales reflejará necesariamente el origen de los estadísticos básicos a partir de los cuales se han compilado.

5. Oportunidad y Frecuencia: Los indicadores sociales deben proporcionar un examen periódico oportuno de las condiciones de vida y problemas de la población y de las actividades de los servicios sociales. La frecuencia con que se desea producir determinados indicadores variará según las necesidades y circunstancias de cada entorno tiempo - espacial.

6. Viabilidad: Cuando se desarrollan indicadores sociales debe prestarse atención a la viabilidad de la reunión y tabulación de las estadísticas básicas requeridas, así como las áreas de preocupación social a las que se destinan los indicadores.

2.3. Métodos para el establecimiento de indicadores sociales y para la evaluación de la pobreza

Los indicadores pueden consistir en compendios simples o compendios sintéticos de datos básicos. El método de compendio escogido para un indicador social dado, dependerá del mensaje que deba transmitir y de los datos básicos de que se dispone para compilarlo.

❖**Indicadores Sociales Simples:** Estos consisten en compendios de datos básicos tales como proporciones de población con una característica dada, tasas de incidencia o de cambio, medidas de gravedad, tiempo transcurrido, etc.; medias, medianas y otros estadísticos de tendencia central; distribución porcentual o grupos cuartiles, distribución de frecuencias y medidas de asimetría y kurtosis. Las medias aritméticas y medianas son las medidas de tendencia central más comúnmente usadas al compilar los indicadores sociales. Aunque las medias aritméticas se computan o manipulan mucho más fácilmente

que las medianas, ellas son medidas ineficientes de tendencia central, cuando se trata de distribución asimétrica.

❖ **Indicadores Sintéticos:** Indicador sintético o global es la función matemática de los indicadores parciales, que reúne un conjunto de condiciones que se juzgan necesarios para concretar una medición expresiva del objetivo buscado. Esta expresión matemática concretiza una medición expresiva del estado en que se encuentra un componente en relación con el objetivo buscado.

Se entiende por “componente” de un indicador sintético o global, aquella propiedad que aporta alguna información con independencia de la apreciación subjetiva acerca del objetivo, si el objetivo es el bienestar, un componente será la situación educativa, salud, ingresos, etc.

2.4. Pasos para el establecimiento de indicadores sintéticos

En el trabajo sobre indicadores sintéticos presentado por **Coral Díaz (4)**, propone los siguientes pasos para el establecimiento de indicadores sintéticos, los cuales pueden ser considerados de una forma general.

- 1.** Se selecciona un conjunto de indicadores parciales representativos de la macromagnitud elegida.

2. Se analiza la congruencia entre el indicador parcial y la macromagnitud elegida; es decir, debe haber una correspondencia biunívoca, con el fin de seleccionar aquellos que mejor representen la evolución de la misma. Este análisis debe llevarse a cabo de tres formas distintas:

a) Estudio de los gráficos de la variable macro junto con cada uno de los indicadores parciales para analizar si presentan la misma evolución.

b) Cálculo de los coeficientes de correlación de la macromagnitud con cada uno de los indicadores parciales y, las correlaciones cruzadas entre todos los indicadores parciales para evitar la selección de indicadores parciales que representen el mismo concepto.

c) Es necesario realizar también un análisis de regresión, de forma que conozcamos cuáles son los indicadores parciales que explican una mayor parte de la variación de la variable macro.

3. Una vez seleccionados, basados en los criterios anteriormente expuestos, los indicadores que van a formar parte del indicador sintético, entramos en la problemática relativa a la ponderación y agregación de los mismos en un único indicador. Ésta última se puede llevar a cabo de dos formas:

a) Con la utilización de una metodología basada en el concepto de distancia, obteniendo de esta forma un indicador de tipo cualitativo. Se define la *distancia* como la diferencia cuadrática entre la situación o estado de un indicador con respecto al mismo tomado como referencia. Para cada indicador parcial se determina un valor base de referencia, que puede ser el mayor valor, el medio, el menor, etc., y de acuerdo a este valor fijado se calcula la distancia.

b) Utilización del método Stepwise, basado en la selección automática de indicadores parciales y cuyo resultado es un indicador sintético cuantitativo. El procedimiento denominado Stepwise consiste en una selección de variables en un modelo de regresión múltiple, mezcla a su vez de otros dos métodos: selección hacia delante y eliminación hacia atrás. En el método selección hacia delante, la primera variable considerada para entrar en la ecuación es la de mayor coeficiente correlación con la variable endógena. Por otro lado, eliminación hacia atrás comienza con todas las variables seleccionadas incluidas en la ecuación y las que no cumplen un determinado criterio se eliminan, estos criterios se basan en valores de test F prefijados.

En definitiva, el método selección de Stepwise mezcla los dos procedimientos anteriores, cada vez que se introduce una nueva variable por el primer procedimiento, se comprueba por el segundo si se debe sacar alguna variable.

Finalmente los pesos de los indicadores parciales pueden ser sus coeficientes de correlación con la variable macro, o bien, se puede obtener realizando un análisis de componentes principales.

4. La siguiente etapa consiste en estudiar la congruencia entre el indicador compuesto y el indicador objetivo, con los mismos métodos que utilizamos para seleccionar los indicadores formantes del mismo. De esta forma, si el resultado ha sido satisfactorio, dispondremos de una herramienta con la que podremos realizar predicciones de la variable macro y, de esta forma anticipar su evolución futura. Para ello, basta con realizar predicciones de los indicadores parciales e introducir esos nuevos valores en el indicador sintético.

2.5. Propiedades de los indicadores sintéticos

Las propiedades usuales que se exigen a un indicador sintético son:

❖ **Existencia y determinación:** La función matemática que define el indicador debe ser tal que éste exista y no sea indeterminado para todo sistema de valores de indicadores.

❖ **Monotonicidad:** Que el indicador responda siempre positivamente (negativamente) a una modificación positiva de los componentes y negativamente (positivamente) a una modificación negativa.

❖ **Unicidad:** Que el indicador sea tal que para una situación dada genere un único valor.

❖ **Invariación:** Que el indicador no varíe respecto a un cambio de origen o de escala en las unidades en que se expresan los valores (estado) de los componentes.

❖ **Homogeneidad:** Que la función matemática que define el indicador sintético en función de los indicadores simples, sea una función homogénea de grado uno de homogeneidad.

❖ **Transitividad:** Que si por ejemplo (1), (2) y (3) son tres situaciones distintas del objeto medible por el indicador sintético, e $I(1)$, $I(2)$, $I(3)$ son los valores del indicador correspondiente a esas tres situaciones, se verifique que:

$$[I(1) > I(2) \wedge I(2) > I(3)] \Rightarrow [I(1) > I(3)]$$

❖ **Exhaustividad:** Que el indicador aproveche al máximo y de forma útil la información suministrada por los indicadores simples.

2.6. Fuentes de datos que pueden usarse para la evaluación de la pobreza

La información que se suministra en los hogares, es importante para conocer como ha ido evolucionando la pobreza a través del tiempo, además de analizar el impacto distributivo de las políticas económicas y sociales utilizadas. En este sentido es necesario conocer que cantidad de personas que viven en estado de pobreza o en que región se concentra.

Es por este motivo que en los últimos años se ha tomado experiencia en la elaboración de encuesta de hogares. Entre las más importantes se destacan las siguientes:

❖ **Encuesta de Hogares o de Empleo:** Se realizan con frecuencia una vez al año. Entre sus objetivos están:

a) Conocer las características demográficas y sociales de la población que permiten el análisis tanto en términos individuales como en los de hogares y su interrelación con las características económicas.

b) Proporcionar datos que puedan ser incorporadas en las Cuentas Nacionales.

c) Conocer la estructura general del Presupuesto de los hogares, esto es, los distintos orígenes de sus recursos, el monto de los mismos y el modo de como se lo emplea en el consumo de los diferentes bienes y servicios.

d) Estimar los niveles de vida por dominios de estudios, determinar las fuentes y niveles de ingreso y la estructura y la composición del gasto.

Se considera como unidades de análisis alternativas, el hogar y el individuo. Se comprende como hogar al conjunto de personas, emparentados o no entre sí, que residan habitualmente en una vivienda y se constituyen o agrupan para atender sus necesidades alimentarias u otras necesidades vitales.

El INEC realiza esta encuesta y tiene como finalidad obtener información sobre los ingresos y gastos de los hogares del área urbana del país, así como determinar la cuantía, origen de los ingresos y estructura de los gastos, relacionándoles con características demográficas, educacionales y ocupacionales de la población.

En resumen, el objetivo de la encuesta de hogares es determinar el nivel de empleo, desempleo y subempleo, características de la vivienda, características demográficas, educacionales, económicas de los miembros del hogar, variables de consumo, y variables de ingreso.

❖ **Encuestas de Ingresos y Gastos:** Se realiza cada cinco o diez años según los países con el objetivo de tener información real para estimar y actualizar las ponderaciones del índice de precios al consumidor.

El ingreso del hogar comprende las percepciones monetarias y en especies que recibe el hogar o los miembros del mismo, y pueden ser corrientes o no regular.

❖ **Encuesta de Condiciones de Vida:** Se focalizan en la evaluación de las condiciones en que vive la población, como: la alimentación, la salud, la educación, la disposición de servicios básicos, entre otros.

El INEC realiza estas encuestas y definen como su principal objetivo el de medir y analizar las condiciones de vida del pueblo del Ecuador, y el producir información en el ámbito de hogares y personas sobre los efectos (impactos) de las políticas de ajuste estructural y de los programas de compensación social en los niveles de bienestar de la población.

Además, tiene como objetivo facilitar la formulación de las políticas y el diseño de acciones destinadas a reducir los niveles de pobreza. En esta perspectiva, es una herramienta de apoyo a las decisiones de política social, priorización y

focalización del gasto, del monitoreo y evaluación de programas y de medición de cambios en los niveles de bienestar de la población.

❖ **Encuestas Urbana de Empleo, Desempleo y Subempleo:** Los objetivos de estas encuesta son:

a) Difundir periódicamente datos sobre aspectos sociales, demográficos y económicos de la población con énfasis en indicadores relacionados con Empleo, Subempleo y Desempleo.

b) Proporcionar información actualizada y periódica sobre la población económicamente activa (ocupados y desocupados).

Y como se dijo en el Capítulo I se utilizará la información del Censo de Población y Vivienda, por lo cual es necesario establecer lo que comprende el mismo.

❖ **Censo de Población y Vivienda:** Tiene como objetivo dotar de información sobre la magnitud, estructura, crecimiento, distribución de la población y de sus características demográficas, económicas y sociales, así como de las condiciones generales en que habita la población, que sirva de base para la formulación de políticas de población y para la elaboración de planes de

desarrollo y de programas y proyectos a cargo de los organismos públicos y privados.

2.7. Selección de indicadores de pobreza

Antes que nada se hablará de la relación que existe entre los indicadores sociales y de pobreza.

Para elaborar indicadores sociales hay que basarse en observaciones sobre un aspecto social que sea de interés. El aspecto social que se quiere tratar en los indicadores de pobreza, como su nombre lo incluye es la pobreza; para lo cual se observará los datos de diferentes variables que están involucrados en la determinación de la misma, en nuestro caso, serán tomados de un censo.

Una vez aclarado este punto, se puede definir los índices de pobreza y de bienestar como estadísticos que utilizan datos de encuestas y de censos que ayudan a analizar la evolución de la pobreza, a la vez que permiten hacer comparaciones entre diferentes grupos sociales.

Hay que señalar que los conceptos de pobreza y bienestar están relacionados, pero presentan diferentes matices. El concepto de pobreza estaría relacionado con el ingreso y el gasto de las personas, mientras el concepto de bienestar tiene

que ver más con las condiciones de vida, que brinda un nivel adecuado de satisfacción a las personas, y que no dependen tan directamente del ingreso. Así, los servicios públicos como agua potable, luz eléctrica, alcantarillado, líneas telefónicas, entre otros, brindan cierto grado de bienestar a las personas, pero dependen más de políticas de desarrollo que del ingreso de los hogares

Para establecer indicadores de pobreza hay que tomar en cuenta los siguientes axiomas que los incluye en su trabajo **Holger Capa (3)**, y que son una particularidad de los indicadores sintéticos.

❖ **Axioma de monoticidad:** Con todo lo demás constante, el estimador de pobreza se incrementa si el ingreso de una familia pobre disminuye.

❖ **Axioma de transferencia:** Con todo lo demás constante, el índice de pobreza se incrementa (respectivamente disminuye) si cualquier transferencia de ingreso se efectúa desde un hogar pobre a uno menos pobre.

❖ **Axioma de sensibilidad de transferencia:** Establece que para transferencias de una persona a otra más pobre dentro de la población pobre, la disminución de la pobreza es mayor para transferencias mayores.

Teniendo presente las especificaciones anteriores se debería pensar en variables de pobreza que estén relacionadas con el nivel de vida, la calidad de los bienes o servicios a las que un individuo o grupo de individuos tienen acceso para satisfacer sus necesidades más relevantes como son alimentación, salud, vivienda, vestuario, educación, etc. Es decir, que los indicadores más directos serán aquellos relacionados con la calidad de los bienes o servicios.

En capítulo IV se analizará más detalladamente las variables a elegir, ya que para ello debemos analizar el método multivariante de las componentes principales.

III. TÉCNICAS MULTIVARIANTES PARA LA CONSTRUCCIÓN DE INDICADORES SINTÉTICOS

3.1. Empleo de Técnicas Estadísticas Multivariantes para el Establecimiento de Indicadores Sintéticos de Pobreza

Las técnicas multivariantes pueden utilizarse de diferentes maneras, sea por la situación que se presenta, así como, por el objetivo del análisis que se quiera realizar.

Cada situación requiere una evaluación particular para utilizar el método de análisis multivariado más adecuado que permita tener la máxima información disponible del conjunto de datos, y a su vez garantice la validez de su aplicación.

En el caso particular del presente trabajo, se quiere explorar la adecuación de las técnicas multivariantes al establecimiento de indicadores sintéticos, sin importar la naturaleza de los datos básicos disponibles. Esta exploración no comenzará con un conjunto de datos a su máximo nivel de desagregación, como sería el individuo, el hogar o el perceptor de ingresos. Ello es así porque no se conoce acerca de los individuos. Para solucionar este problema se trabajará con masas de poblaciones lo que permitirá una buena base experimental.

Está comprobado que el uso de ciertas técnicas es satisfactorio, en el sentido de que las validaciones a que son sometidas a cierto contexto son positivas, queda entonces abierto el camino para su confiable aplicación en otros entornos, con otras condiciones y a otros niveles de agregación que los ensayados en las primeras etapas exploratorias. Estudio realizado por **Fernando Medina (9)**.

La naturaleza multidimensional de los fenómenos sociales en general y de las situaciones y condiciones de pobreza en particular, reclama un tratamiento estadístico que reconozca y respete dichas características. Las técnicas multivariantes logran esto en forma adecuada. Dentro de ellas, la de Componentes Principales, es en principio apropiada para abstraer y revelar estructuras multidimensionales latentes con ciertas propiedades deseadas. Bajo ciertas condiciones dichas construcciones pueden ser directamente empleadas como indicadores sintéticos del fenómeno o de los fenómenos que el conjunto

de variables empleadas busca representar. Un uso juicioso de los resultados que pueden extraerse de su aplicación, brinda la oportunidad para establecer indicadores alternativos de amplia cobertura y capacidad sinterizadora.

A continuación se presentará la parte más importante de la teoría del Análisis de Componentes Principales, que se necesita conocer para la construcción del indicador sintético de pobreza.

3.2. Análisis de Componentes Principales.

El método de Análisis de Componentes Principales (**ACP**), permite la estructuración de un conjunto de datos multivariantes obtenidos de una población, cuya distribución no es necesariamente conocida.

Lo que se propone es hallar una combinación lineal de las variables representativas del fenómeno multidimensional.

3.2.1. Objetivos del Análisis de Componentes Principales

Los objetivos del análisis por el método de componentes principales son:

- ❖ Generar nuevas variables que puedan expresar la información contenida en el conjunto original de datos.

- ❖ Reducir la dimensionalidad del problema que se está estudiando, como paso previo para futuros análisis.

- ❖ Eliminar, cuando sea posible, algunas de las variables originales si ellas aportan poca información.

Las nuevas variables originadas se conocen como *componentes principales*.

3.2.2. Características deseables de las Componentes Principales

- ❖ Los componentes principales no están correlacionadas y si, además, puede suponerse multinormalidad en los datos originales, son independientes.

- ❖ Cada componente principal sintetiza la máxima variabilidad residual contenida en los datos.

3.2.3. Generación de las Componentes Principales

Se estudiará un conjunto de n individuos mediante p -variables, encontrando nuevas variables que se denominan $Y(j)$, $j=1, \dots, p$ que son combinaciones lineales de las variables originales $X(i)$, las cuales deben ser centradas y reducidas, además, que cumplan las características anteriores.

Es decir, se encontrarán $(p \times p)$ variables tales que:

$$Y(j) = \sum_{i=1}^p C(ij)X(i) \quad , \quad j = 1, \dots, p \quad (3.1)$$

donde $C(ij)$ son las constantes de la combinación lineal. Este valor indicará el grado de contribución a la variable definida por la transformación. Cabe resaltar el hecho que en cada variable $Y(j)$ se está incluyendo a todas las variables originales $X(i)$.

Para que se cumpla la condición de que las componentes principales sean **no correlacionadas** es necesario mostrar que:

$$E\langle Y(j), Y(l) \rangle = 0 \quad j, l = 1, \dots, p \quad j \neq l \quad (3.2)$$

para lo que se supone que las variables son centradas, es decir:

$$E\langle X(i) \rangle = 0, \quad i = 1, \dots, p$$

Remplazando la función de la nueva variable -ecuación (3.1) - en la ecuación (3.2) obtenemos:

$$E\left\langle \left(\sum_{i=1}^p C(ij)X(i) \right) \cdot \left(\sum_{k=1}^p C(kl)X(k) \right) \right\rangle = 0 \quad (3.3)$$

Dado que $C(ij)$ y $C(kl)$ son constantes, se puede intercambiar operadores lo que dará una sumatoria doble.

$$E\langle Y(j).Y(l) \rangle = \sum_{i=1}^p \sum_{k=1}^p C(ij)C(kl) E\langle X(i).X(k) \rangle \quad (3.4)$$

Lo que muestra es que $E\langle X(i).X(k) \rangle$ es la covarianza entre las variables originales $X(i)$ y $X(j)$.

La condición de $j \neq l$ muestra la condición que no exista correlación dos a dos entre componentes principales, por lo cual habrá $(1/2)p(p-1)$ restricciones sobre las constantes $C(ij)$ para que el sistema tenga solución única.

Además; por la condición de homogeneidad que establece que las componentes principales son perpendiculares entre sí, se tiene las siguientes restricciones:

$$\begin{aligned} \sum_{i=1}^p C(ij)C(il) &= 0 & j \neq l & \quad j, l=1, \dots, p \\ \sum_{i=1}^p C(ij)C(il) &= 1 & j = l & \end{aligned} \quad (3.5)$$

Expresando en forma matricial las constantes $C(ik)$, cuya matriz tendrá dimensión $(p \times p)$, quedaría así:

$$C = \begin{bmatrix} C(11) & \cdot & \cdot & \cdot & \cdot & C(1p) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & C(ij) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ C(p1) & \cdot & \cdot & \cdot & \cdot & C(pp) \end{bmatrix} \quad (3.6)$$

Para satisfacer la condición de ortogonalidad, C es una matriz ortogonal, que cumple la característica propia de dicha matriz, es decir:

$$CC' = C'C = C^{-1}C = I \quad (3.7)$$

Expresando las componentes principales $Y(j)$ -ecuación (3.1)- en forma matricial, se obtiene:

$$Y = X C \quad (3.8)$$

$(n \times p) \quad (n \times p) \quad (p \times p)$

De aquí en adelante se trabajará en forma matricial, debido a que es más sencilla la explicación del cálculo de las componentes principales.

Si se calcula el valor esperado entre las componentes principales -ecuación (3.4)-, esta vez utilizando la forma matricial, se tendrá:

$$E\langle Y'Y \rangle = E\langle (XC)'(XC) \rangle$$

Aplicando propiedades de valor esperado, tratando a C como matriz de constante se tiene

$$= C' E\langle X'X \rangle C \quad (3.9)$$

Conociendo que $E\langle X'X \rangle$ es la matriz de Covarianza de los datos originales, la cual notaremos por S , queda

$$E\langle Y'Y \rangle = C' S C = L \quad (3.10)$$

donde L será una matriz diagonal, cuyos elementos en la diagonal son las varianzas de las nuevas variables (componentes principales).

Multiplicando ambos miembros por C y tomando en cuenta que $C'C = I$ - ecuación (3.7) -, obtenemos:

$$CL = SC \quad (3.11)$$

En forma matricial esta ecuación quedaría expresada de la siguiente manera:

$$\begin{bmatrix} \alpha_{(1)1} & \dots & \alpha_{(1)p} \\ \dots & \dots & \dots \\ \dots & \alpha_{ij} & \dots \\ \dots & \dots & \dots \\ \alpha_{(p)1} & \dots & \alpha_{(p)p} \end{bmatrix} \begin{bmatrix} \lambda(1) & 0 & 0 & 0 & 0 \\ 0 & \lambda(2) & 0 & 0 & 0 \\ 0 & 0 & \lambda(j) & 0 & 0 \\ 0 & 0 & 0 & \lambda(p-1) & 0 \\ 0 & 0 & 0 & 0 & \lambda(p) \end{bmatrix} \begin{bmatrix} S(1) & \dots & S(1p) \\ \dots & \dots & \dots \\ \dots & S(ii) & \dots \\ \dots & \dots & \dots \\ S(p) & \dots & S(p) \end{bmatrix} \begin{bmatrix} \alpha_{(1)1} & \dots & \alpha_{(1)p} \\ \dots & \dots & \dots \\ \dots & \alpha_{ij} & \dots \\ \dots & \dots & \dots \\ \alpha_{(p)1} & \dots & \alpha_{(p)p} \end{bmatrix} \quad (3.12)$$

Realizando la debida multiplicación, se tiene el siguiente sistema de ecuaciones:

$$\text{Primer término: } C(11) \lambda(1) = S(11) C(11) + S(12) C(21) + \dots + S(1p)C(p1)$$

$$\text{Segundo término: } C(21) \lambda(1) = S(21) C(11) + S(22) C(21) + \dots + S(2p)C(p1).$$

$$P\text{-ésimo término: } C(p1) \lambda(1) = S(p1) C(11) + S(p2) C(21) + \dots + S(pp)C(p1).$$

Igualando las ecuaciones a cero y agrupando los términos de $C(jk)$, se tiene un nuevo sistema de ecuaciones, que es:

$$0 = [S(11) - \lambda(1)] C(11) + S(12) C(21) + \dots + S(1p)C(p1)$$

$$0 = S(21) C(11) + [S(22) \lambda(1)] C(21) + \dots + S(2p)C(p1)$$

.

.

$$0 = S(p1) C(11) + \dots + [S(pp) - \lambda(1)]C(p1)$$

Expresado en forma matricial:

$$\begin{bmatrix} S(11) & \cdot & \cdot & \cdot & S(1p) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & S(ii) & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ S(p1) & \cdot & \cdot & \cdot & S(pp) \end{bmatrix} \begin{bmatrix} \lambda(1) & 0 & 0 & 0 & 0 \\ 0 & \lambda(1) & 0 & 0 & 0 \\ 0 & 0 & \lambda(1) & 0 & 0 \\ 0 & 0 & 0 & \lambda(1) & 0 \\ 0 & 0 & 0 & 0 & \lambda(1) \end{bmatrix} \begin{bmatrix} C(11) \\ C(21) \\ \cdot \\ \cdot \\ C(p1) \end{bmatrix} = \mathbf{0} \quad (3.13)$$

En estas ecuaciones sólo intervino el primer elemento de la diagonal de L , si se sigue multiplicando se obtendrán ecuaciones idénticas, lo que muestra que las p soluciones se encuentran al igualar a cero el determinante de la matriz de covarianza menos λI

$$|S - \lambda I| = 0 \quad (3.14)$$

que se conoce como *polinomio característico* de la matriz S . Debido a que existen $p+1$ incógnitas y p ecuaciones, se aplicará la condición de ortonormalidad de los vectores propios para tener solución única. Es decir, que

$$\text{debe cumplirse que: } \sum_{i=1}^p C(i)^2 = 1.$$

El resto de las ecuaciones se hallarán a partir de la condición de normalización,

$$\text{ya que deberá cumplirse que: } \sum_{i=1}^p C(i)^2 = 1, \text{ para todo } j.$$

Estos p vectores son los que se conocen como *vectores propios* que forman la matriz C , que son las constantes de la combinación lineal hallada de $X(i)$.

Ahora mostremos la condición de *máxima varianza*. Se modificará la nueva variable -ecuación (3.1)- quitándole el subíndice, y luego se hallará la varianza de la misma.

$$Y = \sum_{i=1}^p C(i)X(i) \quad (3.15)$$

$$\text{VAR} (Y) = \text{VAR} \left[\sum_{i=1}^p C(i) X(i) \right] \quad (3.16)$$

Desarrollando los corchetes, utilizando la fórmula de la varianza y recordando que las covarianzas $S(ik)$ y $S(ki)$, se tiene la siguiente expresión:

$$\text{VAR} (Y) = \sum_{j=1}^p C(i)^2 S(ii) + 2 \sum_{\substack{i=1 \\ k=2}}^p C(i)C(k)S(ik) \quad i < k \quad (3.17)$$

Si se considera el segundo término para valores de i y k diferentes, la ecuación puede escribirse de otra manera

$$\text{VAR}(Y) = \sum_{j=1}^p C(i)^2 S(ii) + \sum_{\substack{i=1 \\ k=1}}^p C(i)C(k)S(ik) \quad i \neq k \quad (3.18)$$

Como i y k contienen todas las combinaciones posibles de subíndices, entonces la ecuación quedaría:

$$\text{VAR} (Y) = \sum_{i=1}^p \sum_{k=1}^p C(i)C(k)S(ik) \quad i \neq k \quad (3.19)$$

Esta expresión muestra que para hallar la varianza de la componente principal, intervienen las varianzas y covarianzas de las variables originales $S(ik)$. Las

constantes $C(i)$ y $C(k)$ son los p valores asociados con la nueva variable, es decir, son los valores del vector propio o una columna de la matriz C .

Se procede a calcular el máximo valor de la $VAR(Y)$, tomando en cuenta otra vez la condición de homogeneidad -ecuación (3.5)- con la debida restricción que produce, la cual sea expresada de la forma

$$\sum_{i=k=1}^p C(i)C(k) = 1 \quad (3.20)$$

Si aplicamos el método de Maximización de Lagrange; es decir, la función varianza menos g veces la restricción planteada y seguimos el proceso de maximización, se tendrá la siguiente expresión:

$$\frac{\partial \left[\sum_{i=1}^p \sum_{k=1}^p C(i)C(k)S(ik) - g \left(\sum_{i=1}^p C(i)^2 - 1 \right) \right]}{\partial (C_i)} = 0 \quad (3.21)$$

Derivamos la ecuación para los p posibles valores de $C(i)$.

Particularizando la expresión (3.21), suponiendo $i=1$, derivando se obtendrá:

$$\sum_{k=1}^p C(k)S(1k) + \sum_{k=1}^p C(k)S(k1) - g(2C(1)) = 0 \quad (3.22)$$

Teniendo en cuenta la propiedad de simetría de la matriz de covarianza, puede observarse que:

$$2\sum_{k=1}^p C(k)S(1k) - 2gC(1) = 0 \quad (3.23)$$

lo cual permite eliminar el número 2 en ambos miembros.

Ahora, si se deriva la expresión original con respecto a $C(2) \dots C(p)$, se podrán encontrar ecuaciones similares a la última, que forman un sistema de ecuaciones en que la incógnita es el valor g . Este sistema será:

$$\sum_{k=1}^p C(k)S(1k) - gC(1) = 0$$

$$\sum_{k=1}^p C(k)S(2k) - gC(2) = 0 \quad (3.24)$$

$$\vdots \quad \vdots \quad \ddots \quad = \quad \vdots$$

$$\sum_{k=1}^p C(k)S(pk) - gC(p) = 0$$

Igualando el primer término con el segundo, se tiene:

$$\sum_{k=1}^p C(k)S(1k) = gC(1)$$

$$\sum_{k=1}^p C(k)S(2k) = gC(2) \quad (3.25)$$

$$\vdots \quad = \quad \vdots$$

$$\sum_{k=1}^p C(k)S(pk) = gC(p)$$

Cada ecuación es el producto de dos vectores, los cuales pueden expresarse en forma matricial. Estos vectores para la primera ecuación serían:

$$\langle C(1) \ C(2) \ \dots \ C(p) \rangle \begin{bmatrix} S(11) \\ S(21) \\ \cdot \\ \cdot \\ S(p1) \end{bmatrix} = g \ C(1) \quad (3.26)$$

Colocando los vectores de cada ecuación en forma matricial se tiene:

$$\langle C(1) \ C(2) \ \dots \ C(p) \rangle \begin{bmatrix} S(11) & S(12) & \cdot & \cdot & S(1p) \\ S(21) & S(22) & \cdot & \cdot & S(2p) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ S(p1) & \cdot & \cdot & \cdot & S(pp) \end{bmatrix} = g \langle C(1) \ C(2) \ \dots \ C(p) \rangle$$

Sin restar generalidad, es posible expresar el segundo término de la ecuación en forma matricial y obtener:

$$\langle C(1) \ C(2) \ \dots \ C(p) \rangle \begin{bmatrix} S(11) & S(12) & \cdot & \cdot & S(1p) \\ S(21) & S(22) & \cdot & \cdot & S(2p) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ S(p1) & \cdot & \cdot & \cdot & S(pp) \end{bmatrix} = \langle C(1) \ C(2) \ \dots \ C(p) \rangle \begin{bmatrix} g & 0 & \cdot & \cdot & 0 \\ 0 & g & \cdot & \cdot & \cdot \\ \cdot & \cdot & g & \cdot & \cdot \\ \cdot & \cdot & \cdot & g & \cdot \\ 0 & \cdot & \cdot & \cdot & g \end{bmatrix}$$

Igualando a cero y sacando el vector de constantes como factor común, se obtendrá:

$$\langle C(1) \ C(2) \ \dots \ C(p) \rangle \begin{bmatrix} \mathbf{S}(11) & \mathbf{S}(12) & \cdot & \cdot & \mathbf{S}(1p) \\ \mathbf{S}(21) & \mathbf{S}(22) & \cdot & \cdot & \mathbf{S}(2p) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{S}(p1) & \cdot & \cdot & \cdot & \mathbf{S}(pp) \end{bmatrix} \begin{bmatrix} g & 0 & \cdot & \cdot & 0 \\ 0 & g & \cdot & \cdot & \cdot \\ \cdot & \cdot & g & \cdot & \cdot \\ \cdot & \cdot & \cdot & g & \cdot \\ 0 & \cdot & \cdot & \cdot & g \end{bmatrix} = \mathbf{0} \quad (3.27)$$

$(1xp) \qquad \qquad \qquad (pxp) \qquad \qquad \qquad (pxp) \qquad \qquad \qquad (1xp)$

En este sistema se tiene, en el primer término el vector de las constantes para el cual existen p incógnitas, una matriz conocida ya que es la matriz de varianzas-covarianzas de las variables originales, y la matriz con g como incógnita. Existen p ecuaciones con $(p+1)$ incógnitas, es decir, que para que el sistema tenga solución única debe aumentarse una ecuación más. Esta es la condición que se les puso a las $C(i)$.

Si se transpone el sistema, para hacer más fáciles las operaciones, se obtiene:

$$\begin{array}{c} \left[\begin{array}{cccc} S(11) & S(12) & \cdot & \cdot \\ S(21) & S(22) & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ S(p1) & \cdot & \cdot & \cdot \end{array} \right] \left[\begin{array}{c} g \\ 0 \\ \cdot \\ \cdot \\ 0 \end{array} \right] - \left[\begin{array}{cccc} 0 & g & \cdot & \cdot \\ \cdot & \cdot & g & \cdot \\ \cdot & \cdot & \cdot & g \\ 0 & \cdot & \cdot & g \end{array} \right] \left[\begin{array}{c} C(1) \\ C(2) \\ \cdot \\ \cdot \\ C(p) \end{array} \right] = 0 \end{array} \quad (3.28)$$

$(p \times p)$
 $(p \times p)$
 $(p \times 1)$
 $(p \times 1)$

Si comparamos este sistema -ecuación (3.27)- con el último sistema obtenido para mostrar la no-correlación entre componentes principales -ecuación (3.13)-, se puede notar que son idénticos sólo que en ese sistema las incógnitas eran $\lambda(I)$ y ahora son g . Dichas ecuaciones cumplen la condición de ser transformación lineal para que las variables no esten correlacionadas. Ya se mostró que dichos sistemas tienen solución única y que son valores propios generados de la matriz S , los cuales originan vectores propios que son los valores de $C(i)$.

La transformación lineal que sintetiza la máxima variabilidad, corresponderá a la generada por el valor de $\lambda(i)$ que sea mayor. A esta varianza máxima se ha denominado $\lambda(1)$, tal que se cumpla:

$$\lambda(1) > \lambda(2) \dots > \lambda(p) \quad (3.29)$$

Así, la primera componente principal sintetiza la máxima variabilidad posible en los datos originales. La segunda componente principal, sintetiza la variabilidad residual, sujeta a la condición de no-correlación con el primer componente principal, y así hasta el p-ésimo componente principal.

En resumen, la nueva expresión de los datos representada en las componentes principales cumplen las siguientes propiedades:

- a. $E\langle Y(j) \rangle = E\langle X \rangle l(j)$, donde $l(k)$ es el j -ésimo vector propio
- b. $Var(Y) = \lambda(j)$, donde $\lambda(j)$ es el j -ésimo vector propio
- c. $Cov\langle Y(j), Y(l) \rangle = 0$, para $j \neq l$
- d. $Var[Y(1)] \geq Var[Y(2)] \geq \dots \geq Var[Y(p)] \geq 0$
- e. $\sum_{j=1}^p Var\langle Y(j) \rangle = tr S$
- f. $\prod_{j=1}^p Var\langle Y(j) \rangle = |S|$

3.2.4. Selección del número de componentes principales.

Entre las propiedades que deben cumplir las nuevas variables se encuentra que la suma de las varianzas de las variables originales es igual a la suma de los valores propios, y a su vez cada valor propio es la varianza de cada componente principal, para lo cual debe cumplirse que:

$$\sum_{i=1}^p s(ii) = \sum_{j=1}^p \lambda(j) \quad i, j=1, \dots, p$$

Cada componente principal explicará una proporción de la varianza total, lo cual se conoce como la *importancia relativa de cada componente* en el análisis o proporción de varianza total explicada por el componente jésimo, y se calcula como sigue:

$$\text{Varianza Total Explicada} = \frac{\lambda(j)}{\text{Trasa } S} \quad (3.30)$$

Como los valores propios se ordenan de forma decreciente, es posible seleccionar los primeros q valores propios ($q < p$) y la eficiencia del ajuste de los datos originales por los q nuevos componentes principales estará dada por la proporción de la varianza total explicada por la suma de los q primeros valores propios.

$$\text{Porcentaje de Varianza Total Explicada} = \frac{\sum_{j=1}^q \lambda(j)}{\text{Trasa } S} * 100 \quad (3.31)$$

La varianza explicada por cada componente o por las primeras q primeras componentes, puede ser expresada en función de la suma de los valores propios, considerando las propiedades b y e vistas en la sección anterior. Así se tiene:

$$\text{Porcentaje de Varianza Explicada por el } j\text{-ésimo componente} = \frac{\lambda(j)}{\sum_{j=1}^p \lambda(j)} * 100 \quad (3.32)$$

$$\text{Porcentaje de Varianza Explicada por los } q \text{ primeros componentes} = \frac{\sum_{j=1}^q \lambda(j)}{\sum_{j=1}^p \lambda(j)} * 100 \quad (3.33)$$

Cuando se toman $q=p$ componentes principales la proporción de la varianza explicada es 1 y en porcentaje 100%.

No existe la manera de determinar cuantos componentes principales son necesarios para alcanzar un porcentaje de varianza explicada satisfactorio, pero se han establecido criterios útiles que han sido validados para la determinación del número de componentes **Abascal (1)** y **Pla (10)**.

Entre estos criterios para seleccionar el número de componentes se encuentran:

1. Elaboración de un gráfico de los porcentajes de variación explicada por cada componente en las ordenadas y los componentes (correspondientes a cada valor propio) en orden decreciente en las abscisas, se pueden eliminar los ejes cuyo número de orden es posterior al “codo” que se produce en la curva.
2. Fijar un porcentaje mínimo de varianza explicada que se quiere conservar y retener el número de ejes necesario para ello.
3. Incluir aquellas componentes principales cuyos valores propios sean superiores al promedio. Si se utiliza la matriz R , se incluirá las componentes cuyos valores propios sean mayores a I .
4. Si la nube inicial de variables no presenta una dirección privilegiada, los valores propios serán próximos. Se conservará entonces un eje cuyo porcentaje de varianza sea superior a $\frac{I}{p} * 100$ que es el valor que le correspondería si explicasen la misma cantidad de información.

3.2.5. Correlación entre variables originales y componentes principales

La correlación entre dos variables se calcula como sigue:

$$r = \frac{cov(x,y)}{var(x)var(y)} \quad (3.34)$$

El cuadrado de (3.34), es lo que conocemos como *coeficientes de determinación* y constituye una medida de la asociación entre las dos variables.

Para el estudio de la correlación entre las variables originales y las componentes principales habrá que calcular las correlaciones de cada variable original con cada nueva variable.

Teniendo en cuenta la transformación (3.8), se puede encontrar una expresión del valor esperado de los vectores, lo cual quedaría así:

$$E \langle X, Y' \rangle = E \langle X, X' C \rangle$$

Considerando la ecuación (3.11) y recordando que $E \langle X, X' \rangle$ es la matriz de Covarianza de los datos originales Σ y cuyo estimador es S , tenemos:

$$E \langle X, Y' \rangle = E \langle X, X' \rangle C = \Sigma C = CC' \Sigma C = CL$$

En la matriz C en la posición (i) se encuentra el i -ésimo elemento del vector propio, mientras en L –matriz diagonal– los valores propios de la matriz S . Por tanto en la matriz CL en la posición (ij) estará ubicada la covarianza entre $X(i)$ e $Y(j)$, es decir, que:

$$Cov\langle X(i), Y(j) \rangle = C(ij) \lambda(j) \quad i, j = 1, \dots, p \quad (3.35)$$

Para calcular la correlación deberá dividirse por la raíz de las varianzas de las variables por los valores propios.

$$r(ij) = \frac{C(ij) \lambda(j)}{\sqrt{S(ii) \lambda(j)}} = C(ij) \left\langle \frac{\lambda(j)}{S(ii)} \right\rangle^{1/2} \quad (3.36)$$

Si se utiliza la matriz de correlación R de los datos originales, las varianzas serán unitarias y se tendría:

$$r(ij) = C(ij) \langle \lambda(j) \rangle^{1/2} \quad (3.37)$$

Si se suman en j , es decir, se suman las proporciones de la varianza explicada por los p componentes principales para la variable original $X(i)$, se obtendrá el valor 1. Puede efectuarse la sumatoria para los primeros q primeros componentes seleccionados para el análisis del conjunto de datos y determinar cual es la proporción de la varianza de cada variable original considerada en el nuevo subconjunto. Esto es posible ya que los componentes no están

correlacionados entre sí; de lo contrario, no se aplicaría lo anterior por no tomar en cuenta las covarianzas. Así en términos de matriz S será:

$$r^2(ij) = \frac{\lambda(j) C^2(ij)}{S(ii)} \quad i,j=1,\dots,p \quad (3.38)$$

para el j -ésimo componente, y sumando para los q primeros :

$$\sum_{j=1}^q r^2(i, j) = \frac{1}{S(ii)} \sum_{j=1}^q \lambda(j) C^2(i, j)$$

Analizando estas proporciones y los elementos del vector propio podrán inferirse las conclusiones necesarias para explicar los datos.

IV. APLICACIÓN DE COMPONENTES PRINCIPALES PARA LA CONSTRUCCIÓN DE INDICADORES SINTÉTICOS

4. 1. Definición de Observaciones

La base de datos usada para el desarrollo del trabajo fue el Resumen de las Principales Variables y Categorías que obtuvo el Instituto de Estadísticas y Censos de Ecuador (INEC), para el V Censo de Población y IV Censo de Vivienda realizado en el año de 1.990.

Se consideraron las 21 provincias existentes hasta el momento censal y se incluye las zonas no delimitadas. Es decir que existirán 22 observaciones, las cuales estarán ordenadas alfabéticamente, como se muestra a continuación:

1. Azuay

2. Bolivar
3. Cañar
4. Carchi
5. Cotopaxi
6. Chimborazo
7. El Oro
8. Esmeraldas
9. Guayas
10. Imbabura
11. Loja
12. Los Rios
13. Manabí
14. Morona Santiago
15. Napo
16. Pastaza
17. Pichincha
18. Tungurahua
19. Zamora Chinchipe
20. Galapagos
21. Sucumbios
22. Zonas No Delimitadas

4. 2. Definición de Variables

La información de la base de datos obtenida ha sido sintetizada por medio de 15 variables, las cuales se muestran a continuación con su respectiva codificación. En el **APÉNDICE A** podrá encontrarse definiciones sobre las variables censales.

X_1 Proporción de viviendas por provincias.

X_2 Indicador de tipo de vivienda.

X_3 Indicador de abastecimiento de agua.

X_4 Indicador del sistema de eliminación de aguas servidas.

X_5 Proporción de las viviendas que disponen de servicio eléctrico del total de viviendas provinciales.

X_6 Proporción de las viviendas que disponen de servicio telefónico del total de viviendas provinciales.

X_7 Indicador del sistema de eliminación de basura.

X_8 Indicador de disponibilidad de servicio higiénico.

X_9 Indicador de disponibilidad de ducha.

X_{10} Indicador de tenencia de vivienda.

X_{11} Proporción de la población por provincias

X_{12} Proporción de la población rural por provincias.

X_{13} Indicador nivel de instrucción.

X_{14} Proporción de la población analfabeta de la población total provincial.

X_{15} Proporción de la población económicamente activa de la población total provincial.

En el **APÉNDICE B**, se puede apreciar las tablas de datos de cada variable.

Las variables expresadas en proporciones se deben a que no hay homogeneidad entre provincias respecto al tamaño, lo cual de no ser tomado en cuenta podría afectar el análisis.

4. 3. Escala de Likert

Indicadores porcentuales se utilizarán como variables, cuando existan varias subcategorías dentro de una categoría.

Para calcular estos indicadores porcentuales, se utilizó Escalas de Likert, cuya utilización permite ponderar los datos de varias opciones en una sola variable.

Los pasos básicos en el desarrollo de una Escala de Likert o Escala Acumulada son los siguientes:

1. Definir el centro del problema.- Como en todos los métodos de ajuste, el primer paso necesario es definir que se medirá. Debido a que este es un método

unidimensional de ajuste, se supone que el concepto que se mide es de una dimensión en la naturaleza. La definición depende de lo que se va a crear o generar del conjunto inicial de ítems candidatos en la escala.

2. Generador de ítems. - Luego, se debe crear el conjunto de ítems en la escala potencial. Estos deben ser ítems que puedan clasificarse en una escala de juicio. Algunas veces pueden crearse los ítems basados en un conocimiento íntimo de la materia que se está tratando.

3. Clasificar los ítems. - El próximo paso es tener un grupo de juicio para clasificar los ítems, cuyo uso define los ítems favorables y desfavorables respecto a la construcción de interés.

4. Seleccionar los ítems. - Una vez que se tiene el grupo de juicio, se crea una nueva variable que es la suma de todas de los ítems individuales por su respectiva escala de juicio.

5. Administrar la escala. - Ahora se puede usar la escala de Likert. Cada respuesta (observación) puesta en juicio, debe ser clasificada dentro de la escala.

A continuación se presenta un ejemplo con el tipo de vivienda.

Indicador de tipo de vivienda

El problema que existe es que la categoría tipo de vivienda se encuentra dividida en ocho subcategorías, las cuales constituyen los ítems que serán medidos por medio de la escala de juicio.

Dichas subcategorías tendrán la escala de juicio como sigue:

1. Otro
2. Choza
3. Covacha
4. Rancho
5. Mediagua
6. Cuarto en casa de inquilinato
7. Departamento
8. Casa o villa

Observando que la mejor subcategoría es casa o villa hasta llegar que la peor es otro (ver **APÉNDICE A**), se designó una escala de juicio a cada subcategoría; así, casa o villa tendrá ocho, departamento siete, hasta llegar a otro que tiene uno.

Una vez designado la escala de juicio, se procede a crear la nueva variable Tipo de Vivienda, para lo cual multiplicamos la escala de juicio por la proporción de viviendas que corresponden a dicha subcategoría.

A continuación se presentan los cálculos para la construcción del indicador en el ámbito nacional.

Escala de juicio	8		7		6		5	
	Casa	%prov	Depart	%prov	Alquiler	%prov	Mediagua	%prov
Nacional	1'520.465	64,997	182.393	7,80	159.637	6,8242	267.818	11,449

4		3		2		1		
Rancho	%prov	Covacha	%prov	Choza	%prov	Otro	%prov	Indicador
157.235	6,7215	9.194	0,393	37117	1,58668	1.692	0,0723	7,04055

El Indicador porcentual de tipo de vivienda se obtendrá como sigue:

$$\text{Indicador} = \frac{(64,997*8) + (7,80*7) + (6,8242*6) + (11,449*5) + (6,7215*4) + (0,393*3) + (1,58668*2) + (0,0723*1)}{100} = 7,04055$$

Si utilizamos la escala de juicio propuesta, el indicador nos muestra que en el ámbito nacional el tipo de vivienda se sitúa entre las clases departamento y casa.

Este mismo cálculo se realizará para cada provincia; así mismo para cada variable en la cual existan subcategorías.

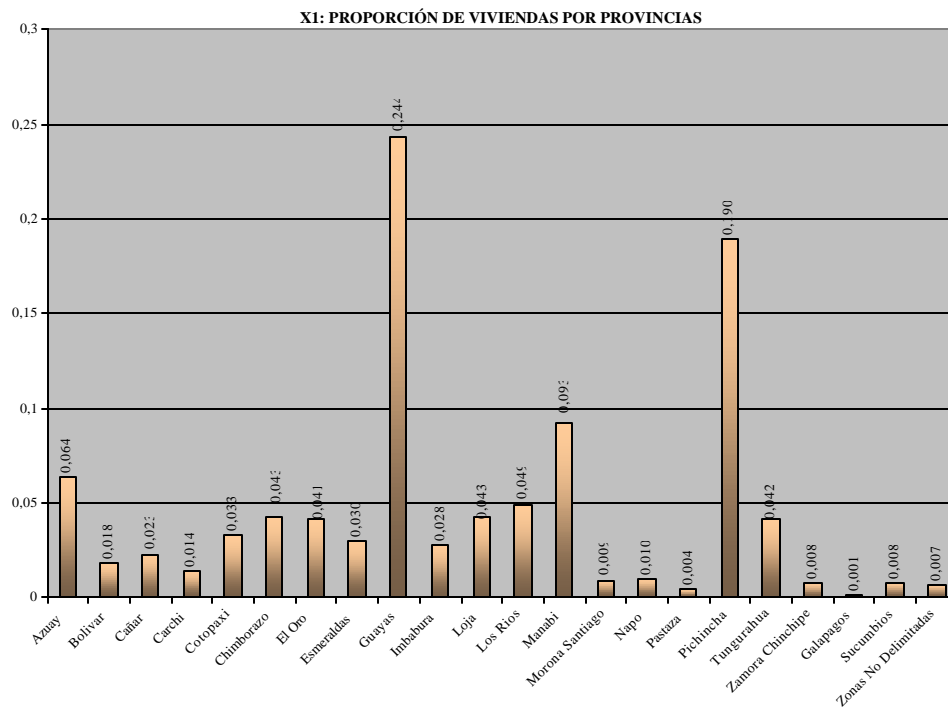
Las tablas de cada variable –indicadores y proporciones- se muestran en el **APÉNDICE B.**

4. 4. Estadística descriptiva de las variables

Una vez hechos los cálculos para las diferentes variables, los datos obtenidos se los recopilan en una tabla general de datos, la cual se muestra en el **APÉNDICE C.**

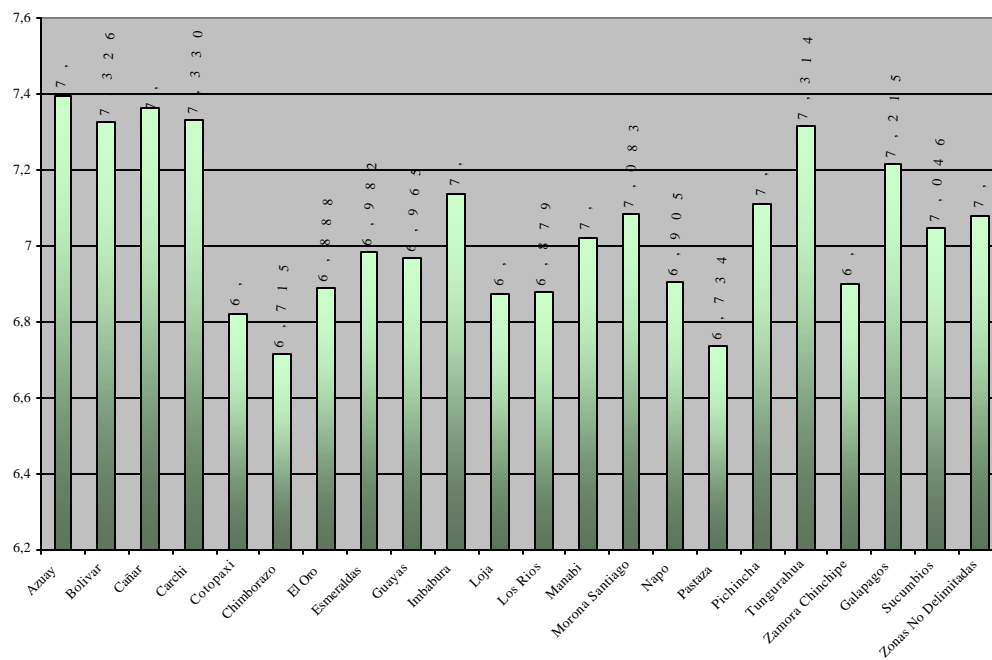
A partir de dichos datos, se calcula la estadística descriptiva; así como también, se grafican las variables. A continuación se muestran los resultados obtenidos.

FIGURA 4.1



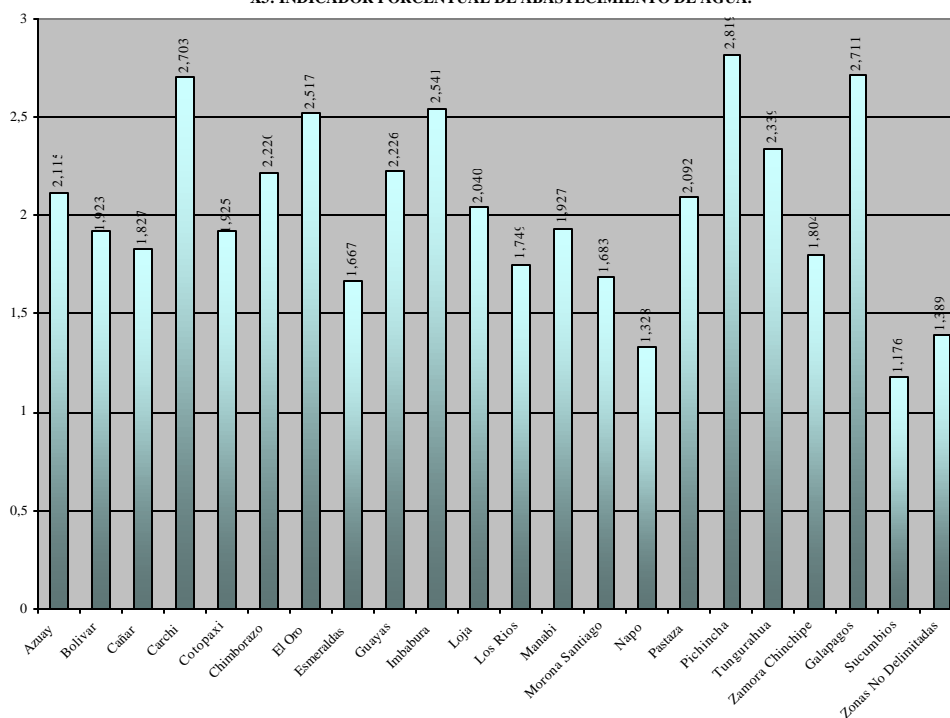
Considerando las viviendas que existen en cada provincia, se observa que encabeza la lista la provincia del Guayas con un 24,4 % del total de viviendas existentes en el país, seguida por Pichincha con un 19% y Manabí con un 9,03%.

FIGURA 4.2
X2: INDICADOR PORCENTUAL DE TIPO DE VIVIENDA



Con respecto al indicador tipo de vivienda, las provincias que tienen mejor ubicación dentro de la escala son: Azuay con un valor de 7,398, seguido por Cañar con un 7,362 y Bolívar con 7,326.

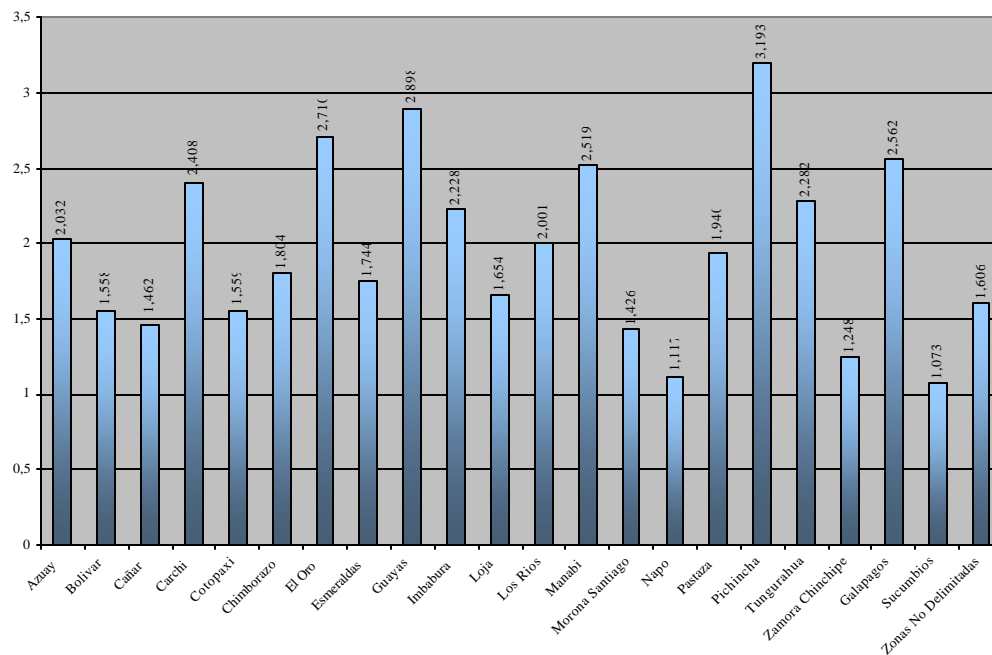
FIGURA 4.3
X3: INDICADOR PORCENTUAL DE ABASTECIMIENTO DE AGUA.



Dentro del indicador de abastecimiento de agua, sobresalen Pichincha con un valor de 2,819 seguido por Galapagos con 2,711 y Cañar con 2,103. Sucumbios es la provincia que cuenta con el indicador más bajo en el sistema de abastecimiento de agua.

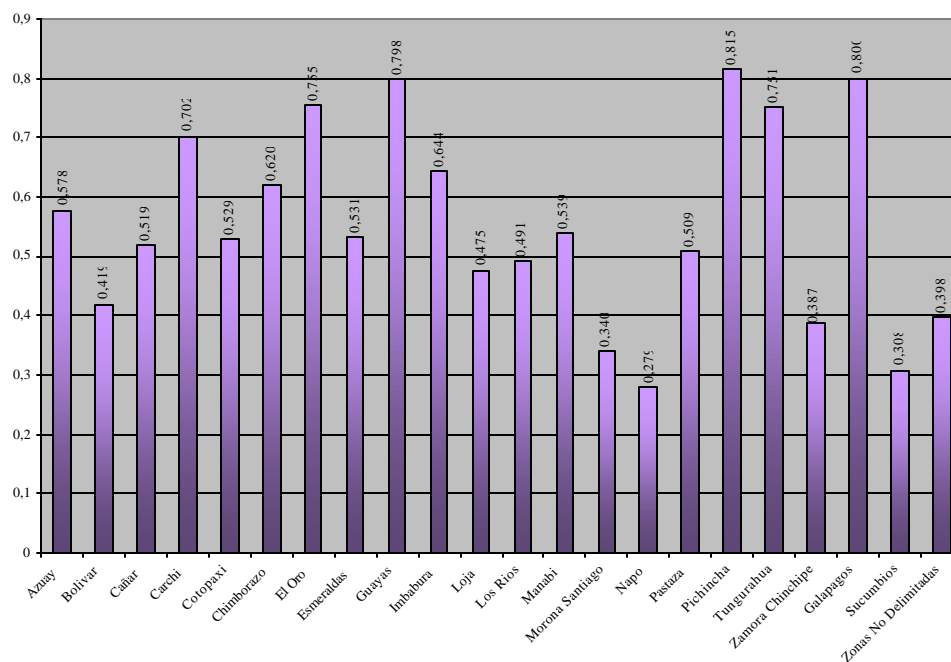
FIGURA 4.4

X4: INDICADOR PORCENTUAL DEL SISTEMA DE ELIMINACIÓN DE AGUAS SERVIDAS.



En la categoría de sistema de eliminación de aguas servidas, la provincia que mejor indicador dentro de la escala tiene es Pichincha con un 3,193 seguido por Guayas con 2,898 y El Oro con 2.71

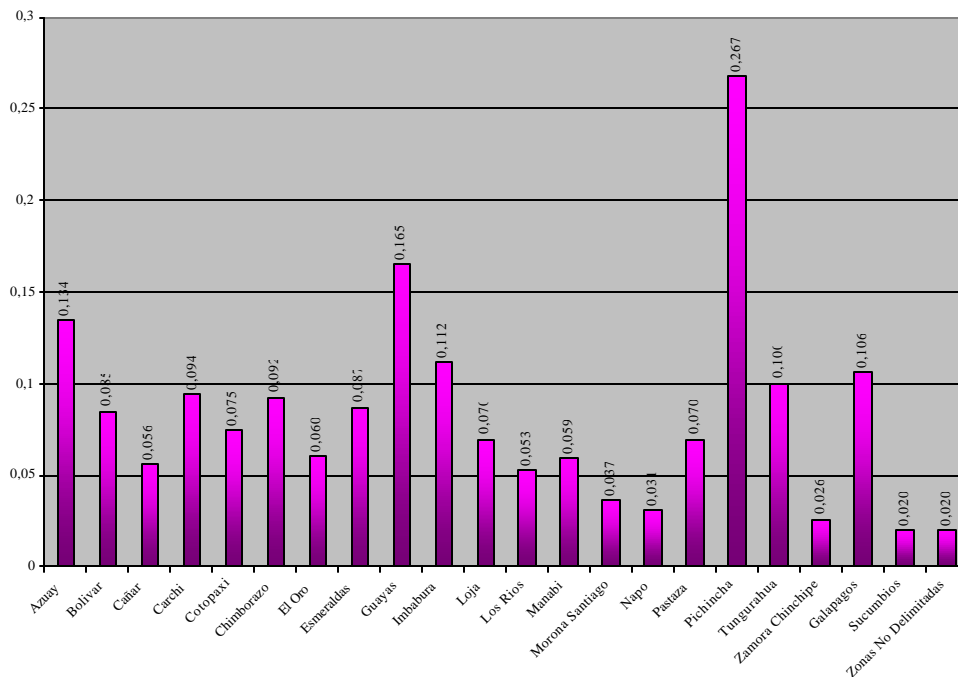
FIGURA 4.5
X5: PROPORCIÓN DE LAS VIVIENDAS QUE DISPONEN DE SERVICIO ELÉCTRICO DEL TOTAL DE VIVIENDAS PROVINCIALES.



La provincia que cuenta con mayor porcentaje de viviendas que tienen servicio eléctrico es Pichincha con 81,5% del total de viviendas que posee esta provincia, le siguen en porcentaje Galapagos con un 80% y Guayas con un 79,8%. La provincia que tiene un menor porcentaje de disponibilidad de servicio eléctrico es Napo con un 27,9%.

FIGURA 4.6

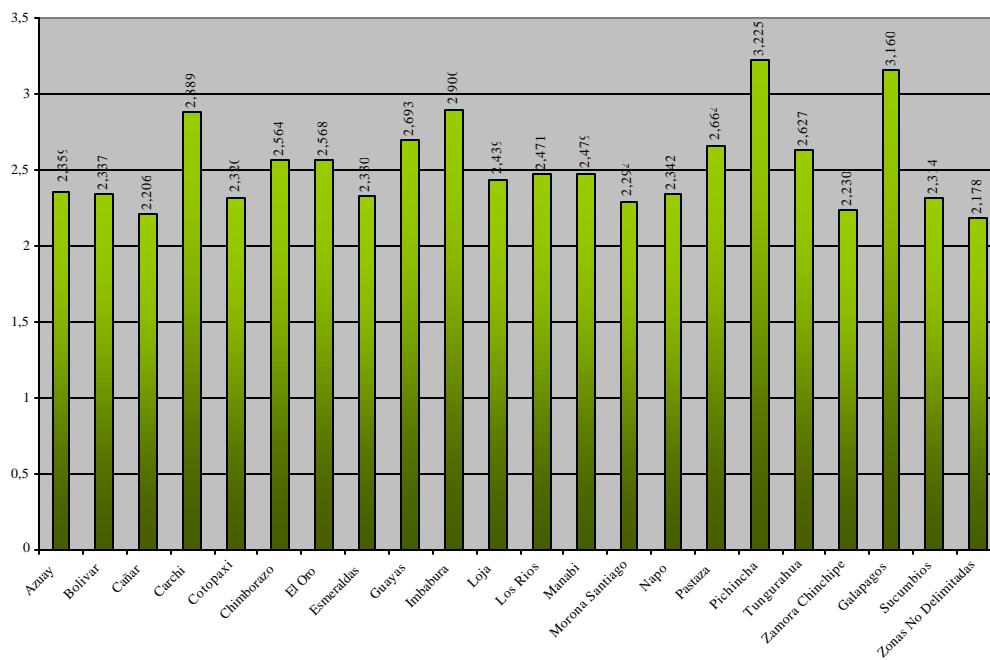
X6 PROPORCIÓN DE LAS VIVIENDAS QUE DISPONEN DE SERVICIO TELEFÓNICO DEL TOTAL DE VIVIENDAS PROVINCIALES.



A diferencia de otras variables, ésta es la que cuenta con menor porcentaje en cada provincia; así, Pichincha posee un 26,7% de disponibilidad de servicio telefónico, seguido por Guayas con un 16,5% y Azuay con un 13,4%. Sucumbios y las zonas no delimitadas cuentan con el menor porcentaje de disponibilidad de servicio telefónico que es el 2%.

FIGURA 4.7

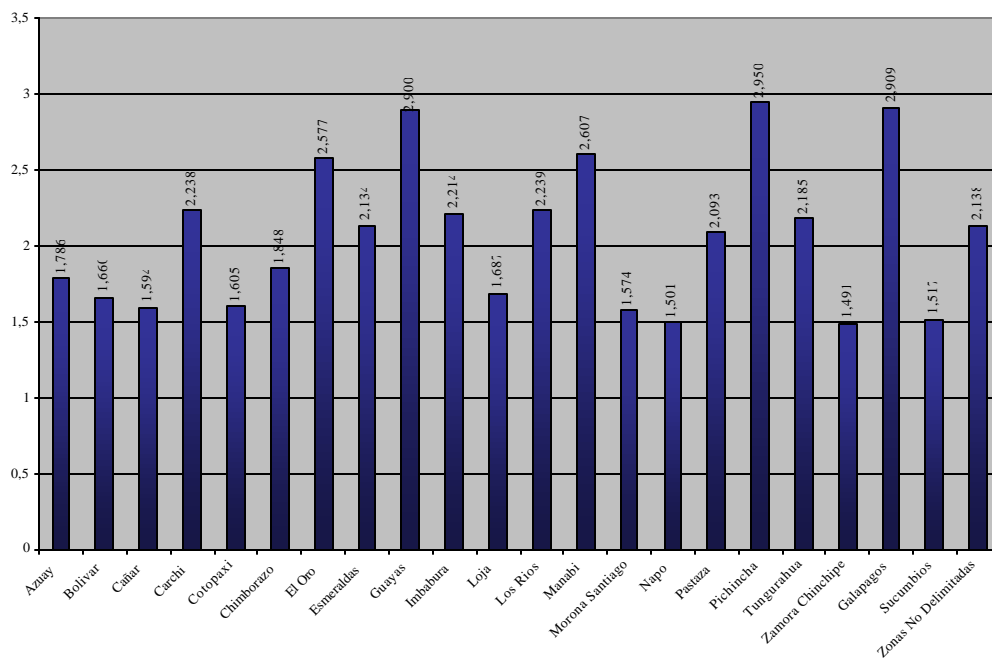
X7: INDICADOR PORCENTUAL DEL SISTEMA DE ELIMINACIÓN DE BASURA.



Esta variable tiene un comportamiento homogéneo entre las provincias, es así, que los porcentajes no tienen una gran variación. La provincia con mayor ubicación en la escala es Pichincha con 3,225, luego sigue Galapagos con 3,16 e Imbabura con 2,9.

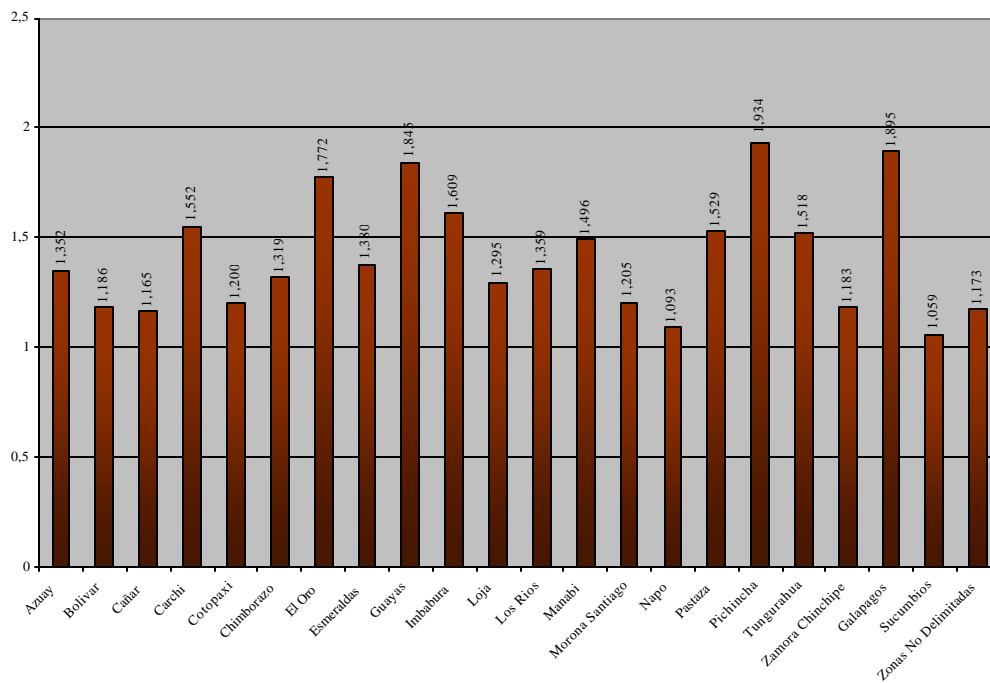
FIGURA 4.8

X8: INDICADOR PORCENTUAL DE DISPONIBILIDAD DE SERVICIO HIGIÉNICO.

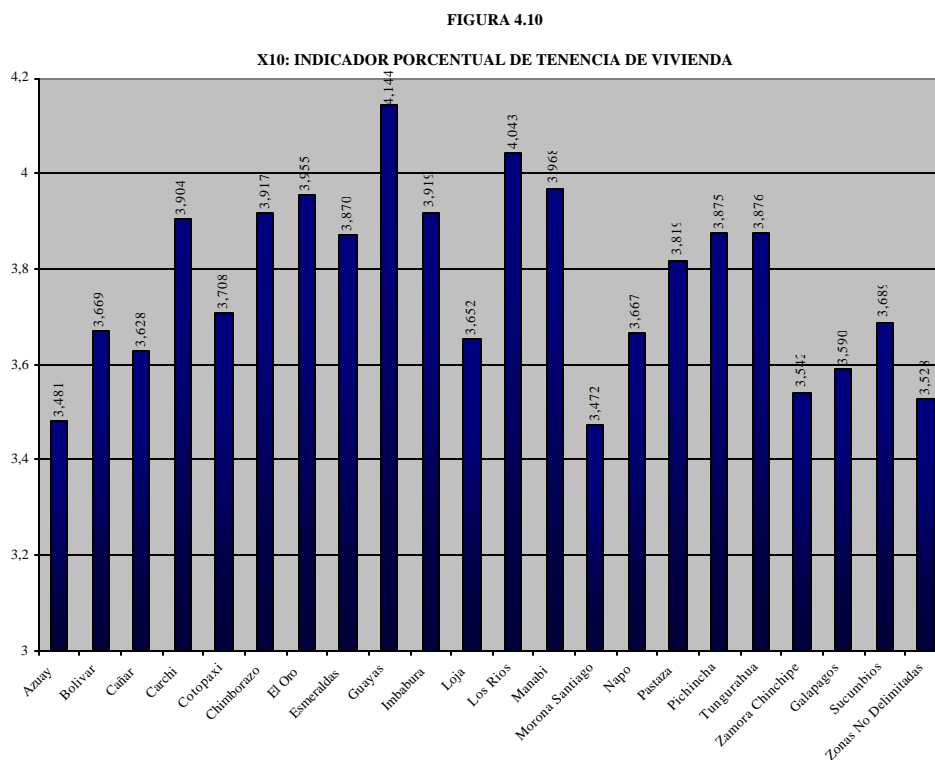


En la escala de servicio higiénico ocupa el primer Pichincha con un valor de 2,95 luego Galapagos con 2,909 y Guayas con 2,9

FIGURA 49
X9: INDICADOR PORCENTUAL DE DISPONIBILIDAD DE DUCHA.

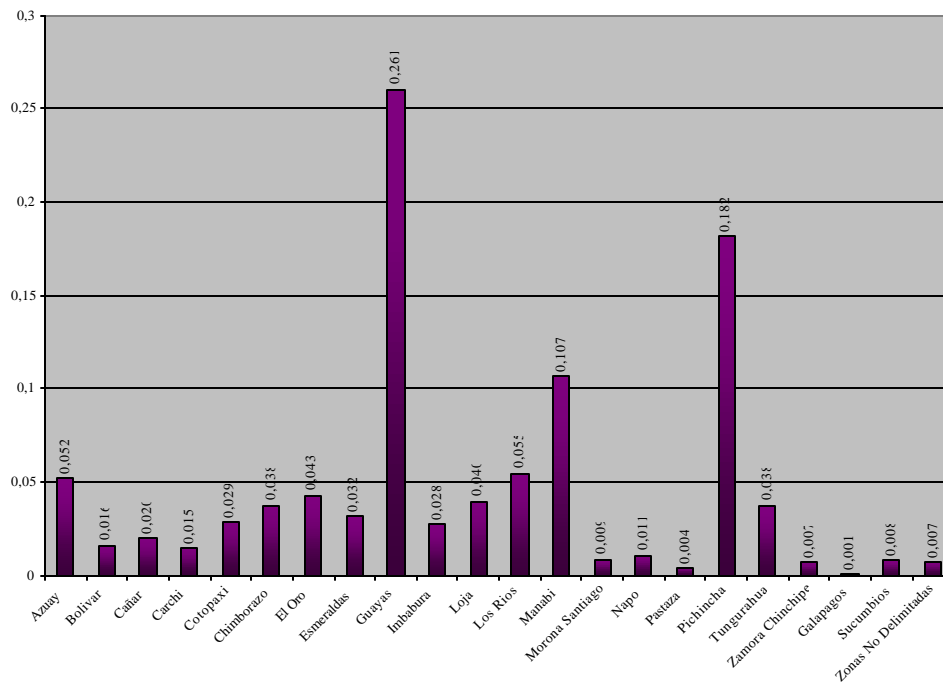


La provincia que se ubica en la primera posición dentro de la escala de disponibilidad de ducha es Pichincha con 1,934, seguido por Galapagos con 1,895 y Guayas con 1,845.



En la el indicador de tenencia de vivienda sobresale la provincia de Guayas con un valor de 4,144 seguido por Los Ríos con 4,043 y Manabí con 3,968. Los más bajos en esta categoría son Azuay y Morona Santiago con valores de 3,481 y 3,472 respectivamente.

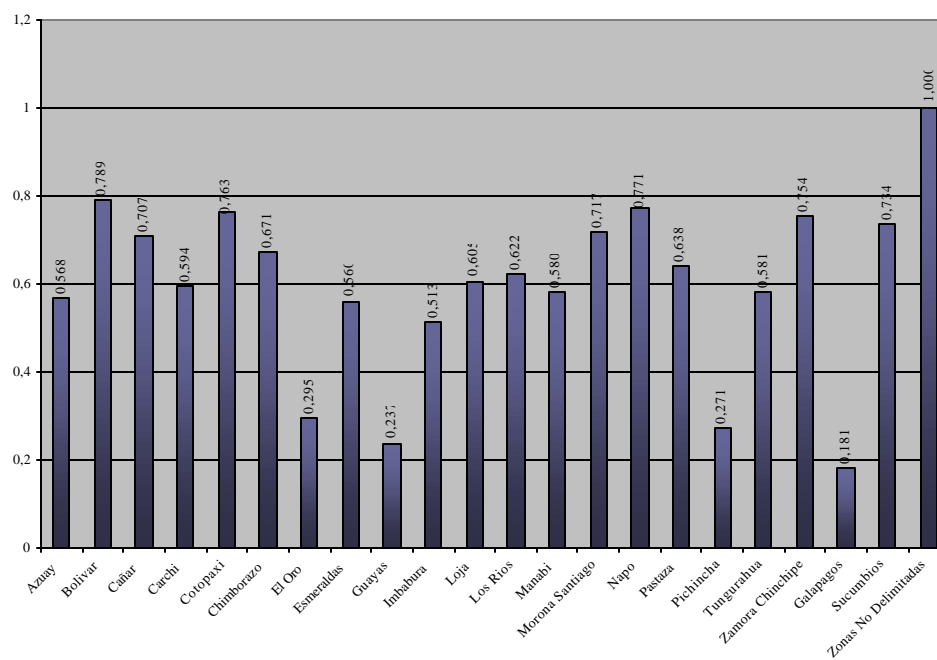
FIGURA 4.11
XII: PROPORCIÓN DE LA POBLACIÓN POR PROVINCIAS



La provincia que cuenta con un mayor porcentaje de pobladores en el territorio nacional, es la provincia del Guayas con un 28,1% de la población total, seguida por Pichincha con 18,2% y Manabí con 10,7%. La provincia que con menor porcentaje de población es Galapagos con un 0,1%.

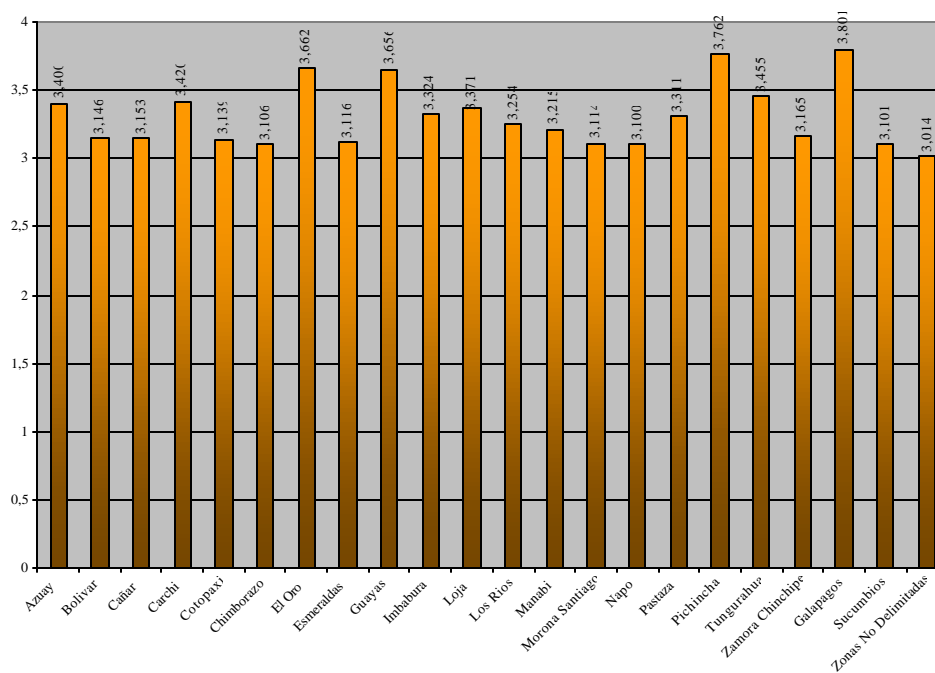
FIGURA 4.12

X12 : PROPORCIÓN DE LA POBLACIÓN RURAL POR PROVINCIAS.



Las zonas no delimitadas son puramente rurales, por ende el 100% de ellas poseen población rural. Dentro del resto de las provincias se encuentran Bolívar con un 78,9% y Napo con un 77,1%, con un mayor porcentaje de población rural.

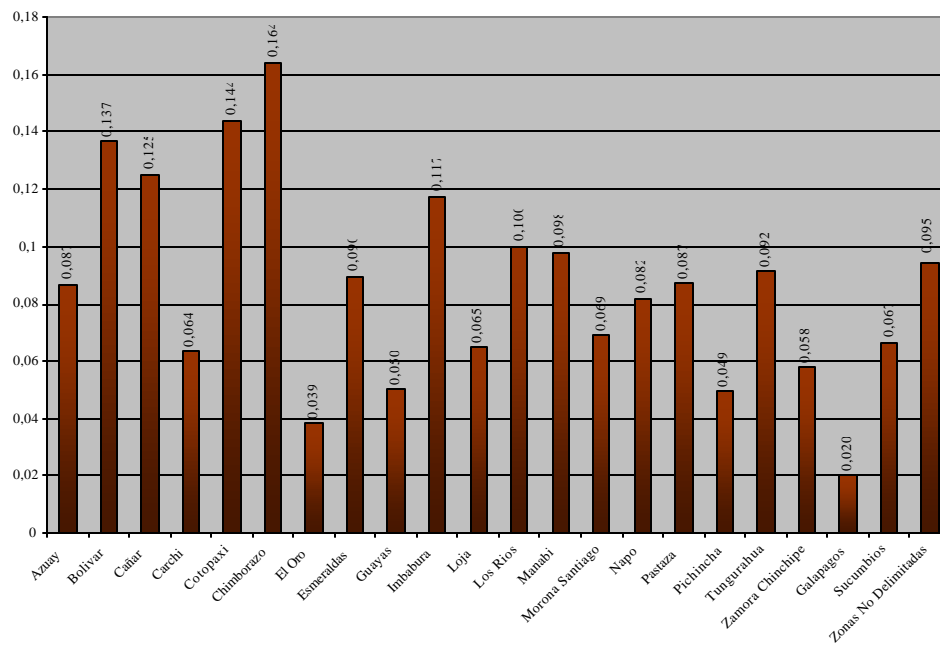
FIGURA 4.13
X13: INDICADOR PORCENTUAL NIVEL DE INSTRUCCIÓN.



Dentro de la variable educacional -nivel de instrucción-, se nota la similitud entre las provincias, sobresaliendo en la escala Galapagos con 3,801, Pichincha con 3,762 y El Oro con 3,662.

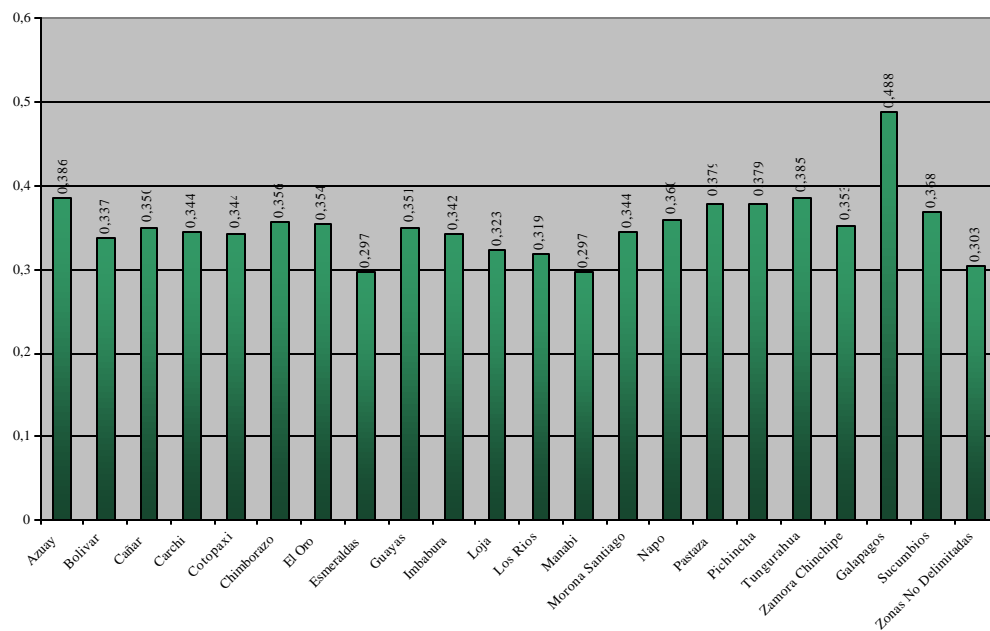
FIGURA 4.14

X14: PROPORCIÓN DE LA POBLACIÓN ANAFABETA DE LA POBLACIÓN TOTAL PROVINCIAL



En contraste a la variable nivel de instrucción, se encuentra analfabetismo, el cual se presenta en mayor porcentaje en Chimborazo con 16,4%, Cotopaxi con 14,4% y Bolívar con 13,7%

FIGURA 4.15
X15: PROPORCIÓN DE LA POBLACIÓN ECONÓMICAMENTE ACTIVA DE LA POBLACIÓN TOTAL PROVINCIAL.



La población económicamente activa en el Ecuador tiene mayor concentración en las provincias de Galapagos, Azuay y Tungurahua, con porcentajes de 48,8%, 38,6% y 38,5% respectivamente.

	X10	X11	X12	X13	X14	X15
Número de casos	22	22	22	22	22	22
Mínimo	-1,5531	-0,7073	-2,0786	-1,2617	-1,8504	-1,4035
Máximo	1,9662	3,4258	2,0051	2,1129	2,1891	3,3928
Rango	3,5193	4,1331	4,0836	3,3746	4,0395	4,7963
Suma	0	0	0	0	0	0
Mediana	-0,0288	-0,2766	0,0789	-0,3177	0,0238	-0,0536
Media	0	0	0	0	0	0
CI Superior 95%	0,4434	0,4434	0,4434	0,4434	0,4434	0,4434
CI Inferior 95%	-0,4434	-0,4434	-0,4434	-0,4434	-0,4434	-0,4434
Error Estándar	0,2132	0,2132	0,2132	0,2132	0,2132	0,2132
Desviación Estándar	1	1	1	1	1	1
Varianza	1	1	1	1	1	1
Covarianza	321,7E+12	-9,0E+15	4,0E+15	623,9E+12	-3,4E+15	3,1E+15
Sesgo (G1)	0,0886	2,5877	-0,5916	0,9165	0,3932	1,6900
SE Sesgo	0,4910	0,4910	0,4910	0,4910	0,4910	0,4910
Kurtosis (G2)	-0,9742	6,8016	0,3366	-0,2507	-0,0622	5,6936
SE Kurtosis	0,9528	0,9528	0,9528	0,9528	0,9528	0,9528

4. 5. Análisis de Componentes Principales

Antes de entrar al desarrollo de las componentes mediante las cuales hallaremos el indicador porcentual sintético, se debe analizar que clase de *ACP* se va a realizar.

El *ACP* propuesto en el trabajo es normalizado, en el cual las variables serán centradas y reducidas; por lo que las componentes se constituirán con los vectores propios de la matriz de correlación R , y no de la matriz S de varianza-covarianza como se trató en el capítulo anterior.

Cuando se usa la matriz R , algunas propiedades de los valores y vectores propios varían, es así que:

$$\ast \sum_{j=1}^p \text{Var}\langle Y(j) \rangle = \text{tr } R = p$$

$$\ast \prod_{j=1}^p \text{Var}\langle Y(j) \rangle = |R|$$

❖ Traza R ($=p$), por ende la importancia de cada componente j estará dada por

$$\frac{\lambda(j)}{p}$$

Estas características de los valores y vectores propios determinan que el análisis de componentes principales es sensible a los cambios de escala para las variables.

Teniendo estas consideraciones, se puede empezar el análisis de la tabla de datos encontrando la matriz de correlación de las variables.

Note que se tienen $n=22$ observaciones-individuos- y $p=15$ variables.

Calculando los valores propios de la matriz R , y la varianza total explicada por la componente asociada- ecuación (3.30)- se tiene el siguiente resultado.

Tabla IV.I

**TABLA DE VALORES PROPIOS OBTENIDOS DE LA MATRIZ R Y
PORCENTAJE DE LA VARIANZA TOTAL EXPLICADA**

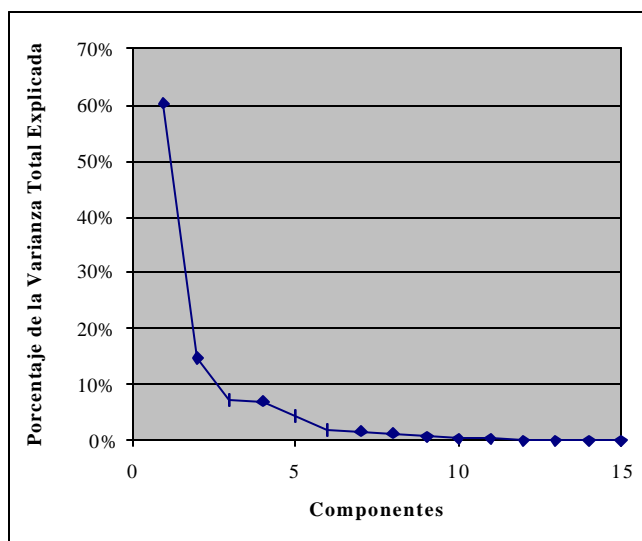
	Valor Propio	Proporción de la Varianza Total Explicada	
		Absoluta%	Acumulada%
1	9,067	60,450%	60,450%
2	2,197	14,647%	75,097%
3	1,113	7,420%	82,518%
4	1,032	6,880%	89,398%
5	0,628	4,187%	93,585%
6	0,29	1,933%	95,518%
7	0,242	1,613%	97,132%
8	0,202	1,347%	98,478%
9	0,096	0,640%	99,118%
10	0,058	0,387%	99,505%
11	0,038	0,253%	99,758%
12	0,019	0,127%	99,885%
13	0,012	0,080%	99,965%
14	0,005	0,033%	99,998%
15	2,31665E-04	0,002%	100,000%

Note que la suma de los valores propios es igual a 15, el número total de variables consideradas y varianza total explicada si se consideran todos los componentes.

El gráfico de la varianza explicada por cada componente, servirá de ayuda para determinar el número de componentes a considerar en el análisis.

Figura 4.16

VARIANZA EXPLICADA POR CADA COMPONENTE



Si consideramos los criterios 1 y 3 para seleccionar el número de componentes, se puede observar en el gráfico que se produce el “codo” alrededor del cuarto componente, y a su vez estos valores son mayores que 1.

Este número de componentes ($q=4$), recoge el siguiente porcentaje de cantidad de información ecuación (3.33)-:

$$\begin{aligned} \text{Porcentaje de Varianza Explicada por los } q \text{ primeros componente} &= \frac{\sum_{j=1}^q \lambda(j)}{\sum_{j=1}^p \lambda(j)} * 100 \\ &= \frac{13,409}{15} = 89,398\% \end{aligned}$$

Se calcula los vectores propios correspondientes a los cuatro primeros valores propios, para poder realizar la combinación lineal que será la función de la nueva variable.

Tabla IV.II

**TABLA DE VALORES PROPIOS CON SUS RESPECTIVOS VECTORES
PROPIOS, BASADOS EN EL NÚMERO DE FACTORES**

Valor Propio				
	9,067	2,197	1,113	1,032
Vectores Propios				
	1	2	3	4
X1	0,230	0,38	0,103	-0,393
X2	0,045	-0,312	0,666	-0,294
X3	0,28	-0,163	0,182	0,319
X4	0,319	0,071	0,05	0,068
X5	0,307	-0,038	0,143	0,203
X6	0,273	0,066	0,355	-0,128
X7	0,288	-0,162	-0,005	0,231
X8	0,298	0,07	-0,125	0,04
X9	0,324	-0,039	-0,107	0,094
X10	0,187	0,416	-0,107	0,375
X11	0,226	0,399	0,061	-0,387
X12	-0,307	0,072	0,177	0,049
X13	0,309	-0,176	-0,11	-0,102
X14	-0,159	0,25	0,522	0,477
X15	0,139	-0,507	-0,076	-0,033

Haciendo uso de la fórmula (3.14), donde $C(i)$ es el vector propio para el i -ésimo valor propio y $X(i)$ las variables originales, obtenemos $Y(j)$ que es la j -ésimo variable nueva; es decir, la componente principal.

$$Y(j) = \sum_{i=1}^p C(i)X(i)$$

Los resultados de esta combinación se muestran en la siguiente tabla.

Tabla IV.III

TABLA DE COMPONENTES PRINCIPALES O FACTORES

	Comp. I	Comp. II	Comp. III	Comp. IV
1	0,145	-1,397	1,696	-1,492
2	-2,275	-0,146	1,904	0,015
3	-2,305	-0,496	1,672	-0,362
4	1,812	-1,068	0,778	0,937
5	-2,101	0,798	0,408	0,878
6	-0,689	1,318	0,393	2,134
7	3,037	-0,273	-1,760	0,699
8	-0,984	1,241	-0,261	0,168
9	5,742	2,997	-0,436	-1,774
10	1,478	-0,071	0,917	1,599
11	-0,956	0,119	-0,822	-0,453
12	-0,438	1,664	-0,755	0,690
13	0,731	2,198	-0,084	0,045
14	-2,865	-0,836	-0,314	-1,037
15	-3,468	-0,134	-0,979	-0,599
16	-0,125	-0,366	-1,317	1,238
17	6,932	0,557	0,813	-1,056
18	1,544	-0,877	1,075	0,464
19	-2,885	-0,708	-1,076	-0,709
20	4,635	-4,402	-1,039	0,196
21	-3,459	-0,498	-0,914	-1,031
22	-3,506	0,383	0,101	-0,549

La varianza explicada por cada componente será:

Comp. I	9,067
Comp. II	2,197
Comp.III	1,113
Comp. IV	1,032

cuyo porcentaje de varianza explicada –ecuación (3.32) – es:

Comp. I	60,450
Comp. II	14,650
Comp.III	7,417
Comp. IV	6,877

La primera componente sintetiza el 60,45% de la varianza.

4.5.1. Interpretación de las componentes principales en función de su correlación con las variables originales.

Una vez calculados las nuevas variables, es importante analizar la correlación existente entre las variables originales y las nuevas variables.

Debido a que puede haber variables que resultan redundantes en el problema, éstas se agrupan en componentes principales o factores y se interpretan a partir de su correlación con las variables originales.

La correlación es la proyección de la variable sobre la componente, es así, que mientras una variable este correlacionada con una componente se ubicará más cerca de la coordenada. Mientras mayor este correlacionada una variable sobre el componente, mayor peso tendrá en la explicación del mismo.

Los cálculos de las correlaciones se la realiza con la ecuación (3.37)

Tabla IV.IV
CORRELACIONES DE LAS VARIABLES CON LAS COMPONENTES
PRINCIPALES O FACTORES

	Comp.I	Comp.II	Comp.III	Comp.IV
X1	0,691	0,563	0,109	-0,399
X2	0,134	-0,463	0,703	-0,299
X3	0,844	-0,242	0,192	0,324
X4	0,959	0,105	0,052	0,069
X5	0,925	-0,056	0,151	0,206
X6	0,821	0,098	0,374	-0,13
X7	0,866	-0,241	-0,005	0,235
X8	0,898	0,104	-0,132	0,041
X9	0,975	-0,058	-0,112	0,095
X10	0,564	0,616	-0,113	0,381
X11	0,681	0,592	0,064	-0,393
X12	-0,923	0,106	0,186	0,049
X13	0,93	-0,261	-0,116	-0,103
X14	-0,477	0,37	0,55	0,485
X15	0,418	-0,752	-0,08	-0,033

Note que la primera componente principal tiene una correlación positiva con X_9 y negativamente con X_{12} ,

Asi mismo, la componente II con X_{10} y X_{15} , la componente III con X_2 y X_{15} y por último la componente IV con X_{14} y X_1 .

Las variables que sean mejor representadas serán importantes al nombrar las componentes.

4.5.2. Interpretación de la nube de variables y gráficos de correlación.

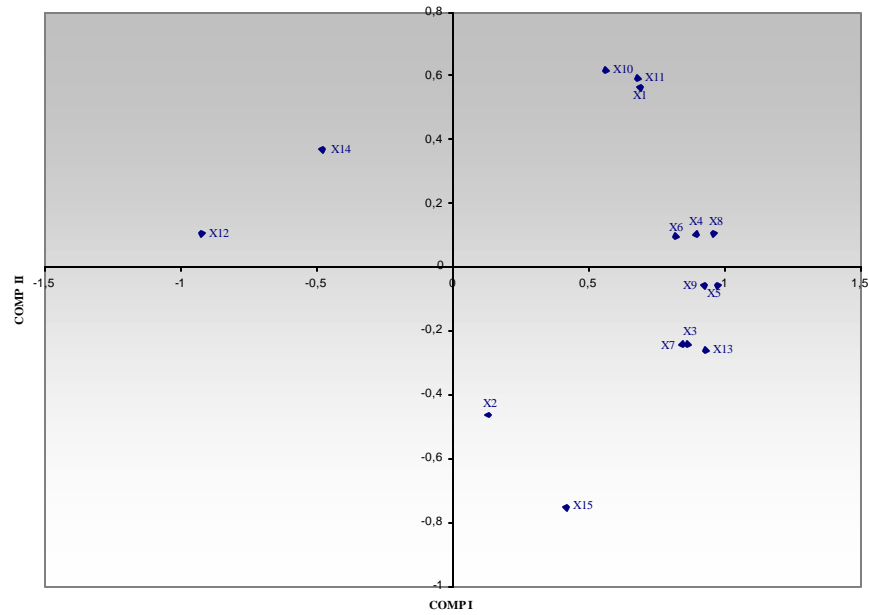
Para interpretar la nube de datos, es necesario graficar la correlación $r(i, j)$ de cada variables original con las componentes principales, las cuales serán los ejes del gráfico que se utilizará.

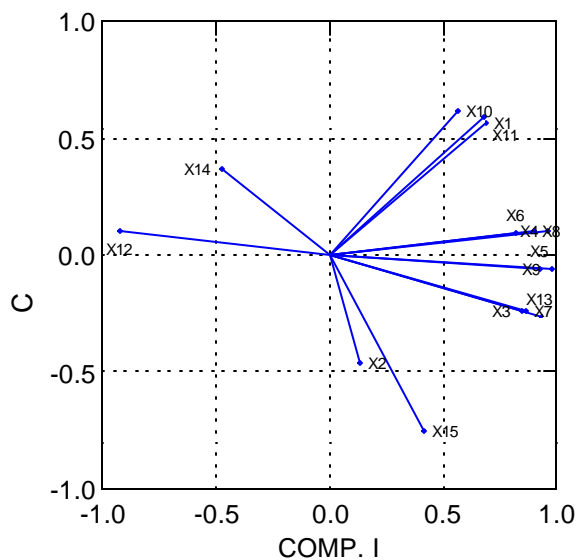
El círculo de radio unitario con centro en el origen, ayudará a identificar las variables cuya correlación sea mayor. Así, las variables mejor representadas se situarán en ubicarán cerca del círculo y las menos explicadas hacia el origen.

A continuación se presentan los gráficos de las correlaciones entre las componetes principales y las variables.

Figura 4.17

GRÁFICOS DE LAS VARIABLES SOBRE EL PLANO FORMADO POR LA COMPONENTE I Y LA COMPONENTE II Y SU CORRELACIÓN.





Las dos primeras componentes explican el 71% de la varianza total.

Si analizamos la componente I se nota que la variable mejor representada es X_9 con respecto a la correlación positiva que posee con ella y X_{12} con la correlación negativa más alta es la que se ubica en el plano sobre el lado negativo del eje. Esto se puede notar por la cercanía de las variables al círculo unitario.

En la matriz R se puede analizar la correlación entre las variables; y, en el gráfico verificar que las variables X_8 , X_6 , X_4 están correlacionadas positivamente entre ellas, así también con la componente principal I. Igual

ocurre X_5 , X_9 y X_3 , X_7 , X_{13} . Básicamente variables que representan servicios que ofrecen las viviendas, salvo X_{13} que es el indicador de nivel de instrucción.

Es curioso el hecho que la variable X_{13} , está correlacionada positivamente con la mayoría de las variables, excepuando la de analfabetismo X_{14} , PEA X_{15} y población rural X_{12} .

Una vez analizada la componente I y su relación con las variables, se puede proponer un nombre a la nueva variable. Por la información que posee sobre servicios básicos e instrucción, una manera de sintetizar en un nombre la incidencia de ambos aspectos en la sociedad sería como la **calidad de vida**

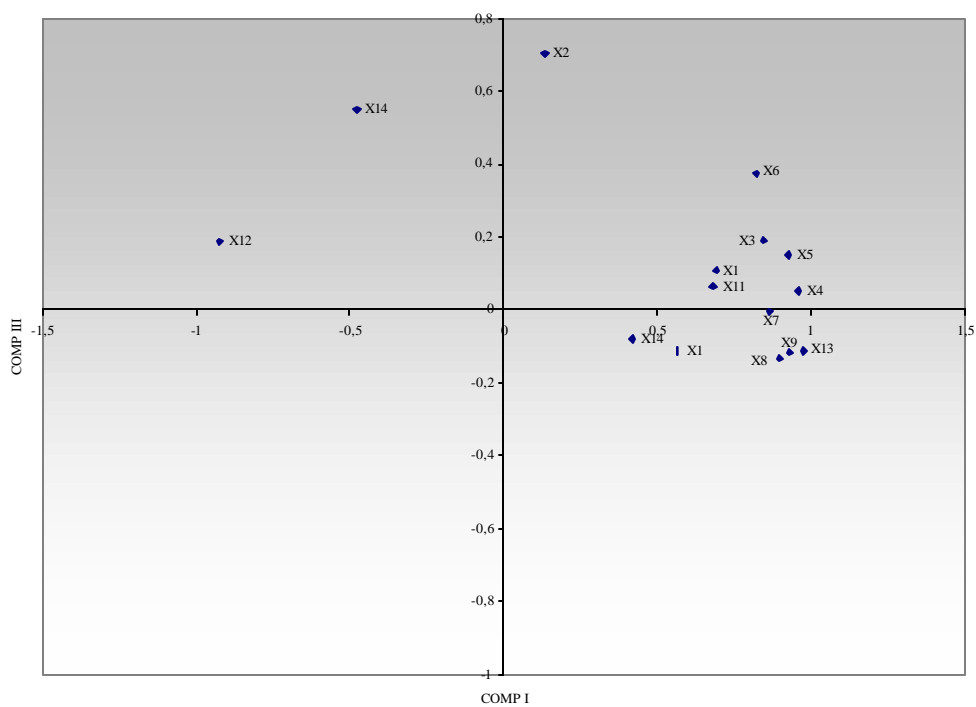
Ahora analizando la componente II, se observa que las variables con mayor representatividad son X_{10} y X_1 en el eje positivo, mientras X_{15} y X_2 en el eje negativo.

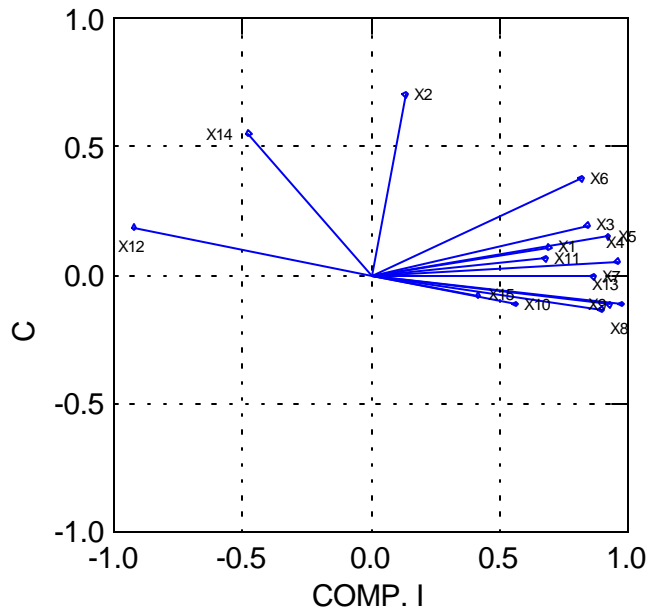
Las variables X_{10} y X_1 están correlacionadas positivamente al igual que X_{15} con X_2 , X_{14} con X_{12} .

Una vez analizadas las variables que conforman la componente II, se puede evidenciar una influencia de **pérfil económico de la población** en la misma.

Figura 4.18

GRÁFICOS DE LAS VARIABLES SOBRE EL PLANO FORMADO POR LA COMPONENTE I Y LA COMPONENTE III Y SU CORRELACIÓN.





Ya habiendo analizado la componente principal I, resta analizar las otras componentes basados en que la componente I es la que mayor porcentaje de la varianza total explica.

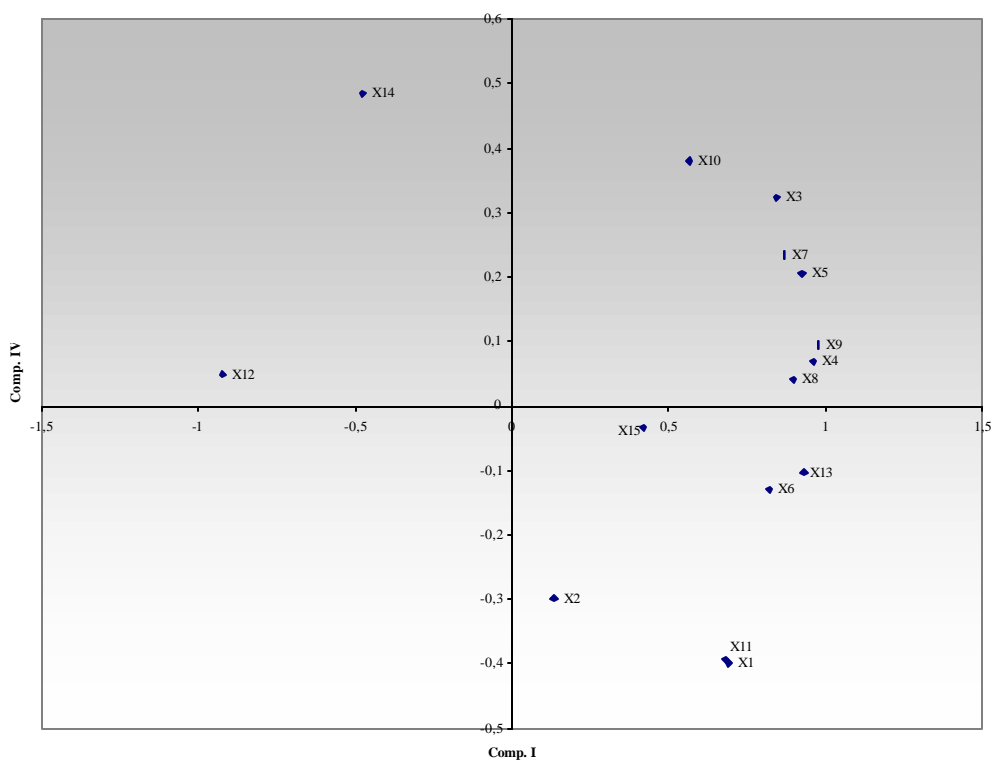
○

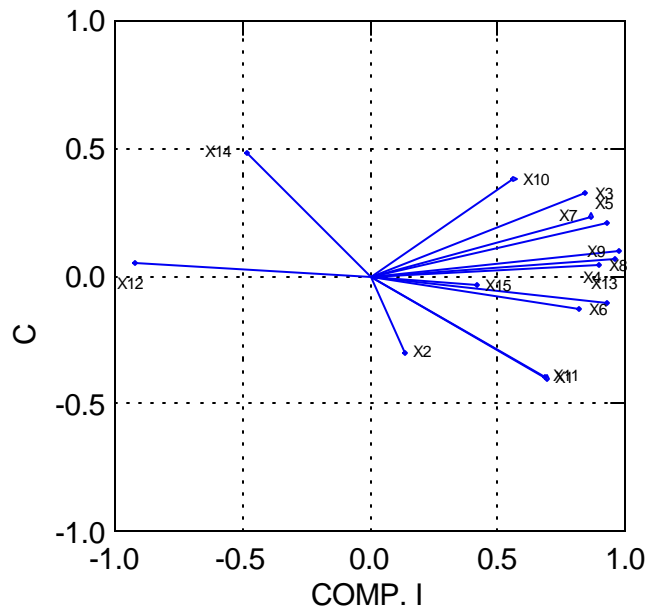
Tomando en cuenta el eje de la componente III, se observa que X_2 y X_{14} son las mejor representadas sobre el eje positivo y X_{13} , X_{10} y X_9 en el eje negativo; revisando la tabla R se observa que X_2 y X_{14} están correlacionadas negativamente.

Revisando la información más representativa, la nueva variable se puede decir representa una **dualidad entre características de vivienda y educación.**

Figura 4.19

GRÁFICOS DE LAS VARIABLES SOBRE EL PLANO FORMADO POR LA COMPONENTE I Y LA COMPONENTE IV Y SU CORRELACIÓN.





La variable que mejor está representada por la última componente es X_{14} en el eje positivo y X_2 y X_{11} en eje negativo.

○

La proyección de X_{11} y X_2 en el plano muestra que, estas variables están muy correlacionadas, tanto que no se distinguen en la proyección.

Las variables que más sobresalen en esta componente son las variables que enfocan los aspectos generales del censo como, la población y vivienda total dentro de la provincia. Por ende un nombre de acuerdo a con la información obtenida será **características generales del censado**.

≡

4.5.3. Interpretación de la nube de individuos.

Una vez interpretada la nube de variables, se puede seguir el análisis con la interpretación de la nube de individuos representando sus coordenadas sobre las componentes halladas.

En el gráfico de los individuos, no se puede notar la calidad de la representación en el plano, debido a que la distancia al origen varía.

La proyección de los puntos de observaciones sobre el plano se obtienen como sigue.

$$F_j = \sqrt{\frac{n}{p}} X v_{-j} \quad j = 1, \dots, q \quad (4.1)$$

donde X es la matriz de datos de las variables, v_{-j} es el vector unitario correspondiente al j -ésimo componente principal.

Haciendo uso de esta fórmula, se calculan las proyecciones de cada individuo sobre las componentes principales.

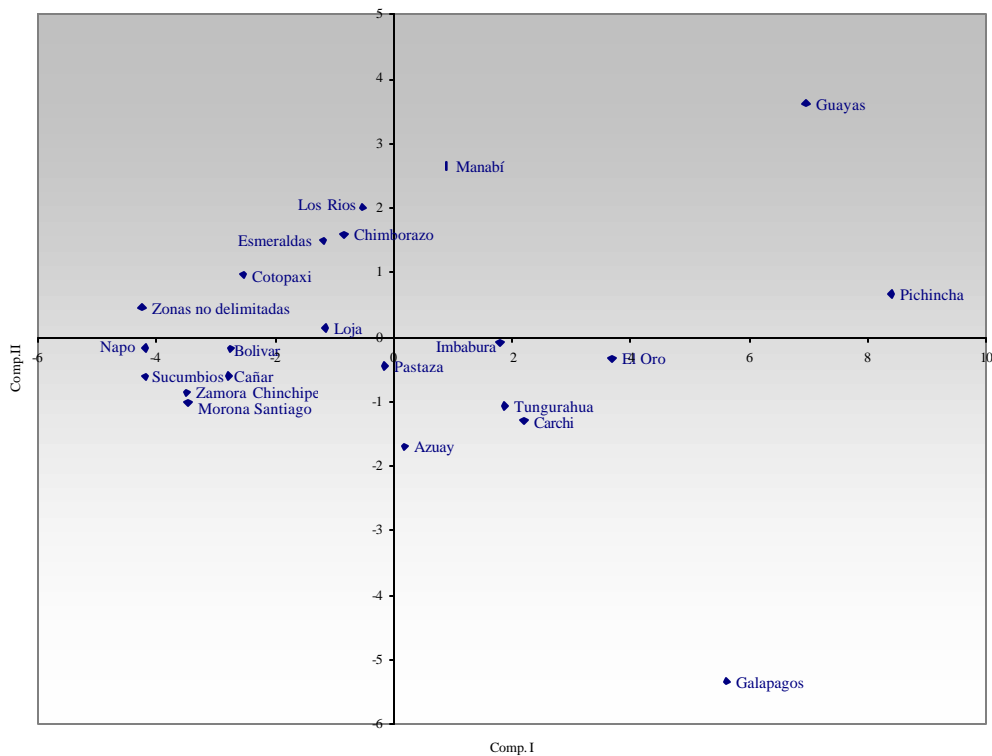
TABLA IV.V

**PROYECCIÓN DE LAS PROVINCIAS SOBRE LAS COMPONENTES
PRINCIPALES**

Provincias	Comp.I	Comp.II	Comp.III	Comp.IV
Azuay	0,1751	-1,6918	2,0542	-0,6004
Bolivar	-2,7554	-0,1775	2,3060	0,9726
Cañar	-2,7915	-0,6004	2,0250	0,6422
Carchi	2,1951	-1,2935	0,9425	2,1069
Cotopaxi	-2,5436	0,9662	0,4932	0,2725
Chimborazo	-0,8341	1,5949	0,4750	1,4237
El Oro	3,6784	-0,3308	-2,1304	0,2878
Esmeraldas	-1,1918	1,5025	-0,3162	-0,0307
Guayas	6,9536	3,6297	-0,5288	-2,4385
Imbabura	1,7898	-0,0868	1,1111	2,2417
Loja	-1,1578	0,1450	-0,9950	-1,1571
Los Rios	-0,5312	2,0148	-0,9158	0,2442
Manabi	0,8843	2,6611	-0,1022	-0,0419
Morona Santiago	-3,4705	-1,0117	-0,3799	-1,1425
Napo	-4,1994	-0,1624	-1,1866	-1,2261
Pastaza	-0,1508	-0,4434	-1,5955	0,4056
Pichincha	8,3940	0,6738	0,9844	-1,0608
Tungurahua	1,8700	-1,0622	1,3036	1,4794
Zamora Chinchipe	-3,4927	-0,8571	-1,3045	-1,3703
Galapagos	5,6141	-5,3308	-1,2551	0,8134
Sucumbios	-4,1889	-0,6031	-1,1069	-1,2625
Zonas No Delimitadas	-4,2457	0,4640	0,1226	-0,5584

Figura 4.20

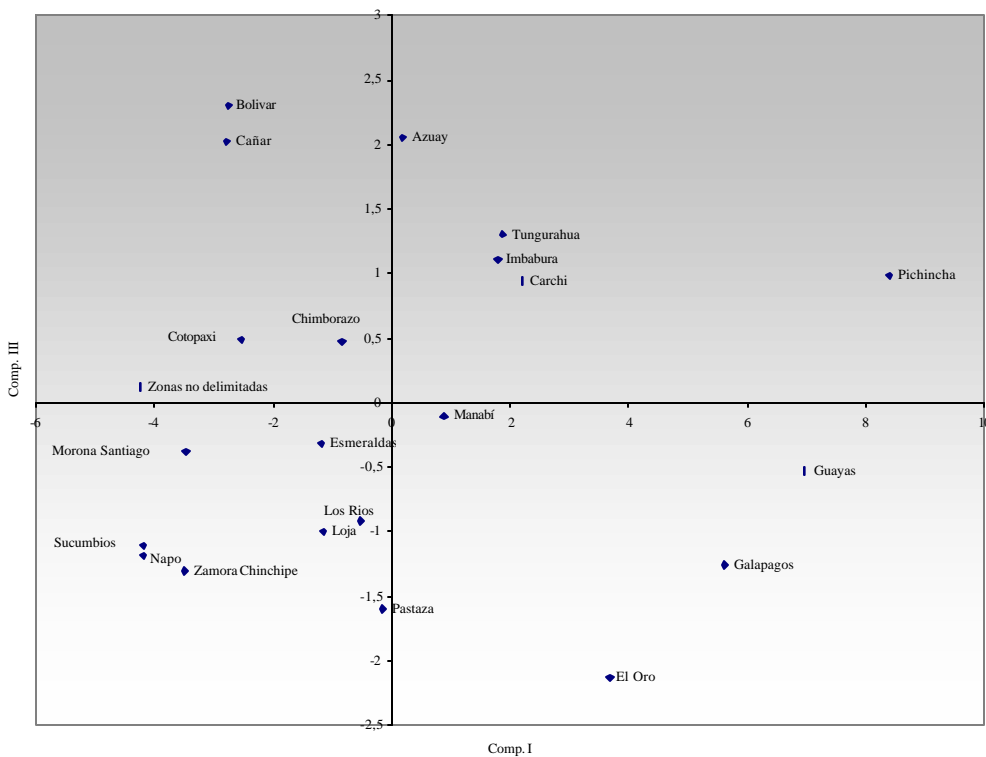
**GRÁFICOS DE LAS PROVINCIAS SOBRE EL PLANO FORMADO POR LA
COMPONENTE I Y LA COMPONENTE II.**



Con respecto a la ubicación de los individuos dentro del plano de calidad de vida y perfil económico de la población, las provincias que se ubican en una situación favorable son Pichincha, Guayas y Galapagos, mientras Napo, Sucumbios y las zonas no delimitadas ocupan los últimos niveles.

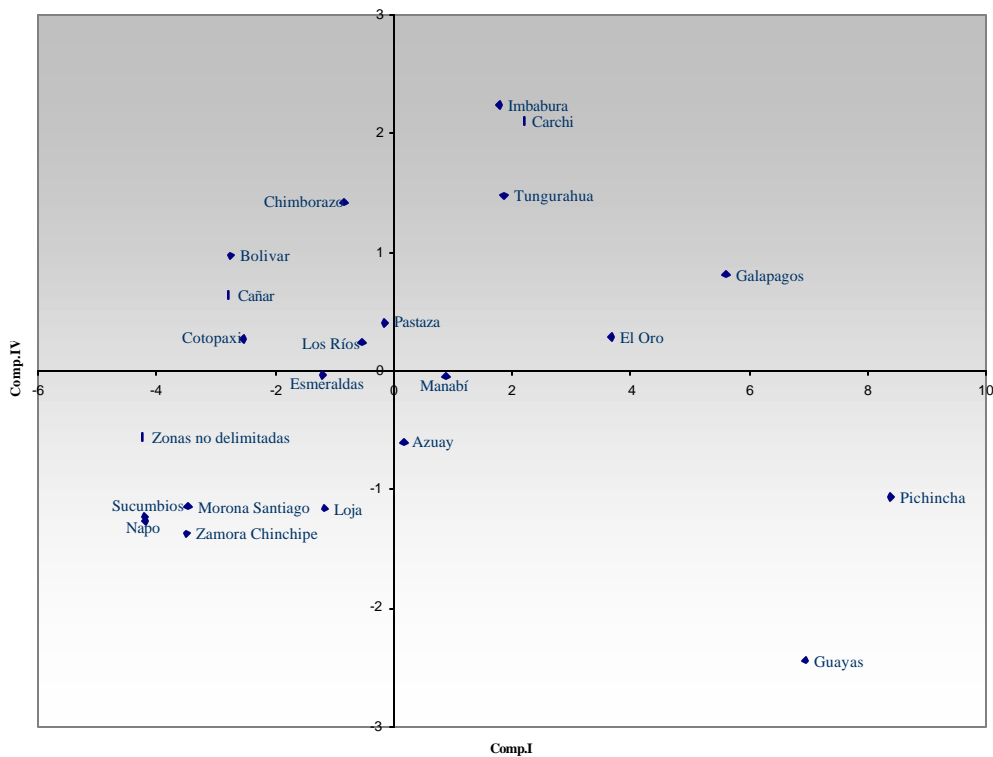
Figura 4.21

GRÁFICOS DE LAS PROVINCIAS SOBRE EL PLANO FORMADO POR LA COMPONENTE I Y LA COMPONENTE III.



Dentro de las variables calidad de vida y dualidad entre características de vivienda y educación, las provincias que en mejor situación se ubican son Pichincha, Guayas y Galapagos. Las que ocupan los lugares más bajos son Napo Sucumbios y las Zonas no delimitadas.

Figura 4.22
GRÁFICOS DE LAS PROVINCIAS SOBRE EL PLANO FORMADO POR LA
COMPONENTE I Y LA COMPONENTE IV.



En la escala de características generales del censo, junto con la variable calidad de vida, sigue manteniéndose el orden que se mostró en gráficos anteriores; así Pichincha, Guayas y Galapagos se ubican en las primeras posiciones, mientras Sucumbios Napo y las zonas no delimitadas vuelven a quedar en últimos puestos.

4. 6. Construcción del Indicador Sintético de Pobreza.

Como se dijo un indicador social debe tratar sobre el aspecto de la vida social de la población en que se está interesado, en el caso de específico del trabajo se consideró la pobreza.

Ya se estableció que el método para construir el indicador social va a ser indicadores sintéticos, para construirlo se aplicará el criterio propuesto por **KAMINSKY (7)**, el cual propone escoger la componente principal de mayor explicación como el indicador sintético. Es decir, que la Componente Principal I será el indicador sintético de pobreza y las variables constituyen los indicadores parciales.

Los criterios para establecer los indicadores sociales son desarrollados a continuación, con una breve explicación de como las componentes principales se adaptan para constituirse como tal.

- ❖ **Adecuación:** El uso de las componentes principales ayuda a establecer una tendencia, la cual dependerá de las características que presenten en cada una de ellas. La preocupación social relevante en el análisis resultaría en las variables que mayor incidencia presenten en la Componente Principal I que serían de los

servicios de la vivienda y el nivel de instrucción, que son catalogados como calidad de vida.

- ❖ **Capacidad de Compendiación:** Es importante notar que los objetivos del análisis de componentes principales van relacionados con este aspecto; así se tiene, generar nuevas variables que puedan expresar la información contenida en el conjunto original de datos, reducir la dimensionalidad del problema y eliminar algunas de las variables originales si ellas aportan poca información.
- ❖ **Coordinación:** Las componentes principales que se obtuvieron del análisis de los datos, son combinaciones lineales que involucran varias facetas de la vida socioeconómica del país.
- ❖ **Exactitud:** Es importante comparar los resultados de cada variable analizada en la estadística descriptiva con los datos obtenidos por la componente principal que va a ser considerada como indicador social.
- ❖ **Oportunidad y Frecuencia:** Al adquirir experiencia en el análisis de componentes principales y una vez establecidas las variables que inciden en el indicador social, puede resultar sencillo su cálculo periódico. Como método alternativo se puede utilizar información de otras encuestas o censos.

- ❖ **Viabilidad:** Los datos de las variables que están siendo utilizadas para el análisis de la pobreza, son parte de las estadísticas que se encuentran en el INEC. Esta información se encuentra reunida y su tabulación por ser variables cuantitativas es sencilla.

Siguiendo los pasos para el establecimiento de los indicadores sintéticos vistos en el Capítulo II tenemos:

1. Las variables son los indicadores parciales y a su vez determinantes de la pobreza.
2. Es parte fundamental del análisis de componentes de principales, la verificación de la congruencia entre las variables originales y las componentes principales. Esto se verifica de dos formas, gráficos de componentes principales (variable macro), correlación de cada variable con las componentes principales.
3. Los indicadores parciales van a formar totalmente parte del indicador sintético, esto se comprueba por la utilización de un modelo de regresión lineal por el método Stepwise, pero con el uso de componentes principales podemos analizar la varianza; además, los pesos de los indicadores parciales son los coeficientes de correlación de las variables originales con las componentes principales.

Una vez tratados los puntos más importantes sobre el establecimiento de los indicadores sociales por el método de indicadores sintéticos, puede mostrarse el indicador de pobreza como la primera componente principal encontrada en el análisis de componentes principales.

TABLA IV.VI
INDICADOR SINTÉTICO DE POBREZA

Provincias	Indicador de Pobreza
Pichincha	6,93200
Guayas	5,74200
Galapagos	4,63500
El Oro	3,03700
Carchi	1,81200
Tungurahua	1,54400
Imbabura	1,47800
Manabi	0,73100
Azuay	0,14500
Pastaza	-0,12500
Los Rios	-0,43800
Chimborazo	-0,68900
Loja	-0,95600
Esmeraldas	-0,98400
Cotopaxi	-2,10100
Bolivar	-2,27500
Cañar	-2,30500
Morona Santiago	-2,86500
Zamora Chinchipe	-2,88500
Sucumbios	-3,45900
Napo	-3,46800
Zonas No Delimitadas	-3,50600

Tomado en cuenta la escala, podemos concluir que los que presentan un mayor valor son las provincias con un menor índice de pobreza y viceversa.

V. CONCLUSIONES Y RECOMENDACIONES

Una vez que se cumplió con los objetivos planteados se puede obtener conclusiones y dar recomendaciones.

Conclusiones:

1. En el presente trabajo se ha expuesto la utilidad de la técnica multivariante de análisis de componentes principales (*ACP*), como una herramienta para la construcción de indicadores sintéticos que reflejan una situación socio-económica relevante, la pobreza. Esto se comprueba por su idoneidad con las características de un indicador social, de tipo sintético.
2. Se realizó la aplicación del *ACP*, y se pudo notar como su utilización permite la desagregación, particularizando en indicadores parciales (variables) que reflejan diferentes aspectos que logran caracterizar la pobreza.

3. El indicador de pobreza, permite colocar a cada provincia en una posición que dependerá de las otras para interpretarse como “buena” o “mala”, que ayudará a identificar donde existe retraso en la dotación de servicios, y que evita el desarrollo de la provincia. Este indicador debe ser netamente objetivo para la comunidad y autoridades interesadas, porque aún cuando en el país existen problemas internos, es importante analizar esta posición con la realidad nacional que se vive.

4. La combinación lineal para expresar el Indicador Sintético de Pobreza está dada por:

$$\text{Indicador} = 0.23 * X_1 + 0.045 * X_2 + 0.28 * X_3 + 0.319 * X_4 + 0.307 * X_5 + 0.273 * X_6 + 0.288 * X_7 + 0.298 * X_8 + 0.324 * X_9 + 0.187 * X_{10} + 0.226 * X_{11} - 0.307 * X_{12} + 0.309 * X_{13} - 0.159 * X_{14} + 0.139 * X_{15}$$

5. La correlación existente entre los indicadores parciales (variables) y la componente principal I, y dado que la primera componente principal explica el 60.45% de la información muestral, lleva a la conclusión que puede servir como un buen indicador sintético. Las variables que mayor representación tienen en esta componente son las relacionadas con los servicios básicos y el nivel de instrucción, lo cual ayudó para nombrar a la variable como **calidad de vida**.

6. Las provincias que obtuvieron una mejor posición fueron Pichincha, Guayas y Galapagos y las últimas Sucumbios, Napo y las zonas no delimitadas. El análisis subjetivo del indicador les corresponde a los sociólogos y entendidos de la materia.

Recomendaciones

1. Aun cuando ha existido desagregación en los datos, resultaría interesante realizar al nivel de cantones y obtener una escala en la cual se pueda notar la concordancia con los indicadores obtenidos en el ámbito provincial. Así también, dividirla en grandes áreas como son: urbana y rural, o por regiones.

2. Anteriormente se dijo que la construcción de los indicadores sociales debería contemplar información referente a la calidad de bienes y servicios que se adquieren. Pero características de interés no están contemplados en el Censo de Población y Vivienda como son salud, alimentos y vestimenta, por lo cual debería incluirse el diseño de futuros censos.

3. En el Censo de Población y Vivienda del año 1.990 se han notado falencias de información, que podrían llevar a conclusiones erróneas. Este es el caso por ejemplo de los datos de Tenencia de Vivienda, el cual refleja que la mayoría de los pobladores poseen viviendas propias, pero no hay especificaciones sobre que tipo de vivienda propia, si formal o informal, tomando como formal las viviendas sobre las cuales se posean escrituras, caso contrario sería informal.

4. Existe redundancia de información en las variables censales, por lo cual sería recomendable que al presentar los resultados de las principales categorías se trate de ser breve y brindar otro tipo de información, como lo expuesto en el punto anterior.

5. Sería interesante realizar una rotación oblicua para obtener cargas mínimas del sistema de variables originales, con el fin de ortogonalizar las correlaciones entre las mismas.

APÉNDICES

APÉNDICE A

CONCEPTOS REFERENTES AL CENSO DE POBLACIÓN Y VIVIENDA DE 1.990

Censo.- Es el inventario de los recursos y características de un territorio definido en un momento dado; los recursos pueden ser humanos, económicos, agropecuarios, etc.

Tipo de censo.- Se distinguen dos tipos de censo de población y vivienda: el *censo de hecho o de facto*, que se caracteriza porque las personas son empadronadas con respecto al lugar donde se encuentren en el momento censal, y *el censo de jure o derecho*, que consiste en empadronar a cada persona con referencia al lugar geográfico donde reside habitualmente. El censo de 1990 es un censo de hecho

Momento censal.- Para el V Censo de Población y IV de Vivienda se consideró las 00h00 del día 25 de noviembre de 1990, como momento censal.

Día de censo.- Es el día calendario (domingo 25 de noviembre de 1990), fijado para efectuar el empadronamiento de viviendas, hogares y población existentes en el territorio nacional.

Población del área rural.- Es aquella que fue empadronada en las parroquias rurales (cabeceras provinciales y restos de parroquias). Incluye además la población empadronada en la periferia de las capitales provinciales y cabeceras cantonales.

Nivel de instrucción.- Se refiere al nivel más alto de instrucción que asisten o asistieron las personas, dentro del sistema regular de enseñanza, sea éste centro de alfabetización, nivel primario, secundario, superior y postgrado.

Analfabetismo.- Se considera analfabetos a las personas que declararon no saber leer y escribir, se tabula para la población de 10 años y más.

Población económicamente activa.- Se considera población económicamente activa (PEA) a aquella que interviene en la producción de bienes y servicios.

Vivienda.- El local o recinto de alojamiento con acceso independiente, construido, edificado, transformado o dispuesto para ser habitado por una o más personas siempre que en el momento del Censo no se utilice totalmente con finalidad distinta. También se consideran como viviendas aquellas móviles e improvisadas y locales no destinados para vivir, que se hallan habitados en el momento de levantarse el Censo, como: barcasas, cuevas, carpas, vagones, etc. (Una vivienda tiene acceso

independiente, cuando para llegar a ella no se pasa por el interior de cuartos de otras viviendas, puede tener acceso directo desde la calle o pasando por patios, corredores, escaleras, etc. de uso común).

Tipo de vivienda.- De acuerdo con las características de las viviendas se clasifican en los siguientes tipos:

- 1. Casa o villa:** Es toda construcción permante hecha con materiales resistentes, tales como: hormigón, piedra, ladrillo, adobe, caña o madera. Generalmente tiene abastecimiento de agua y servicio higiénico de uso exclusivo.
- 2. Departamento:** Conjunto de cuartos que forma parte de un edificio de uno o más pisos. Se caracteriza por ser independiente, tiene abastecimiento de agua y servicio higiénico de uso exclusivo.
- 3. Cuarto en casa de inquilinato:** Comprende uno o varios cuartos pertenecientes a una casa, con entrada común y directa desde un pasillo, patio, corredor o calle y que por lo general, no cuenta con servicio exclusivo de agua o servicio higiénico, siendo estos servicios de uso común para todo los hogares o viviendas.

- 4. Mediagua:** Es una construcción de un solo piso con paredes de ladrillo, adobe, bloque o madera, con techo de paja, asbesto(eternit) o zinc. Tiene una sola caída de agua y no más de dos cuartos.
- 5. Rancho:** Es aquella construcción rústica, cubierta con palma, paja o cualquier otro material similar, con paredes de caña y con piso de madera, caña o tierra.
- 6. Covacha:** Es aquella construcción en la que se utilizan materiales rústicos tales como: ramas, cartones, restos de asbesto, latas, plásticos, etc., con piso de madera o tierra.
- 7. Choza:** Es aquella construcción que tiene paredes de adobe o paja, piso de tierra y techo de paja.
- 8. Otro:** Constituyen aquellas viviendas improvisadas, así como locales que no han sido construidos para viviendas tales como barcos, balsas, vagones, carpas, casetas, etc., que al momento del censo se encuentran habitadas.

Abastecimiento de agua.- Es la forma como se abastece de agua la vivienda, diferenciando el sistema y el medio de abastecimiento. En el trabajo solo se utiliza los datos del sistema de abastecimiento.

Sistema de abastecimiento.- Tiene las siguientes categorías:

- ❖ **Por tubería dentro de la vivienda.-** Si la tubería esta dentro de la vivienda y puede abastecerse de agua en forma directa.

- ❖ **Por tubería fuera de la vivienda, pero dentro del edificio, lote o terreno.-** Cuando para abastecerse de agua tiene que salir de la vivienda al edificio o lote, en los que se encuentran ubicadas las llaves o grifos de agua.

- ❖ **Por tubería fuera del edificio, lote o terreno.-** Cuando para abastecerse de agua, debe desplazarse a otros lugares distintos del lote o terreno en el que está ubicada la vivienda.

- ❖ **No recibe agua por tubería.-** Cuando la vivienda se abastece de agua que no es entubada, por ejemplo: aprovisionamiento en forma manual y directa de un río, acequia, pozo, carro repartidor, etc.

Eliminación de aguas servidas.- Es la forma de evacuación de los excrementos y aguas servidas de la vivienda. Se distinguen las categorías:

- ❖ **Conectado a red pública de alcantarillado.-** Si la eliminación de las aguas servidas, es a través de un sumidero subterráneo público.
- ❖ **Pozo ciego.-** Si las aguas servidas se eliminan en una excavación.
- ❖ **Otra forma.-** Si la vivienda tiene infraestructura de eliminación de aguas servidas que no están conectadas a red pública o pozo ciego.
- ❖ **Ninguno.-** Si la vivienda no dispone de ningún sistema de eliminación de aguas servidas enunciadas anteriormente.

Servicio eléctrico.- Es la disponibilidad o no de servicio eléctrico, sea proveniente de una red de servicio público, cooperativas, empresas particulares o sistema de alumbrado eléctrico de uso exclusivo.

Servicio telefónico.- Se refiere a la disponibilidad de servicio telefónico dentro de la vivienda.

Eliminación de la basura.- Es la forma de evacuar o eliminar la basura de la vivienda, comprende las siguientes categorías:

- ❖ Por carro recolector
- ❖ Eliminación en terreno baldío o quebrada
- ❖ Por incineración o entierro
- ❖ Otra forma, diferente a las anteriores

Servicio higiénico.- Para conocer la disponibilidad de servicio higiénico dentro de la vivienda, se categorizó por medio de:

- ❖ **Escusado de uso exclusivo.-** Si es utilizado en forma exclusiva por los miembros del hogar o de la vivienda que se investiga.
- ❖ **Escusado de uso común.-** Cuando el servicio es utilizado por varios hogares.
- ❖ **Letrina.-** Si el lugar de depósitos de excrementos es un orificio generalmente cubierto por una caseta.

❖ **Ninguno.**- Si no dispone de los servicios o categorías enunciadas.

Servicio de ducha.- Si la vivienda dispone de ducha en forma exclusiva o comparte con otros hogares o viviendas.

APÉNDICE B

TABLAS DE DATOS DE VARIABLES A SER UTILIZADAS

TABLA B.I

X₁: PROPORCIÓN DE VIVIENDAS POR PROVINCIAS

	Vivendas	%prov
Nacional	2339281	100
Azuay	148855	6,36
Bolivar	42623	1,82
Cañar	53420	2,28
Carchi	33223	1,42
Cotopaxi	76693	3,28
Chimborazo	100068	4,28
El Oro	96953	4,14
Esmeraldas	69698	2,98
Guayas	570124	24,37
Imbabura	65211	2,79
Loja	99974	4,27
Los Rios	115427	4,93
Manabi	216989	9,28
Morona Santiago	21468	0,92
Napo	23365	1,00
Pastaza	9772	0,42
Pichincha	443340	18,95
Tungurahua	97445	4,17
Zamora Chinchipe	17596	0,75
Galapagos	2573	0,11
Sucumbios	18378	0,79
Zonas No Delimitadas	16086	0,69

TABLA B.V

X₅ PROPORCIÓN DE LAS VIVIENDAS QUE DISPONEN DE SERVICIO ELÉCTRICO DEL TOTAL DE VIVIENDAS PROVINCIALES.

DISPONIBILIDAD DE SERVICIO ELECTRICO		
	dispone	%prov
Nacional	1559822	66,6795
Azuay	86002	57,7757
Bolivar	17855	41,8905
Cañar	27716	51,8832
Carchi	23323	70,2014
Cotopaxi	40586	52,9201
Chimborazo	62037	61,9948
El Oro	73196	75,4964
Esmeraldas	37044	53,1493
Guayas	454992	79,8058
Imbabura	41985	64,3833
Loja	47521	47,5334
Los Rios	56684	49,1081
Manabi	116853	53,8520
Morona Santiago	7291	33,9622
Napo	6514	27,8793
Pastaza	4974	50,9005
Pichincha	361185	81,4691
Tungurahua	73139	75,0567
Zamora Chinchipe	6803	38,6622
Galapagos	2058	79,9845
Sucumbios	5655	30,7705
Zonas No Delimitadas	6409	39,8421

TABLA B.VI

**X₆ PROPORCIÓN DE LAS VIVIENDAS QUE DISPONEN DE SERVICIO
TELEFÓNICO DEL TOTAL DE VIVIENDAS PROVINCIALES.**

DISPONIBILIDAD DE SERVICIO TELEFONICO		
	dispone	%prov
Nacional	315739	13,4973
Azuay	20006	13,4399
Bolivar	3614	8,4790
Cañar	2990	5,5972
Carchi	3124	9,4031
Cotopaxi	5736	7,4792
Chimborazo	9254	9,2477
El Oro	5839	6,0225
Esmeraldas	6054	8,6860
Guayas	94034	16,4936
Imbabura	7292	11,1822
Loja	6951	6,9528
Los Rios	6104	5,2882
Manabi	12848	5,9210
Morona Santiago	784	3,6519
Napo	724	3,0987
Pastaza	681	6,9689
Pichincha	118532	26,7361
Tungurahua	9751	10,0067
Zamora Chinchipe	456	2,5915
Galapagos	274	10,6490
Sucumbios	368	2,0024
Zonas No Delimitadas	323	2,0080

TABLA B. XI**X₁₁ PROPORCIÓN DE LA POBLACIÓN POR PROVINCIAS**

	Población	%
Nacional	9.648.189	100
Azuay	506.090	5,2
Bolivar	155.088	1,6
Cañar	189.347	2,0
Carchi	141.482	1,5
Cotopaxi	276.324	2,9
Chimborazo	364.682	3,8
El Oro	412.572	4,3
Esmeraldas	306.628	3,2
Guayas	2.515.146	26,1
Imbabura	265.499	2,8
Loja	384.698	4,0
Los Rios	527.559	5,5
Manabi	1.031.927	10,7
Morona Santiago	84.216	0,9
Napo	103.387	1,1
Pastaza	41.811	0,4
Pichincha	1.756.228	18,2
Tungurahua	361.980	3,8
Zamora Chinchipe	66.167	0,7
Galapagos	9.785	0,1
Sucumbios	76.952	0,8
Zonas No Delimitadas	70.621	0,7

TABLA B. XII

X₁₂ PROPORCIÓN DE LA POBLACIÓN RURAL POR PROVINCIAS.

	Rural	%prov
Nacional	4.302.331	45
Azuay	287.471	56,8
Bolivar	122.438	78,9
Cañar	133.828	70,7
Carchi	83.974	59,4
Cotopaxi	210.905	76,3
Chimborazo	244.869	67,1
El Oro	121.823	29,5
Esmeraldas	171.668	56,0
Guayas	596.876	23,7
Imbabura	136.325	51,3
Loja	232.899	60,5
Los Rios	328.185	62,2
Manabi	598.836	58,0
Morona Santiago	60.417	71,7
Napo	79.758	77,1
Pastaza	26.684	63,8
Pichincha	476.231	27,1
Tungurahua	210.428	58,1
Zamora Chinchipe	49.863	75,4
Galapagos	1.772	18,1
Sucumbios	56.460	73,4
Zonas No Delimitadas	70.621	100,0

TABLA B. XIV**X₁₄ PROPORCIÓN DE LA POBLACIÓN ANALFABETA DE LA POBLACIÓN
TOTAL PROVINCIAL.**

	analfabetismo	%prov
Nacional	731093	7,578
Azuay	43995	8,693
Bolivar	21182	13,658
Cañar	23711	12,523
Carchi	8990	6,354
Cotopaxi	39804	14,405
Chimborazo	59890	16,423
El Oro	15924	3,860
Esmeraldas	27484	8,963
Guayas	126153	5,016
Imbabura	31185	11,746
Loja	24930	6,480
Los Rios	52798	10,008
Manabi	101182	9,805
Morona Santiago	5829	6,921
Napo	8468	8,191
Pastaza	3653	8,737
Pichincha	86889	4,947
Tungurahua	33175	9,165
Zamora Chinchipe	3842	5,807
Galapagos	200	2,044
Sucumbios	5132	6,669
Zonas No Delimitadas	6677	9,455

TABLA B. XV

**X₁₅ PROPORCIÓN DE LA POBLACIÓN ECONÓMICAMENTE ACTIVA DE
LA POBLACIÓN TOTAL PROVINCIAL.**

	PEA	%prov
Nacional	3359767	34,823
Azuay	195270	38,584
Bolivar	52256	33,694
Cañar	66355	35,044
Carchi	48734	34,445
Cotopaxi	94961	34,366
Chimborazo	129920	35,626
El Oro	145952	35,376
Esmeraldas	91027	29,686
Guayas	881612	35,052
Imbabura	90831	34,211
Loja	124135	32,268
Los Rios	168306	31,903
Manabi	306049	29,658
Morona Santiago	28974	34,404
Napo	37244	36,024
Pastaza	15846	37,899
Pichincha	665231	37,878
Tungurahua	139235	38,465
Zamora Chinchipe	23331	35,261
Galapagos	4776	48,809
Sucumbios	28316	36,797
Zonas No Delimitadas	21406	30,311

TABLA B. II

X₂ INDICADOR DE TIPO DE VIVIENDA.

Escala de juicio	TIPO DE VIVIENDA																Indicador
	8		7		6		5		4		3		2		1		
	Casa	%prov	Depart	%prov	Alquiler	%prov	Mediagua	%prov	Rancho	%prov	Covacha	%prov	Choza	%prov	Otro	%prov	
Nacional	1520465	64,997	182393	7,80	159637	6,82419	267818	11,449	157235	6,7215	9194	0,393	37117	1,58668	1692	0,0723	7,040553
Azuay	116158	78,0343287	7433	4,99	7933	5,32934735	10912	7,33062376	2121	1,42487656	1426	0,95797924	2457	1,650599577	106	0,0712102	7,398038359
Bolivar	34678	81,3598292	369	0,87	1195	2,80365061	2975	6,97979964	1622	3,80545715	246	0,57715318	1471	3,451188326	18	0,0422307	7,325575394
Cañar	43196	80,8611007	528	0,99	1728	3,23474354	4997	9,35417447	1106	2,07038562	679	1,27105953	1089	2,038562336	35	0,0655185	7,362242606
Carchi	25198	75,8450471	722	2,17	2802	8,43391626	3347	10,0743461	547	1,64644975	65	0,19564759	472	1,420702525	21	0,0632092	7,330253138
Cotopaxi	52833	68,8889468	1342	1,75	2787	3,6339692	10423	13,5905493	1754	2,28704054	332	0,43289479	7129	9,295502849	20	0,026078	6,821809031
Chimborazo	66350	66,3049127	3296	3,29	4039	4,03625535	12954	12,9451973	880	0,87940201	622	0,62157733	11821	11,81296718	28	0,027981	6,714753967
El Oro	60070	61,9578559	2660	2,74	6897	7,11375615	20993	21,6527596	4828	4,97973245	437	0,45073386	781	0,805544955	73	0,0752942	6,887718792
Esmeraldas	48700	69,8728801	1859	2,67	2550	3,65864157	4329	6,2110821	11149	15,9961548	718	1,03015868	143	0,20517088	56	0,0803466	6,982266349
Guayas	349754	61,3470052	52326	9,18	26731	4,68862914	86154	15,1114494	53089	9,31183392	965	0,16926142	89	0,015610639	286	0,0501645	6,965475932
Imbabura	45475	69,7351674	2408	3,69	4944	7,58154299	9667	14,8241861	753	1,15471316	143	0,21928816	1663	2,550183251	20	0,0306697	7,137476806
Loja	64274	64,2907156	3675	3,68	7448	7,44993698	17428	17,4325325	718	0,71818673	195	0,19505071	5973	5,974553384	61	0,0610159	6,873877208
Los Rios	76892	66,6152633	3213	2,78	2847	2,46649397	10621	9,20148665	21386	18,5277275	135	0,11695704	27	0,023391408	181	0,1568091	6,878789191
Manabi	153644	70,8072759	7801	3,60	2572	1,18531354	13997	6,4505574	38303	17,6520469	359	0,16544617	84	0,03871164	36	0,0165907	7,021872076
Morona Santiago	15700	73,1321036	184	0,86	1177	5,48257872	1313	6,11607975	2718	12,6607043	135	0,62884293	96	0,447177194	67	0,3120924	7,0826812
Napo	16131	69,0391611	137	0,59	1434	6,13738498	1403	6,0047079	3694	15,8099722	145	0,62058635	240	1,027177402	22	0,0941579	6,905157287
Pastaza	5983	61,2259517	215	2,20	1015	10,3868195	688	7,04052395	1631	16,6905444	32	0,32746623	137	1,401964797	7	0,0716332	6,733524355
Pichincha	235408	53,0987504	87953	19,84	70154	15,8239726	41999	9,47331619	3832	0,8643479	1714	0,38661073	1139	0,25691343	536	0,1209004	7,112234402
Tungurahua	71662	73,5409718	5659	5,81	7237	7,42675355	9934	10,1944687	356	0,36533429	373	0,38278003	2074	2,128380112	56	0,0574683	7,314361948
Zamora Chinchipe	11642	66,1627643	287	1,63	1690	9,60445556	1131	6,42759718	2389	13,5769493	238	1,35258013	102	0,579677199	34	0,1932257	6,902023187
Galapagos	1790	69,568597	146	5,67	399	15,5071901	115	4,46949087	62	2,40963855	0	0	0	0	4	0,1554606	7,214535562
Sucumbios	12990	70,6823376	120	0,65	1628	8,85841767	1377	7,49265426	1848	10,0555011	160	0,87060616	92	0,500598542	17	0,0925019	7,045706823
Zonas No Delimitadas	11937	74,2073853	60	0,37	430	2,67313192	1061	6,59579759	2449	15,2244187	75	0,46624394	38	0,236230262	8	0,0497327	7,081064279

TABLA B. III

X₃ INDICADOR DE ABASTECIMIENTO DE AGUA.

ABASTECIMIENTO DE AGUA									
<i>Escala de juicio</i>	4		3		2		1		Indicador
	dentro de la vivienda	%prov	fuera de la vivienda	%prov	fuera del edificio	%prov	no recibe	%prov	
Nacional	815284	34,85191	322046	13,7669	122298	5,228017	749027	32,020	2,23183833
Azuay	52619	35,34917	18930	12,7171	6414	4,308891	34759	23,351	2,115165765
Bolivar	10758	25,23989	6547	15,3603	3064	7,188607	13149	30,850	1,922670858
Cañar	13130	24,57881	6858	12,8379	3046	5,701984	18398	34,440	1,826731561
Carchi	13562	40,82112	8966	26,9873	1790	5,387834	5068	15,254	2,702766156
Cotacachi	15749	20,53512	17021	22,1937	6039	7,874252	21465	27,988	1,924582426
Chimborazo	30576	30,55522	17950	17,9378	11100	11,09246	23794	23,778	2,21997042
El Oro	38473	39,68211	19671	20,2376	4778	4,928161	21679	22,360	2,516580199
Esmeraldas	14822	21,26603	4866	6,98155	2981	4,277024	36364	52,174	1,667364917
Guayas	202583	35,53315	61117	10,7199	26396	4,62987	222568	39,039	2,225906996
Imbabura	27393	42,00672	10533	16,1522	5843	8,960145	12867	19,731	2,541350386
Loreto	27179	27,18607	18012	18,0167	7745	7,747014	25778	25,735	2,04023046
Los Rios	24748	21,44039	10491	9,08886	4415	3,824928	62552	54,192	1,748698311
Manabi	60133	27,71246	19352	8,91842	11949	5,506731	95685	44,097	1,927152989
Morona Santiago	4365	20,33259	2761	12,861	1057	4,923607	8281	38,574	1,68334265
Napo	2575	11,02076	1973	8,44425	854	3,65504	13110	56,110	1,328354376
Pastaza	3132	32,05076	1288	13,1805	266	2,722063	3515	35,970	2,091588211
Pichincha	231661	52,25358	69382	15,6498	17731	3,999414	79554	17,944	2,819068886
Tunja	34127	35,02181	19760	20,2781	4288	4,400431	23557	24,175	2,338970701
Zamora Chinchipe	3312	18,82246	3323	18,885	1527	8,678109	5469	31,081	1,80381905
Galapagos	1045	40,61407	793	30,8201	85	3,303537	247	9,600	2,711232025
Sucumbios	1298	7,062792	1219	6,63293	447	2,432256	11868	64,577	1,175916857
Zonas No Delimitadas	2044	12,7067	1283	7,97588	483	3,002611	9350	58,125	1,388847445

TABLA B. IV

X₄ INDICADOR DEL SISTEMA DE ELIMINACIÓN DE AGUAS SERVIDAS.

ELIMINACIÓN DE AGUAS SERVIDAS									
<i>Escala de juicio</i>	4		3		2		1		
	red publica	%prov	pozo ciego	%prov	otra forma	%prov	ninguna	%prov	Indicador
Nacional	793178	39,425	503252	25,014	155953	6,66671	556272	23,78	2,69855921
Azuay	47166	41,756	9801	8,6768	5973	4,01263	49782	2,1281	2,03207346
Bolívar	8688	25,891	4524	13,482	2356	5,52753	17950	0,7673	1,5582993
Cañar	8509	20,513	7685	18,527	2019	3,77948	23219	0,9926	1,46186069
Carchi	14254	48,435	2717	9,2323	3150	9,48138	9265	0,3961	2,40793819
Cotopaxi	14320	23,736	9370	15,531	4972	6,48299	31612	1,3514	1,55852197
Chimborazo	27123	32,488	10182	12,196	6071	6,06687	40044	1,7118	1,80387902
El Oro	36067	42,563	25996	30,678	3836	3,95656	18652	0,7973	2,70996438
Esmeraldas	13803	23,317	13325	22,509	4338	6,22399	27567	1,1784	1,74421328
Guayas	204474	39,834	194948	37,978	37961	6,65838	75281	3,2181	2,89802339
Imbabura	25208	44,415	5520	9,7258	4927	7,55547	20981	0,8969	2,22843736
Loja	25342	32,15	7861	9,9729	2522	2,52266	42939	1,8356	1,65401334
Los Rios	17711	17,31	36866	36,031	12241	10,605	35388	1,5128	2,00056247
Manabí	43019	22,97	83657	44,668	26634	12,2744	33809	1,4453	2,51877363
Morona Santiago	3503	21,199	2241	13,562	1793	8,35197	8927	0,3816	1,42569408
Napo	2664	14,293	2317	12,431	1960	8,38862	11571	0,4946	1,11735912
Pastaza	2925	35,433	824	9,9818	1083	11,0827	3369	0,144	1,93985806
Pichincha	260059	65,199	60249	15,105	24044	5,42338	53976	2,3074	3,19262863
Tungurahua	33014	40,354	14573	17,813	5900	6,0547	28245	1,2074	2,28172002
Zamora Chinchipe	3237	23,635	666	4,8629	1348	7,66083	8380	0,3582	1,24809818
Galapagos	335	15,096	1208	54,435	418	16,2456	209	0,0089	2,56188294
Sucumbios	1290	8,6283	2767	18,507	1548	8,42311	9227	0,3944	1,07276144
Zonas No Delimitadas	467	3,5425	5955	45,172	859	5,34005	5879	0,2513	1,60617216

TABLA B.VII

X₇ INDICADOR DEL SISTEMA DE ELIMINACIÓN DE BASURA.

ELIMINACIÓN DE LA BASURA									
<i>Escala de juicio</i>	4		3		2		1		
	por carro recolector	% prov	en terreno baldío	% prov	por incineración	% prov	otra forma	% prov	Indicador
Nacional	867226	37,0723	590689	25,2509	450928	19,2764	99812	4,26678	2,668614
Azuay	43126	28,9718	47987	32,2374	13024	8,74945	8585	5,76736	2,358658
Bolivar	6725	15,7779	21112	49,5319	3707	8,69718	1974	4,6313	2,33733
Cañar	8540	15,9865	21694	40,6103	7397	13,8469	3801	7,11531	2,205859
Carchi	13350	40,183	12387	37,2844	1784	5,36977	1865	5,61358	2,889384
Cotopaxi	13512	17,6183	33662	43,8919	9782	12,7547	3318	4,32634	2,319847
Chimborazo	24864	24,8471	45717	45,6859	7147	7,14214	5692	5,68813	2,564186
El Oro	29111	30,0259	25076	25,8641	26896	27,7413	3468	3,57699	2,567553
Esmeraldas	13987	20,068	24004	34,44	13408	19,2373	7634	10,953	2,330196
Guayas	237926	41,7323	56603	9,92819	195823	34,3474	22312	3,91353	2,693223
Imbabura	26237	40,234	25503	39,1084	2786	4,27229	2110	3,23565	2,900416
Loja	25891	25,8977	40394	40,4045	6716	6,71775	5663	5,66447	2,439044
Los Rios	28155	24,392	28697	24,8616	41124	35,6277	4230	3,66465	2,47073
Manabi	60261	27,7715	49051	22,6053	71811	33,0943	5996	2,76327	2,478536
Morona Santiago	4070	18,9584	9961	46,3993	653	3,04174	1780	8,29141	2,294066
Napo	4286	18,3437	11167	47,7937	1011	4,32698	2048	8,76525	2,34175
Pastaza	3203	32,7773	3942	40,3397	340	3,47933	716	7,32706	2,664142
Pichincha	281252	63,4393	80209	18,092	27336	6,16592	9531	2,14982	3,22515
Tungurahua	32658	33,5143	31022	31,8354	14262	14,6359	3790	3,88937	2,627246
Zamora Chinchipe	3190	18,1291	7703	43,777	639	3,63151	2099	11,9288	2,230393
Galapagos	1828	71,0455	154	5,98523	169	6,56821	19	0,73844	3,160124
Sucumbios	2929	15,9375	8793	47,8452	1315	7,15529	1795	9,76711	2,313636
Zonas No Delimitadas	2125	13,2102	5851	36,3732	3798	23,6106	1386	8,61619	2,177981

TABLA B.VIII

X₈ INDICADOR DE DISPONIBILIDAD DE SERVICIO HIGIÉNICO

DISPONIBILIDAD DE SERVICIO HIGIENICO									
<i>Escala de juicio</i>	4		3		2		1		
	de uso exclusivo	%prov	de uso común	%prov	letrina	%prov	ninguno	%prov	Indicador
Nacional	996916	42,61634	206576	8,8307	286503	12,2475	518660	22,1718	2,43624344
Azuay	40891	27,47036	10776	7,2393	8939	6,00517	52116	35,0113	1,78620805
Bolivar	8869	20,80801	1958	4,5938	6701	15,7216	15990	37,515	1,65971424
Cañar	11398	21,33658	2313	4,3298	4894	9,16136	22827	42,7312	1,59389742
Carchi	11128	33,49487	4118	12,395	3337	10,0442	10803	32,5166	2,23769678
Cotopaxi	15038	19,60805	2764	3,604	12148	15,8398	30324	39,5395	1,60463145
Chimborazo	25637	25,61958	5613	5,6092	13403	13,3939	38767	38,7407	1,84834313
El Oro	44501	45,89956	10176	10,496	11471	11,8315	18403	18,9814	2,57730034
Esmeraldas	23013	33,01816	4653	6,6759	11349	16,2831	20018	28,7211	2,13387759
Guayas	328496	57,61834	39719	6,9667	75542	13,2501	68907	12,0863	2,89960079
Imbabura	22616	34,68127	6528	10,011	6813	10,4476	20679	31,7109	2,2136296
Loja	23009	23,01498	8006	8,0081	4904	4,90528	42745	42,7561	1,68650849
Los Rios	42641	36,94196	4398	3,8102	19492	16,8869	35675	30,907	2,23879162
Manabi	106570	49,11309	7530	3,4702	43744	20,1595	29275	13,4915	2,60673583
Morona Santiago	3978	18,5299	1772	8,2541	1853	8,63145	8861	41,2754	1,57420347
Napo	3388	14,50032	2065	8,838	2269	9,71111	10790	46,1802	1,50117697
Pastaza	3008	30,78183	1191	12,188	844	8,63692	3158	32,3168	2,09281621
Pichincha	239640	54,05332	79470	17,925	31750	7,16155	47468	10,7069	2,95019173
Tungurahua	31732	32,56401	8189	8,4037	19626	20,1406	22185	22,7667	2,1851506
Zamora Chinchipe	2466	14,01455	1953	11,099	1304	7,41078	7908	44,942	1,49119118
Galapagos	1478	57,44267	390	15,157	100	3,88651	202	7,85076	2,90866693
Sucumbios	2194	11,93819	1951	10,616	2556	13,9079	8131	44,2431	1,51659593
Zonas No Delimitadas	5225	32,48166	1043	6,4839	3464	21,5343	3428	21,3105	2,13757304

TABLA B. IX

X₉ INDICADOR DE DISPONIBILIDAD DE DUCHA.

DISPONIBILIDAD DE DUCHA							
Escala de juicio	3		2		1		Indicador
	exclusivo	%prov	comun	%prov	no tiene	%prov	
Nacional	797043	34,0721	158266	6,76558	1053346	45,0286	1,607762
Azuay	39664	26,6461	9147	6,14491	63911	42,9351	1,3516308
Bolivar	7716	18,1029	1585	3,71865	24217	56,8167	1,1856275
Cañar	9612	17,9933	1592	2,98016	30228	56,5855	1,1652565
Carchi	9475	28,5194	3233	9,73121	16678	50,2002	1,5522078
Cotopaxi	14433	18,8192	2855	3,72263	42986	56,0494	1,1995228
Chimborazo	22171	22,1559	4271	4,2681	56978	56,9393	1,3194328
El Oro	38866	40,0875	9520	9,81919	36165	37,3016	1,7720236
Esmeraldas	16821	24,1341	3523	5,05466	38689	55,5095	1,3802118
Guayas	252346	44,2616	34440	6,04079	225878	39,6191	1,8448548
Imbabura	21238	32,5681	5834	8,94634	29564	45,3359	1,6093297
Loja	21083	21,0885	8671	8,67326	48910	48,9227	1,2953468
Los Rios	25186	21,8199	4278	3,70624	72742	63,0199	1,3589195
Manabi	65469	30,1716	6485	2,98863	115165	53,0741	1,4956611
Morona Santiago	3858	17,9709	1695	7,89547	10911	50,8245	1,2052823
Napo	2715	11,6199	1603	6,86069	14194	60,749	1,0933019
Pastaza	2745	28,0905	1253	12,8223	4203	43,0106	1,5292673
Pichincha	205881	46,4386	47372	10,6853	145075	32,7232	1,9340957
Tungurahua	29723	30,5023	6763	6,94033	45246	46,4323	1,5182
Zamora Chinchipe	2646	15,0375	1901	10,8036	9084	51,6254	1,1834508
Galapagos	1186	46,0941	333	12,9421	651	25,3012	1,8946755
Sucumbios	1625	8,84209	1378	7,4981	11829	64,365	1,0588747
Zonas No Delimitadas	2584	16,0637	534	3,31966	10042	62,427	1,1725724

TABLA X

X₁₀ INDICADOR DE TENENCIA DE VIVIENDA.

TENENCIA DE VIVIENDA											
<i>Escala de juicio</i>	5		4		3		2		1		
	propia	%prov	arrendada	%prov	gratuita	%prov	por servicio	%prov	otra	%prov	Indicador
Nacional	1367045	58,43868	454782	19,4411	110116	4,70726	66306	2,83446	10406	0,444838	3,90193354
Azuay	80969	54,39455	21732	14,59944	6665	4,47751	2979	2,00128	377	0,253267	3,48058849
Bolivar	27006	63,36016	3333	7,819722	1795	4,21134	1262	2,96084	122	0,28623	3,66921615
Cañar	33325	62,383	4303	8,055036	2468	4,61999	1211	2,26694	125	0,233995	3,6276301
Carchi	19093	57,46922	5849	17,60527	2408	7,24799	1585	4,77079	451	1,357493	3,90410258
Cotopaxi	49519	64,56782	6446	8,404939	2554	3,33016	1576	2,05495	179	0,233398	3,70792641
Chimborazo	66201	66,15601	11717	11,70904	3423	3,42067	1751	1,74981	328	0,327777	3,9170564
El Oro	57015	58,80684	19387	19,99629	4880	5,03337	2895	2,98598	374	0,385754	3,9547719
Esmeraldas	42866	61,50248	10038	14,40213	3174	4,55393	2795	4,01016	160	0,229562	3,87032626
Guayas	363732	63,79875	113356	19,88269	21108	3,70235	12754	2,23706	1714	0,300636	4,1440634
Imbabura	38744	59,41329	11569	17,74087	3504	5,37333	2205	3,38133	614	0,941559	3,91854135
Loja	57787	57,80203	15226	15,22996	4109	4,11007	1440	1,44037	102	0,102027	3,65242963
Los Rios	76315	66,11538	13646	11,82219	6231	5,39822	5805	5,02915	209	0,181067	4,04299687
Manabi	140860	64,91573	24760	11,41072	14911	6,87178	6270	2,88955	318	0,146551	3,96762509
Morona Santiago	11901	55,436	2437	11,35178	1100	5,12391	968	4,50904	58	0,27017	3,47247065
Napo	13921	59,58057	3005	12,86112	924	3,95463	623	2,66638	39	0,166916	3,66710892
Pastaza	5565	56,94842	1949	19,94474	341	3,48956	327	3,3463	19	0,194433	3,81876791
Pichincha	194374	43,8431	159840	36,05359	23390	5,27586	15849	3,57491	4875	1,099608	3,8750688
Tungurahua	58484	60,01745	17384	17,83981	4248	4,35938	1361	1,39669	255	0,261686	3,8757966
Zamora Chinchipe	9399	53,41555	3085	17,53239	736	4,18277	368	2,09138	43	0,244374	3,54182769
Galapagos	1015	39,44812	871	33,85154	112	4,3529	170	6,60707	2	0,07773	3,58997279
Sucumbios	10499	57,12809	2917	15,87224	822	4,47274	569	3,09609	25	0,136032	3,6887583
Zonas No Delimitadas	8455	52,56123	1932	12,01044	1213	7,54072	1543	9,59219	17	0,105682	3,52760164

TABLA B. XIII

X₁₃ INDICADOR NIVEL DE INSTRUCCIÓN.

Escala de juicio	NIVEL DE INSTRUCCION														
	2		3		4		5		6		7		8		Indicador
	ninguno	%prov	Centro de Alfabetización	%prov	primario	%prov	secundario	%prov	superior	%prov	postgrado	%prov	no declarado	%prov	
Nacional	795272	8,243	99380	1,030038	4139447	42,904	2105815	21,826	658094	6,8209	30345	0,315	306342	3,17512	3.4662327
Azuay	46867	9,261	5027	0,993302	244251	48,262	86351	17,062	30324	5,9918	1381	0,273	11317	2,23616	3,3995969
Bolivar	22670	14,618	2108	1,359228	75482	48,670	20508	13,223	4452	2,8706	125	0,081	4134	2,66558	3,1456528
Cañar	24420	12,897	1596	0,842897	95464	50,417	24859	13,129	5007	2,6444	145	0,077	6220	3,28497	3,1532372
Carchi	9610	6,792	837	0,591595	76820	54,297	24072	17,014	5182	3,6627	123	0,087	2566	1,81366	3,4201524
Cotopaxi	41146	14,890	3023	1,094006	128394	46,465	39353	14,242	9115	3,2987	210	0,076	9503	3,43908	3,1389420
Chimborazo	60379	16,557	8294	2,27431	151299	41,488	52438	14,379	17040	4,6726	610	0,167	13095	3,5908	3,1058045
El Oro	18419	4,464	3252	0,788226	187637	45,480	109023	26,425	26295	6,3734	866	0,210	4859	1,17773	3,6622650
Esmeraldas	31342	10,222	3191	1,040675	130147	42,445	54228	17,685	11471	3,741	656	0,214	17913	5,84193	3,1155537
Guayas	139529	5,548	20726	0,824048	991380	39,416	691269	27,484	214752	8,5384	9456	0,376	78556	3,12332	3,6563925
Imbabura	32068	12,078	1688	0,635784	123827	46,639	45169	17,013	13801	5,1981	432	0,163	6457	2,43202	3,3244570
Loja	28309	7,359	10431	2,711478	187696	48,790	62119	16,147	22055	5,7331	741	0,193	9847	2,55967	3,3705764
Los Rios	58953	11,175	9283	1,759614	238552	45,218	92773	17,585	21137	4,0066	729	0,138	20662	3,91653	3,2535034
Manabi	116012	11,242	13399	1,298445	477081	46,232	161346	15,635	45179	4,3781	2137	0,207	44679	4,32967	3,2153311
Morona Santiago	6477	7,691	1176	1,396409	40495	48,085	14303	16,984	1512	1,7954	58	0,069	2811	3,33785	3,1142063
Napo	9564	9,251	1207	1,167458	51136	49,461	15809	15,291	1854	1,7933	81	0,078	2500	2,4181	3,1002834
Pastaza	4024	9,624	643	1,537873	16931	40,494	9954	23,807	1569	3,7526	64	0,153	1119	2,67633	3,3113774
Pichincha	93458	5,322	7870	0,448119	638900	36,379	501600	28,561	200652	11,425	11493	0,654	48028	2,73472	3,7617650
Tungurahua	33504	9,256	3517	0,971601	171919	47,494	67767	18,721	21618	5,9722	755	0,209	11547	3,18996	3,4549146
Zamora Chinchipe	4396	6,644	600	0,906796	33936	51,288	9978	15,080	1745	2,6373	35	0,053	2501	3,77983	3,1653543
Galapagos	220	2,248	70	0,715381	3505	35,820	3226	32,969	917	9,3715	66	0,675	428	4,37404	3,8009198
Sucumbios	5876	7,636	680	0,883668	39537	51,379	10912	14,180	1392	1,8089	40	0,052	3531	4,58857	3,1014529
Zonas No Delimitadas	8029	11,369	762	1,078999	35058	49,642	8758	12,401	1025	1,4514	42	0,059	4069	5,76174	3,0143867

APÉNDICE C

TABLA GENERAL DE DATOS DE VARIABLES

TABLA C. I

TABLA GENERAL DE DATOS DE VARIABLES

Provincias	VARIABLES														
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
Azuay	0,30173	1,69218	0,181985	0,13092	0,145035	0,933545	-0,57038	-0,57793	-0,23707	-1,51052	0,111414	-0,14907	0,391022	0,017619	0,831987
Bolivar	-0,45204	1,34026	-0,243026	-0,67994	-0,8196	0,038008	-0,64273	-0,83953	-0,86269	-0,52208	-0,46763	0,955228	-0,69853	1,412458	-0,39258
Cañar	-0,37543	1,51834	-0,454851	-0,84499	-0,21279	-0,48221	-1,08873	-0,97565	-0,93947	-0,74	-0,41112	0,542895	-0,66598	1,093442	-0,05456
Carchi	-0,51874	1,36297	1,479354	0,774208	0,89959	0,204831	1,23005	0,355789	0,518859	0,708758	-0,49008	-0,02187	0,479215	-0,63948	-0,20451
Cotopaxi	-0,2103	-1,1063	-0,238805	-0,67956	-0,14982	-0,14248	-0,70204	-0,95345	-0,81032	-0,31924	-0,26763	0,824469	-0,72732	1,622256	-0,22443
Chimborazo	-0,04444	-1,6263	0,413385	-0,25963	0,401245	0,176776	0,126852	-0,44943	-0,35841	0,776638	-0,12187	0,366726	-0,86949	2,189105	0,091058
El Oro	-0,06654	-0,7862	1,068272	1,291123	1,221132	-0,40543	0,138274	1,05812	1,347291	0,974273	-0,04286	-1,50916	1,517999	-1,34028	0,028591
Esmeraldas	-0,25993	-0,3271	-0,806718	-0,36175	-0,1359	0,075386	-0,66693	0,141082	-0,12935	0,531764	-0,21764	-0,18979	-0,82767	0,093524	-1,39634
Guayas	3,29088	-0,4086	0,426492	1,612984	1,482824	1,484789	0,564595	1,724666	1,621775	1,966191	3,425756	-1,79821	1,492803	-1,0155	-0,05255
Imbabura	-0,29177	0,42673	1,122963	0,466995	0,546286	0,525979	1,267474	0,306016	0,734138	0,784419	-0,28549	-0,42112	0,068634	0,875236	-0,2631
Loja	-0,04511	-0,8535	0,016535	-0,51613	-0,47694	-0,23749	-0,29768	-0,78412	-0,44919	-0,61005	-0,08885	0,03735	0,266509	-0,60402	-0,74977
Los Rios	0,06454	-0,8296	-0,627142	0,076989	-0,38131	-0,53799	-0,19019	0,358053	-0,2096	1,436586	0,146832	0,120501	-0,23579	0,387014	-0,84128
Manabi	0,78518	-0,1347	-0,23313	0,963902	-0,09323	-0,42375	-0,16371	1,118995	0,305749	1,041626	0,978888	-0,08781	-0,39957	0,330031	-1,40346
Morona Santiago	-0,60215	0,16061	-0,771441	-0,90689	-1,30105	-0,83336	-0,7895	-1,01637	-0,78862	-1,55306	-0,58455	0,595844	-0,83345	-0,4801	-0,21477
Napo	-0,58869	-0,7016	-1,555223	-1,4346	-1,67044	-0,93324	-0,62774	-1,1674	-1,21065	-0,53313	-0,55293	0,865352	-0,89318	-0,12356	0,190812
Pastaza	-0,68514	-1,5351	0,129928	-0,02691	-0,27246	-0,23459	0,465942	0,056163	0,432402	0,26159	-0,65451	0,200901	0,012516	0,029929	0,660452
Pichincha	2,39127	0,30414	1,73614	2,117198	1,583827	3,333748	2,3691	1,829292	1,958102	0,556616	2,173768	-1,62939	1,944904	-1,03467	0,655263
Tungurahua	-0,06305	1,2858	0,676127	0,558187	1,194433	0,313782	0,340776	0,247119	0,390692	0,56043	-0,12632	-0,08274	0,628363	0,150152	0,802131
Zamora Chinchipe	-0,62963	-0,7168	-0,50544	-1,21085	-1,01564	-1,02479	-1,0055	-1,18805	-0,8709	-1,18962	-0,61433	0,776301	-0,614	-0,79334	-0,0003
Galapagos	-0,73622	0,80098	1,498046	1,037683	1,493673	0,429742	2,148508	1,743415	1,809537	-0,93733	-0,70734	-2,07856	2,112898	-1,85039	3,392841
Sucumbios	-0,62408	-0,019	-1,891791	-1,51093	-1,49487	-1,13113	-0,72311	-1,13551	-1,34039	-0,41968	-0,59654	0,677121	-0,88817	-0,55101	0,384427
Zonas No Delimitadas	-0,64034	0,15276	-1,42166	-0,59801	-0,94399	-1,13013	-1,18331	0,148724	-0,91189	-1,26417	-0,60698	2,005047	-1,26172	0,231575	-1,2399

MATRIZ DE CORRELACIÓN DE LAS VARIABLES

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
X1	1
X2	-0,035	1
X3	0,347	0,224	1
X4	0,673	0,145	0,821	1
X5	0,542	0,201	0,896	0,907	1
X6	0,746	0,233	0,711	0,756	0,744	1
X7	0,38	0,126	0,85	0,792	0,785	0,725	1
X8	0,606	0,057	0,657	0,933	0,812	0,621	0,759	1
X9	0,576	0,07	0,843	0,95	0,905	0,726	0,871	0,926	1
X10	0,551	-0,286	0,381	0,614	0,547	0,364	0,428	0,599	0,548	1
X11	0,996	-0,059	0,32	0,67	0,525	0,701	0,365	0,622	0,574	0,583	1
X12	-0,578	-0,064	-0,745	-0,832	-0,83	-0,68	-0,774	-0,798	-0,916	-0,482	-0,574	1	.	.	.
X13	0,529	0,187	0,814	0,851	0,852	0,699	0,81	0,783	0,915	0,343	0,513	-0,921	1	.	.
X14	-0,212	-0,081	-0,254	-0,388	-0,28	-0,174	-0,4	-0,465	-0,513	0,038	-0,224	0,579	-0,652	1	.
X15	-0,053	0,241	0,452	0,244	0,409	0,304	0,552	0,243	0,421	-0,229	-0,083	-0,5	0,594	-0,407	1

BIBLIOGRAFÍA

1. **ABASCAL, Elena.** Métodos Multivariantes para la Investigación Comercial, Ariel Economía, España, 1.989
2. **AGUINAGA, Consuelo.** Determinación de los Niveles de Pobreza Area Urbana del Ecuador, Instituto Nacional de Estadísticas y Censos (INEC), Ecuador, 1.991.
3. **CAPA, Holger.** Modelo para la Evaluación de la Pobreza - Aplicaciones al Ecuador.
4. **DIAZ, Coral.** Indicadores Sintéticos, España.
<http://www.festadistica.fguam.es/>
5. **INSTITUTO NACIONAL DE ESTADÍSTICAS Y CENSOS (INEC).** Compendio de las Necesidades Básicas Insatisfechas de la Población Ecuatoriana, Guayaquil, 1.991
6. **INSTITUTO NACIONAL DE ESTADÍSTICAS Y CENSOS (INEC).** Resultados del V Censo de Población y IV de Vivienda de 1.990, Ecuador.
7. **KAMINSKY, Mario.** Vinculación de la Información del Censo de Población y del Censo de Viviendas en Indicadores Sintéticos Alternativos de Condiciones de

Pobreza para la Caracterización de Micro Regiones Geográficas - Análisis Cuantitativo Comparativo Multivariante Intra - País e Inter - países: Brasil, Costa Rica, Chile, México, Lima, 1.979

8. **LEE, Haeduck.** Metodología de la Investigación de la Pobreza y la Distribución del Ingreso, Instituto Nacional de Estadísticas y Censos (INEC), Ecuador, 1.994
9. **MEDINA, Fernando.** La evaluación del Bienestar: Alternativas de medición para los Estados de la República mexicana, Estadísticas, México, 1.995, 81 p. FMedina@cap.inegi.gob.mx
10. **PLA, Laura.** Análisis multivariado: Método de componentes principales, Secretaría General de la Organización de los Estados Americanos, Estados Unidos, 1986