

UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES
ESCUELA DE ESTADÍSTICA
MÉRIDA-EDO MÉRIDA

**ANÁLISIS ESTADÍSTICO MULTIVARIADO DE LOS
PRESUPUESTOS FAMILIARES DE LA REGIÓN DE LOS ANDES,
1988/89**

AUTOR: Nieves Juana

TUTOR: Prof. Giampaolo Orlandoni Merli

ASESOR: Prof. Daniel Anido

Mérida, Octubre de 2005

**ANÁLISIS ESTADÍSTICO MULTIVARIADO DE LOS
PRESUPUESTOS FAMILIARES DE LA REGIÓN DE LOS ANDES,
1988/89**

AUTOR: Nieves Juana

TUTOR: Prof. Giampaolo Orlandoni Merli

ASESOR: Prof. Daniel Anido

TRABAJO DE GRADO

Presentado ante la ilustre Universidad de Los Andes como requisito final para optar
Al título de Licenciado en Estadística.

Mérida, Octubre de 2005

RESUMEN

Una de las necesidades que impulsó la realización de este trabajo fue el rescatar los datos de la I Encuesta de Presupuestos Familiares y almacenarlos en un dispositivo adecuado, para luego aprovechar al máximo esta información. Por ende, se diseñó e implementó una base de datos, con tecnología referencial, que permitiera de forma eficiente realizar cualquier tipo de consulta a los datos, así como también que sirviera de materia prima para análisis estadísticos de mayor envergadura. Es decir, se trataba de retribuir en investigación el alto costo económico que significó la realización de esta encuesta y evitar a toda costa la pérdida de esta valiosa información.

Por otro lado, aprovechando las bondades de la base de datos, se realizó una serie de análisis estadísticos como: análisis de correspondencias múltiples, los cuales permitieron describir de forma clara la relación existente entre distintas variables socioeconómicas de carácter cualitativo. Por último, con la idea de relacionar las variables cualitativas antes mencionadas con variables cuantitativas de gastos del hogar, se realizó análisis de componentes principales sobre estas últimas variables, categorizando a los hogares con las variables socioeconómicas del tipo categóricas o cualitativas.

ÍNDICE GENERAL

Índice del contenido	i
Índice de tablas	ii
Índice de figuras y gráficos	iv
Introducción	v
Resumen	vi
Exposición y justificación del tema a investigar	vii
Muestra	viii
Limitaciones de la investigación	xi
Antecedentes de la investigación	xi
Objetivos	xiii
Capítulo I: Metodología estadística	1
1.1 Análisis de Componentes Principales (ACP)	1
1.1.1 Objetivos del Análisis de componentes principales	2
1.1.2 ¿Cuándo es apropiado el ACP ?	2
1.1.3 Propiedades del Análisis de Componentes Principales	2
1.1.4 Cuantos componentes vamos a retener	3
1.1.4.1 Calidad de representación (CR)	3
1.1.5 Ajuste en R^p (n puntos individuos)	4
1.1.6 Ajuste en R^n (p puntos variables)	9
1.1.7 Relación entre los ajustes	11
1.1.8 Interpretación geométrica	13
1.1.8.1 Variabilidad captada por las componentes	13
1.1.8.2 Los n-individuos en R^p	15
1.1.8.3 Las p-variables en R^n	15
1.1.8.4 Biplot	15
1.2 Análisis Factorial de Correspondencia Múltiple (A.F.C.M)	16
1.2.1 Objetivos del Análisis de Factorial de correspondencias múltiples	16
1.2.2 ¿Cuándo es apropiado el AFCM ?	17
1.2.3 Descripción de la técnica de A.F.C.M.	17
1.2.4 Generalización del Análisis Factorial de Correspondencias Múltiples.	18
1.2.5 Propiedades del Análisis del A.F.C.M	20
1.2.6 Interpretación geométrica	22
1.2.6.1 Porción de varianza explicada por los ejes	22
1.2.6.2 Distancia entre puntos	22
1.2.6.3 Centros de gravedad G	23
1.2.6.4 Contribuciones absolutas	23
1.2.6.5 Coseno cuadrado	23
Capítulo II: Diseño de la base de datos	26
2.1 Marco teórico	26
2.1.1 Diseño conceptual	26
2.1.2 Modelo Entidad Relación	26
2.2 Descripción de las tablas	28

Capítulo III: Análisis estadístico	32
3.1 Estadísticas básicas	32
3.2 Análisis Factorial de correspondencias Múltiples de las características de las viviendas vs las variables suplementarias ingreso y número de miembros	43
3.3 Análisis Factorial de Correspondencias Múltiples de las características del jefe de la familia vs el tipo de vivienda	54
3.4 Análisis de componentes principales gastos vs ingresos familiares	64
Anexo 1 Definiciones básicas de la encuesta de prepuestos familiares	
Anexo 2 Manual del usuario	
Anexo 3 Algunas salidas de los análisis estadísticos	
Anexo 4 Programas en SAS	
Conclusiones y recomendaciones	
Bibliografía	115

ÍNDICE DE TABLAS

Capitulo II	
Tabla 2.1 Categoría	28
Tabla 2.2 Desconsulta	28
Tabla 2.3 Entidad	28
Tabla 2.4 Gastos	29
Tabla 2.5 Hogar	29
Tabla 2.6 Ingresos	29
Tabla 2.7 Miembros	30
Tabla 2.8 Vivienda	30
Capitulo III	
Tabla 3.1 Distribución de frecuencias del tipo de vivienda	32
Tabla 3.2 Distribución de frecuencias del tenencia de la vivienda principal	33
Tabla 3.3 Distribución de frecuencias del tenencia de la vivienda ocasional	33
Tabla 3.4 Distribución de frecuencias del número de habitaciones que poseen las viviendas	34
Tabla 3.5 Distribución de frecuencias del número de habitaciones para dormir	34
Tabla 3.6 Distribución de frecuencias del material predominante en paredes exteriores	35
Tabla 3.7 Distribución de frecuencias del número material predominante en el piso	36
Tabla 3.8 Distribución de frecuencias del material predominante en el techo	36
Tabla 3.9 Distribución de frecuencias de la ubicación de la instalación para agua corriente por tubería	36
Tabla 3.10 Distribución de frecuencias del abastecimiento de agua	37
Tabla 3.11 Distribución de frecuencias de las habitaciones de baño con ducha o regadera que posee la viviendas	37
Tabla 3.12 Distribución de frecuencias del tipo de excusado que tienen las viviendas	38
Tabla 3.13 Distribución de frecuencias de la clase de alumbrado	38
Tabla 3.14 Distribución de frecuencias del combustible que se utiliza para cocinar	38
Tabla 3.15 Distribución de frecuencias del estado civil de los miembros de la familias	39
Tabla 3.16 Distribución de frecuencias del Género	39
Tabla 3.17 Distribución de frecuencias de la nacionalidad	39
Tabla 3.18 Distribución de frecuencias del nivel educativo	40
Tabla 3.19 Distribución de frecuencias de la dedicación de las personas	40
Tabla 3.20 Distribución de frecuencias de las personas con más de un trabajo remunerado	41
Tabla 3.21 Distribución de frecuencias de la categoría ocupacional	41
Tabla 3.22 Distribución de frecuencias de la actividad de las personas	42
Tabla 3.23 Codificación de las variables para el Análisis Factorial de Correspondencia Múltiples de las características de las	

viviendas vs las variables suplementarias ingreso y número de miembros.	43
Tabla 3.24 Inercia y descomposición Chi-cuadrada para las características de las viviendas vs las variables suplementarias ingreso y número de miembros.	45
Tabla 3.25 Calidad y masa para las características de las viviendas vs las variables suplementarias ingreso y número de miembros.	46
Tabla 3.26 Contribuciones parciales para las características de las viviendas vs las variables suplementarias ingreso y número de miembros	49
Tabla 3.27 Codificación de las variables para el Análisis Factorial de Correspondencia Múltiples de las características del jefe de la familia vs el tipo de vivienda.	54
Tabla 3.28 Inercia y descomposición Chi-cuadrada para las características del jefe de la familia vs el tipo de vivienda.	56
Tabla 3.29 Calidad y masa para las características del jefe de la familia vs el tipo de vivienda.	57
Tabla 3.30 Contribuciones parciales para las características del jefe de la familia vs el tipo de vivienda.	59
Tabla 3.31 Codificación de las variables para el análisis de Componentes Principales de los gastos vs los ingresos familiares en la Región de Los Andes en base al período 1988/89.	64
Tabla 3.32 Autovalores de la matriz de correlación	
Tabla 3.33 Autovectores propios de la matriz de correlación	
Tabla 3.34 Factores del modelo de la matriz de correlación	
Tabla 3.35 Comunalidad Final Estimada de la matriz de correlación	
Tabla 3.36 Autovalores de la matriz de varianzas-covarianzas	
Tabla 3.37 Autovectores propios de la matriz de varianzas-covarianzas	
Tabla 3.38 Factores del modelo de la matriz de varianzas-covarianzas	

ÍNDICE DE FIGURAS Y GRÁFICOS

Capítulo I	
Figura 1.1 Proyección en R^3	5
Figura 1.2 Representación gráfica del coseno cuadrado	23
Figura 1.3 Representación gráfica del coseno cuadrado cuando el $\text{Cos}^2\theta \rightarrow 1$	24
Figura 1.4 Representación gráfica del coseno cuadrado cuando el $\text{Cos}^2\theta \rightarrow 0$	25
Capítulo II	
Figura 2.1 Modelo Entidad-Relación de la base de datos para la I Encuesta de Presupuestos Familiares en la Región de Los Andes.	27
Capítulo III	
Gráfico 3.1 Plano factorial: Características de las viviendas vs las variables suplementarias ingreso y número de miembros	52
Gráfico 3.2 Plano factorial: Características del jefe de la familia vs el tipo de vivienda	62
Gráfico 3.3 SCREE (codo) de los autovalores o % de varianza explicada con la matriz de correlación	
Gráfico 3.4 SCREE (codo) de los autovalores o % de varianza explicada con la matriz de varianzas-covarianzas.	

INTRODUCCIÓN

La Encuesta sobre presupuestos Familiares o de gastos familiares, es una investigación estadística por muestreo que se realiza a los hogares con el fin de obtener información sobre sus ingresos, egresos, características de las viviendas que habitan, composición del hogar y otras variables económicas y sociales de los miembros. “Diversos han sido los autores y sus nacionalidades quienes, partiendo de datos tomados de este tipo de encuestas, se han abocado al estudio de la conducta del consumidor, intentando establecer parámetros para medir el grado de respuesta de los consumidores ante cambios en diversos tipos de variables: económicas, como los precios y el ingreso; demográficas, como el tamaño del hogar o familia, el sexo del jefe de la misma, ubicación geográfica/localización de esa familia o del consumidor individual, etc.”

En esta investigación, se tiene como propósito fundamental una base de datos con tecnología referencial que permita aprovechar al máximo la información recogida en la encuesta. Además nos permitirá integrar toda la información y así realizar análisis estadísticos de mayor alcance o envergadura que facilite la consulta de la información recolectada, la cual nos permitirá hacer comparaciones con otros estudios y realizar un seguimiento a posteriori.

El estudio se ha estructurado en cinco capítulos. El capítulo 1 corresponde al problema de investigación, el cual comprende la exposición y justificación del tema a investigar, la muestra, las limitaciones, los antecedentes y objetivos. En el capítulo 2 se muestran los aspectos teóricos de la base de datos, así como el diagrama entidad-relación. El capítulo 3 corresponde a la metodología estadística. En el capítulo 4 se realizan los análisis estadísticos, lo cual constituye el aporte más importante para investigaciones futuras. Por último en el capítulo 5 las conclusiones y recomendaciones. Se incluyen así mismo 4 anexos el primero relacionado con la encuesta, el segundo con la base de datos, el tercero con el análisis y el cuarto con los programas con los cuales se realiza los análisis multivariantes.

CAPÍTULO I. PLANTEAMIENTO DEL PROBLEMA

1.1. Exposición y justificación del tema a investigar

El Centro de Investigaciones Agroalimentarias (CIAAL) tiene como misión, *“producir y difundir conocimientos que apoyen los esfuerzos de la sociedad venezolana dirigidos a incrementar su capacidad para comprender y mejorar la realidad agroalimentaria y nutricional del país”*.

En este sentido, se requiere el análisis estadístico de los datos aportados por el Banco Central de Venezuela (BCV) a través de la I Encuesta de Presupuesto Familiar (EPF), ya que le permitirá al Centro de investigaciones Agroalimentarias difundir conocimientos, realizar estudios posteriores y hacer un seguimiento de esta información. Así contribuir al mejoramiento de la calidad de vida relacionada con la agroalimentación de las familias venezolanas. A partir de esto se propone condensar los aspectos teóricos y resultados empíricos más importantes en una publicación y una versión electrónica para difundir el estado del arte en el tema a nivel nacional, y proporcionar una herramienta didáctica a los estudiantes, tesistas, entre otros, sobre este tema de investigación de escasa difusión en el país. Debido a que esto corresponde a un proyecto más amplio, el cual es llevado a cabo por el CIAAL, para hacer comparaciones con otros años.

Otro motivo de la realización de este trabajo es la falta de una base de datos con tecnología referencial que permita aprovechar al máximo la información recogida en la encuesta. Además nos permitirá integrar toda la información y así realizar análisis estadísticos de mayor alcance o envergadura (análisis multivariante punto importante de esta investigación). También será una herramienta que facilite la consulta de la información recolectada para así atacar problemas sociales, tales como: pobreza y desigualdad, nutrición, equipamiento familiar, educación, economía informal, etc.

La Encuesta sobre presupuestos Familiares es una investigación estadística por muestreo que se realiza a los hogares con el fin de obtener información sobre sus ingresos, egresos, características de las viviendas que habitan, composición del hogar y otras variables económicas y sociales de los miembros que lo integran a los fines de:

- Actualizar el índice de precios al consumidor (IPC), en cuanto a la composición de la canasta y estructura de ponderaciones de los bienes y servicios que la componen.
- Elaborar las principales cuentas del sector hogares, según los requerimientos del sistema de cuentas nacionales.
- Estimar el nivel y estructura del ingreso de los hogares según su origen y naturaleza.

1.2. Muestra

La recolección de la información se realiza para cada región de Venezuela (región Capital, Central, Centro Occidental, Zuliana, Los Andes, Insular, Nor Oriental, Guayana y Los Llanos) tomando en cuenta el número de familias, que allí habitan. Aunque en esta investigación sólo nos interesa la Región de los Andes para 1988/89.

La encuesta en la región de los Andes fue aplicada a 1.354 familias quedando de ellas 1.179 y fue dividida en 4 Zonas:

1 = Mérida

2 = San Cristóbal

3 y 4 = Resto de Los Andes (en la cual el 4 se recodificó a 3 para un mejor manejo de la data).

Para obtener el tamaño de muestra se utilizó el muestreo bietápico, ya que las dos variables que explican principalmente el comportamiento del gasto son el ingreso familiar y el número de miembros. En la primera Fase se recoge información, entre otros aspectos, del ingreso familiar y el número de miembros del hogar. Estas variables se utilizan para estratificar bidimensionalmente a las familias, constituyendo el marco de referencia para la selección de la muestra de la segunda Fase.

La población que habita en los centros poblados de 10.000 y más habitantes en cada región se dividió en dos (2) dominios de estudio: uno constituido por la(s) ciudad(es) considerada(s) como cabeza(s) de región y el otro por el Resto de la Región. Los centros poblados localizados en este segundo dominio de estudio se clasificaron de acuerdo al nivel de población, estructurándose tres (3) estratos: E₁ (10.000 a 20.000 habitantes), E₂ (20.000 a 50.000 habitantes) y E₃ (más de 50.000 habitantes). De cada estrato se seleccionó una muestra aleatoria de centros poblados, aplicándose luego muestreo bifásico a nivel de cada estrato, pero con la condición adicional que la muestra prevista para todo el dominio de estudio resultara autoponderada. La selección de las unidades últimas correspondiente a la muestra de Primera Fase se realiza siguiendo el diseño y utilizando el marco muestral de la Encuesta Nacional de Hogares.

Para el cálculo de los tamaños de muestra, a nivel de cada dominio de estudio, se utilizó un estimador de razón de los gastos en alimentos a los gastos totales. Las variaciones y covarianzas que entran en la fórmula fueron calculadas basándose en la Encuesta de Presupuestos Familiares levantada por el. BCV en el Área Metropolitana de Caracas durante 1986. Previamente a la aplicación de la fórmula se clasificaron las regiones en cuatro (4) grupos y a cada grupo se le hizo corresponder una determinada precisión para la cabeza de región y otra para el resto.

Es conveniente señalar que la muestra de Primera Fase equivale a tres (3) veces la Segunda, de esta manera se cuenta con cierta holgura al realizar la selección de la Segunda Fase. En el caso del Área Metropolitana de Caracas, las 536 familias constituyen una sub- muestra de la utilizada en la encuesta levantada en 1986 (2.000 familias). Si bien esta muestra produce resultados con mayor precisión, se refiere a la realidad prevaleciente en 1986 y el resto de las encuestas caracterizan la situación de 1988 y principios de 1989. El procedimiento de la sub-muestra permite obtener mediciones para un mismo grupo de familias en dos ocasiones diferentes (1986 y 1988). La metodología que se utilizó para la actualización de las Ponderaciones correspondientes al Área Metropolitana de Caracas fue el de aplicar a los resultados obtenidos en 1986 la tasa de cambio observada entre esta fecha y 1988.

De acuerdo al diseño de muestra descrito, la Encuesta de Primera Fase abarca un total de 23.280 familias y a la de Segunda Fase un total de 8.272: la muestra está distribuida en doce (12) ciudades cabezas de región (4.260 encuestas) y 52 centros poblados (3.245 encuestas).

El relevamiento de información se desarrolló en forma conjunta entre el Banco Central de Venezuela (BCV) y la Oficina Central de Estadística e Informática (OCEI), actualmente Instituto Nacional de Estadística (INE), que contaba con el apoyo de las corporaciones regionales y universidades. Específicamente estaban incorporadas al estudio las siguientes instituciones: BCV, OCEI, CORPOCENTRO, FUDECO, CORPOZULIA, CONZUPLAN, UNIVERSIDAD DE LOS ANDES, UPESUROESTE, CORPORIENTE, CORPOLLANOS y CVG. Para la Región de Los Andes correspondiendo a la OCEI y a la Universidad de Los Andes la responsabilidad de la recolección de los datos y al BCV el procesamiento de información.

1.3. Limitaciones de la investigación

La falta de información del diseño completo del cuestionario establecido para la recolección de la muestra, constituye una limitación, pues debido a esto no se conoce algunas de las codificaciones de las variables importantes para llevar a cabo el estudio. Por ser la I Encuesta de Presupuesto Familiar diseñada nacionalmente hay fallos en la misma en cuanto a como se definen las variables y se recoge la información, lo cual dificulta el análisis de la información allí contenida.

1.4. Antecedentes

En el año 1993, la profesora Elsy Garnica Olmos -Universidad de los Andes-Mérida, Venezuela, desarrolló un trabajo donde resume algunos resultados de la Encuesta de Presupuestos Familiares, efectuada en la ciudad de Mérida en el año 1986 por el Instituto de Investigaciones Económicas y Sociales (IIES) de la Facultad de Ciencias Económicas y Sociales (FACES), y la Corporación de los Andes. Los objetivos generales de este trabajo son: a) determinar la estructura del gasto familiar, de acuerdo a las características demográficas y económicas y b) actualizar el sistema de ponderaciones de los índices de precios al consumidor, los cuales se transmiten mensualmente al Banco Central de Venezuela (BCV). Los objetivos específicos están referidos al estudio de la estructura de los gastos familiares de acuerdo al nivel de ingreso (Tinto Arandes, 1988).

La recolección de la información para esta encuesta consistió en calcular un tamaño de la muestra de 487 familias, con un diseño de muestreo estratificado y polietápico. Para obtener el tamaño de muestra se utilizó el muestreo aleatorio estratificado con afijación proporcional, omitiendo la corrección para población finita y tomando en cuenta la variabilidad del ingreso para cada estrato. Los cuatro estratos fueron: (1) zona residencial (manzanas con viviendas del tipo quinta o casa); (2) zona comercial (manzanas donde predominan casas y corresponden a la parte vieja de la ciudad); (3) zonas de transición (manzanas donde existen mejoramiento urbano y predominan casas construidas por el Instituto Nacional de la vivienda (INAVI), actualmente INREVI, y similares); (4) zona de ranchos (incluye a barrios y áreas sin mejoras). Las dos etapas del diseño fueron manzanas y familias.

El estudio de la profesora Garnica (1993), presenta análisis fundamentalmente descriptivos y se realizan primero con las variables cualitativas de la encuesta (características demográficas y de vivienda), y en segundo término, sobre los gastos de la canasta familiar además de observar las interrelaciones entre las variables que muestran la estructura del gasto en las familias merideñas.

Los análisis multivariantes aplicados al trabajo de Garnica fueron: Análisis Factorial de Correspondencias Múltiples (AFCM), Análisis de Componentes Principales (ACP), Análisis Factorial Discriminante (AFD) y el Análisis de Regresión Múltiple (ARM).

Otra de las investigaciones realizadas al considerar el gasto familiar fue en 1962 por el Instituto de Investigaciones Económicas y Sociales (IIES) de la Facultad de Ciencias Económicas y Sociales (FACES). En este estudio se calcula el Índice de Precios al Consumidor (IPC) para la ciudad de Mérida con una canasta de mercado familiar fija. Para obtener este indicador se utilizaron dos factores esenciales: en primer lugar las ponderaciones utilizadas en cada rubro, sub-grupo o grupo de la canasta familiar, las cuales son fijas, y en segundo lugar, los precios recabados mensualmente, de los diferentes rubros de esa canasta familiar. Las ponderaciones calculadas a partir de un Estudio de Presupuestos Familiares en los cuales se obtienen los ingresos de cada unidad familiar y la distribución del gasto. Se supone que la distribución del gasto familiar se mantiene en el tiempo (Vielma, 1990).

1.5. OBJETIVOS

1.5.1. Objetivos generales

- Realizar un análisis estadístico sobre la información recolectada en la I Encuesta Nacional de Presupuestos Familiares para la región de Los Andes, con base en el período 1988/89.
- Sistematizar las encuestas de los Presupuestos Familiares de la región de los Andes para el período 1988/89.

1.5.2. Objetivos específicos

- Diseñar e implementar un Sistema de Base de Datos con la información obtenida en la Encuesta Nacional de Presupuestos Familiares, tomando solo la región de los Andes para el período 1988/89.
- Realizar una completa descripción estadística de la información en base a las variables más importantes del estudio.
- Hacer un análisis estadístico exploratorio en base a las variables más importantes relacionadas con el objetivo en estudio.
- Determinar la estructura de las variables socioeconómicas, cuales están asociadas y cuales son aquellas que van en el mismo sentido o sentido opuesto.

CAPÍTULO II. DISEÑO DE LA BASE DE DATOS

En el sentido más simple, una base de datos es un conjunto de registros y archivos que están organizados para un uso determinado, los cuales deben ser eficientes y de fácil actualización. Actualmente todos los sistemas de gestión de bases de datos modernos almacenan y tratan la información utilizando el método de gestión de bases de datos relacional. El nombre relacional procede del hecho de que cada registro de la base de datos contiene información relacionada con un tema y sólo con este tema (Korth y Silberthatz, 1993).

2.1. Diseño conceptual

Su objetivo es crear un esquema conceptual, independiente de cualquier sistema de administración de base de datos, partiendo de especificaciones de requerimientos que describan la realidad. El propósito del diseño conceptual es describir el contenido de información de la base de datos, más que las estructuras de almacenamiento que se necesitarán para manejar esta información. (Batini, 1994)

2.2. Modelo Entidad Relación

El modelo entidad relación provee las herramientas metodológicas para diseñar o construir el nivel lógico conceptual de la base de datos, basándose en los conceptos de entidades, conjunto de entidades y las relaciones que existe entre esas entidades. Fue introducido por Peter Chen en 1976 y es el más ampliamente usado.

Entidad es cualquier objeto, acerca del cual se recolectan datos. Se representa gráficamente mediante un rectángulo.

Atributo es una característica de una entidad. Pueden existir muchos atributos para cada entidad. Por ejemplo, una Vivienda (entidad) puede tener numerosos atributos, tales como: piso, techo, etc. Los atributos identifican los datos. De hecho, un dato es la unidad más pequeña de una tabla.

Los datos pueden tener valores, los cuales pueden ser de longitud variable o fija y de tamaño fijo; pueden ser también alfabéticos, numéricos o alfanuméricos.

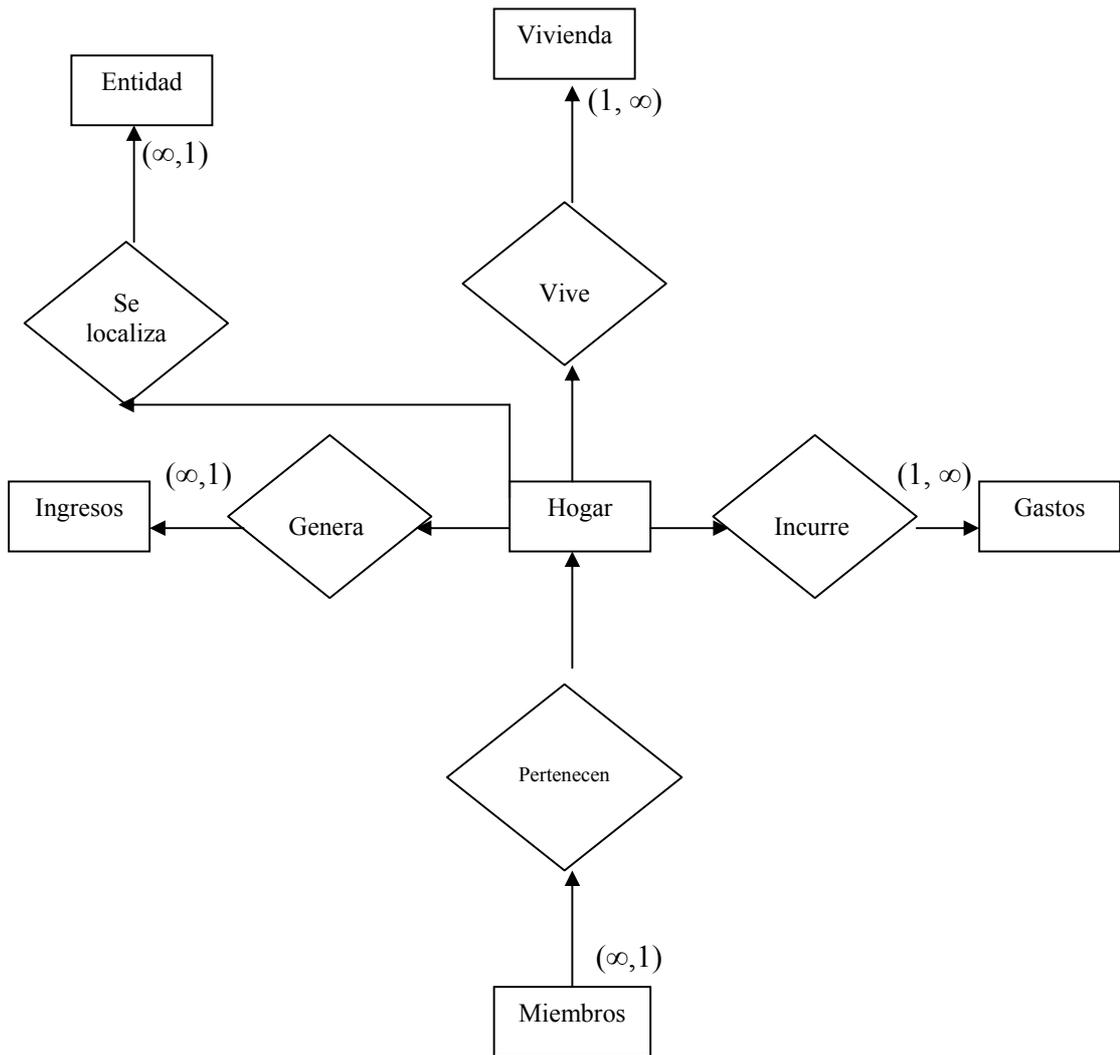
Registro es una colección de datos elementales que tienen algo en común con la entidad descrita. Esta constituido por el nombre del campo, Tipo de dato y longitud del dato.

Llave o clave es un dato o una combinación de ellos en un registro, que se utiliza como criterio de identificación para éste. Cuando una llave identifica de manera exclusiva a un registro se le denomina llave primaria.

Las relaciones son asociaciones entre entidades, (en algunos casos pueden referirse a asociaciones de datos) cada relación tiene un significado específico; se representa gráficamente por medio de un rombo (Kendall, 1991).

Cardinalidad de mapeo tipo de asociación que se establece entre los conjuntos de entidades, considerando el número de entidades individuales que están involucradas en la relación. La cardinalidad de mapeo entre dos entidades (o de una asociación), puede ser uno a uno o en símbolos $(1,1)$, una a varias $(1,\infty)$, varias a una $(\infty,1)$, y varias a varias (∞,∞) . (Ponsot, 1998).

Figura 1
Modelo Entidad-Relación de la base de datos para la I Encuesta de Presupuestos
Familiares en la Región de los Andes.



Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

2.3. Descripción de las tablas

Las tablas son los objetos principales de una base de datos, ya que contiene la información que desea organizarse. La primera fase en un diseño de una base de datos es definir las tablas y los tipos de datos que guardará cada una. Los datos almacenados en una tabla pueden relacionarse con los de otra y permiten crear aplicaciones personalizadas. Una tabla está conformada por campos (columnas) y registros (filas); la intersección de fila con columna es una celda.

La base de datos origen de este trabajo, consiste en presentar la información recogida por el Banco Central de Venezuela (BCV) y la Oficina Central de Estadística e Informática (OCEI), actualmente Instituto Nacional de Estadística (INE) de la encuesta de Presupuestos Familiares de la Región de los Andes en base al período 1988/89. Las tablas que a continuación se muestran, forman parte del diseño de la base de datos, donde cada una de ellas es esencial para lograr presentar la información que se recoge en la encuesta.

Tabla 1 Categoría

En esta tabla se identifica las siguientes tablas de la base de datos:

La unidad básica de muestreo, las características de las viviendas, las características generales del hogar, los gastos e ingresos de las familias en la Región.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Codigotabla	Código de la tabla	Texto	1
Categoría	Categoría	Texto	50

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

En el campo Categoría están los nombres de las tablas antes mencionadas.

El campo Codigotabla es la clave principal que identifica en forma unívoca las diferentes modalidades del campo Categoría de la variable de este campo.

Tabla 2 Desconsulta

En esta tabla, se especifican todas aquellas relacionadas con las identificadas en la tabla categoría de la base de datos.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Codigoconsulta	Código de la consulta	Texto	5
Descripción	Descripción de la consulta	Texto	250
Codigotabla	Código de la tabla	Texto	1

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

El campo Codigoconsulta es la clave principal que identifica en forma unívoca las diferentes modalidades del campo Descripción de la variable (Desconsulta) Descripción de la consulta.

Tabla 3 Entidad

En esta tabla, se mencionan las entidades de la región de Los Andes.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Codigo_entidad	Código de la entidad	Texto	5
Entidad	Nombre de la entidad	Texto	255

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

El campo Codigo_entidad es la clave principal que identifica en forma unívoca las diferentes entidades de la región de Los Andes del campo (Entidad) Nombre de la Entidad.

Tabla 4 Gastos

En esta tabla, se da a conocer todo lo relacionado al gasto en el hogar.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Num_gasto	Número del gasto	Autonúmerico	Entero largo
Codgasto	Código del gasto específico	Texto	255
CodEPF	Código de la encuesta	Texto	255
Hogar	Número del hogar	Texto	50
OG	Código del gasto general	Texto	255
Tigas	Tipo de gasto	Texto	255
Unimed	Unidad de medida	Texto	255
Forad	Forma de adquisición	Texto	255
Tipest	Tipo de establecimiento	Texto	255
Cantidad	Cantidad del gasto	Númerico	Doble
Valor	Valor del gasto	Númerico	Doble

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Los campos (Num_gasto) Número del gasto y (Codgasto) Código del gasto específico es la clave principal que identifica en forma unívoca los diferentes gastos en la región de Los Andes.

Tabla 5 Hogar

Esta tabla representa el número del hogar, donde está localizado y el número de miembros que lo componen.

Nombre del campo	Descripción	Tipo de dato	Tamaño
CodEPF	Código de la encuesta	Texto	255
Hogar	Número del hogar	Texto	50
Codigo_Entidad	Código de la Entidad	Texto	255
Numiemb	Número de miembros que posee el hogar	Texto	255

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Los campos (CodEPF) Código de la encuesta y (Hogar) Número del hogar es la clave principal que identifica en forma unívoca los diferentes hogares en la región de Los Andes.

Tabla 6 Ingresos

Esta tabla se refiere a todo lo relacionado con el ingreso familiar

Nombre del campo	Descripción	Tipo de dato	Tamaño
Numing	Numero del ingreso	Autonumérico	Entero largo
CodEPF	Código de la encuesta	Texto	255
Coding	Código del ingreso	Númerico	Doble
Hogar	Número del hogar	Texto	50
FMT	Código relacionado con el ingreso	Númerico	Doble
Ingreso	Valor del ingreso	Númerico	Doble

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Los campos (Numing) Número del ingreso, (CodEPF) Código de la encuesta y (Coding) Código del ingreso representan la clave principal que identifica en forma unívoca los diferentes ingresos en la región de Los Andes.

Tabla 7 Miembros

Esta tabla se refiere a todo lo relacionado a los miembros que componen el hogar.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Numiemb	Numero del miembro en el hogar	Autonúmerico	Entero largo
CodEPF	Código de la encuesta	Texto	255
Hogar	Número del hogar	Texto	50
Miembro	Identificación del miembro	Texto	255
PCJHogar	Parentesco con el jefe del hogar	Texto	255
Genero	Género	Texto	255
Ecivil	Estado civil	Texto	255
Edad	Edad	Texto	50
Neduca	Nivel educativo	Texto	255
Nacionalidad	Nacionalidad	Texto	255
Dedica	Dedicación	Texto	255
Mtremu	Más de un trabajo remunerado	Texto	255
Ocup	Ocupación	Texto	255
Catocup	Categoría ocupacional	Texto	255
Activi	Actividad	Texto	255

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Los campos (Numiemb) Número del miembro en el hogar, y (CodEPF) Código de la encuesta representan la clave principal que identifica en forma unívoca los diferentes Miembros en el hogar en la región de Los Andes.

Tabla 8 Vivienda

Esta tabla se refiere a todo lo relacionado con la vivienda en general.

Nombre del campo	Descripción	Tipo de dato	Tamaño
Numviv	Número de la vivienda	Autonúmero	Entero largo
CodEPF	Código de la encuesta	Texto	255
Hogar	Número del hogar	Texto	50
Tipviv	Tipo de vivienda	Texto	255
Tenviv1	Tenencia de la vivienda 1	Texto	255
Alqmvpl	Alquiler mensual de la vivienda principal 1	Número	Entero largo
Alqmvo1	Alquiler mensual de la vivienda ocasional 1	Número	Entero largo
Tenviv2	Tenencia de la vivienda 2	Texto	50
Alqmvpl2	Alquiler mensual de la vivienda principal 2	Número	Entero largo
Alqmvo2	Alquiler mensual de la vivienda ocasional 2	Número	Entero largo
Habita	Número de habitaciones que tiene la vivienda	Texto	255
DHabita	Número de habitaciones que se utilizan para dormir	Texto	255
paredex	Material en las paredes exteriores	Texto	255
Techo	Material en el techo	Texto	255
Piso	Material en el piso	Texto	255
Utuba	Ubicación de la instalación para agua corriente por tubería	Texto	255
Abagua	Abastecimiento normalmente del agua	Texto	255
Nbadu	Número de habitaciones de baño con ducha que posee la vivienda	Texto	255
Excusado	Tipo de excusado que tiene la vivienda	Texto	255

Continuación de la tabla Vivienda

Nombre del campo	Descripción	Tipo de dato	Tamaño
Alumbrado	Clase de alumbrado que se utiliza	Texto	255
Combustible	Combustible que se utiliza para cocinar	Texto	255

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Los campos (Numviv) Número de la vivienda, y (CodEPF) Código de la encuesta representan la clave principal que identifica en forma unívoca las diferentes viviendas en la región de Los Andes.

CAPÍTULO III. METODOLOGÍA ESTADÍSTICA

El análisis estadístico multivariante utilizado en la presente investigación es el análisis de componentes principales (ACP) y el análisis Factorial de Correspondencias Múltiples (AFCM).

3.1 Análisis de Componentes Principales (ACP)

El análisis de componentes principales (ACP) tuvo sus inicios en el año 1901, cuando Karl Pearson publica un trabajo sobre el ajuste de un sistema de puntos a un plano. Más tarde Harold Hotelling retoma las ideas de Pearson y presenta la técnica como hoy la conocemos, como sintetizador de la mayor variabilidad del sistema de puntos. El ACP es una técnica para tratar con matrices de datos de grandes dimensiones; se puede ver como una técnica exploratoria interdependiente que busca condensar la información en un número reducido de nuevas variables que explican la máxima variabilidad total de los datos (Johnson y Dear, 1992).

El análisis de componentes principales transforma un conjunto de variables correlacionadas en un nuevo conjunto de variables incorrelacionadas llamadas componentes principales (Dallas, 2000). Dichas componentes son combinaciones lineales derivados en forma decreciente de importancia de tal manera que el mayor componente es el que explica la mayor variabilidad posible de la variación total contenida en los datos originales. La segunda componente principal se elige de forma que explique la mayor cantidad posible de variación que resta sin explicar por la primera componente principal, sujeta a la condición de estar incorrelada con la primera componente principal. La tercera componente no está correlacionada ni con la primera ni con la segunda y tiene la tercera varianza más grande, y así sucesivamente. “Si las variables ya están casi no correlacionadas, entonces nada se puede ganar al llevar a cabo un ACP. En este caso, la dimensionalidad real de los datos es igual al número de

variables respuesta medidas y no es posible examinar los datos en un espacio con un número reducido de dimensiones” (Dallas, 2000).

3.1.1 Objetivos del Análisis de componentes principales

Los objetivos de esta técnica son:

- Reducir la dimensionalidad del problema que se está estudiando
- Eliminar cuando sea posible algunas de las variables originales si ellas aportan poca información
- Generar un nuevo conjunto de variables incorrelacionadas, que pueden expresar la información contenida en el conjunto original de datos con la mínima pérdida posible.

El ACP no requiere de un modelo estadístico para explicar la estructura probabilística de los errores y no se necesita hacer supuestos acerca de la distribución de probabilidad de las variables originales (Johnson y Dear, 1992).

3.1.2 ¿Cuándo es apropiado el ACP?

El análisis de componentes principales sólo es apropiado en aquellos casos en donde:

- Todas las variables deben estar medidas en las mismas unidades o, por lo menos en unidades comparables.
- Las variables deben tener varianzas que tengan tamaños aproximadamente muy semejantes. (Dallas, 2000).

3.1.3 Propiedades del Análisis de Componentes Principales

Las componentes principales, poseen las siguientes propiedades:

1. Son ortogonales (cada componente representa una dirección del espacio de las variables originales).
2. Cada puntuación está incorrelada con la anterior.
3. La primera componente es la que más varianza contiene y la j-ésima tiene más varianza que la j-1 ésima.

Las relaciones estocásticas entre las p variables originales se pueden condensar en otras m variables que explican sólo la variación del sistema, descartando la información redundante. Geométricamente el subespacio que se crea con las m primeras componentes da el mejor ajuste posible al conjunto de datos medidos mediante la suma de los cuadrados de las distancias perpendiculares desde cada punto al subespacio. El subespacio de menor dimensionalidad será $m=1$ componente para hacer la representación en un sólo eje, pero el conjunto inicial se puede distorsionar. Por eso se introducirá un nuevo eje para definir un subespacio $m=2$, con lo cual se pierde menos información. Cuando $m=p$ existe el mismo número de variables y no se reduce la dimensión; sólo se hace una rotación rígida del conjunto de datos para su representación.

3.1.4 Cuantos componentes vamos a retener

No existe un criterio específico, sin embargo algunos aspectos pueden ayudar al investigador a tomar la decisión de cuantos componentes principales retener, como son:

- El porcentaje de variación explicada por los componentes: frecuentemente se considera satisfactorio un porcentaje que explique por lo menos el 70% de la varianza total (Kaiser 1958).
- Contribución de un componente en la explicación de la variación de una variable determinada (contribución relativa o calidad de representación).

3.1.4.1 Calidad de representación (CR)

Se define como la suma de los primeras K correlación al cuadrado, es decir:

$$CR = \sum_{j=1}^K r_{X_i Y_j}^2$$

3.1.5 Ajuste en R^p (n puntos individuales)

La nube de puntos en R^p , que se denotará por $N(I)$, está definida por el conjunto de las n filas X_1, X_2, \dots, X_n , dotadas de la métrica euclídea. El objetivo del ajuste es encontrar el subespacio de R^p de dimensión q ($q < p$), con direcciones ortonormalizadas, que mejor se ajuste a la nube $N(I)$ en el sentido de los mínimos cuadrados. Esto es, se trata de la búsqueda de un subespacio de representación de datos que garantice el mayor parecido global de las filas X_i a sus aproximaciones sobre tal subespacio, de forma que se pierda la menor información posible.

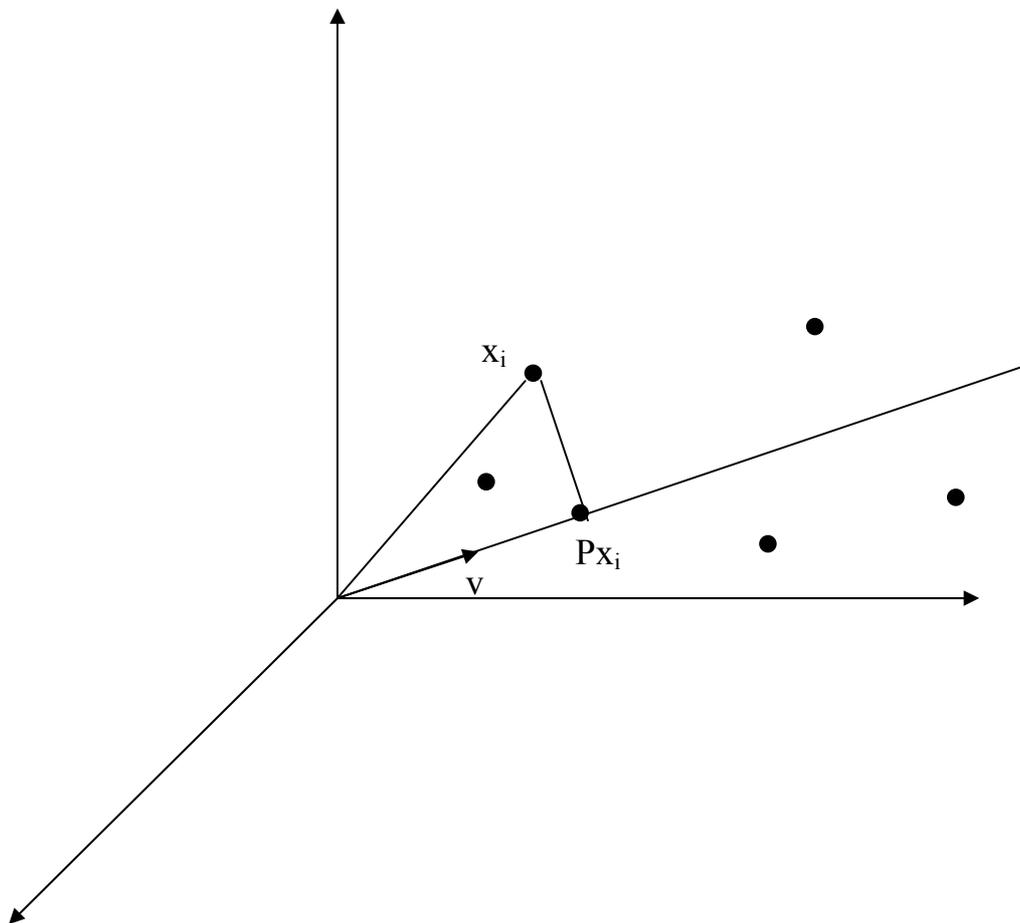
El procedimiento a seguir para la búsqueda del subespacio que se denominará como de “mejor ajuste”, consistirá en una sucesión de etapas, en cada una de las cuales se construye una nueva dirección del subespacio. En una primera etapa se hallará el subespacio de dimensión 1, es decir, la recta que mejor se ajusta a la nube. El vector director de esta recta será denotado por v , del cual se exigirá además que sea unitario. Matemáticamente, se trata de un problema de mínimos cuadrados.

$$\text{Min } \sum d^2(x_i, Px_i)$$

donde Px_i es la proyección ortogonal de x_i sobre la recta.

Figura 2

Gráficamente en \mathbb{R}^3



Fuente: Rao (citado por Márquez, 2004)

Con el objeto de facilitar el procedimiento de minimización conviene tomar en cuenta que:

- La distancia cuadrado del origen a un punto x_i , que es un valor fijo, es igual a:

$$d^2(O, x_i) = \| x_i \|^2 = x_i^t x_i$$

- La distancia cuadrado del origen a la proyección Px_i , que depende del vector v , es igual a

$$\begin{aligned} d^2(O, Px_i) &= \| Px_i \|^2 = (x_i^t v)^2 \\ &= (x_i^t v)^t (x_i^t v) = v^t x_i x_i^t v \end{aligned}$$

- En virtud del teorema de Pitágoras se establece la relación fundamental:

$$d^2(O, x_i) = d^2(O, Px_i) + d^2(x_i, Px_i)$$

de donde:

$$d^2(x_i, Px_i) = d^2(O, x_i) - d^2(O, Px_i)$$

- De acuerdo con lo anterior, minimizar la suma $\sum d^2(x_i, Px_i)$, resulta equivalente a maximizar la suma:

$$\sum d^2(O, Px_i) = \sum v^t x_i x_i^t v = v^t \sum x_i x_i^t v = v^t X^t X v$$

Para maximizar esta forma cuadrática dado la restricción de que v sea un vector unitario, se construye el lagrangiano:

$$L = v^t X^t X v - \lambda (v^t v - 1)$$

derivando respecto de v e igualando al vector nulo:

$$2 X^t X v - 2 \lambda v = 0$$

Se obtiene la siguiente ecuación:

$$X^t X v = \lambda v$$

lo que determina que el vector buscado v sea un autovector de la matriz $X^t X$ asociado con el autovalor λ .

Esto quiere decir que una condición necesaria para que v defina un máximo en la forma cuadrática en referencia, es que sea un autovector de $X^t X$. Dentro del conjunto de autovectores de esa matriz se centrará el interés en aquel que efectivamente maximice la forma cuadrática:

$$\text{Max}(v^t X^t X v) = \text{Max}(v^t \lambda v) = \text{Max}(\lambda v^t v) = \text{Max}(\lambda)$$

De modo que el vector director de la recta que mejor se ajusta a la nube de puntos es el autovector de $X^t X$, que denotamos por v^1 , asociado con el mayor autovalor, que denotamos por λ_1

En la segunda etapa del procedimiento se hallará el subespacio de dimensión 2, es decir, el plano que mejor se ajuste a la nube. Este plano tendrá como primera dirección la definida por v^1 , mientras que la segunda quedará determinada por el vector unitario v^2 , ortogonal a v^1 que minimice la suma de cuadrados $\sum d^2(x_i, P_{x_i})$, donde ahora P_{x_i} es la proyección ortogonal de x_i sobre el plano.

El Lagrangiano que toma en cuenta tales restricciones es ahora:

$$L = v^t X^t X v - \lambda (v^t v - 1) - \gamma (v^t v^1)$$

Derivando respecto de v e igualando al vector nulo:

$$2 X^t X v - 2 \lambda v - \gamma v^1 = 0$$

Premultiplicando por $(v^1)^t$ se obtiene que:

$$2(v^1)^t X^t X v - 2 \lambda (v^1)^t v - \gamma (v^1)^t v^1 = -\gamma = 0$$

ya que $(v^1)^t v = 0$. Entonces, al igual que en la primera etapa, la expresión queda reducida a que $X^t X v = \lambda v$. Por lo tanto, las direcciones del plano que mejor se ajusta a la nube de puntos en R^p están definidas por los vectores v^1 y v^2 , autovectores ortonormalizados de la matriz $X^t X$ asociados con sus dos mayores autovalores λ_1 y λ_2 .

Reiterando q veces el mismo procedimiento se obtiene que:

“Una base ortonormalizada de dimensión q que mejor se ajusta a la nube de puntos en R^p , está constituida por los q autovectores v^1, v^2, \dots, v^q correspondientes a los q mayores autovalores $\lambda_1, \lambda_2, \dots, \lambda_q$ de la matriz $X^t X$ ”.

Las coordenadas de proyección de los n puntos fila sobre el eje α conforman el

vector:

$$\psi^\alpha = \begin{pmatrix} X_1^t v^\alpha \\ X_2^t v^\alpha \\ \cdot \\ X_n^t v^\alpha \end{pmatrix} = \begin{pmatrix} X_1^t \\ X_2^t \\ \cdot \\ X_n^t \end{pmatrix} v^\alpha = X v^\alpha = \sum_{j=1}^p v_{j\alpha} X^j$$

que constituye una combinación lineal de las columnas (variables) de la matriz X, y

además es tal que:

$$\sum_{i=1}^n \psi_{i\alpha}^2 = (\psi^\alpha)^t \psi^\alpha = (X v^\alpha)^t X v^\alpha = (v^\alpha)^t X^t X v^\alpha$$

$$\sum_{i=1}^n \psi_{i\alpha}^2 = (v^\alpha)^t \lambda_\alpha v^\alpha = \lambda_\alpha (v^\alpha)^t v^\alpha = \lambda_\alpha$$

y

$$(\psi^\alpha)^t \psi^\beta = \lambda_\beta (v^\alpha)^t v^\beta = 0$$

Las coordenadas de los n puntos filas sobre q ejes pueden organizarse en la matriz

$$\Psi(q) = (\psi^1, \psi^2, \dots, \psi^q) = X V(q)$$

siendo V(q) la matriz cuyas columnas son los q autovectores ortonormalizados asociados con los q mayores autovalores de la matriz $X^t X$.

En resumen:

El vector ψ^α de coordenadas de proyección de los puntos fila sobre el eje α es tal que:

- Es una combinación lineal de las columnas (variables) de la matriz X
- Su norma cuadrado es igual a λ_α
- Es ortogonal a ψ^β , para cualquier $\beta \neq \alpha$

3.1.6 Ajuste en \mathbb{R}^n

La nube de puntos en \mathbb{R}^n , que se denotara por $N(J)$, está definida por el conjunto de las p columnas X^1, X^2, \dots, X^p , dotadas de la métrica euclidea. El objetivo del ajuste es encontrar el subespacio de \mathbb{R}^n de dimensión q ($q < p$), con direcciones ortonormalizadas, que mejor se ajuste a la nube $N(J)$ en el sentido de los mínimos cuadrados.

Procediendo en forma similar al caso anterior se encuentra que:

“Una base ortonormalizada de dimensión q que mejor se ajusta a la nube de puntos en \mathbb{R}^n , está constituida por los q autovectores u^1, u^2, \dots, u^q correspondientes a los q mayores autovalores $\lambda_1, \lambda_2, \dots, \lambda_q$ de la matriz XX^t ”.

Las coordenadas de proyección de los p puntos columna sobre el eje α conforman el vector:

$$\varphi^\alpha = \begin{pmatrix} (X^1)^t u^\alpha \\ (X^2)^t u^\alpha \\ \cdot \\ \cdot \\ (X^p)^t u^\alpha \end{pmatrix} = \begin{pmatrix} (X^1)^t \\ (X^2)^t \\ \cdot \\ \cdot \\ (X^p)^t \end{pmatrix} u^\alpha = X u^\alpha = \sum_{i=1}^n u_{i\alpha} X_i$$

que constituye una combinación lineal de las filas (individuos) de la matriz X, y además es tal que:

$$\sum_{j=1}^p \varphi_{i\alpha}^2 = (\varphi^\alpha)^t \varphi^\alpha = (X\mathbf{u}^\alpha)^t X\mathbf{u}^\alpha = (\mathbf{u}^\alpha)^t XX^t \mathbf{u}^\alpha$$

$$\sum_{j=1}^p \varphi_{i\alpha}^2 = (\mathbf{u}^\alpha)^t \lambda_\alpha \mathbf{u}^\alpha = \lambda_\alpha (\mathbf{u}^\alpha)^t \mathbf{u}^\alpha = \lambda_\alpha$$

$$\text{y } (\varphi^\alpha)^t \varphi^\beta = \lambda_\beta (\mathbf{u}^\alpha)^t \mathbf{u}^\beta = 0$$

Las coordenadas de los p puntos columna sobre q ejes pueden organizarse en la matriz

$$\varphi(\mathbf{q}) = (\varphi^1, \varphi^2, \dots, \varphi^q) = X^t U(\mathbf{q})$$

siendo U(q) la matriz cuyas columnas son los q autovectores ortonormalizados asociados con los q mayores autovalores de la matriz XX^t .

En resumen:

El vector φ^α de coordenadas de proyección de los puntos fila sobre el eje α es tal que:

- Es una combinación lineal de las filas (individuos) de la matriz X
- Su norma cuadrado es igual a λ_α
- Es ortogonal a φ^β , para cualquier $\beta \neq \alpha$

3.1.7 Relación entre los ajustes

Existen unas relaciones entre los autovectores v de la matriz $X^t X$ y los u de la matriz XX^t , a partir de los cuales se definen las siguientes ecuaciones matriciales, conocidas como “relaciones de doble transición”:

$$V = X^t U D^{-1}$$

$$U = X V D^{-1}$$

donde D es la matriz diagonal cuyos elementos son los r valores singulares de la matriz X . Por consiguiente, basta solamente con diagonalizar una de las dos matrices ($X^t X$) o (XX^t), para hallar las coordenadas de las filas y de las columnas sobre el subespacio de mejor ajuste. Por lo general se diagonaliza $X^t X$ ya que usualmente el número de variables es menor que el número de individuos.

A continuación estudiaremos las relaciones entre las coordenadas de los dos análisis, en R^p y en R^n .

- Si se ha diagonalizado la matriz $X^t X$ y se han obtenido los autovectores v^α y los autovalores λ_α , se pueden calcular las coordenadas de los puntos columna por medio de:

$$\varphi = X^t U = V D$$

de modo que:

$$\varphi^\alpha = \sqrt{\lambda_\alpha} v^\alpha$$

- Si se ha diagonalizado la matriz XX^t y se han obtenido los autovectores u^α y los autovalores λ_α , se pueden calcular las coordenadas de los puntos fila por medio de:

$$\psi = X V = U D$$

de modo que:

$$\varphi^\alpha = \sqrt{\lambda_\alpha} u^\alpha$$

- Por otro lado, como:

$$V = (X^t U) D^{-1} = \varphi D^{-1} \text{ y } U = (XV) D^{-1} = \psi D^{-1}$$

entonces:

$$\varphi = X^t U = X^t \psi D^{-1}$$

luego:

$$\varphi^\alpha = X^t \psi^\alpha \frac{1}{\sqrt{\lambda_\alpha}}$$

y por tanto:

$$\varphi_{j\alpha} = (X^j)^t \psi^\alpha \frac{1}{\sqrt{\lambda_\alpha}} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n x_{ij} \psi_{i\alpha} \quad j = 1, 2, 3, \dots, p$$

- En forma similar, de la relación:

$$\psi = XV = X\varphi D^{-1}$$

se obtiene:

$$\psi^\alpha = X\varphi^\alpha \frac{1}{\sqrt{\lambda_\alpha}}$$

de donde: $\psi_{i\alpha} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p x_{ij} \varphi_{j\alpha} \quad i = 1, 2, 3, \dots, n$

3.1.8 Interpretación geométrica

3.1.8.1 Variabilidad captada por las componentes

Hemos señalado que uno de los aspectos importantes del ACP es que permite resumir la variabilidad contenida en los datos originales, en unas pocas variables. A este respecto es conveniente tomar en cuenta los siguientes elementos:

- La traza de la matriz de varianza y covarianza es una medida de variabilidad total. Por su parte:

$$\text{tr}(\mathbf{S}) = \sum_{j=1}^p S_j^2 = \frac{1}{n-1} \text{tr}(\hat{\mathbf{X}}^t \hat{\mathbf{X}}) = \frac{1}{n-1} \sum_{j=1}^p \lambda_j = \sum_{j=1}^p \lambda_j^*$$

y además :

$$\begin{aligned} \frac{1}{n-1} \sum_{i=1}^n d^2(x_i, \bar{x}) &= \frac{1}{n-1} \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - \bar{x}_{.j})^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{(x_{ij} - \bar{x}_{.j})^2}{n-1} \\ &= \sum_{j=1}^p S_j^2 = \text{tr}(\mathbf{S}) \end{aligned}$$

Es decir, $\text{tr}(\mathbf{S})$ es un indicador de la variabilidad total de la nube de puntos, que mide globalmente el grado de semejanza entre cada fila y su centro de gravedad, al cual se denomina Inercia total.

Esta expresión puede escribirse también como:

$$\text{Tr}(\mathbf{S}) = \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{s=1}^n d^2(x_i, x_s)$$

Es decir, se trata también de una medida de las desemejanzas entre filas de una matriz de datos, dos a dos.

- La variabilidad captada por las q primeras componentes es la mejor aproximación a $\text{tr}(\mathbf{S})$

Si denotamos por $\psi_{i(q)}$ al vector que contiene las coordenadas de proyección del i -ésimo individuo sobre el subespacio q -dimensional de mejor ajuste,

entonces la variabilidad de la nube proyectada sobre el subespacio viene dada por:

$$\frac{1}{n-1} \sum_{i=1}^n d^2(\psi_{i(q)}, \bar{\psi}_{(q)}) = \frac{1}{n-1} \sum_{i=1}^n \sum_{\alpha=1}^q \psi_{i\alpha}^2 = \frac{1}{n-1} \sum_{\alpha=1}^q \lambda_{\alpha} = \sum_{\alpha=1}^q \lambda_{\alpha}^*$$

expresión que toma un valor máximo (RAO, 1964)

- La proporción de variabilidad total captada por las q primeras componentes principales viene dada por la:

$$\text{Tasa de inercia} = \frac{\sum_{\alpha=1}^q \lambda_{\alpha}}{\sum_{\alpha=1}^p \lambda_{\alpha}} \cdot 100$$

- Un indicador de la importancia de la α -ésima componente sería:

$$\frac{\lambda_{\alpha}}{\sum \lambda_{\alpha}} \cdot 100 = \% \text{ de la variabilidad total captada por la } \alpha\text{-ésima componente.}$$

- Los primeros dos componentes principales constituyen la mejor representación posible de la variabilidad total, entonces el gráfico de estos componentes en el plano factorial, es una buena representación de la proyección de los datos originales.

3.1.8.2 Los n-individuos en R^p

Para el análisis, las proximidades entre puntos se interpreta en términos de similitud de comportamiento con respecto a las variables. Al decir que dos puntos están cercanos, esto tan sólo significa que ellos tienen valores similares en algunas variables; no necesariamente en todas.

3.1.8.3 Las p-variables en R^n

En este espacio euclideo, las proximidades entre puntos se interpreta en términos de correlación. Para representar las variables se usa las correlaciones entre la variable principal y los componentes principales. Una distancia pequeña entre dos columnas o variables (puntos de R^n), registradas sobre el conjunto de individuos, significa que ellas miden casi lo mismo.

3.1.8.4 Biplot

Es un gráfico que nos permite representar tanto a individuos como a variables simultáneamente. Es posible comparar a 2 individuos con todas las nubes de puntos de las variables y a 2 variables con el grupo total de individuos; el biplot ayuda a la interpretación de los componentes principales, pero hay que tener en cuenta para su interpretación que los individuos y las variables pertenecen a distintas nubes de puntos.

“Se utiliza la posibilidad de introducir individuos y/o variables suplementarias cuando todos los individuos y/o todas las variables no presenten el mismo interés, es decir, jueguen papeles diferentes. Estos “elementos suplementarios” podrán ser examinados correctamente sobre los gráficos; su única particularidad será la de no contribuir en nada a la formación de los ejes.” (Dervin1991).

3.2 Análisis Factorial de Correspondencia Múltiple (A.F.C.M)

El AFCM fue ideado por el profesor J. P. Benzecri, a principios de los años 60, posteriormente actualizada por su equipo de investigación de la Universidad de Paris VI.

El Análisis Factorial de Correspondencia Múltiples (A.F.C.M.) es una extensión del análisis de Correspondencia Simple al análisis de más de dos conjuntos de características.

Esta técnica está especialmente diseñada para analizar datos cualitativos (nominales u ordinales) y cualitativos (clasificados en categorías) y representados en tablas de contingencia formadas por números positivos, resultado de contar frecuencias. También se utiliza en archivos de encuestas, cruzando un conjunto de filas, con el conjunto de modalidades de respuestas a varias preguntas. Los datos de encuestas incluyen respuestas en forma disyuntiva completa, es decir, las diferentes categorías de respuestas son mutuamente exclusivas y sólo se selecciona una categoría.

3.2.1 Objetivos del Análisis de Correspondencia Múltiple

Los objetivos de esta técnica entre otros son:

- Transformar la tabla de datos en perfiles fila y perfiles columnas.
- Ajustar los datos o puntos ponderados por sus perfiles marginales (fila o columnas).
- El objetivo esencial es reducir las dimensiones del espacio de representación de los datos cualitativos a, por lo general, un plano (formado por los dos primeros ejes principales), tratando de no perder mucha información en esta reducción.

3.2.2 ¿Cuándo es apropiado el A.F.C.M.?

Este análisis es útil para el análisis de encuestas donde se cumplan que:

- El número de categorías de cada pregunta debe ser similar para que así cada pregunta tenga pesos equivalentes.
- Las diferentes modalidades deben tener efectivos similares.
- Se deben evitar modalidades de respuesta varias.

El análisis de correspondencia múltiple (A.F.C.M.) no solamente se aplica a datos de encuestas, también se aplica a otros tipos de datos: continuos, discretos, etc.; siempre que transformen los datos en la forma disyuntiva completa, tomando en consideración los aspectos antes mencionados sobre el número de categorías de cada variable y el número de efectivos en cada categoría.

3.2.3 Descripción de la técnica de A.F.C.M.

Se tiene un archivo con Q preguntas ($q=1,2,\dots,Q$) hechas a n individuos, es decir, n filas. Cada pregunta presenta P_q modalidades de respuesta, bajo forma disyuntiva completa:

$$P = \sum_{q=1}^Q P_q$$

El archivo de datos (respuestas codificadas en forma binaria) se presenta en una matriz Z formada por la yuxtaposición de Q submatrices:

$$Z = [Z_1, Z_2, \dots, Z_q, \dots, Z_Q]$$

$$\text{Cada } Z_{ij} = \begin{cases} 1, & i^{\text{mo}} \text{ individuo selecciona la modalida } j \text{ de la pregunta } q. \\ 0, & \text{Otro caso.} \end{cases}$$

Cada Z_q es de orden $n \times p_q$, la i^{ma} fila contiene $(P_q - 1)$ ceros y sólo un 1, en la columna correspondiente a la modalidad de la pregunta q .

3.2.4 Generalización del Análisis Factorial de Correspondencias Múltiples.

La matriz de datos Z tiene p columnas, a las cuales corresponden p puntos de R^n .

Esto es:

$$Z = [Z_1, Z_2, \dots, Z_q, \dots, Z_Q] \rightarrow \left(\frac{1}{n} I_n \right)$$
$$\downarrow$$
$$\left(\frac{1}{nQ} \right) D$$

Siendo:

$$D = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D_Q \end{pmatrix}$$

Q Preguntas.

P_q Modalidades para cada pregunta q .

$$P = \sum_{q=1}^Q P_q$$

n Individuos.

El rango es : $Z \leq [P - (Q - 1)]$.

En R^p , para calcular los factores del A.C. se debe maximizar la expresión:

$$\Psi' N \Psi = U' M X' N X M U, \quad U' M U = 1$$

En este caso del A.F.C.M. se tiene lo siguiente:

La matriz de datos es : $X = \left(\frac{1}{Q}\right)Z$.

La métrica es : $M = nQD^{-1}$.

El criterio de Ponderación anterior queda : $N = \left(\frac{1}{n}\right)I_n$.

Sustituyendo en la expresión anterior queda:

$$\frac{1}{Q}D^{-1}BU_h = \lambda_h U_h, \quad h = (1, 2, \dots, q)$$

$$\left(\frac{1}{Q}D^{-1}B - \lambda_h I\right)U_h = 0$$

Los factores se obtienen así:

$$\varphi_h = MU_h \rightarrow \varphi_h = (nQD^{-1})U_h$$

Las proyecciones de los puntos filas sobre los nuevos ejes es:

$$\Psi_h = XMU_h = \frac{1}{Q}(nQD^{-1}U_h)Z \rightarrow \hat{\varphi}_h = \left(\frac{1}{Q}\right)Z\varphi_h$$

3.2.5 Propiedades del Análisis Factorial de Correspondencias Múltiples

1. Las Q sub-nubes de puntos correspondientes a las P_q modalidades de una pregunta q, tienen el mismo centro de gravedad ($1/n$), que coincide con el centro de gravedad de la nube de puntos total. Por tanto, todos los factores están centrados:

I. Las coordenadas del subconjunto de puntos relativos a la pregunta q son las columnas de: $Z_q D_q^{-1}$.

II. Las masas relativas de los P_q puntos del subconjunto q son los elementos diagonales de: $(1/n)D_q$.

III. Por tanto las coordenadas del centro de gravedad son:

$$G_{qi} = \sum_{j \in J_q} \left(\frac{d_{ij}}{n} \right) \left(\frac{z_{ij}}{d_{ij}} \right) = \frac{1}{n}, \text{ pues } \sum_{j \in J_q} Z_{ij} = 1$$

J_q : Subconjunto de los p valores con índice correspondiente a q.

2. La inercia total de la nube de puntos es:

$$I = \sum_j \lambda_j - 1 = \text{tr} \left(\frac{1}{Q} (D^{-1} B) \right) - 1 = \left(\frac{P}{Q} \right) - 1$$

Donde: Q = Número de preguntas.

P = Número total de modalidades para las q preguntas

$$\left(P = \sum_q P_q \right)$$

3. El cuadrado de la distancia de un punto modalidad j al centro de gravedad G, en R^n es:

$$d^2(j, G) = \sum_{i=1}^n \left(\frac{1}{f_i} \right) \left(\frac{f_{ij}}{f \cdot j} - f_i \right)^2$$

Donde:

$$\left\{ \begin{array}{l} d_{jj} = \sum_{i=1}^n Z_{ij} \\ \sum_j d_{jj} = nQ \end{array} \right. , \sum_{i=1}^n Z_{ij} \quad \text{Que es el número de efectivos de la modalidad j.}$$

4. Contribución a la inercia total:

Contribución de la modalidad j a la inercia total:

$$C_j = \left(\begin{array}{c} \text{Masa o Peso} \\ \text{Relativo} \\ \text{de la} \\ \text{Modalidad j} \end{array} \right) * d^2(G, j) = \left[\frac{d_{jj}}{\sum_j d_{jj}} \right] * d^2(G, j) = \frac{d_{jj}}{nQ} n \left(\frac{1}{d_{jj}} - \frac{1}{n} \right) = \frac{1}{Q} \left(1 - \frac{d_{jj}}{n} \right) = C_j$$

La proporción de inercia debida a una modalidad o categoría de respuesta j, está en función inversa al número de efectivos en tal modalidad. Esta proporción aumenta cuando el número de efectivos disminuye. El máximo valor se obtiene cuando una modalidad tiene cero efectivos.

$$d_{jj} = 0 \Rightarrow C_j = 1/Q$$

Deben evitarse modalidades con bajas tasas de respuesta.

II Contribución de la pregunta q a la inercia total:

$$C_q = \sum_{j=1}^{P_q} C_j = \sum_{j=1}^{P_q} \frac{1}{Q} \left(1 - \frac{d_{jj}}{n} \right) = \frac{1}{Q} (P_q - 1)$$

La proporción de inercia debida a una pregunta es una función creciente del número de modalidades de esa pregunta.

Cuando una pregunta sólo tiene dos categorías entonces C_q toma su mínimo valor:

$$(1/Q)$$

Se cumple que:

$$\sum_{q=1}^Q C_q = \sum_{q=1}^Q \frac{1}{Q} (P_q - 1) = \left(\frac{P}{Q} - 1 \right)$$

3.2.6. Interpretación geométrica

Similar al análisis de componentes principales (ACP). Los resultados se presentan en un gráfico, conformando la nube de puntos proyectados sobre los ejes principales.

3.2.6.1. Porción de varianza explicada por los ejes

La porción de varianza explicada por el α -ésimo eje la mide la magnitud del α -ésimo autovalor. Es decir $\frac{\lambda_\alpha}{\sum \lambda_\alpha} \cdot 100 \%$

3.2.6.2. Distancia entre puntos

Es correcto interpretar las distancias entre los puntos de una misma categoría. Si la distancia es pequeña significa que sus contribuciones son similares también es legítimo interpretar la posición de un punto de una categoría con respecto de todos los

puntos de la otra categoría. Pero es muy peligroso comparar dos puntos de diferentes categorías.

3.2.6.3. Centros de gravedad G

El centro de gravedad G, localizado en el centro de los ejes corresponde al promedio de ambos conjuntos de puntos. Es decir, los $n + p$ puntos están centrados en el origen G.

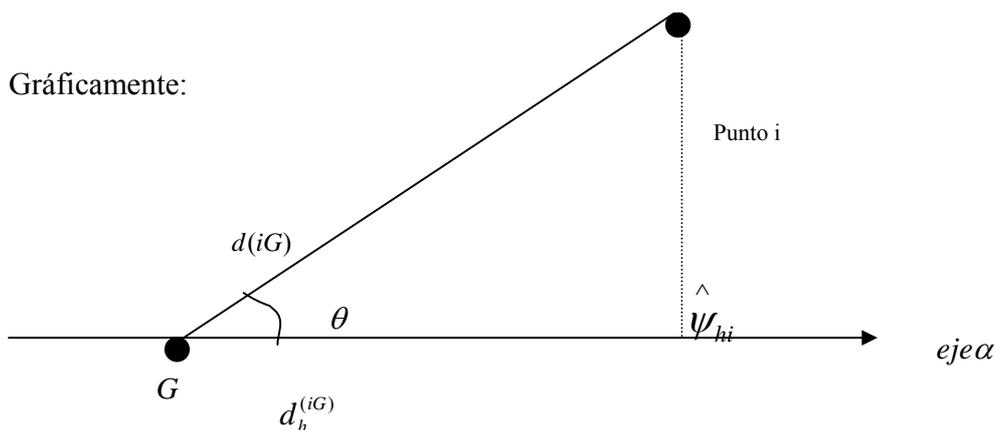
3.2.6.4. Contribuciones absolutas

Indica la proporción de varianza explicada por cada categoría en relación a cada eje principal. Permite saber las categorías que más contribuyen en la formación del eje.

3.2.6.5. Coseno cuadrado

Mide la calidad de representación sobre el eje considerado. Indica la parte de varianza de la categoría explicada por el eje principal.

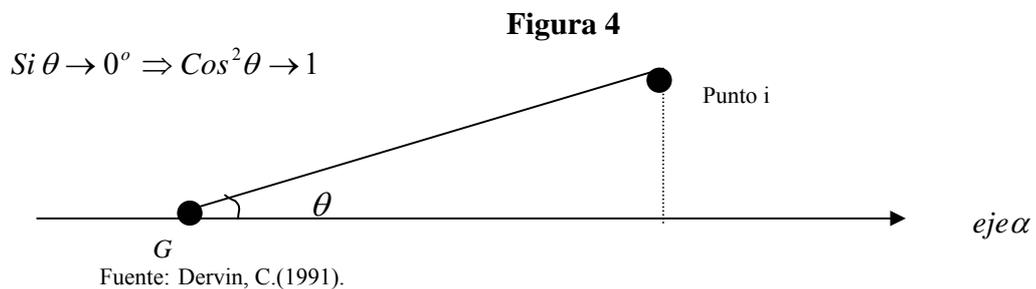
Figura 3



Fuente: Dervin, C.(1991).

$$\text{Cos}^2\theta = \frac{d_h^2(i, G)}{d^2(i, G)} = \frac{\psi_{h\alpha}^2}{d^2(i, G)} = (\text{Correlación entre } i \text{ y } h)^2$$

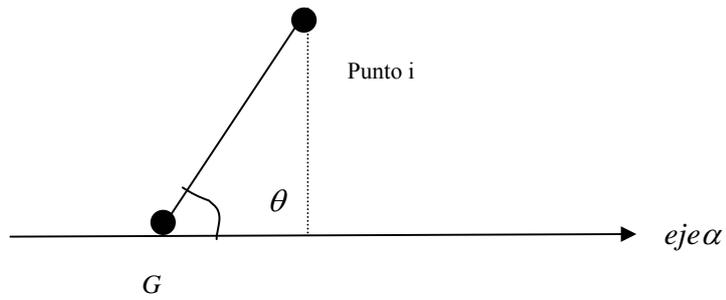
Si un punto i es cercano del eje, es decir se sitúa en una dirección cuyo ángulo con el eje es pequeño su coseno cuadrado será cercano a 1. En este caso será poco deformadora y el punto estará bien representado sobre ese eje.



Si un punto i está alejado del eje ya que se sitúa en una dirección cuyo ángulo con el eje es grande su coseno cuadrado será cercano a 0. En este caso, su proyección será muy deformadora y el punto estará mal representado sobre este eje.

Figura 5

$Si \theta \rightarrow 90^\circ \Rightarrow \text{Cos}^2 \theta \rightarrow 0$



Fuente: Dervin, C.(1991).

Sin embargo, no se debe interpretar a la ligera ya que se puede concurrir en equivocaciones, por ello se debe utilizar otro indicador llamado las contribuciones relativas. (Dervin, 1991).

CAPÍTULO IV. RESULTADOS Y DISCUSIÓN- ANÁLISIS ESTADÍSTICO

4.1 Estadísticas básicas

Con una muestra de 1.187 viviendas que comprenden 1.179 hogares y 6.146 miembros de la región de Los Andes de la I Encuesta de Presupuestos Familiares en base al período 1988/89, se realizó un análisis estadístico a las principales variables, con el fin de obtener una primera descripción de las observaciones más relevantes.

En los siguientes cuadros se describen, las principales características físicas de la vivienda principal, ya que sólo se conoce la tenencia de la vivienda ocasional, la cual se da a conocer. Por ejemplo, podemos observar que un 3,5 % de las viviendas son piezas o ranchos urbanos, y un 96,5 % son viviendas en condiciones más aptas para vivir (Tabla 9).

Tabla 9. Distribución de frecuencias del tipo de vivienda

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
Casa o quinta	900	76,3	76,3	76,3
Apto, en edificio, casa o quinta	238	20,2	20,2	96,5
Pieza en casa, quinta o apartamento	3	0,3	0,3	96,8
Pieza en casa de vecindad	4	0,3	0,3	97,1
Rancho urbano	34	2,9	2,9	100,0
Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 10. Distribución de frecuencias de la tenencia de la vivienda principal

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Alquilada	217	18,4	18,4	18,4
	Propia pagando	213	18,1	18,1	36,5
	Propia pagada	693	58,8	58,8	95,3
	Gratuita	46	3,9	3,9	99,2
	Ocupante de hecho	5	0,4	0,4	99,6
	Cedida P/Empleador	5	0,4	0,4	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Considerando la tenencia de la vivienda principal, en la tabla 10 se observa que el 58,8 % de las mismas son propias, el 18,4 % es alquilada y el 18,1 % es propia pagándose. Mientras que en la vivienda ocasional predomina la propia pagada (Tabla 11).

Tabla 11. Distribución de frecuencias de la tenencia de la vivienda ocasional

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Propia pagando	1	0,1	0,1	99,4
	Propia pagada	7	0,6	0,6	100,0
	Total	8	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 12. Distribución de frecuencias del número de habitaciones que poseen las viviendas

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Validos	1	6	0,5	0,5	0,5
	2	38	3,2	3,2	3,7
	3	137	11,6	11,6	15,4
	4	330	28,0	28,0	43,3
	5	314	26,6	26,6	70,0
	6	170	14,4	14,4	84,4
	7	88	7,5	7,5	91,9
	8	45	3,8	3,8	95,7
	9	24	2,0	2,0	97,7
	10	12	1,0	1,0	98,7
	11	6	0,5	0,5	99,2
	12	5	0,4	0,4	99,7
	14	1	0,1	0,1	99,7
	15	1	0,1	0,1	99,8
	17	2	0,2	0,2	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 13. Distribución de frecuencias del número de habitaciones para dormir

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	72	6,1	6,1	6,1
	2	217	18,4	18,4	24,5
	3	457	38,8	38,8	63,3
	4	273	23,2	23,2	86,4
	5	109	9,2	9,2	95,7
	6	28	2,4	2,4	98,0
	7	13	1,1	1,1	99,2
	8	6	0,5	0,5	99,7
	9	1	0,1	0,1	99,7
	10	1	0,1	0,1	99,8
	11	2	0,2	0,2	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

En la tabla 12 se observa que el número de habitaciones que poseen las viviendas es de 4, la siguiente frecuencia más alta corresponde a las viviendas que poseen 5 habitaciones. Respecto a las habitaciones para dormir predominantes en las viviendas es de 3, la siguiente frecuencia más alta es 4 habitaciones (Tabla 13).

Las estadísticas descriptivas que se muestran a continuación indican: el número promedio de habitaciones que poseen las viviendas es 5, el promedio de aquellas que se utilizan para dormir es 3, el mínimo que poseen y que se utilizan para dormir es 1, mientras el máximo que poseen es 17 y que se utilizan para dormir es 11.

	Habitaciones	Habitaciones para dormir
Media	5,01	3,28
Desviación estándar	1,809	1,286
Mínimo	1	1
Máximo	17	11

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Entre los materiales predominantes en paredes exteriores se tiene el bloque o ladrillo frisado (90,7 %) (Tabla 14). Para el piso existe una división entre materiales de mosaico (55 %) y cemento(44,5 %) (Tabla 15) y en materiales de techo predomina la platabanda con un 54,1% (Tabla 16).

Tabla 14. Distribución de frecuencias del material predominante en paredes exteriores

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Bloques o ladrillos frisados	1069	90,7	90,7	90,7
	Bloques o ladrillos sin frisar	38	3,2	3,2	93,9
	Concreto	3	0,3	0,3	94,1
	Madera aserrada	1	0,1	0,1	94,2
	Adobe, tapia o bahareque frisados	46	3,9	3,9	98,1
	Adobe, tapia o bahareque sin frisar	8	0,7	0,7	98,8
	Otros (caña, palma, lata, cartón y similares)	14	1,2	1,2	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 15. Distribución de frecuencias del material predominante en el piso

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Mosaico	648	55,0	55,0	55,0
	Cemento	525	44,5	44,5	99,5
	Tierra	6	0,5	0,5	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 16. Distribución de frecuencias del material predominante en el techo

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Dato faltante	2	,2	,2	,2
	Platabanda	638	54,1	54,1	54,3
	Teja	152	12,9	12,9	67,2
	Asbestos y similares	114	9,7	9,7	76,8
	Lámina metálica (zinc y similares)	270	22,9	22,9	99,7
	Otros (caña, palma, tablas, lata y similares)	3	0,3	0,3	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 17. Distribución de frecuencias de ubicación de la instalación para agua corriente por tubería

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Dentro de la vivienda	1140	96,7	96,7	96,7
	Fuera de la vivienda	34	2,9	2,9	99,6
	No tiene	5	0,4	0,4	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

En la tabla 17 se muestra que mayormente (96,7 %) la ubicación de la instalación para agua corriente por tubería se encuentra dentro de la vivienda.

Tabla 18. Distribución de frecuencias del abastecimiento de agua

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Acueducto	1167	99,0	99,0	99,0
	Pila pública	2	0,2	0,2	99,2
	Camión tanque	2	0,2	0,2	99,3
	Otros medios	8	0,7	0,7	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Como lo muestra la tabla 18 el abastecimiento predominante de agua en las viviendas es en 99 % por acueducto.

Tabla 19. Distribución de frecuencias de las habitaciones de baño con ducha o regadera que posee las viviendas

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	686	58,2	58,2	58,2
	2	307	26,0	26,0	84,2
	3	162	13,7	13,7	98,0
	4	24	2,0	2,0	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Como se aprecia en la (Tabla 19) predominan en las viviendas una habitación que tiene baño con ducha o regadera en un 58,2 %, la siguiente frecuencia más alta es de dos habitaciones que tiene baño con ducha o regadera en un 26 %.

Tabla 20. Distribución de frecuencias del tipo de excusado que tienen las viviendas

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Excusado (w.c.) a cloaca	1133	96,1	96,1	96,1
	Excusado (w.c.) a pozo séptico	29	2,5	2,5	98,6
	Excusado de hoyo o letrina	12	1,0	1,0	99,6
	No tiene excusado	5	0,4	0,4	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

En la tabla 20 se observa que el tipo de excusado predominante es el excusado (W.C.) a cloaca con un 96,1%, seguido del excusado (W.C.) a pozo séptico con un 2,5%.

Tabla 21. Distribución de frecuencias de la clase de alumbrado

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Servicio eléctrico público	1175	99,7	99,7	99,7
	Otro tipo	4	0,3	0,3	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Tabla 22. Distribución de frecuencias del combustible que se utiliza para cocinar

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Gas	1152	97,7	97,7	97,7
	Kerosene o gasolina	12	1,0	1,0	98,7
	Electricidad	9	0,8	0,8	99,5
	Otra clase (leña, carbón, etc.)	2	0,2	0,2	99,7
	No cocina	4	0,3	0,3	100,0
	Total	1179	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

En las tablas 21 y 22 respectivamente se observa que la clase de alumbrado más utilizada es el servicio eléctrico público (99,7%) y el combustible que se utiliza para cocinar es el gas (97,7%).

Tabla 23. Distribución de frecuencias del estado civil

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Unido	327	5,3	5,3	5,3
Casado	1522	24,8	24,8	30,0
Separado	102	1,7	1,7	31,8
Viudo	212	3,4	3,4	35,2
Soltero	3857	62,8	62,8	98,0
Divorciado	126	2,1	2,1	100,0
Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

La tabla 23 muestra que de la población encuestada el 62,8 % son solteros, seguido de casados con un 24,8 % y la mínima frecuencia es separado 1,7%.

Tabla 24. Distribución de frecuencias del género

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	NRGénero	2	0,0	0,0	0,0
	M	2784	45,3	45,3	45,3
	F	3360	54,7	54,7	100,0
	Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Como se muestra en la tabla 24 el género predominante es femenino (54,7%).

Tabla 25. Distribución de frecuencias de la nacionalidad

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Venezolana	5821	94,7	94,7	94,7
	Extranjera	325	5,3	5,3	100,0
	Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Considerando la nacionalidad se observa que el 94,7% es Venezolana y un 5,3% es extranjera.

Tabla 26. Distribución de frecuencias del nivel educativo

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Analfabeta	273	4,4	4,4	4,4
	Educación no formal	83	1,4	1,4	5,8
	Primaria no completa	1449	23,6	23,6	29,4
	Primaria completa	750	12,2	12,2	41,6
	Secundaria no completa	1516	24,7	24,7	66,3
	Secundaria completa	481	7,8	7,8	74,1
	Superior no completa	589	9,6	9,6	83,7
	Superior completa	221	3,6	3,6	87,3
	NRNivel Educativo	792	12,8	12,8	100,0
	Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Se observa en la tabla 26 que el nivel educativo de los miembros es secundaria no completa (24,7%), seguido de primaria no completa (23,6%), con la mínima frecuencia observada educación no formal (1,4%).

Tabla 27. Distribución de frecuencias de la dedicación de las personas

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Esta trabajando	1936	31,5	31,5	61,4
Buscando trabajo	195	3,2	3,2	64,5
Asiste a centro de enseñanza	1004	16,3	16,3	80,9
Oficios del hogar	848	13,8	13,8	94,7
Jubilado o pensionado	104	1,7	1,7	96,4
Vive de las rentas	36	0,6	0,6	97,0
Otra situación	186	3,0	3,0	100,0
No esta trabajando	1837	29,9	29,9	100,0
Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

La dedicación más predominante de las personas es el trabajo (31,5%), los que no trabajan con un 29,9 % seguidos de los que asisten a centros de enseñanza con un 16,3%. (Tabla 27)

Tabla 28. Distribución de frecuencias de las personas con más de un trabajo remunerado

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 0	4113	66,9	66,9	66,9
1	56	0,9	0,9	67,8
2	1977	32,2	32,2	100,0
Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

En la tabla 28 se observa que predominan las personas que tienen cero trabajos remunerados (66,9%), seguido de las personas con dos trabajos remunerado (32,2%) y por ultimo las personas que tienen un trabajo remunerado con (0,9%).

Tabla 29. Distribución de frecuencias de la categoría ocupacional

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Otras categorías	4022	65,4	65,4	65,4
	Empleado u obrero	1413	23,0	23,0	88,4
	Patrón o empleador	120	2,0	2,0	90,4
	Trabajador por cuenta propia	533	8,7	8,7	99,1
	Trabajador familiar no remunerado	2	0,0	0,0	99,1
	Buscando trabajo por primera vez	56	0,9	0,9	100,0
	Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

La categoría ocupacional más predominante es otra categoría (65,4%), seguido de empleado u obrero con un 23%. (Tabla 29)

Tabla 30. Distribución de frecuencias de la actividad de las personas

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Actividades no especificadas	4080	66,4	66,4	66,4
	Agricultura, caza, silvicultura y pesca	71	1,2	1,2	67,5
	Explotación de minas y canteras	14	0,2	0,2	67,8
	Industrias manufactureras	210	3,4	3,4	71,2
	Electricidad, gas y agua	27	0,4	0,4	71,6
	Construcción	140	2,3	2,3	73,9
	Comercio al mayor y detal, restaurantes y hoteles	498	8,1	8,1	82,0
	Transporte almacenamiento y comunicación	111	1,8	1,8	83,8
	Establecimientos financieros seguros bienes inmuebles y servicios	81	1,3	1,3	85,1
	Servicios comunales, sociales y personales	914	14,9	14,9	100,0
	Total	6146	100,0	100,0	

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 30 se observa que predominan las personas que tienen actividades no especificadas (66,4%), seguido de las personas que prestan Servicios comunales, sociales y personales (14,9%) y la mínima actividad que realizan las personas es Explotación de minas y canteras (0,4%).

4.2 Análisis Factorial de Correspondencia Múltiple de las características de las viviendas vs las variables suplementarias ingresos y número de miembros.

Tabla 31. Inercia y descomposición Chi-cuadrada

Val ores Si ngul ares	Inerci a Pri nci pal	Chi - Cuadrado	Porcentaj e	Porcentaj e Acumul ado	2	4	6	8	10
0. 54165	0. 29339	7014. 9	8. 93	8. 93	*****	*****	*****	*****	*****
0. 42452	0. 18022	4309. 0	5. 48	14. 41	*****	*****	*****	*****	*****
0. 39009	0. 15217	3638. 4	4. 63	19. 05	*****	*****	*****	*****	*****
0. 35888	0. 12880	3079. 6	3. 92	22. 97	*****	*****	*****	*****	*****
0. 35076	0. 12303	2941. 8	3. 74	26. 71	*****	*****	*****	*****	*****
0. 33644	0. 11319	2706. 4	3. 44	30. 15	*****	*****	*****	*****	*****
0. 31817	0. 10123	2420. 5	3. 08	33. 24	*****	*****	*****	*****	*****
0. 31119	0. 09684	2315. 5	2. 95	36. 18	*****	*****	*****	*****	*****
0. 30663	0. 09402	2248. 0	2. 86	39. 04	*****	*****	*****	*****	*****
0. 29914	0. 08948	2139. 6	2. 72	41. 77	*****	*****	*****	*****	*****
0. 29190	0. 08520	2037. 3	2. 59	44. 36	*****	*****	*****	*****	*****
0. 28665	0. 08217	1964. 6	2. 50	46. 86	*****	*****	*****	*****	*****
0. 28298	0. 08008	1914. 7	2. 44	49. 30	*****	*****	*****	*****	*****
0. 28046	0. 07866	1880. 7	2. 39	51. 69	*****	*****	*****	*****	*****
0. 27271	0. 07437	1778. 2	2. 26	53. 96	*****	*****	*****	*****	*****
0. 27045	0. 07314	1748. 8	2. 23	56. 18	*****	*****	*****	*****	*****
0. 26965	0. 07271	1738. 5	2. 21	58. 40	*****	*****	*****	*****	*****
0. 26879	0. 07225	1727. 4	2. 20	60. 59	*****	*****	*****	*****	*****
0. 26796	0. 07180	1716. 8	2. 19	62. 78	*****	*****	*****	*****	*****
0. 26723	0. 07141	1707. 5	2. 17	64. 95	*****	*****	*****	*****	*****
0. 26463	0. 07003	1674. 4	2. 13	67. 08	*****	*****	*****	*****	*****
0. 26296	0. 06915	1653. 4	2. 10	69. 19	*****	*****	*****	*****	*****
0. 26103	0. 06814	1629. 2	2. 07	71. 26	*****	*****	*****	*****	*****
0. 25816	0. 06665	1593. 6	2. 03	73. 29	*****	*****	*****	*****	*****
0. 25165	0. 06333	1514. 2	1. 93	75. 22	*****	*****	*****	*****	*****
0. 24456	0. 05981	1430. 1	1. 82	77. 04	*****	*****	*****	*****	*****
0. 23955	0. 05739	1372. 1	1. 75	78. 78	*****	*****	*****	*****	*****
0. 23792	0. 05661	1353. 5	1. 72	80. 51	*****	*****	*****	*****	*****
0. 23537	0. 05540	1324. 7	1. 69	82. 19	*****	*****	*****	*****	*****
0. 23047	0. 05312	1270. 1	1. 62	83. 81	*****	*****	*****	*****	*****
0. 22743	0. 05172	1236. 7	1. 57	85. 38	*****	*****	*****	*****	*****
0. 22361	0. 05000	1195. 6	1. 52	86. 91	*****	*****	*****	*****	*****
0. 22189	0. 04923	1177. 2	1. 50	88. 40	*****	*****	*****	*****	*****
0. 21281	0. 04529	1082. 9	1. 38	89. 78	*****	*****	*****	*****	*****
0. 20581	0. 04236	1012. 8	1. 29	91. 07	*****	*****	*****	*****	*****
0. 20351	0. 04142	990. 3	1. 26	92. 33	*****	*****	*****	*****	*****
0. 18884	0. 03566	852. 7	1. 09	93. 42	*****	*****	*****	*****	*****
0. 18675	0. 03487	833. 9	1. 06	94. 48	*****	*****	*****	*****	*****
0. 17514	0. 03067	733. 4	0. 93	95. 41	*****	*****	*****	*****	*****
0. 17209	0. 02961	708. 1	0. 90	96. 31	*****	*****	*****	*****	*****
0. 16845	0. 02837	678. 4	0. 86	97. 18	*****	*****	*****	*****	*****
0. 15997	0. 02559	611. 9	0. 78	97. 96	*****	*****	*****	*****	*****
0. 15531	0. 02412	576. 8	0. 73	98. 69	*****	*****	*****	*****	*****
0. 14874	0. 02213	529. 0	0. 67	99. 36	*****	*****	*****	*****	*****
0. 12195	0. 01487	355. 6	0. 45	99. 82	*****	*****	*****	*****	*****
0. 07752	0. 00601	143. 7	0. 18	100. 00	*****	*****	*****	*****	*****
Total	3. 28571	78. 562, 5	100. 00						

Degrees of Freedom = 3. 481

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 31 se muestran los resultados de los autovalores, porcentajes de variación, además del valor de la inercia y la distancia Chi-cuadrada, donde, podemos observar que los dos primeros autovalores explican un 14,41% de la variación total (tasa de inercia acumulada), siendo esto el mejor ajuste para la variación total de los datos, considérense los primeros 20 autovalores, los cuales explican el 64,95 % de la variación total lo que es significativo tomando en cuenta que el análisis lo que trata es reducir la dimensionalidad, para facilitar la interpretación de los resultados, obsérvese la figura que muestra, que el mejor ajuste se obtiene considerando los dos primeros autovalores y

siendo el eje 1 el que mejor explica la variación total de los datos por mostrar características predominantes. Además se puede observar que la distancia total Chi-cuadrado es de 78.562,5 con 3.481 grados de libertad y la variación total 3.28571 inercia resultante que fue descompuesta a lo largo de 46 direcciones principales de alargamiento de las nubes de puntos modalidades e individuos.

Tabla 32. Calidad y masa para las Características de las viviendas vs las variables suplementarias ingreso y número de miembros.

Estadísticas básicas para los puntos columnas

	Cal i dad	Masa	I nerci a
apt	0. 6012	0. 0144	0. 0174
cas	0. 5687	0. 0545	0. 0051
pv	0. 0325	0. 0002	0. 0217
pza	0. 0233	0. 0002	0. 0217
rub	0. 5943	0. 0021	0. 0211
al q	0. 1818	0. 0131	0. 0177
ced	0. 0355	0. 0003	0. 0216
gra	0. 0546	0. 0028	0. 0209
oc	0. 0747	0. 0003	0. 0216
pga	0. 2646	0. 0129	0. 0178
pgo	0. 3590	0. 0420	0. 0090
not2	0. 0495	0. 0709	0. 0001

Estadísticas básicas para los puntos columnas (continuación)

	Cal i dad	Masa	I nerci a
pago	0. 0089	0. 0001	0. 0217
pgo2	0. 0602	0. 0004	0. 0216
ci s	0. 0762	0. 0293	0. 0128
si m	0. 4602	0. 0111	0. 0183
tcu	0. 3429	0. 0283	0. 0131
ud	0. 2262	0. 0027	0. 0209
ci	0. 2754	0. 0083	0. 0192
si t	0. 1228	0. 0014	0. 0213
tcu2	0. 1189	0. 0442	0. 0083
und	0. 3891	0. 0175	0. 0164
ads	0. 2710	0. 0005	0. 0216
adt	0. 0866	0. 0028	0. 0209
bl o	0. 2160	0. 0648	0. 0020
bl s	0. 1398	0. 0023	0. 0210
con	0. 0052	0. 0002	0. 0217
mad	0. 0034	0. 0001	0. 0217
otp	0. 2037	0. 0008	0. 0215
abs	0. 0941	0. 0069	0. 0196

l am	0. 4368	0. 0164	0. 0168
nreg	0. 0022	0. 0001	0. 0217
otr	0. 1079	0. 0002	0. 0217
pl a	0. 5544	0. 0387	0. 0100
tej	0. 1075	0. 0092	0. 0189
cem	0. 6359	0. 0318	0. 0121
mos	0. 6361	0. 0393	0. 0098
ti e	0. 4677	0. 0004	0. 0216
dvi	0. 4105	0. 0691	0. 0007
fvi	0. 3865	0. 0020	0. 0211
not	0. 5726	0. 0003	0. 0216
acu	0. 4972	0. 0708	0. 0002
cam	0. 4037	0. 0001	0. 0217
otm	0. 2899	0. 0005	0. 0216
pi l	0. 0509	0. 0001	0. 0217
dos	0. 1369	0. 0186	0. 0161
nob	0. 7116	0. 0015	0. 0213
tre	0. 2956	0. 0098	0. 0188
uno	0. 4546	0. 0416	0. 0091
cl oa	0. 4067	0. 0686	0. 0008
hoyo	0. 4114	0. 0007	0. 0215
noex	0. 3815	0. 0003	0. 0216
sept	0. 1002	0. 0018	0. 0212
otra	0. 0431	0. 0002	0. 0217
serl	0. 0431	0. 0712	0. 0001

Estadísticas básicas para los puntos columnas (continuación)

	Cal i dad	Masa	I nerci a
el ec	0. 1556	0. 0005	0. 0216
gasc	0. 1371	0. 0699	0. 0005
kerg	0. 0529	0. 0007	0. 0215
noct	0. 0390	0. 0002	0. 0217
otco	0. 1681	0. 0001	0. 0217

Calidad de representación para los puntos columnas suplementarios

four	0. 0568
one	0. 0291
tree	0. 0066
two	0. 0236
l 1	0. 1642
l 2	0. 1018
l 3	0. 0529
l 4	0. 0104

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

Calidad: Indica que los cuatro primeros ejes explican la formación de las categorías, por ejemplo, los cuatro primeros ejes explican el 60,12 % de la categoría (apt) tipo de vivienda apartamento en edificio, casa o quinta, el 59,43 % de la categoría (rub) tipo de vivienda rancho urbano, el 46,02 % de la categoría (sim) número de habitaciones igual o mayores de siete, el 55,44 % de la categoría (pla) material del techo platabanda, el 63,61 % de la categoría (mos) material del piso mosaico, cerámica, ladrillo, alfombra y similares, el 3, y así de forma similar para las demás categorías.

Masa: Indica el peso de una categoría, dentro de todas las modalidades de todas las preguntas, por ejemplo, dentro de todas las modalidades la opción (serl) servicio eléctrico público en la clase de alumbrado tiene un peso del 7,12 %, de forma similar ocurre con los demás pesos de las categorías. (Tabla 32).

Tabla 33 Contribuciones parciales para las Características de las viviendas vs las variables suplementarias ingreso y número de miembros.

Contribución parcial de la Inercia para los puntos columnas

	Di m1	Di m2	Di m3	Di m4
apt	0.0075	0.0042	0.2058	0.0001
cas	0.0007	0.0019	0.0591	0.0005
pv	0.0009	0.0057	0.0006	0.0073
pza	0.0000	0.0018	0.0040	0.0056
rub	0.1260	0.0081	0.0006	0.0211
al q	0.0002	0.0132	0.0413	0.0144
ced	0.0000	0.0002	0.0151	0.0015
gra	0.0038	0.0144	0.0002	0.0001
oc	0.0122	0.0013	0.0005	0.0111
pga	0.0074	0.0264	0.0535	0.0032
pgo	0.0010	0.0001	0.0666	0.0010
not2	0.0000	0.0001	0.0000	0.0000
pago	0.0001	0.0003	0.0034	0.0002
pgo2	0.0001	0.0139	0.0068	0.0054
ci s	0.0032	0.0060	0.0001	0.0092
si m	0.0059	0.0504	0.1106	0.0009
tcu	0.0018	0.0398	0.0414	0.0063
ud	0.0423	0.0044	0.0033	0.0143
ci	0.0049	0.0336	0.0578	0.0084
si t	0.0005	0.0090	0.0430	0.0023
tcu2	0.0049	0.0025	0.0071	0.0020
und	0.0276	0.0541	0.0082	0.0147
otp	0.0236	0.0152	0.0007	0.0358
ads	0.0352	0.0077	0.0006	0.0575
adt	0.0008	0.0096	0.0262	0.0000
bl o	0.0040	0.0008	0.0008	0.0000
bl s	0.0163	0.0271	0.0000	0.0001
con	0.0002	0.0006	0.0013	0.0002
mad	0.0000	0.0005	0.0010	0.0000
abs	0.0002	0.0138	0.0068	0.0194
l am	0.0413	0.0525	0.0162	0.0001
nreg	0.0001	0.0000	0.0008	0.0001
otr	0.0100	0.0016	0.0060	0.0277
pl a	0.0177	0.0336	0.0444	0.0013
tej	0.0005	0.0007	0.0420	0.0002
ceto	0.0164	0.0799	0.0378	0.0018
mos	0.0206	0.0552	0.0284	0.0011
ti e	0.0851	0.0406	0.0043	0.0024
dvi	0.0028	0.0006	0.0000	0.0000

Contribución parcial de la Inercia para los puntos columnas (continuación)

	Di m1	Di m2	Di m3	Di m4
fvi	0.0768	0.0057	0.0004	0.0250
not	0.0088	0.0281	0.0001	0.2566
acu	0.0003	0.0003	0.0000	0.0016
cam	0.0123	0.0413	0.0033	0.1338
otm	0.0123	0.0098	0.0000	0.1178
pi l	0.0049	0.0002	0.0004	0.0164
dos	0.0082	0.0264	0.0003	0.0002
nob	0.1445	0.0256	0.0015	0.0197
tre	0.0080	0.0610	0.0274	0.0054
uno	0.0011	0.0669	0.0037	0.0049
cl oa	0.0035	0.0000	0.0001	0.0006
hoyo	0.0718	0.0106	0.0015	0.0457
noex	0.0577	0.0368	0.0014	0.0260
sept	0.0100	0.0109	0.0082	0.0066
otra	0.0089	0.0009	0.0004	0.0019
serl	0.0000	0.0000	0.0000	0.0000
el ec	0.0020	0.0235	0.0020	0.0458
gasc	0.0005	0.0000	0.0000	0.0005
kerp	0.0098	0.0034	0.0000	0.0020
noct	0.0006	0.0082	0.0004	0.0082
otco	0.0323	0.0091	0.0023	0.0041

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

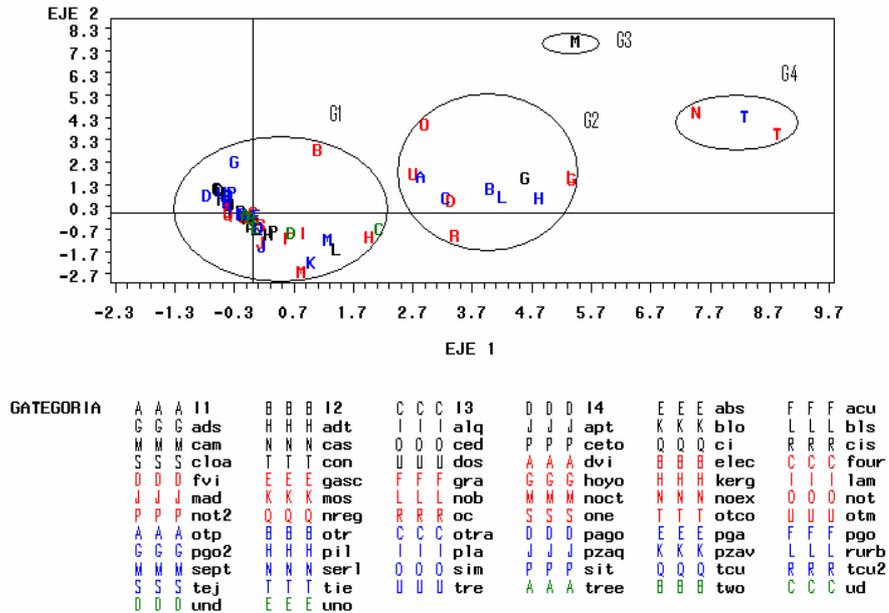
Las contribuciones parciales de la inercia para los puntos columnas nos dice que categorías son las que aportan más a la formación de los ejes. Los valores más grandes de Dim 1 de acuerdo a la categoría (nob) no tiene número de habitaciones de baño con ducha, del tipo de vivienda rancho urbano (rub) son las que más contribuyen a la formación del primer eje las cuales pertenecen a las variables números de cuartos de baño con ducha y tipo de vivienda respectivamente, contribuye más a la formación del segundo eje la categoría (ceto) cemento de la variable material en el piso, contribuyen más a la formación del tercer eje la categoría (apt) Apto, en edificio, casa o quinta y (sim) mayor o igual a 7 habitaciones de las variables tipo de vivienda y número de habitaciones totales respectivamente, y así de forma similar contribuyen más a la formación del cuarto eje (Tabla 33).

A través de los índices de coordenadas (anexo 3) podemos observar cuales o qué categorías están mejor representadas por qué eje, lo cual lo indica la columna best. Por ejemplo el eje 1 representa mejor la inercia a la variación de las variables tipo de vivienda rancho urbano (rub), tenencia de la vivienda uno ocupante de hecho (oc), del número de habitaciones totales uno a dos habitaciones de la vivienda (ud), material en paredes exteriores bloques o ladrillos frisados (blo), material en el piso tierra (tie), donde esta ubicada la instalación para agua corriente por tubería dentro de la vivienda (dvi), donde esta ubicada la instalación para agua corriente por tubería fuera de la vivienda (fvi), no tiene números de cuartos de baño con ducha (nob), tiene esta vivienda excusado que corresponde a la cloaca (cloa), tiene esta vivienda excusado que corresponde al hoyo o letrina (hoyo), no tiene la vivienda excusado (noex), la clase de alumbrado que utiliza la vivienda es otro (otra) o servicio eléctrico público (serl), combustible que se utiliza para cocinar es el gas (gasc), o kerosene o gasolina (kerg), o otra (otco). Mientras que los demás ejes recogen más variación de las categorías restantes.

En el anexo 3 también se pueden observar los cosenos cuadrados para los puntos columnas que no son más que la parte de la varianza de la categoría explicada por el eje. Por ejemplo, el 60,61 % de la variación de la variable no tiene número de habitaciones de baño con ducha (nob) es recogida por el primer eje, el eje 2 recoge 6,59 %, el eje 3 recoge 0,32 % y el eje 4 recoge 3,63 %. Otro ejemplo se puede ver con la categoría (rub) rancho urbano de la variable tipo de vivienda uno el 53,29 % de la variación de esta categoría es explicada por el primer eje. De igual manera hacemos el análisis para los demás ejes.

Gráfico 1 Plano factorial

Características de las viviendas vs las variables suplementarias ingreso y numeros de miembros



Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

El gráfico 1 muestra la disposición de las modalidades en el primer plano factorial. Se observan cuatro grupos y el primer eje factorial está determinado por las variables no tiene números de habitaciones de baño con ducha (nob), el tipo de vivienda es rancho urbano (rurb), situación que se corrobora con los cosenos cuadrados, recuérdese que un valor de estos cercanos a 1.0 indica un ángulo de la modalidad con el respectivo eje próximo a 0.0, es decir una alta modalidad entre la modalidad y el eje. Se destaca la diferenciación mostrada entre el grupo ingreso más alto a mejor posesión de las características de las viviendas, lo que sugiere una relación directa entre el grupo vivienda e ingreso. En cambio, no se podría decir que se observa bastante diferenciación entre la posesión características de las viviendas con respecto al número de miembros, en este sentido, se observa que existe un mayor número de miembros por una parte, a que no posean una mejor calidad en las características de las viviendas.

Sin embargo en este gráfico se observa claramente la diferenciación que hay, ya que en el lado derecho, predominan las variables de peor características de las viviendas, entre las que están: el tipo de vivienda es rancho urbano (rub), ó pieza en casa de vecindad (pv), la tenencia de la vivienda uno es propia pagada (pgo), gratuita (gra), ocupante de hecho (oc), o cedida por empleador (ced), la ubicación de agua por tubería esta fuera de la vivienda (fvi) o no tiene (not), el excusado es hoyo, letrina (hoyo) o no tiene (noex), el abastecimiento de agua es por medio de camión tanque (cam) o pila pública (pil), el material en las paredes exteriores es de adobe, tapia, o bahareque frisados (adt) o sin frisar (ads), el material en el piso es tierra (tie) o cemento (ceto), el número de habitaciones de baño con ducha es dos (dos) o no tiene (nob), entre la clase de alumbrado es el servicio eléctrico público (serl) u otro (otra) el que predomina, utilizan cualquier tipo de combustible para cocinar o no cocinan, aunque es extraño observar que esta la electricidad (elec) como otro combustible que se utiliza para cocinar y el ingreso es el más bajo (I1). En el lado izquierdo del gráfico predominan las variables de mejores características de las viviendas, entre las que están: el tipo de vivienda es casa o quinta (cas), la tenencia de la vivienda 2 es propia pagada (pgo2) o pagando (pago), el número de habitaciones totales es igual o mayor de tres (tcu, cis, sim), el material en las paredes exteriores es concreto (con), madera aserrada (mad), bloques o ladrillos frisados (blo), el material en el techo es teja (tej), el número de habitaciones de baño con ducha es de tres (tre), la clase de alumbrado es planta eléctrica (plan), la instalación de agua por tubería está dentro de la vivienda (dvi), el abastecimiento de agua es el acueducto (acu) y con los mejores ingresos observados.

4.3 Análisis Factorial de Correspondencias Múltiples de las características del jefe de la familia vs el tipo de vivienda

Tabla 34 Inercia y descomposición Chi-cuadrada para las Características del jefe de la familia vs el tipo de vivienda.

Descomposición Chi-cuadrado e Inercia					
Val ores Si ngul ares	I nerci a Pri nci pal	Chi - Cuadrado	Porcentaj e	Porcentaj e Acumul ado	2 4 6 8 10 -+-----+-----+-----+-----+-----
0. 56652	0. 32095	3680. 3	7. 34	7. 34	*****
0. 53144	0. 28242	3238. 5	6. 46	13. 79	*****
0. 43661	0. 19063	2185. 9	4. 36	18. 15	*****
0. 43242	0. 18699	2144. 2	4. 27	22. 42	*****
0. 41085	0. 16879	1935. 6	3. 86	26. 28	*****
0. 39831	0. 15865	1819. 2	3. 63	29. 91	*****
0. 38885	0. 15120	1733. 8	3. 46	33. 36	*****
0. 38421	0. 14762	1692. 7	3. 37	36. 74	*****
0. 37726	0. 14233	1632. 1	3. 25	39. 99	*****
0. 37325	0. 13932	1597. 5	3. 18	43. 17	*****
0. 36843	0. 13574	1556. 5	3. 10	46. 28	*****
0. 36523	0. 13339	1529. 6	3. 05	49. 33	*****
0. 36362	0. 13222	1516. 2	3. 02	52. 35	*****
0. 36021	0. 12975	1487. 9	2. 97	55. 31	*****
0. 35647	0. 12707	1457. 1	2. 90	58. 22	*****
0. 35276	0. 12444	1426. 9	2. 84	61. 06	*****
0. 35166	0. 12367	1418. 1	2. 83	63. 89	*****
0. 34791	0. 12104	1388. 0	2. 77	66. 66	*****
0. 34275	0. 11748	1347. 1	2. 69	69. 34	*****
0. 34156	0. 11666	1337. 7	2. 67	72. 01	*****
0. 33566	0. 11267	1292. 0	2. 58	74. 58	*****
0. 33368	0. 11134	1276. 8	2. 54	77. 13	*****
0. 32674	0. 10676	1224. 2	2. 44	79. 57	*****
0. 32166	0. 10346	1186. 4	2. 36	81. 93	*****
0. 31835	0. 10134	1162. 1	2. 32	84. 25	*****
0. 31325	0. 09813	1125. 2	2. 24	86. 49	*****
0. 30473	0. 09286	1064. 8	2. 12	88. 62	*****
0. 30177	0. 09106	1044. 2	2. 08	90. 70	*****
0. 29416	0. 08653	992. 2	1. 98	92. 67	*****
0. 29059	0. 08445	968. 3	1. 93	94. 60	*****
0. 28086	0. 07888	904. 5	1. 80	96. 41	*****
0. 25104	0. 06302	722. 7	1. 44	97. 85	****
0. 22580	0. 05098	584. 6	1. 17	99. 01	***
0. 20772	0. 04315	494. 8	0. 99	100. 00	**
Total	4. 37500	50167. 8	100. 00		
Degrees of Freedom = 1764					

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 34 se muestran los resultados de los autovalores, porcentajes de variación, además del valor de la inercia y la distancia Chi-cuadrada, donde podemos observar que los tres primeros autovalores explican un 18,15 % de la variación total (tasa de inercia acumulada), siendo esto el mejor ajuste para la variación total de los datos, considérense los primeros 23 autovalores, los cuales explican el 79,57 % de la variación total lo que es significativo tomando en cuenta que el análisis lo que trata es reducir la dimensionalidad, para facilitar la interpretación de los resultados, obsérvese la

figura que muestra, que el mejor ajuste se obtiene considerando los tres primeros autovalores y siendo el eje 1 el que mejor explica la variación total de los datos por mostrar características predominantes. Además se puede observar que la distancia total Chi-cuadrado es de 50.167,8 con 1.764 grados de libertad y la variación total 4.37500 inercia resultante que fue descompuesta a lo largo de 34 direcciones principales de alargamiento de las nubes de puntos modalidades e individuos.

Tabla 35 Calidad y masa para las Características del jefe de la familia vs el tipo de vivienda.

Estadísticas básicas para los puntos columnas

	Cal i dad	Masa	I nerci a
apt	0. 2066	0. 0252	0. 0228
cas	0. 1293	0. 0954	0. 0068
pv	0. 0192	0. 0004	0. 0285
pza	0. 0057	0. 0003	0. 0285
rub	0. 1471	0. 0036	0. 0277
F	0. 6265	0. 0422	0. 0189
M	0. 6265	0. 0828	0. 0096
casd	0. 6078	0. 0649	0. 0137
di vor	0. 1794	0. 0090	0. 0265
noi d	0. 0174	0. 0001	0. 0285
sepd	0. 0253	0. 0043	0. 0276
sol t	0. 4179	0. 0245	0. 0230
uni	0. 2644	0. 0139	0. 0254
vi u	0. 3182	0. 0083	0. 0267

Estadísticas básicas para los puntos columnas (continuación)

	Cal i dad	Masa	I nerci a
16-30	0. 5874	0. 0236	0. 0232
31-45	0. 2177	0. 0473	0. 0178
46-60	0. 1744	0. 0373	0. 0200
61-75	0. 3003	0. 0139	0. 0254
76-90	0. 1120	0. 0029	0. 0279
1ncomp	0. 1801	0. 0243	0. 0230
2ncomp	0. 0909	0. 0271	0. 0224
anal f	0. 1075	0. 0054	0. 0273
noform	0. 0768	0. 0031	0. 0279
pri m	0. 0614	0. 0269	0. 0224
secun	0. 0988	0. 0171	0. 0247
sncomp	0. 4088	0. 0086	0. 0266
super	0. 2261	0. 0125	0. 0257
E	0. 1215	0. 0130	0. 0256
V	0. 1215	0. 1120	0. 0030
empl	0. 5243	0. 0624	0. 0143
otro	0. 8745	0. 0179	0. 0245
patron	0. 2738	0. 0101	0. 0263
tprop	0. 3572	0. 0346	0. 0207
Agri	0. 1181	0. 0047	0. 0275
comerc	0. 1669	0. 0265	0. 0225
const	0. 0881	0. 0091	0. 0265
el ect	0. 0681	0. 0018	0. 0282
estf	0. 0593	0. 0038	0. 0277
expl	0. 0404	0. 0010	0. 0284
i ndut	0. 0731	0. 0090	0. 0265
nespec	0. 8745	0. 0179	0. 0245
serv	0. 4279	0. 0436	0. 0186
transp	0. 0394	0. 0076	0. 0268

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

Calidad: Indica que los cuatro primeros ejes explican la formación de las categorías, por ejemplo, los cuatro primeros ejes explican el 60,78 % de la categoría casado (casd) de la variable estado civil, el 58,74 % de la categoría edad comprendida entre 16 y 30 años (16-30), el 87,45 % y 52,43 % de la categoría empleado (emp) y otras (otro) de la variable categoría ocupacional, el 87,45 % de la categoría (nespec) actividad no especificada, y así de forma similar para las demás categorías.

Masa: Indica el peso de una categoría, dentro de todas las modalidades de todas las preguntas, por ejemplo, dentro de todas las modalidades la opción tipo de vivienda casa o quinta (cas) tiene un peso del 9,54 %, y la opción nacionalidad Venezolana (V) un peso del 11,20 % de forma similar ocurre con los demás pesos de las categorías (Tabla 35).

Tabla 36 Contribuciones parciales para las características del jefe de la familia vs el tipo de vivienda

Contribución parcial de la Inercia para los puntos columnas

	Di m1	Di m2	Di m3	Di m4
apt	0.0000	0.0292	0.0294	0.0361
cas	0.0000	0.0043	0.0010	0.0129
pv	0.0017	0.0000	0.0094	0.0001
pza	0.0003	0.0007	0.0005	0.0017
rub	0.0000	0.0115	0.0713	0.0053
F	0.0259	0.1145	0.0330	0.0263
M	0.0132	0.0583	0.0168	0.0134
casd	0.0045	0.0309	0.1382	0.0000
di vor	0.0006	0.0325	0.0014	0.0599
noi d	0.0002	0.0002	0.0047	0.0061
sepd	0.0004	0.0012	0.0045	0.0092
sol t	0.0000	0.0764	0.0471	0.0612
uni	0.0100	0.0237	0.0972	0.0052
vi u	0.0741	0.0000	0.0342	0.0366
16-30	0.0041	0.0453	0.0298	0.2125
31-45	0.0305	0.0076	0.0041	0.0224
46-60	0.0003	0.0243	0.0052	0.0393
61-75	0.0832	0.0233	0.0004	0.0001
76-90	0.0381	0.0047	0.0005	0.0003
1ncomp	0.0114	0.0497	0.0001	0.0023
2ncomp	0.0145	0.0025	0.0108	0.0079
anal f	0.0134	0.0057	0.0323	0.0040
noform	0.0106	0.0023	0.0201	0.0080
pri m	0.0001	0.0069	0.0144	0.0070
secun	0.0063	0.0209	0.0106	0.0038
sncomp	0.0015	0.0238	0.0150	0.2008
super	0.0028	0.0183	0.0783	0.0238
E	0.0021	0.0324	0.0001	0.0202

Contribución parcial de la Inercia para los puntos columnas (continuación)

	Di m1	Di m2	Di m3	Di m4
V	0.0002	0.0038	0.0000	0.0024
empl	0.0433	0.0640	0.0000	0.0044
otro	0.2736	0.0024	0.0167	0.0104
patron	0.0052	0.0457	0.0726	0.0164
tprop	0.0034	0.0676	0.0608	0.0029
Agri	0.0012	0.0288	0.0196	0.0104
comerc	0.0031	0.0239	0.0455	0.0001
const	0.0062	0.0192	0.0034	0.0115
elect	0.0020	0.0007	0.0109	0.0293
estf	0.0036	0.0055	0.0168	0.0069
expl	0.0024	0.0000	0.0158	0.0064
indut	0.0058	0.0063	0.0074	0.0184
nespec	0.2736	0.0024	0.0167	0.0104
serv	0.0223	0.0677	0.0027	0.0431
transp	0.0041	0.0109	0.0006	0.0005

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

Las contribuciones parciales de la inercia para los puntos columnas nos dice que categorías son las que aportan más a la formación de los ejes.

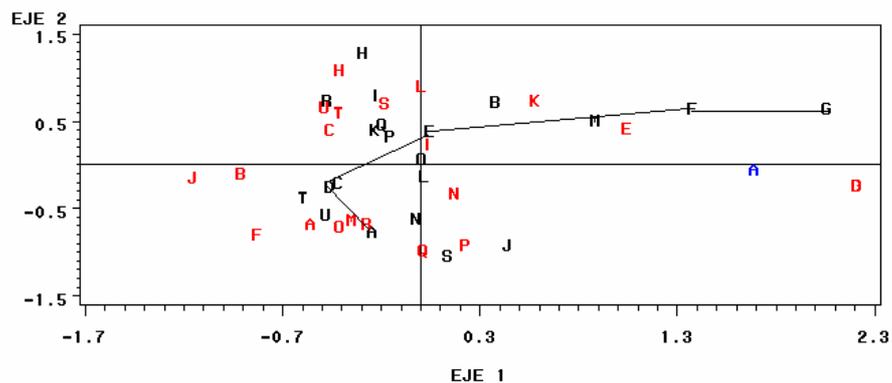
Los valores más grandes de Dim 1 de acuerdo a la categoría ocupacional otras categorías (otro), actividad no especificada (nespec) son las que más contribuyen a la formación del primer eje las cuales pertenecen a las variables categoría ocupacional y actividad que desempeña el jefe de la familia respectivamente, contribuye más a la formación del segundo eje la categoría de la variable género femenino (F) y estado civil soltero (solt), contribuyen más a la formación del tercer eje la categoría de la variable estado civil casado (casd) y unido (uni), y así de forma similar contribuyen más a la formación del cuarto eje (Tabla 36).

A través de los índices de coordenadas (anexo 3) podemos observar cuales o qué categorías están mejor representadas por qué eje, lo cual lo indica la columna best. Por ejemplo el eje 1 representa mejor la inercia a la variación de las categorías de la variable estado civil viudo (viu), de la variable edad entre 31-45 años (31-45), entre 61-75 años (61-75), entre 76-90 años (76-90), nivel educativo secundaria no completa (2ncomp), y otras categorías (otro). Mientras que los demás ejes recogen más variación de las categorías restantes.

En el anexo también se pueden observar los cosenos cuadrados para los puntos columnas que no son más que la parte de la varianza de la categoría explicada por el eje. Por ejemplo, el 82,02 % de la variación de la variable otras categorías ocupacionales (otro) y de la variable actividad no especificada (nespec) es recogida por el primer eje, el eje 3 recoge 2,98 % y el eje 4 recoge 1,81 % de dichas categorías. El eje 2 recoge 39,05 % de ambas categorías del sexo, el eje 3 recoge 43,83% de la variable estado civil casado (casd) y el eje 4 recoge 39,21% de la edad del jefe entre 16 y 30 años (16-30). De igual manera hacemos el análisis para los demás ejes.

Gráfico 2 Plano factorial

Características del jefe de la familia vs el tipo de vivienda



GATEGORIA	A A A 16-30	B B B 1ncomp	C C C 2ncomp	D D D 31-45	E E E 46-60
	F F F 61-75	G G G 76-90	H H H Agri	I I I E	J J J F
	K K K M	L L L V	M M M analf	N N N apt	O O O cas
	P P P casd	Q Q Q comerc	R R R const	S S S divor	T T T elect
	U U U empl	A A A estf	B B B expl	C C C indut	D D D nespec
	E E E noform	F F F noid	G G G otro	H H H patron	I I I prim
	J J J pv	K K K pza	L L L rub	M M M secun	N N N sepd
	O O O serv	P P P sncomp	Q Q Q solt	R R R super	S S S tprop
	T T T transp	U U U uni	A A A viu		

Fuente: Elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

En el gráfico 2 se observa la disposición de las modalidades en el primer plano factorial. El primer eje factorial está determinado por las variables actividad no especificada (nespec), otra categoría ocupacional (otro) situación que se corrobora con

los cosenos cuadrados, recuérdese que un valor de estos cercanos a 1.0 indica un ángulo de la modalidad con el respectivo eje próximo a 0.0, es decir una alta modalidad entre la modalidad y el eje.

En este gráfico se observa claramente la diferenciación que hay, ya que en el lado derecho, predominan la mayoría de las edades del jefe de la familia con los tipos de vivienda en pieza en casa, quinta o apartamento, casa o quinta y rancho urbano, con los niveles educativos no formal, primaria completa, primaria no completa y analfabeta, estado civil divorciado, separado y sexo femenino. En el lado izquierdo predomina la edad del jefe de la familia entre 31 y 45 años, con el mejor nivel educativo, con los estados civil soltero o unido o casado, los tipos de vivienda apartamento, en edificio, casa o quinta y pieza en casa de vecindad, las categorías empleado u obrero, patrón o empleador, y trabajador por cuenta propia, con todas las actividades económicas excepto la no especificada.

4.4 Análisis de Componentes Principales de los gastos vs los ingresos familiares en la Región de Los Andes

Para el análisis de componentes principales la muestra que originalmente comprendía a 1.179 familias se redujo a 442, ya que en este análisis lo que se estudia son los gastos comunes de las familias con respecto al ingreso.

Tabla 37 Matriz de correlación

	ACD	ACP	ALD	AFH	AOA	VSC	HVS
ACD	1.0000	0.4051	0.4955	0.2507	0.3446	0.1892	0.2096
ACP	0.4051	1.0000	0.5112	0.3997	0.3532	0.2971	0.3692
ALD	0.4955	0.5112	1.0000	0.4562	0.4101	0.2677	0.2812
AFH	0.2507	0.3997	0.4562	1.0000	0.2423	0.1967	0.2033
AOA	0.3446	0.3532	0.4101	0.2423	1.0000	0.2279	0.1757
VSC	0.1892	0.2971	0.2677	0.1967	0.2279	1.0000	0.1777
HVS	0.2096	0.3692	0.2812	0.2033	0.1757	0.1777	1.0000
HSED	0.0504	0.1528	0.1316	0.0394	0.1136	0.1453	0.1607
HOG	0.1005	0.1058	0.1336	0.1553	0.1086	0.2042	0.2268
DTR	0.1488	0.4148	0.3126	0.3870	0.2301	0.2379	0.3139
DOD	0.1661	0.2290	0.2679	0.2276	0.2093	0.2097	0.2448
ING1	-0.1634	-0.2292	-0.1516	-0.0543	-0.1157	-0.0602	-0.4312
ING4	0.1241	0.0496	0.0302	-0.0210	0.1193	-0.0013	0.2539

	HSED	HOG	DTR	DOD	ING1	ING2
ACD	0.0504	0.1005	0.1488	0.1661	-0.1634	0.1241
ACP	0.1528	0.1058	0.4148	0.2290	-0.2292	0.0496
ALD	0.1316	0.1336	0.3126	0.2679	-0.1516	0.0302
AFH	0.0394	0.1553	0.3870	0.2276	-0.0543	-0.0210
AOA	0.1136	0.1086	0.2301	0.2093	-0.1157	0.1193
VSC	0.1453	0.2042	0.2379	0.2097	-0.0602	-0.0013
HVS	0.1607	0.2268	0.3139	0.2448	-0.4312	0.2539
HSED	1.0000	0.1425	0.1626	0.1059	-0.1632	0.0026
HOG	0.1425	1.0000	0.1771	0.2046	-0.2923	0.1420
DTR	0.1626	0.1771	1.0000	0.2400	-0.1884	0.0288
DOD	0.1059	0.2046	0.2400	1.0000	-0.1672	0.0824
ING1	-0.1632	-0.2923	-0.1884	-0.1672	1.0000	-0.1932
ING2	0.0026	0.1420	0.0288	0.0824	-0.1932	1.0000

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 37 puede comprobarse que entre las variables de alimentos cereales y derivados (ACD) y leche, quesos y huevos (ALD) la asociación lineal es fuerte, al igual que con las variables leche, quesos y huevos (ALD), carnes y pescados (ACP), la correlación entre ellas es 0,4955 y 0,5112 respectivamente, esto quiere decir que las familias consumen más cereales, leches, quesos, huevos, carnes y pescados. Cuando el ingreso percibido es mayor que 60080,21 Bs (I4) la asociación lineal que existe más alta

es con la variable vivienda y servicios (HVS) igual a 0,2539, lo que quiere decir que cuando el ingreso es mayor que 60080,21 Bs, se invierte más en la vivienda que cuando el ingreso es menor.

Tabla 38

Autovalores de la matriz de correlación: Total = 13 Promedio = 0.86666667

	Autovalores	Diferencia	Proporción	Acumulada
1	3.65661632	2.19483875	0.2813	0.2813
2	1.46177756	0.33847107	0.1124	0.3937
3	1.12330650	0.17986071	0.0864	0.4801
4	0.94344578	0.02011785	0.0726	0.5527
5	0.92332793	0.12924254	0.0710	0.6237
6	0.79408540	0.01801289	0.0611	0.6848
7	0.77607251	0.03325988	0.0597	0.7445
8	0.74281264	0.08332686	0.0571	0.8016
9	0.65948578	0.11268980	0.0507	0.8524
10	0.54679598	0.05132446	0.0421	0.8944
11	0.49547151	0.03565237	0.0381	0.9326
12	0.45981914	0.04283620	0.0354	0.9679
13	0.41698294	0.41698294	0.0321	1.0000
14	0.00000000	0.00000000	0.0000	1.0000
15	0.00000000		0.0000	1.0000

En la tabla 38 podemos observar los autovalores asociado a cada componente principal, cuya suma es 25, lo que indica que se usó la matriz de correlación para los cálculos, en este caso podemos interpretar el porcentaje de variación como por ejemplo, 28,13 % de la variación total la recoge el primer componente principal y el 11,24 % de la variación total recoge el segundo componente, el tercero 8,64 %.

Ahora bien, observando el porcentaje de variación vemos que, por ejemplo, los cuatro primeros componentes principales recogen el 55,27 % de la variación total, mientras los cinco primeros recogen el 62,37 %.

Tabla 39. Autovectores propios

	CP1	CP2	CP3	CP4
ACD	0.300075	0.179581	-0.433684	0.245066
ACP	0.382434	0.160648	-0.045753	-0.033797
ALD	0.378486	0.269519	-0.143704	0.062593
AFH	0.304395	0.286900	0.110514	-0.416413
AOA	0.293596	0.172031	-0.232399	0.327861
VSC	0.247156	0.098541	0.305950	0.272802
HVS	0.308825	-0.364124	-0.067548	-0.178581

HSED	0. 149509	-0. 180252	0. 432826	0. 634478
HOG	0. 199411	-0. 371375	0. 241471	-0. 026398
DTR	0. 311889	0. 035657	0. 305068	-0. 346777
DOD	0. 253946	-0. 075069	0. 158612	-0. 149071
I NG1	-0. 224573	0. 507487	0. 060899	0. 044027
I NG2	0. 104019	-0. 424547	-0. 507152	0. 002115

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 39 se observa los autovectores correspondientes a los autovalores, es decir, presenta los componentes asociados a cada componente principal. El primer autovector presenta dos coeficientes positivos más elevados para ACP (0.382434), y ALD (0.378486) el resto de coeficientes son pequeños ó es menor con respecto a los valores.

Tabla 40. Factores del modelo

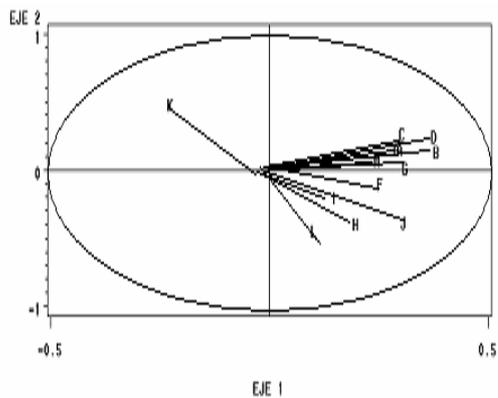
	Factor1	Factor2	Factor3	Factor4
ACP	0. 73130	0. 19423	-0. 04849	-0. 03283
ALD	0. 72375	0. 32586	-0. 15231	0. 06080
DTR	0. 59640	0. 04311	0. 32333	-0. 33683
HVS	0. 59054	-0. 44024	-0. 07159	-0. 17346
AFH	0. 58207	0. 34687	0. 11713	-0. 40447
ACD	0. 57381	0. 21712	-0. 45965	0. 23804
AOA	0. 56142	0. 20799	-0. 24631	0. 31845
DOD	0. 48560	-0. 09076	0. 16811	-0. 14479
VSC	0. 47262	0. 11914	0. 32426	0. 26498
I NG1	-0. 42944	0. 61357	0. 06454	0. 04276
HOG	0. 38132	-0. 44901	0. 25593	-0. 02564
I NG2	0. 19891	-0. 51329	-0. 53751	0. 00205
HSED	0. 28590	-0. 21793	0. 45874	0. 61628

Fuente: Cálculos propios basados en la I Encuesta de Presupuestos Familiares (BCV).

En la tabla 40 se puede observar cuales son las variables que están fuertemente vinculadas a un factor, mientras más elevado es el coseno (en valor absoluto), más vinculada esta la variable al eje. En este caso, se ve que las variables que están

vinculadas más al factor 1 son los alimentos (carnes y pescados (ACP), leches, quesos y huevos (ALD), tomados fuera del hogar (AFH), cereales y derivados (ACD), otros alimentos (AOA)), además del transporte y comunicaciones (DTR), y la vivienda y sus servicios (HVS).

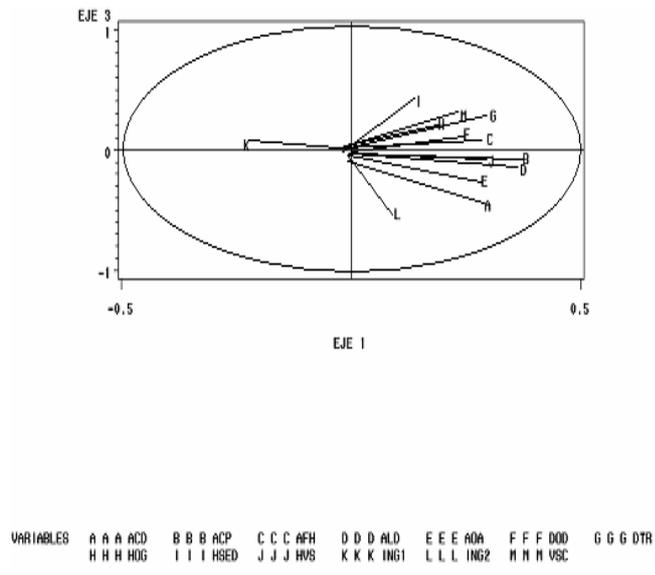
Gráfico 3 Proyección de las variables en el plano factorial eje1*eje2



VARIABLES A A A ACD B B B ACP C C C AFH D D D ALD E E E AOA F F F DOD G G G DTR
 H H H HOG I I I HSED J J J HVS K K K ING1 L L L ING2 N N N VSC

Fuente: elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

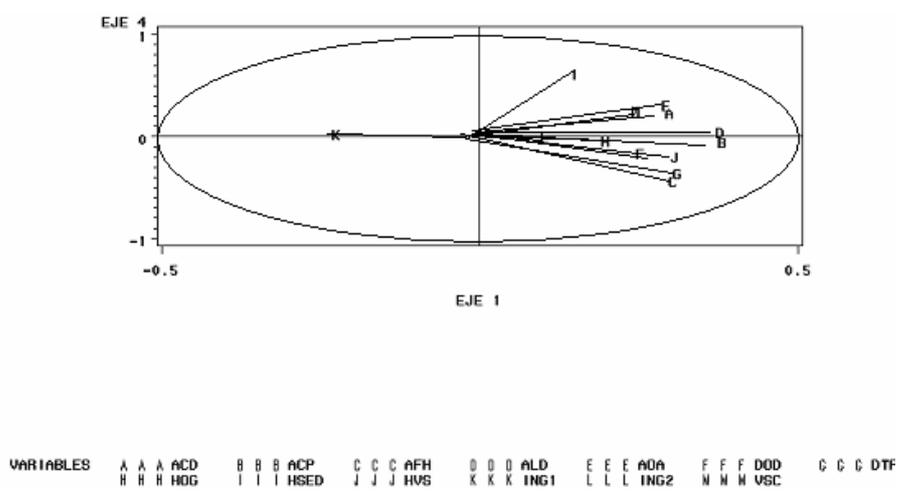
Gráfico 4 Proyección de las variables en el plano factorial eje1*eje3



Fuente: elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

Gráfico 5

Proyección de las variables en el plano factorial eje1*eje4



Fuente: elaboración propia según la I Encuesta de Presupuestos Familiares (BCV).

ANEXO 1 DEFINICIONES BASICAS DE LA ENCUESTA DE PRESUPUESTOS FAMILIARES

Períodos de referencia: Se refiere a los períodos de tiempo, dentro de los cuales va a ser considerada la realización de determinados pagos o la percepción de ciertos ingresos.

En general los periodos de referencia se definen atendiendo principalmente a los siguientes criterios:

- Capacidad de los informantes para recordar
- Importancia del gasto, en términos de valor
- Frecuencia del gasto

Estos tres aspectos se influyen mutuamente. Los gastos que son más frecuentes y/o los que por su bajo monto son más difíciles de recordar, tendrán periodos de referencia menores y viceversa, aquellos que por su monto y características son más fáciles de recordar, tendrán períodos de referencia más largos. En virtud de estos criterios se han establecido para los gastos de consumo, los siguientes periodos de referencia.

- **Semanal:** Se investigan con período de referencia semanal aquellos productos con alta frecuencia de consumo o con precios bajos que pudieran ser fácilmente olvidados. Estos son: alimentos y bebidas consumidas en el hogar, artículos de cuidado personal, periódicos, pequeñas reparaciones, detergentes, etc.
- **Mensual:** Se investigan con período de referencia mensual aquellos productos, cuya frecuencia de compras o uso es más espaciada que la semanal, por ejemplo: prendas de vestir, calzados y accesorios o aquellos servicios cuyo pago se realiza mensualmente: alquiler mensual de la vivienda, mensualidad escolar, pago de la luz, teléfono, agua, etc.
- **Anual:** Se investiga productos de escasa frecuencia de compra y de un valor unitario, que se considera suficiente para evitar olvidos, por ejemplo: muebles,

equipos del hogar, equipos de salud, mantenimiento de vivienda, seguros, pagos de impuestos, etc.

Hogar: un hogar esta constituido por un conjunto de personas que comparten o dependen de una bolsa o fondo común para sus gastos o como mínimo para el gesto en alimentación y que han residido en la misma vivienda la mayor parte de los últimos seis meses o manifiestan tener intenciones de fijar allí su residencia.

Este fondo esta conformado por las aportaciones de los miembros con las cuales se atienden a los gastos que origina la satisfacción de las necesidades, tanto particulares como comunes de los mismos.

La relación de los miembros del hogar puede estar o no fundamentada en nexos consanguíneos o legales y puede depender de diferentes factores, por ejemplo, las costumbres o las estrategias de supervivencia.

Vivienda: Es toda estructura hecha o no para vivir, pero utilizada con ese fin. Esta conformada por una habitación o conjunto de habitaciones y sus dependencias, que constituyen en si misma una edificación o una parte estructuralmente separada de la misma y que, por la forma en que han sido construidas, reconstruidas, transformadas o adaptadas, están habitadas por uno o varios hogares.

Estos ambientes o habitaciones están destinadas o de algún modo se han adaptado para servir como mínimo, de área social, dormitorio, lugar de aseo personal, cocina, etc. Presentan una o mas entradas y están interconectadas entre si. Las personas que la habitan comparten en todo o en parte el uso de la misma.

Un hogar que habite una vivienda puede utilizarla como:

- Vivienda principal: Es la vivienda en la que el hogar reside la mayor parte del tiempo. Es la sede o residencia habitual del hogar
- Vivienda ocasional: Es una vivienda que puede ser utilizada por alguno o varios miembros del hogar con motivo de algunas actividades especiales y que deben desarrollar lejos de su vivienda principal residiendo parte del tiempo en otro lugar, por ejemplo un trabajo fuera de la ciudad de su residencia habitual. También puede ser utilizada para descanso y esparcimiento.

Miembros del hogar: Se consideran miembros del hogar, aquellas personas que comparten y dependen habitualmente del presupuesto de gastos comunes del hogar (colabore económicamente o no en el mismo). Se incluyen los niños y ancianos. Además han residido en la vivienda toda o la mayor parte de los últimos seis meses o han residido en la vivienda encuestada una parte del año mayor que en cualquier otra vivienda del hogar (en caso de que resida alternativamente en distintas viviendas).

Las condiciones que se establecen para determinar si una persona es o no miembro del hogar, trata de evitar las posibilidades de que un mismo individuo pueda clasificarse en más de un hogar o por el contrario, no pueda clasificarse en ninguno.

Jefe del hogar: Se considera jefe del hogar aquel miembro del hogar que periódicamente aporta más al presupuesto o bolsa común que se destina a atender los gastos del hogar. En la mayoría de las veces es el esposo o la esposa y en ocasiones el jefe del hogar puede ser el abuelo(a) u otro miembro del mismo hogar.

Servicio domestico: El tratamiento dado al personal del servicio domestico depende de las condiciones de su trabajo.

- No duerme en la vivienda: Si el servicio domestico no duerme en la vivienda, no existe ninguna duda de que es miembro de un hogar que tiene como vivienda

principal una vivienda diferente a la del hogar en que trabaja. No se considera miembro del hogar.

- Duerme en la vivienda: Si el servicio domestico duerme en la vivienda, se presentan dos casos:
 - Que habite la vivienda donde presta sus servicios, la mayor parte de la semana, pero regularmente o casi todas las semanas se traslada a su hogar de origen y comparte los gastos generales de ese hogar salvo en lo que corresponde a los días que esta trabajando. Compartir los gastos generales, significa que además del aporte que pueda realizar para sufragar los demás gastos, también efectúe una parte relevante de consumo final en dicho hogar. En este caso se considera al servicio domestico miembro de su hogar de origen.
 - Que habite la vivienda donde presta sus servicios la mayor parte de la semana, pero que aun cuando se traslade regularmente a su hogar de origen y ayude en dinero o en especies a ese hogar, no comparte los gastos generales. Se considera en este caso, al servicio domestico como otro hogar separado, que tiene como vivienda principal la vivienda del hogar para el que trabaja. Este hogar estará conformado por el servicio domestico y los familiares del mismo que convivan en la vivienda, independientemente de que también presten servicios domésticos en la misma

Pensionistas: Son aquellas personas que alquilan a otro hogar alguna(s) habitación(es) de la vivienda de dicho hogar y que además de tener el derecho al uso de la misma pueden a veces recibir otros servicios que les son prestados por el hogar arrendador tales como: alimentación, lavado de ropa, etc.

Variables de estudio

Gastos del hogar

Gastos de consumo final

Aunque en este trabajo sólo nos interesa la Región de Los Andes, para los hogares residentes se registraron los gastos en:

- Bienes y servicios adquiridos y utilizados en el país (caso más frecuente).
- Bienes y servicios adquiridos en el país para su utilización fuera de él (paquetes turísticos, viajes de avión al exterior, etc).
- Bienes y servicios adquiridos fuera del país para su utilización en el país (compra de ropa, perfumes, etc.).
- Bienes y servicios adquiridos y utilizados en el exterior (alquiler de vehículos, gastos en hotel y restaurantes durante viajes, etc.)

Gastos de consumo no monetario

- Consumo de la producción por cuenta propia del hogar (autoconsumo), se incluyen todos los bienes y/o servicios producidos por el hogar sin ánimo de ofrecerlos en el mercado y que están destinados para su propio consumo. Abarca los productos primarios de la agricultura, caza, selvicultura y pesca, las manufacturas de esos productos (queso, mantequilla, vinos, telas, calzados, vestidos, muebles, etc.)
- Consumo de bienes y servicios recibidos como ingreso o remuneración en especie. Se incluye el valor de los bienes y/o servicios suministrados por los empleadores a sus asalariados como parte de su remuneración por el trabajo realizado, ejemplo:
 - Alimentos y bebidas
 - Alojamiento de un tipo que puedan utilizarlo todos los miembros del hogar al que pertenece el preceptor
 - Vestimenta utilizable fuera del lugar de trabajo

- Servicio de transporte, estacionamiento de vehículos, servicios médicos, etc.
- Consumo de bienes y servicios por retiro del negocio (autosuministro), se incluye el valor de los bienes y/o servicios producidos o comercializados por empresas no constituidas en sociedad, propiedad del hogar que son retirados del negocio, sin realizar ningún pago, por alguno o algunos de los miembros del hogar propietario para su consumo.
- Consumo de bienes y servicios por donaciones en especie (regalos), se incluye el valor de los bienes y/o servicios que recibe el hogar por parte de otros hogares o de empresas constituidas en sociedad, sin ofrecer a cambio ninguna contraprestación en bienes y/o servicio y/o dinero.

Gastos no imputables al consumo

- Consumo intermedio de bienes y servicios, en su función productiva los hogares utilizan bienes y servicios y factores primarios para obtener bienes y servicios. Los bienes y servicios que los hogares adquieren para el desarrollo de sus actividades productivas constituyen consumo intermedio de las mismas, mientras que los adquiridos para actividades no productivas forman parte de su gasto de consumo.
- Gastos en reparaciones mayores, renovación o ampliación de la vivienda propiedad del hogar.
- Transferencias corrientes en efectivo y/o en especie sin contrapartida, comprende las erogaciones en dinero realizadas por los hogares sin recibir nada como contrapartida, por ejemplo los pagos de cuotas sindicales o para asociaciones políticas, etc.
- Intereses por deudas de los hogares, se refiere a por ejemplo: Impuestos directos (impuesto sobre la renta), Primas netas de seguro de riesgo, contribuciones a la seguridad social y a la caja de pensiones, multas y sanciones, etc.

Ingresos del hogar

Es el conjunto de ingresos percibidos por los miembros del hogar, preceptores de ingresos, cualesquiera sea su origen y naturaleza, en el periodo de referencia, para hacer frente a sus gastos inmediatos, a sus necesidades futuras o para incrementar su patrimonio.

Atendiendo su naturaleza, los ingresos pueden ser monetarios y no monetarios.

Ingresos monetarios: Se considera ingresos monetarios, los ingresos percibidos por los miembros del hogar en los períodos de referencia:

- Ingreso por trabajo asalariado, es la remuneración total en dinero que pagan las empresas a sus asalariados en contraprestación al trabajo realizado.
- Ingreso del trabajo por cuenta propia, Se refiere a los ingresos netos que reciben los trabajadores por el desarrollo de su actividad por cuenta propia
- Rentas del capital y la propiedad, Se consideran rentas de capital y la propiedad, los ingresos percibidos en concepto de intereses de cuentas corrientes, cuentas de ahorro y depósitos a plazo; intereses y dividendos de acciones, obligaciones, bonos, etc.
- Transferencias recibidas, son los ingresos periódicos percibidos por el hogar sin mediar contraprestación alguna.
- Transferencias extraordinarias y otros ingresos ocasionales anuales, se consideran transferencias extraordinarias, los ingresos percibidos de manera extraordinaria por el hogar,
- tanto de la administración pública como de otros hogares o instituciones, sin que exista contraprestación alguna. Se incluye los premios de loterías, kinos, terminales, etc.

VARIABLES A NIVEL DE VIVIENDA

- Números de miembros del hogar que habitan en la vivienda y su desagregación en hogares.

- Identificación de las personas que prestan servicio domestico al hogar

Características de la vivienda:

- Tipo de vivienda
- Materiales predominantes de los pisos, paredes, etc.
- Tipo y número de habitaciones y espacios, en la vivienda.
- Disponibilidad de energía eléctrica, gas directo, etc.
- Instalaciones y forma de abastecimiento relacionadas con el agua

Variables a nivel de hogares y miembros

- Características de los miembros del hogar.
 - Género
 - Edad
 - Parentesco o relación con el jefe del hogar
 - Estado Civil
 - Nacionalidad
 - Nivel de educación
- Relación de los miembros del hogar con la actividad económica
 - Ocupados o trabajando actualmente, desempleados, buscando trabajo por primera vez, etc.
 - Número de trabajos
 - Situación profesional en la ocupación principal y en otras ocupaciones que desarrolle (patrón o empleador, trabajador por cuenta propia, empleado remunerado, obrero remunerado, trabajador familiar, etc.)
 - Ocupación o tipo de trabajo desarrollado
 - Rama de actividad de la empresa donde desarrolla su ocupación principal y otras ocupaciones que realice
- Hábitos de comidas
 - Tipo de comida (desayuno, almuerzo y cena) que realizan habitualmente en la semana
 - Lugares donde realizan sus comidas

- Combustibles utilizados por el hogar con excepción de los usados en los vehículos automotores
- Dotación y equipamiento del hogar
- Vivienda de que dispone el hogar
 - Tipo (principal, secundaria o de descanso)
 - Tenencia: propia, alquilada, ocupada sin pago
 - En el caso de vivienda alquilada: monto del alquiler, quien lo paga
 - En el caso de vivienda propia: información sobre si está o no pagada
 - En el caso de ocupación sin pago: Cual es el motivo: remuneración en especie, donación en especie por parte de instituciones privadas sin fines de lucro u otros hogares
 - Valor que se estima tendría el alquiler de la vivienda si se alquilara en el periodo de referencia.

ANEXO 2 MANUAL DEL USUARIO

Para observar la información que posee la base de datos realice los siguientes pasos:

1. Introduzca el CD que contiene la información en el computador
2. Vaya a mi PC pulse doble clic, encima del icono



Luego doble clic en disco compacto



Y busque el archivo denominado Presupuesto de Microsoft Access



Pulse de nuevo doble clic con lo cual se abrirá después de unos segundos la ventana Región de los andes (formulario1) de la base de datos, este formulario muestra las consultas básicas.

Seleccione una Entidad de la lista despegable como se muestra

The screenshot shows a software interface titled "REGION DE LOS ANDES". Under the heading "Consultas básicas", there is a section "Seleccione una categoría" with a list of categories: "Unidad basica de muestreo", "Caracteristicas de la vivienda", "Caracteristicas generales del hogar", "Gastos diarios del hogar", "Gastos mensuales", "Gastos anuales", and "Ingresos". To the right, there is an "Entidad" dropdown menu showing a list of entities: "1", "2", and "3", with corresponding labels "Mer", "San", and "Rest" visible next to them. Below these sections is a "Seleccionar una tabla" section which is currently empty.

Luego una categoría pulsando sobre ella, con la cual se muestra la tabla o las tablas relacionadas, pulse dos veces o haga doble clic sobre la categoría que quiera observar de este modo vera la información que contiene dicha tabla para la entidad seleccionada.

This screenshot shows the same software interface as the previous one. In the "Seleccione una categoría" list, "Caracteristicas de la vivienda" is now highlighted. In the "Entidad" dropdown menu, the value "1" is selected. The "Seleccionar una tabla" section now displays a list of tables: "Aspectos generales", "Materiales predominantes", and "Dotación y equipamiento".

ANEXO 3

Análisis Factorial de Correspondencias Múltiples de las características de las viviendas vs. las variables suplementarias ingreso y número de miembros

Codificación de las variables

Variable	Codificación	Descripción
Tipo de vivienda	ap.	Apto, en edificio, casa o quinta
	khas	Casa o quinta
	Pb.	Pieza en casa de vecindad
	Pas	Pieza en casa, quinta o apto
	Rut	Rancho urbane
	Otto	Otra clase
Tenencia de la vivienda 1	Ali	Aquacade
	pía	Praia pagan do
	pgo	Praia pagoda
	gras	Gratuity
	Oc.	Ocupante de hecho
	cede	Cedida p/empleador
Tenencia de la vivienda 2	pago	Praia pagan do
	pgo2	Propia pagada
	gra2	Gratuita
	ocu2	Ocupante de hecho
	ced2	Cedida p/empleador
Número de habitaciones totales	ud	1 a 2 habitaciones
	tcu	3 a 4 habitaciones
	cis	5 a 6 habitaciones
	sim	Mayor o igual a 7 habitaciones
Número de habitaciones para dormir	und	1 a 2 habitaciones
	tcu2	3 a 4 habitaciones
	ci	5 a 6 habitaciones
	sit	Mayor o igual a 7 habitaciones
Material en paredes exteriores	blo	Bloques o ladrillos frisados
	bls	Bloques o ladrillos sin frisar
	con	Concreto
	mad	Madera aserrada
	adt	Adobe, tapia o bahareque frisados
	ads	Adobe, tapia o bahareque sin frisar
	otp	Otros (caña, palma, lata, cartón y similares)
Material en el techo	nreg	No registrado
	pla	Platabanda

(continuación)

Variable	Codificación	Descripción
Material en el techo	tej	Teja
	abs	Asbestos y similares
	lam	Lámina metálica
	otr	Otros (caña, palma, tablas, lata, y similares)
Material en el piso	mos	Mosaico, cerámica, ladrillo, alfombra y similares
	ceto	Cemento
	tie	Tierra
Donde esta ubicada la instalación para agua corriente por tubería	dvi	Dentro de la vivienda
	fvi	Fuera de la vivienda
	not	No tiene
Como se abastece normalmente el agua	acu	Acueducto
	pil	Pila pública
	cam	Camión tanque
	otm	Otros medios
Números de cuartos de baño con ducha	uno	Uno
	dos	Dos
	tre	Tres
	nob	No tiene
Tiene esta vivienda excusado	cloa	Cloaca
	sept	Pozo séptico
	hoyo	Hoyo o letrina
	noex	No tiene
Clase de alumbrado que utiliza	serl	Servicio eléctrico público
	plan	Planta eléctrica propia
	otra	Otro tipo
Combustible que utiliza para cocinar	gasc	Gas
	ker	kerosene o gasolina
	elec	Electricidad
	otco	Otra clase
	noct	No cocina
Números de miembros familiares	one	1 a 2 miembros
	two	3 a 4 miembros
	tree	5 a 6 miembros
	four	Mayor a 7 miembros
Ingreso familiar	I1	Menores o igual a 20240,57 Bs
	I2	Entre 20241,57 y 40160,39 Bs
	I3	Entre 40161,39 y 60080,21 Bs
	I4	Mayores que 60080,21 Bs

Fuente: Elaboración propia basada en la I Encuesta de presupuestos Familiares (BCV).

Índices de coordenadas que contribuyen más a la inercia de los puntos columnas

	Di m1	Di m2	Di m3	Di m4	Best
apt	0	0	3	0	3
cas	0	0	3	0	3
pv	0	0	0	0	4
pza	0	0	0	0	4
rub	1	0	0	1	1
alq	0	0	3	0	3
ced	0	0	0	0	3
gra	0	0	0	0	2
oc	0	0	0	0	1
pga	0	3	3	0	3
pgo	0	0	3	0	3
not2	0	0	0	0	2
pago	0	0	0	0	3
pgo2	0	0	0	0	2
cis	0	0	0	0	4
sim	0	3	3	0	3
tcu	0	3	3	0	3
ud	1	0	0	0	1
ci	0	3	3	0	3
sit	0	0	3	0	3
tcu2	0	0	0	0	3
und	2	2	0	0	2
Otp	4	0	0	4	4
ads	4	0	0	4	4
adt	0	0	0	0	3
blo	0	0	0	0	1
bls	0	2	0	0	2
con	0	0	0	0	3
mad	0	0	0	0	3
abs	0	0	0	0	4
lam	2	2	0	0	2
nreg	0	0	0	0	3
otr	0	0	0	4	4
pla	3	3	3	0	3
tej	0	0	3	0	3
ceto	0	2	2	0	2
mos	2	2	0	0	2
tie	1	1	0	0	1
dvi	0	0	0	0	1
fvi	1	0	0	1	1

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Índices de coordenadas que contribuyen más a la inercia de los puntos columnas
(Continuación)

	Di m1	Di m2	Di m3	Di m4	Best
not	0	4	0	4	4
acu	0	0	0	0	4
cam	0	4	0	4	4
otm	0	0	0	4	4
pi l	0	0	0	0	4
dos	0	2	0	0	2
nob	1	1	0	1	1
tre	0	2	0	0	2
uno	0	2	0	0	2
cl oa	0	0	0	0	1
hoyo	1	0	0	1	1
noex	1	1	0	1	1
sept	0	0	0	0	2
otra	0	0	0	0	1
serl	0	0	0	0	1
elec	0	4	0	4	4
gasc	0	0	0	0	1
kerf	0	0	0	0	1
noct	0	0	0	0	2
otco	1	0	0	0	1

Cosenos cuadrados para los puntos columnas

	Di m1	Di m2	Di m3	Di m4
apt	0. 0385	0. 0132	0. 5493	0. 0002
cas	0. 0124	0. 0199	0. 5324	0. 0040
pv	0. 0035	0. 0145	0. 0013	0. 0132
pza	0. 0001	0. 0045	0. 0086	0. 0102
rub	0. 5329	0. 0211	0. 0012	0. 0391
al q	0. 0012	0. 0409	0. 1078	0. 0319
ced	0. 0000	0. 0005	0. 0323	0. 0027
gra	0. 0160	0. 0379	0. 0005	0. 0001
oc	0. 0503	0. 0033	0. 0010	0. 0200
pga	0. 0369	0. 0814	0. 1392	0. 0071
pgo	0. 0097	0. 0005	0. 3444	0. 0044
not2	0. 0008	0. 0345	0. 0068	0. 0073
pago	0. 0005	0. 0008	0. 0073	0. 0004
pgo2	0. 0005	0. 0353	0. 0146	0. 0098
ci s	0. 0221	0. 0256	0. 0005	0. 0280
sim	0. 0285	0. 1507	0. 2792	0. 0018
tcu	0. 0121	0. 1661	0. 1460	0. 0187
ud	0. 1807	0. 0114	0. 0073	0. 0268

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Cosenos cuadrados para los puntos columnas (continuación)

	Di m1	Di m2	Di m3	Di m4
ci	0.0230	0.0959	0.1394	0.0172
si t	0.0020	0.0231	0.0934	0.0042
tcu2	0.0530	0.0167	0.0396	0.0096
und	0.1500	0.1810	0.0230	0.0351
otp	0.0981	0.0388	0.0016	0.0653
ads	0.1457	0.0197	0.0013	0.1043
adt	0.0033	0.0251	0.0581	0.0001
blo	0.1746	0.0228	0.0184	0.0001
bls	0.0691	0.0706	0.0001	0.0002
con	0.0008	0.0014	0.0027	0.0003
mad	0.0000	0.0012	0.0021	0.0001
abs	0.0007	0.0387	0.0160	0.0388
lam	0.2201	0.1719	0.0447	0.0001
nreg	0.0002	0.0000	0.0018	0.0001
otr	0.0410	0.0040	0.0128	0.0501
pla	0.1588	0.1846	0.2061	0.0050
tej	0.0023	0.0020	0.1028	0.0004
ceto	0.1216	0.3633	0.1451	0.0060
mos	0.1876	0.3094	0.1345	0.0046
tie	0.3513	0.1029	0.0091	0.0044
dvi	0.3627	0.0449	0.0005	0.0025
fvi	0.3244	0.0148	0.0010	0.0463
not	0.0365	0.0711	0.0003	0.4647
acu	0.1101	0.0785	0.0025	0.3061
cam	0.0505	0.1044	0.0070	0.2417
otm	0.0510	0.0250	0.0000	0.2138
pil	0.0200	0.0005	0.0009	0.0295
dos	0.0455	0.0901	0.0009	0.0004
nob	0.6061	0.0659	0.0032	0.0363
tre	0.0382	0.1784	0.0677	0.0113
uno	0.0107	0.4036	0.0191	0.0213
cloa	0.3731	0.0029	0.0035	0.0273
hoyo	0.2979	0.0271	0.0033	0.0832
noex	0.2382	0.0933	0.0030	0.0470
sept	0.0421	0.0281	0.0179	0.0122
otra	0.0366	0.0022	0.0009	0.0034
serl	0.0366	0.0022	0.0009	0.0034
elec	0.0084	0.0597	0.0043	0.0832
gasc	0.0967	0.0034	0.0000	0.0369
kerq	0.0407	0.0085	0.0000	0.0037
noct	0.0024	0.0209	0.0008	0.0148
otco	0.1328	0.0230	0.0050	0.0073

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Cosenos cuadrados para los puntos columnas suplementarios

	Di m1	Di m2	Di m3	Di m4
four	0.0007	0.0093	0.0448	0.0019
one	0.0030	0.0147	0.0104	0.0010
tree	0.0034	0.0032	0.0000	0.0000
two	0.0000	0.0046	0.0185	0.0005
l1	0.0264	0.1216	0.0153	0.0009
l2	0.0174	0.0787	0.0055	0.0001
l3	0.0063	0.0305	0.0139	0.0022
l4	0.0016	0.0082	0.0005	0.0000

Coordenadas de las columnas

	Di m1	Di m2	Di m3	Di m4
apt	-0.3902	0.2282	1.4737	-0.0270
cas	-0.0620	-0.0786	-0.4062	0.0351
pv	1.0186	-2.0638	0.6265	1.9656
pza	0.1912	-1.3247	1.8321	2.0006
rub	4.2362	0.8432	0.2021	-1.1481
alq	-0.0744	-0.4256	0.6912	0.3762
ced	0.0061	-0.3435	2.7537	0.7932
gra	0.6285	-0.9664	0.1115	0.0484
oc	3.4356	-0.8842	0.4887	-2.1696
pga	-0.4092	0.6075	0.7945	-0.1800
pgo	0.0825	0.0195	-0.4914	-0.0558
not2	0.0023	-0.0153	0.0068	-0.0071
pago	-0.7359	0.9518	2.9360	-0.6784
pgo2	-0.2814	2.4318	-1.5636	1.2816
ci s	-0.1781	0.1918	-0.0276	-0.2005
sim	-0.3924	0.9027	-1.2287	-0.1000
tcu	0.1358	-0.5033	0.4718	0.1689
ud	2.1588	-0.5431	0.4340	0.8316
ci	-0.4179	0.8540	-1.0295	-0.3617
si t	-0.3207	1.0773	-2.1664	0.4618
tcu2	-0.1805	0.1013	0.1560	-0.0769
und	0.6797	-0.7465	0.2664	0.3289
otp	2.8569	1.7964	0.3611	2.3306
ads	4.6188	1.6974	0.4313	-3.9077
adt	0.2854	-0.7866	-1.1965	-0.0443
blo	-0.1340	0.0484	0.0435	0.0038
bl s	1.4399	-1.4558	-0.0400	-0.0676
con	-0.5575	0.7386	1.0333	-0.3405
mad	0.1780	-1.1907	-1.5765	0.2451
abs	0.0831	-0.6010	-0.3862	0.6018
lam	0.8608	-0.7606	-0.3879	-0.0205
nreg	-0.3808	0.1167	1.0289	-0.2659
otr	4.0104	1.2445	2.2372	-4.4330

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Coordenadas de las columnas (continuación)

	Di m1	Di m2	Di m3	Di m4
pl a	-0.3670	0.3956	0.4180	-0.0648
tej	-0.1251	0.1152	-0.8335	-0.0520
ceto	0.3892	-0.6727	-0.4251	0.0863
mos	-0.3920	0.5035	0.3320	-0.0614
tie	8.2869	4.4859	1.3343	-0.9271
dvi	-0.1099	-0.0387	-0.0042	-0.0091
fvi	3.3563	0.7175	0.1837	-1.2682
not	2.9276	4.0851	-0.2497	10.4461
acu	-0.0322	-0.0272	0.0048	-0.0537
cam	5.4505	7.8400	-2.0324	11.9269
otm	2.7330	1.9111	-0.0660	5.5947
pil	4.8494	0.7797	-1.0510	-5.8979
dos	-0.3594	0.5058	0.0515	-0.0324
nob	5.4006	1.7813	0.3950	-1.3223
tre	-0.4899	1.0583	-0.6517	-0.2659
uno	0.0876	-0.5386	0.1170	0.1236
cl oa	-0.1231	-0.0108	0.0119	0.0333
hoyo	5.3828	1.6221	0.5627	-2.8440
noex	7.4787	4.6803	0.8394	3.3229
sept	1.2916	-1.0553	-0.8422	-0.6956
otra	3.2778	0.7975	0.5237	0.9969
serl	-0.0112	-0.0027	-0.0018	-0.0034
el ec	1.1115	2.9565	-0.7910	3.4900
gasc	-0.0467	-0.0088	0.0009	-0.0288
ker g	1.9899	-0.9115	-0.0075	-0.5987
noct	0.8478	-2.4749	0.4981	2.0877
otco	8.8408	3.6756	1.7140	2.0788

Coordenadas de las columnas suplementarias

	Di m1	Di m2	Di m3	Di m4
four	0.0476	0.1693	-0.3716	-0.0770
one	0.1648	-0.3639	0.3053	0.0954
tree	-0.0825	0.0806	-0.0091	-0.0041
two	-0.0025	-0.0985	0.1966	0.0331
I 1	0.0577	-0.1238	0.0439	0.0106
I 2	-0.4289	0.9116	-0.2410	-0.0369
I 3	-0.5646	1.2372	-0.8369	-0.3324
I 4	-0.5222	1.1727	-0.3008	0.0484

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Análisis Factorial de Correspondencias Múltiples de las características del jefe de la familia vs el tipo de vivienda.

Codificación de las variables

Variable	Codificación	Descripción
Tipo de vivienda	apt	Apto, en edificio, casa o quinta
	cas	Casa o quinta
	pv	Pieza en casa de vecindad
	pza	Pieza en casa, quinta o apto
	rub	Rancho urbano
Sexo	M	Masculino
	F	Femenino
Estado civil (civil)	noid	No identificado
	solt	Soltero
	uni	Unido
	casd	Casado
	divor	Divorciado
	sepd	Separado
	viu	Viudo
Edad	16-30	Entre 16 y 30 años
	31-45	Entre 31 y 45 años
	46-60	Entre 46 y 60 años
	61-75	Entre 61 y 75 años
	76-90	Entre 76 y 90 años
Nivel educativo (Educ)	analf	Analfabeta
	noform	Educación no formal
	1ncomp	Primaria no completa
	prim	Primaria completa
	2ncomp	Secundaria no completa
	secun	Secundaria completa
	sncomp	Superior no completa
	super	Superior completa
Nacionalidad (nacion)	V	Venezolano
	E	Extranjero
Categoría ocupacional (catego)	otro	Otras categorías
	empl	Empleado u obrero
	patron	Patron o empleador
	tprop	Trabajador por cuenta propia

Fuente: Elaboración propia basada en la I Encuesta de presupuestos Familiares (BCV).

Índices de coordenadas que contribuyen más a la inercia de los puntos columnas

	Di m1	Di m2	Di m3	Di m4	Best
apt	0	4	4	4	4
cas	0	0	0	0	4
pv	0	0	0	0	3
pza	0	0	0	0	4

rub	0	0	3	0	3
F	0	2	2	2	2
M	0	2	0	0	2
casd	0	3	3	0	3
di vor	0	4	0	4	4
noi d	0	0	0	0	4
sepd	0	0	0	0	4
sol t	0	2	2	2	2
uni	0	0	3	0	3
vi u	1	0	1	1	1
16-30	0	4	4	4	4
31-45	1	0	0	1	1
46-60	0	4	0	4	4
61-75	1	0	0	0	1
76-90	1	0	0	0	1
1ncomp	0	2	0	0	2
2ncomp	0	0	0	0	1
anal f	0	0	3	0	3
noform	0	0	3	0	3
prim	0	0	0	0	3
secun	0	0	0	0	2
sncomp	0	4	0	4	4
super	0	0	3	3	3
E	0	2	0	2	2
V	0	0	0	0	2
empl	2	2	0	0	2
otro	1	0	0	0	1
patron	0	3	3	0	3
tprop	0	2	2	0	2
Agri	0	2	2	0	2
comerc	0	3	3	0	3
const	0	0	0	0	2
elect	0	0	0	4	4
estf	0	0	0	0	3
expl	0	0	0	0	3
i ndut	0	0	0	0	4
nespec	1	0	0	0	1
serv	0	2	0	2	2
transp	0	0	0	0	2

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Cosenos cuadrados para los puntos columnas

	Di m1	Di m2	Di m3	Di m4
apt	0.0001	0.0826	0.0563	0.0677
cas	0.0002	0.0413	0.0062	0.0816
pv	0.0045	0.0000	0.0145	0.0002
pza	0.0009	0.0015	0.0008	0.0026
rub	0.0000	0.0268	0.1120	0.0082
F	0.1006	0.3905	0.0760	0.0595
M	0.1006	0.3905	0.0760	0.0595
casd	0.0242	0.1453	0.4383	0.0000
di vor	0.0015	0.0792	0.0022	0.0965
noi d	0.0006	0.0005	0.0071	0.0092
sepd	0.0011	0.0029	0.0070	0.0143
sol t	0.0001	0.2146	0.0893	0.1139
uni	0.0288	0.0602	0.1668	0.0087
vi u	0.2038	0.0000	0.0558	0.0586
16-30	0.0131	0.1262	0.0561	0.3921
31-45	0.1260	0.0278	0.0100	0.0539
46-60	0.0012	0.0781	0.0112	0.0839
61-75	0.2403	0.0592	0.0007	0.0001
76-90	0.1001	0.0108	0.0007	0.0004
1ncomp	0.0362	0.1394	0.0003	0.0042
2ncomp	0.0475	0.0073	0.0211	0.0150
anal f	0.0361	0.0136	0.0515	0.0063
noform	0.0278	0.0052	0.0314	0.0123
pri m	0.0003	0.0198	0.0280	0.0133
secun	0.0187	0.0547	0.0187	0.0067
sncomp	0.0041	0.0577	0.0245	0.3226
super	0.0080	0.0458	0.1326	0.0396
E	0.0059	0.0817	0.0001	0.0338
V	0.0059	0.0817	0.0001	0.0338
empl	0.2224	0.2888	0.0001	0.0131
otro	0.8202	0.0064	0.0298	0.0181
patron	0.0144	0.1123	0.1204	0.0267
tprop	0.0119	0.2112	0.1281	0.0060
Agri	0.0033	0.0675	0.0311	0.0162
comerc	0.0100	0.0686	0.0881	0.0001
const	0.0171	0.0468	0.0056	0.0185
elect	0.0051	0.0016	0.0169	0.0445
estf	0.0095	0.0128	0.0264	0.0106
expl	0.0063	0.0000	0.0243	0.0097
i ndut	0.0159	0.0154	0.0122	0.0296
nespec	0.8202	0.0064	0.0298	0.0181
serv	0.0878	0.2347	0.0064	0.0989
transp	0.0112	0.0263	0.0010	0.0009

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Coordenadas de las Columnas

	Di m1	Di m2	Di m3	Di m4
apt	-0.0177	-0.5715	-0.4717	0.5172
cas	0.0075	0.1131	0.0440	-0.1590
pv	-1.1508	-0.1077	2.0609	-0.2106
pza	0.5840	0.7768	-0.5445	1.0049
rub	0.0087	0.9503	1.9424	0.5258
F	0.4442	-0.8753	0.3861	-0.3417
M	-0.2264	0.4461	-0.1968	0.1741
casd	-0.1498	0.3669	-0.6372	0.0037
di vor	0.1407	-1.0094	0.1691	-1.1144
noi d	-0.8198	-0.7560	-2.8999	-3.2926
sepd	0.1752	-0.2824	0.4418	-0.6295
sol t	0.0173	-0.9386	0.6053	0.6838
uni	-0.4803	0.6938	1.1550	0.2636
vi u	1.6960	-0.0065	0.8879	-0.9096
16-30	-0.2368	-0.7356	0.4902	1.2964
31-45	-0.4551	-0.2137	-0.1283	-0.2976
46-60	0.0522	0.4284	-0.1624	-0.4439
61-75	1.3865	0.6884	0.0743	0.0274
76-90	2.0662	0.6800	-0.1735	-0.1374
1ncomp	0.3878	0.7605	-0.0329	-0.1325
2ncomp	-0.4140	-0.1617	0.2756	0.2328
anal f	0.8930	0.5477	1.0676	-0.3737
noform	1.0498	0.4558	1.1163	-0.6984
pri m	0.0357	0.2684	0.3193	-0.2199
secun	-0.3436	-0.5882	-0.3442	-0.2052
sncomp	0.2350	-0.8844	-0.5762	2.0911
super	-0.2677	-0.6421	-1.0919	-0.5970
E	-0.2255	0.8373	-0.0356	0.5385
V	0.0263	-0.0975	0.0041	-0.0627
empl	-0.4720	-0.5378	-0.0104	-0.1143
otro	2.2139	-0.1959	-0.4217	0.3292
patron	-0.4056	1.1320	-1.1722	0.5515
tprop	-0.1768	0.7434	0.5789	-0.1248
Agri	-0.2914	1.3194	-0.8960	0.6472
comerc	-0.1932	0.5050	0.5720	0.0214
const	-0.4662	0.7713	0.2679	0.4854
elect	-0.5931	-0.3261	-1.0748	-1.7444
estf	-0.5494	-0.6367	-0.9156	0.5801
expl	-0.9066	-0.0542	1.7777	1.1232
i ndut	-0.4528	0.4448	0.3962	0.6174
nespec	2.2139	-0.1959	-0.4217	0.3292
serv	-0.4051	-0.6623	-0.1092	-0.4300
transp	-0.4158	0.6362	-0.1238	-0.1153

Fuente: Cálculos propios basados en la I Encuesta de presupuestos Familiares (BCV).

Análisis de Componentes Principales de los gastos vs la variable suplementaria ingresos familiares en la Región de Los Andes

Codificación de las variables

Variable	Codificación	Descripción
Gastos		
Cereales y derivados	ACD	
Carnes y pescados	ACP	
Leche, quesos y huevos	ALD	
Fuera del hogar	AFH	
Otros alimentos	AOA	
Vestido y calzado	VSC	
Vivienda y servicios	HVS	
Educación, esparcimiento y salud	HSED	
Equipos del hogar	HOG	
Transporte y comunicaciones	DTR	
Otros bienes y servicios varios	DOD	
Ingreso familiar	I1	Menores o igual a 20240,57 Bs
	I2	Mayores que 60080,21 Bs

Fuente: Elaboración propia basada en la I Encuesta de presupuestos Familiares (BCV).

Estadísticas básicas

	Promedio	Desviación estándar	Mínimo	Máximo
ACD	407.2488688	303.0353647	17.00	2469.00
ACP	1146.687783	969.421808	21.00	9365.00
ALD	484.2918552	360.2244682	6.00	3299.00
AFH	700.805430	1176.327770	9.00	17612.00
AOA	1182.223982	956.896610	31.00	8989.00
VSC	2108.549774	2818.889495	18.00	26400.00
HVS	2295.097285	1575.605159	200.00	10851.00
HSED	910.595023	2327.632009	2.00	42319.00
HOG	1100.977376	2169.768234	4.00	37068.00
DTR	1710.409502	3831.880696	4.00	36340.00
DOD	834.085973	1648.401681	8.00	19707.00
ING	11724.022624	9303.099783	322.00	80000.00

ANEXO 4

- 1.- Programa que realiza el análisis factorial de correspondencias múltiples para las características de las viviendas vs las variables suplementarias ingreso y números de miembros familiares.

```
DATA PRUEBA;
input tip $ tn1 $ tn2 $ nha $ dha $ pare $ te $ piso $ tub $ ab $ nb $
exc $ luz $ co $ mi $ ing $;
cards;
acá van los datos
data ACML;
set prueba;
If tip='1' then tip='cas';
else If tip='2' then tip='apt';
else If tip='3' then tip='pza';
else If tip='4' then tip='pv';
else If tip='5' then tip='rub';
else If tip='6' then tip='otv';
If tn1='21' then tn1='alq';
else If tn1='22' then tn1='pga';
else If tn1='23' then tn1='pgo';
else If tn1='24' then tn1='gra';
else If tn1='25' then tn1='oc';
else If tn1='26' then tn1='ced';
else If tn1='00' then tn1='nol';
If tn2='21' then tn2='al2';
else if tn2='22' then tn2='pago';
else If tn2='23' then tn2='pgo2';
else If tn2='24' then tn2='gra2';
else If tn2='25' then tn2='ocu2';
else If tn2='26' then tn2='ced2';
else If tn2='00' then tn2='not2';
If nha='1' then nha='ud';
else If nha='2' then nha='tcu';
else If nha='3' then nha='cis';
else If nha='4' then nha='sim';
If dha='1' then dha='und';
else If dha='2' then dha='tcu2';
else If dha='3' then dha='ci';
else If dha='4' then dha='sit';
If pare='1' then pare='blo';
else If pare='2' then pare='bls';
else If pare='3' then pare='con';
else If pare='4' then pare='mad';
else If pare='5' then pare='adt';
else If pare='6' then pare='ads';
else If pare='7' then pare='otp';
If te='0' then te='nreg';
else If te='1' then te='pla';
else If te='2' then te='tej';
else If te='3' then te='abs';
else If te='4' then te='lam';
else If te='5' then te='otr';
```

```

If piso='1' then piso='mos';
else If piso='2' then piso='cem';
else If piso='3' then piso='tie';
If tub='1' then tub='dvi';
else If tub='2' then tub='fvi';
else If tub='3' then tub='not';
If ab='1' then ab='acu';
else If ab='2' then ab='pil';
else If ab='3' then ab='cam';
else If ab='4' then ab='otm';
If nb='1' then nb='uno';
else If nb='2' then nb='dos';
else If nb='3' then nb='tre';
else If nb='4' then nb='nob';
If exc='1' then exc='cloa';
else If exc='2' then exc='sept';
else If exc='3' then exc='hoyo';
else If exc='4' then exc='noex';
If luz='1' then luz='ser1';
else If luz='2' then luz='plan';
else If luz='3' then luz='otra';
If co='1' then co='gasc';
else If co='2' then co='kerg';
else If co='3' then co='elec';
else If co='4' then co='otco';
else If co='5' then co='noco';
If mi='1' then mi='one';
else If mi='2' then mi='two';
else If mi='3' then mi='tree';
else If mi='4' then mi='four';
If ing='1' then ing='I1';
else If ing='2' then ing='I2';
else If ing='3' then ing='I3';
else If ing='4' then ing='I4';
Proc corresp mca observed data=acml outc=graf cp dimens=4;
TABLES tip tn1 tn2 nha dha pare te piso tub ab nb exc luz co mi ing;
Supplementary mi ing;
RUN;

```

2.-Programa que realiza el gráfico del análisis factorial de correspondencias múltiples para las características de las viviendas vs las variables suplementarias ingreso y números de miembros familiares.

```

DATA VARIA;
INPUT GATEGORIA $ EJE1 EJE2;
CARDS;
Acá van los datos de las coordenadas de las variables citadas en 1
;
SYMBOL1 C=BLACK      V='A';
SYMBOL2 C=BLACK      V='B';
SYMBOL3 C=BLACK      V='C';
SYMBOL4 C=BLACK      V='D';
SYMBOL5 C=BLACK      V='E';
SYMBOL6 C=BLACK      V='F';

```

SYMBOL7	C=BLACK	V='G' ;
SYMBOL8	C=BLACK	V='H' ;
SYMBOL9	C=BLACK	V='I' ;
SYMBOL10	C=BLACK	V='J' ;
SYMBOL11	C=BLACK	V='K' ;
SYMBOL12	C=BLACK	V='L' ;
SYMBOL13	C=BLACK	V='M' ;
SYMBOL14	C=BLACK	V='N' ;
SYMBOL15	C=BLACK	V='O' ;
SYMBOL16	C=BLACK	V='P' ;
SYMBOL17	C=BLACK	V='Q' ;
SYMBOL18	C=BLACK	V='R' ;
SYMBOL19	C=BLACK	V='S' ;
SYMBOL20	C=BLACK	V='T' ;
SYMBOL21	C=BLACK	V='U' ;
SYMBOL22	C=RED	V='A' ;
SYMBOL23	C=RED	V='B' ;
SYMBOL24	C=RED	V='C' ;
SYMBOL25	C=RED	V='D' ;
SYMBOL26	C=RED	V='E' ;
SYMBOL27	C=RED	V='F' ;
SYMBOL28	C=RED	V='G' ;
SYMBOL29	C=RED	V='H' ;
SYMBOL30	C=RED	V='I' ;
SYMBOL31	C=RED	V='J' ;
SYMBOL32	C=RED	V='K' ;
SYMBOL33	C=RED	V='L' ;
SYMBOL34	C=RED	V='M' ;
SYMBOL35	C=RED	V='N' ;
SYMBOL36	C=RED	V='O' ;
SYMBOL37	C=RED	V='P' ;
SYMBOL38	C=RED	V='Q' ;
SYMBOL39	C=RED	V='R' ;
SYMBOL40	C=RED	V='S' ;
SYMBOL41	C=RED	V='T' ;
SYMBOL42	C=RED	V='U' ;
SYMBOL43	C=BLUE	V='A' ;
SYMBOL44	C=BLUE	V='B' ;
SYMBOL45	C=BLUE	V='C' ;
SYMBOL46	C=BLUE	V='D' ;
SYMBOL47	C=BLUE	V='E' ;
SYMBOL48	C=BLUE	V='F' ;
SYMBOL49	C=BLUE	V='G' ;
SYMBOL50	C=BLUE	V='H' ;
SYMBOL51	C=BLUE	V='I' ;
SYMBOL52	C=BLUE	V='J' ;
SYMBOL53	C=BLUE	V='K' ;
SYMBOL54	C=BLUE	V='L' ;
SYMBOL55	C=BLUE	V='M' ;
SYMBOL56	C=BLUE	V='N' ;
SYMBOL57	C=BLUE	V='O' ;
SYMBOL58	C=BLUE	V='P' ;
SYMBOL59	C=BLUE	V='Q' ;
SYMBOL60	C=BLUE	V='R' ;
SYMBOL61	C=BLUE	V='S' ;
SYMBOL62	C=BLUE	V='T' ;

```

SYMBOL63 C=BLUE      V='U';
SYMBOL64 C=GREEN     V='A';
SYMBOL65 C=GREEN     V='B';
SYMBOL66 C=GREEN     V='C';
SYMBOL67 C=GREEN     V='D';
SYMBOL68 C=GREEN     V='E';
GOPTIONS COLORS=(BLACK,BLACK) CBACK=WHITE;
PROC GPLOT DATA=VARIA;
TITLE1'Características de las viviendas vs las variables suplementarias
ingreso y numeros de miembros';
LABEL EJE1='EJE 1'
      EJE2='EJE 2';
AXIS1 LENGTH=2 IN ORDER=-2.7 TO 8.5;
AXIS2 LENGTH=6 IN ORDER=-2.3 TO 9.7;
PLOT EJE2*EJE1=GATEGORIA/ FRAME HAXIS=AXIS2 VAXIS=AXIS1
      HREF=0 VREF=0;

RUN;
PROC PRINT;
RUN;

```

3.-Programa que realiza el análisis factorial de correspondencias múltiples para las características del jefe de la familia vs el tipo de vivienda

```

DATA jefe;
input tip $ sexo $ civil $ edad $ educ $ nacion $ catego $ activi $;
cards;
Acá van los datos
;
data jefel;
set jefe;
If tip='1' then tip='cas';
else If tip='2' then tip='apt';
else If tip='3' then tip='pza';
else If tip='4' then tip='pv';
else If tip='5' then tip='rub';
else If tip='6' then tip='otv';
If sexo='1' then sexo='M';
else If sexo='2' then sexo='F';
If civil='0' then civil='noid';
else if civil='1' then civil='uni';
else if civil='2' then civil='casd';
else if civil='3' then civil='sepd';
else if civil='4' then civil='viu';
else if civil='5' then civil='solt';
else if civil='6' then civil='divor';
If edad='1' then edad='16-30';
else If edad='2' then edad='31-45';
else If edad='3' then edad='46-60';
else If edad='4' then edad='61-75';
else If edad='5' then edad='76-90';
If educ='1' then educ='analf';
else If educ='2' then educ='noform';
else If educ='3' then educ='lncomp';
else If educ='4' then educ='prim';

```

```

else If educ='5' then educ='2ncomp';
else If educ='6' then educ='secun';
else If educ='7' then educ='sncomp';
else If educ='8' then educ='super';
else If educ='9' then educ='ndeclara';
If nacion='10' then nacion='V';
else If nacion='20' then nacion='E';
If catego='0' then catego='otro';
else if catego='1' then catego='empl';
else If catego='2' then catego='patron';
else If catego='3' then catego='tprop';
If activi='0' then activi='nespec';
else If activi='1' then activi='Agri';
else If activi='2' then activi='expl';
else If activi='3' then activi='indut';
else If activi='4' then activi='elect';
else If activi='5' then activi='const ';
else If activi='6' then activi='comerc';
else If activi='7' then activi='transp';
else If activi='8' then activi='estf';
else If activi='9' then activi='serv';
Proc corresp mca observed data=jefel outc=jefe2 cp rp dimens=4;
TABLES tip sexo civil edad educ nacion catego activi;
RUN;

```

4.-Programa que realiza el gráfico del análisis factorial de correspondencias múltiples para las características del jefe de la familia vs el tipo de vivienda

```

DATA VARIA;
INPUT GATEGORIA $ EJE1 EJE2;
CARDS;
Acá van los datos de las coordenadas de las variables citadas en 3
;
SYMBOL1 C=BLACK V='A';
SYMBOL2 C=BLACK V='B';
SYMBOL3 C=BLACK V='C';
SYMBOL4 C=BLACK V='D';
SYMBOL5 C=BLACK V='E';
SYMBOL6 C=BLACK V='F';
SYMBOL7 C=BLACK V='G';
SYMBOL8 C=BLACK V='H';
SYMBOL9 C=BLACK V='I';
SYMBOL10 C=BLACK V='J';
SYMBOL11 C=BLACK V='K';
SYMBOL12 C=BLACK V='L';
SYMBOL13 C=BLACK V='M';
SYMBOL14 C=BLACK V='N';
SYMBOL15 C=BLACK V='O';
SYMBOL16 C=BLACK V='P';
SYMBOL17 C=BLACK V='Q';
SYMBOL18 C=BLACK V='R';
SYMBOL19 C=BLACK V='S';
SYMBOL20 C=BLACK V='T';
SYMBOL21 C=BLACK V='U';

```

```

SYMBOL22 C=RED      V='A';
SYMBOL23 C=RED      V='B';
SYMBOL24 C=RED      V='C';
SYMBOL25 C=RED      V='D';
SYMBOL26 C=RED      V='E';
SYMBOL27 C=RED      V='F';
SYMBOL28 C=RED      V='G';
SYMBOL29 C=RED      V='H';
SYMBOL30 C=RED      V='I';
SYMBOL31 C=RED      V='J';
SYMBOL32 C=RED      V='K';
SYMBOL33 C=RED      V='L';
SYMBOL34 C=RED      V='M';
SYMBOL35 C=RED      V='N';
SYMBOL36 C=RED      V='O';
SYMBOL37 C=RED      V='P';
SYMBOL38 C=RED      V='Q';
SYMBOL39 C=RED      V='R';
SYMBOL40 C=RED      V='S';
SYMBOL41 C=RED      V='T';
SYMBOL42 C=RED      V='U';
SYMBOL43 C=BLUE     V='A';
GOPTIONS COLORS=(BLACK,BLACK) CBACK=WHITE;
PROC GPLOT DATA=VARIA;
TITLE1 'Características del jefe de la familia vs el tipo de vivienda';
LABEL EJE1='EJE 1'
      EJE2='EJE 2';
AXIS1 LENGTH=2 IN ORDER=-1.0 TO 2.5;
AXIS2 LENGTH=6 IN ORDER=-1.5 TO 2.5;
PLOT EJE2*EJE1=GATEGORIA/ FRAME HAXIS=AXIS2 VAXIS=AXIS1
      HREF=0 VREF=0;

RUN;
PROC PRINT;
RUN;

```

5.-Programa que realiza el análisis de componentes principales de los gastos vs el ingreso familiar

```

FILENAME ARCH 'C:\DOCUMENTS AND
SETTINGS\AUXILIARES_2\ESCRITORIO\juana\artesis\dato.DBF';
PROC DBF DB3=ARCH OUT=PRUEBA;
DATA CATEGO;
  SET PRUEBA;
  ING1=0;
  ING2=0;
  IF ING<=20240.57 THEN ING1+1;
  ELSE IF ING>60080.21 THEN ING2+1;
PROC PRINCOMP DATA=CATEGO OUT=PRUEBA STANDARD N=4 PREFIX=CP;
  VAR ACD ACP ALD AFH AOA VSC HVS HSED HOG DTR DOD ING1 ING2;
PROC PRINT;
RUN;

```

BIBLIOGRAFÍA

- ANIDO, Daniel (1998). *Sistema Lineal del Gasto*. Mérida-Venezuela: Universidad de los Andes. Tesis de maestría.
- BATÍNI (1994). *Diseño Conceptual de Base de Datos*. Ed. Addison Wesley Díaz de Santos.
- DALLAS, Johnson. Edición en Español 2000. *Métodos Multivariados Aplicados al Análisis de Datos*. México: Internacional Thomson Editores.
- DERVIN, C. (1991). **¿Cómo interpretar los resultados de un Análisis Factorial de Correspondencias?** A. Tomassone y J. onson-Alvarez (traductores). Paris: Institut Technique des Céréales et des Fourrages.
- DÍAZ, Luis (2002). *Estadística Multivariada: Inferencia y métodos*. Colombia: Facultad de Ciencias. Universidad Nacional de Colombia.
- FERRÁN A., Magdalena. (2002). *SPSS para Windows. Análisis Estadístico*. México. Soborne-MC Gras Hill.
- GARNICA, Eloy (1995). **Análisis de Componentes Principales en los Presupuestos Familiares**, en *Economía*, 11: 55-91.
- _____ (1996). **Análisis Multivariante sobre la vivienda**, en *Economía*, 10: 21-55.
- KENDALL, Kenneth (1991). *Análisis y Diseño de Sistemas*. México: Prentice Hall Hispanoamericana.
- KORTH, Henry; A. Silberthatz (1993). *Fundamentos de Base de Datos*. España: McGraw-Hill.
- LEBART; MORINEAU; WARWICK. (1977). **Multivariate Descriptive Statistical Analysis**. John Wiley&Sons, Inc.
- MÁRQUEZ, Victor (2004). *Análisis Estadístico de la Encuesta de Presupuesto Familiares año 1997*. Mérida-Venezuela: Universidad de los Andes. Instituto de Estadística Aplicada y Computación. Tesis de maestría.

MORRISON, Donald (1976) *Multivariate Statistical Methods*. Tokyo: Mc Graw Hill.

NEIBAUER, Alan. *Acces 2000 para gente ocupada*. México: Mc Graw Hill.

PONSOT, Ernesto A. (1998). *Apuntes de Bases de Datos*. Mérida, Venezuela: Universidad de los Andes.

JOHNSON, Richard; Dear Wichern. (1992). *Applied Multivariate Statistical Analysis*. Englewood Cliffs, N. J. Prentice Hall, Inc.