

## **ANÁLISIS DE COMPONENTES PRINCIPALES EN LOS PRESUPUESTOS FAMILIARES**

**Elsy Garnica Olmos**  
Instituto de Investigaciones  
Económicas y Sociales  
Universidad de Los Andes

**RESUMEN.-** En este trabajo se describen algunas características de los presupuestos familiares, en la ciudad de Mérida, Venezuela. Con la técnica del Análisis de los Componentes Principales se destacan algunas rasgos del gasto familiar que incluyen perfiles del hogar, aspecto personal, equipamiento de la vivienda, patrón de consumo, rubros alimenticios más importantes y la escala de necesidades.

### **0 INTRODUCCIÓN**

Una Encuesta de Presupuestos Familiares (EPF) es un instrumento para recabar información básica sobre condiciones de vida, ingreso y consumo de las familias. Existen organismos internacionales, tales como la CEPAL, Comisión Económica para América Latina y el Caribe (1986) y el Banco Mundial (Altimir, 1974), que elaboran métodos racionales de la encuesta para lograr clasificaciones uniformes que favorezcan la comparación internacional (Naciones Unidas, 1964).

La EPF arroja un gran volumen de información demográfica, sobre la situación y características de la vivienda, educación, empleo, fuentes de ingreso y modalidades de consumo. Es complicado el proceso de organizar el gran volumen de información que aporta una EPF, pero aún más difícil es tener una visión global acerca de las características de las familias observadas. Los análisis multivariantes se convierten en una alternativa de investigación porque ellos permiten manejar la información, en forma global, sin perder detalles importantes que escaparían de un simple estudio (parcial) univariante. Innumerables análisis

cuantitativos multivariantes pueden efectuarse a partir de datos provenientes de este tipo de encuestas (Garnica, 1993), aunque la EPF se utiliza, básicamente, con el objeto de determinar, o actualizar, las ponderaciones de los índices de precios al consumidor (Garnica, 1992).

El interés de esta investigación se centra en los resultados de la EPF realizada en la ciudad de Mérida en 1986 (Vielma, 1992), acerca de los gastos familiares en los subgrupos de la canasta familiar y el ingreso.

Se emplea la técnica multivariante del Análisis de Componentes Principales, para resaltar la información más importante y para estudiar las interrelaciones entre todas las variables continuas observadas. Este artículo está basado en los resultados parciales de la mimeografía de Garnica (1993).

## **1 OBJETIVOS**

Aunque la estructura de presupuestos familiares de Mérida, en el año 1986, no es inmutable en el tiempo y, además, ésta puede ser distinta a otras regiones del país, el Análisis de Componentes Principales se utiliza en esta investigación, básicamente, con fines exploratorios. Por ello, no se establecen hipótesis cuantitativas ni modelos preconcebidos acerca de la estructura de los gastos familiares, ni siquiera se realizan inferencias estadísticas, proyectadas en el tiempo o en otro espacio geográfico.

En esta investigación sólo se intenta reflejar la realidad que plasman, para 1986, las 487 familias observadas en la ciudad de Mérida, tomando el ingreso, como variable ilustrativa del análisis. Por lo antes descrito, no se pretende ser riguroso en cuanto a la generalización de los resultados.

Los **objetivos** de la investigación, son:

1. Presentar, en forma gráfica, la información más importante de los gastos efectuados por las familias merideñas, en 1986.
2. Describir el patrón de consumo de las familias observadas.

3. Describir el papel que juega el ingreso, en el presupuesto familiar.

Estos objetivos se esperan alcanzar utilizando la técnica multivariante de Análisis de Componentes Principales. Este método se describe en la Sección 4.

## 2 CARACTERÍSTICAS DE LA ENCUESTA

En 1986, se realiza una Encuesta de Presupuestos Familiares (EPF) en Mérida, utilizando las técnicas y parámetros fijados por organizaciones nacionales e internacionales, al respecto. La EPF se efectúa en dos periodos, con seis meses de intervalo. Cada uno de estos periodos comprende dos semanas continuas, en la recolección de información (Vielma, 1992 y Tinto A., 1988).

El tamaño de la muestra, con un diseño de muestreo estratificado, con afijación proporcional y polietápico, abarca un número de 487 familias. Los estratos se definen como: (1) Zona Residencial, (2) Zona Comercial, (3) Zona de Mejoramiento Urbano, con casas del Instituto Nacional de la Vivienda, predominantemente, y (4) Zona de Ranchos. Las unidades estadísticas, en las dos etapas del muestreo, son las cuadras y las familias.

En la encuesta se recaban datos demográficos, socioeconómicos y acerca de la vivienda. La EPF aporta una gran masa de datos que, analizados en su justa dimensión, producen información importante para diversas áreas del conocimiento.

Con respecto a los **gastos** debe aclararse que los diferentes tipos de erogaciones que realiza la familia se dividen, formalmente, en cuatro grupos: gastos en alimentos, bebidas y tabaco (Grupo 1), gastos en vestido y calzado (Grupo 2), gastos del hogar (Grupo 3) y gastos diversos (Grupo 4). Las familias encuestadas deben anotar, con todo detalle (sitio de compra, marca, volumen y cualquier otra información valiosa), en unas planillas diseñadas especialmente para ello.

Por otra parte, el dato de **ingreso** se obtiene al precisar el monto percibido por la familia, el mes anterior. Otros ingresos, que deben ser anotados por la familia, se refieren a los aguinaldos, bonos, prestaciones, donaciones, etc.; las familias deben recordar el monto percibido, por estos conceptos, durante el año anterior.

### **3 LAS VARIABLES ESTUDIADAS**

Como se mencionó anteriormente, la EPF arroja datos de muchas clases, pero en este trabajo sólo se analizan las variables: ingreso y gasto familiar.

Los respectivos gastos en cada uno de los cuarenta subgrupos de productos o servicios de usual consumo familiar y el ingreso familiar, son las variables a analizar en esta investigación. A cada nombre de variable de gasto, se le antepone la letra correspondiente a su grupo, por ejemplo, el gasto en aceite se denota con **A**ceite, la primera **A** se refiere a grupo **Alimentos, Bebidas y Tabaco** (Grupo 1), otro ejemplo es el gasto en educación, **D**educación, donde la primera **D** se relaciona con el grupo de **Diversos** (Grupo 4). Los grupos de la canasta familiar están constituidos como sigue:

1. **Grupo Alimento, Bebidas y Tabaco (A)**, 19 variables: **A**cereales, **A**raíces, **A**leguminosas, **A**zúcar, **A**hortalizas, **A**frutas, **A**carne, **A**aves, **A**pescado, **A**preparados, **A**huevos, **A**lácteos, **A**ceites, **A**varios, **A**niño, **A**refresco, **A**bebidas, **A**tabaco y **A**fuerahogar.
2. **Grupo Vestido (V)**, 5 variables: **V**hombre, **V**dama, **V**niño, **V**niña y **V**bebe.
3. **Grupo Hogar (H)**, 7 variables: **H**vivienda, **H**construcción-Ferreter, **H**combustible, **H**enseres, **H**equipos, **H**vehículos y **H**diversos.
4. **Grupo Diversos (D)**, 9 variables: **D**médico, **D**cuidado, **D**educación, **D**diversión, **D**transporte, **D**servicio, **D**seguro, **D**impuestos y **D**otros.

Como puede observarse, cada grupo está formado por varios subgrupos, y éstos constituidos por numerosos rubros, productos

o servicios. Para mayor información acerca de la constitución de los subgrupos aquí nombrados el lector puede consultar el Anexo ofrecido al final de este artículo.

#### **4 METODOLOGÍA ESTADÍSTICA: ACP**

Esta Sección se orienta a los lectores que conociendo los principios de la técnica del Análisis de Componentes Principales (ACP), desean algunas indicaciones para comprender mejor la aplicación de la técnica e interpretar los gráficos resultantes. El aspecto teórico formal del ACP puede ser consultado en Morrison (1976), Mardia, Kent y Bibby (1979), Anderson (1984), Johnson y Wichern (1982), Pla (1986), Batista y Martínez (1989), entre otros.

Planteamientos didácticos sencillos de esta técnica pueden conseguirse en el trabajo de Márquez (1989). Aplicaciones interesante del ACP en casos de estudio, el lector puede encontrar en la revista *Applied Statistics*<sup>1</sup>, de articulistas como Jeffers (1967), Ahamad (1967) y Walker (1967). Estudios recientes de casos, en los cuales se aplica e interpretan los resultados del ACP, pueden encontrarse en González y otros (1994), Pachano A. (1995) y Díaz (1995).

El ACP tuvo su origen cuando Karl Pearson ideó sus líneas generales en una publicación (Pearson, 1901), sobre el ajuste de un sistema de puntos en un multiespacio, a una línea o a un plano. Pearson trabajó sobre las combinaciones lineales (componentes) de variables originales, para las cuales la varianza no explicada, fuera mínima. Años más tarde, H. Hotelling (Hotelling, 1933) retomó la idea de Pearson, pero formuló el ACP como sintetizador de la mayor variabilidad del sistema de puntos; éste es el enfoque utilizado actualmente, donde los **componentes principales** resumen la mayor proporción posible de la variabilidad total de la matriz de datos. Según este criterio, el ACP puede servir de instrumento para clasificar u observar relaciones entre los puntos proyectados en un plano (Pla, 1986). Desde la aparición de las microcomputadoras, en la década de los '80, que facilita el manejo

<sup>1</sup> Applied Statistics, Journal of the Royal Statistical Society (Series C), 21 Bentinck St., London W.1.

personal de una gran masa de información, el ACP se hace, cada día, más popular.

#### **4.1 OBJETIVOS DEL ACP**

El ACP consiste en la obtención de los componentes o ejes principales, no correlacionados entre sí. El cálculo del primer componente, como una combinación lineal de las  $p$  variables iniciales, se hace en base a la consecución de la máxima varianza de los datos; así, el primer eje representa una proporción de la varianza explicada de los individuos. Luego, el cálculo del segundo eje, no correlacionado con el primero, es también una combinación lineal de las variables y representa otra parte de la variación explicada de los individuos. El proceso se desarrolla hasta conseguir el  $p$ -ésimo componente o eje. Cada uno de los componentes recoge una parte de la variación de la nube original de los individuos.

La máxima información posible contenida en una tabla de datos (variables medidas en escala de intervalo o de razón), usualmente, se presentan en forma gráfica, en planos, formados con la combinación de los primeros componentes principales. En estos planos se proyectan las nubes de puntos-variables o nubes de puntos-individuos con el fin de describir las interrelaciones existentes entre esas variables o entre los individuos y las variables.

El ACP es esencialmente descriptivo. La interpretación de un plano principal del ACP origina una situación similar a la del fotógrafo que observa un paisaje y trata de sacarle el mejor provecho con su cámara, tomando una foto desde el “mejor” ángulo.

El primer plano del ACP está formado por los componentes, o ejes, 1 y 2. Cada componente debe ser interpretado de acuerdo a la importancia de las variables en cada uno de ellos. Se debe tener cuidado en la explicación de los ejes 3, 4, etc., cuando éstos resulten importantes, porque demasiada información entorpece el objetivo primordial: un resumen claro del fenómeno y de las interrelaciones entre las variables que lo conforman. En la

investigación que se presenta, se interpreta la nube de puntos variable e individuos, en el primer plano, formado por los ejes 1 y 2, y se realiza una breve descripción del tercer eje.

El ACP puede ser utilizado en varias situaciones como por ejemplo, como primera etapa de la investigación, en estudios exploratorios; o bien, cuando se quiera seleccionar las variables adecuadas con el fin de limitar el número de medidas en estudios similares posteriores; o cuando se desea observar, en gráficos (por lo general, en uno o dos planos principales), las semejanzas o diferencias de grupos de variables o grupos de individuos.

El ACP también puede utilizarse cuando la base teórica, del fenómeno bajo estudio, sustenta la creencia o sospecha de que una de las variables es dependiente, o es "explicada", por las otras - como en el caso del Análisis de Regresión-. En este contexto, en un análisis exploratorio con el ACP, esa variable en particular o "explicada", puede agregarse al análisis como variable ilustrativa, que no aportará nada a la formación de los componentes principales, pero su proyección en el plano principal puede ser de gran utilidad en el análisis.

#### **4.2 TABLA DE DATOS**

El método procesa, inicialmente, una matriz de datos con la finalidad de formar la estructura e interrelaciones entre las características bajo consideración. Para el proceso se requiere de una tabla  $n \times p$ , donde  $n$  es el número de individuos (o familias, en el caso de esta investigación) y  $p$ , el número de variables cuantitativas, medidas en escala de intervalo o razón. El ACP no requiere de supuestos tales como la normalidad multivariante y la igualdad de matrices de varianzas-covarianzas, sólo tres requisitos son necesarios para un adecuado ACP:

- 1) las variables deben ser continuas,
- 2) el número de individuos  $n$ , debe ser mayor que  $p$ , el número de variables,
- 3) las variables deben presentar altas correlaciones de manera que se puedan "agrupar" en componentes.

La información contenida en esa tabla de datos, se presenta, después de efectuado el proceso del ACP, en forma resumida y gráfica, lo menos deformada posible.

Cuando las variables tienen diferentes unidades de medida se trabaja con los datos centrados y reducidos. Centrados, con relación a la media de cada variable y reducidos, con relación a la desviación típica. El ACP, entonces, se realiza a partir de la matriz de correlaciones.

Cuando las variables tienen igual unidad de medida, se recomienda sólo centrar, haciendo el ACP a partir de la matriz de varianzas-covarianzas; en este tipo de análisis, las variables de mayores variaciones quedan proyectadas a los extremos del plano (Márquez, 1989).

#### **4.3 DISTANCIA EUCLIDIANA**

Por lo general, los puntos proyectados en un plano principal están separados entre sí, por medio de "distancias". Este trecho, entre un punto y otro, está medido por la distancia euclidiana:  $d^2(i,i') = \sum (x_{ij} - x_{i'j})^2$ , donde,  $i$  e  $i'$ , son dos filas de la tabla de datos (para  $i$  e  $i' = 1, 2, \dots, n$ ) y  $j$  es la columna correspondiente a la variable  $j$ -ésima (para  $j = 1, 2, \dots, p$ ).

Cuando dos puntos-variables, proyectados en un plano, están cercanos, puede decirse que esas dos variables revelan un comportamiento semejante, en el grupo de individuos. Lo mismo es válido para dos puntos-individuos cercanos: esos dos individuos tienen un comportamiento semejante para el grupo de variables importantes en ese plano. El plano de la nube de puntos-variables no se superpone al plano de la nube de puntos-individuos; pero lo que sí es factible de interpretación es la similar ubicación de esos puntos (variables o individuos) en sus respectivos planos.

#### **4.4 GRÁFICOS**

Los puntos-variables se proyectan en el plano principal formado por los dos primeros componentes (si se desea retener



sólo dos componentes). En este gráfico se analizan las cercanías entre puntos-variables y su interpretación debe estar en un todo de acuerdo a los indicadores descritos más adelante.

El nombramiento de los ejes es un paso relativamente subjetivo; se hace de acuerdo a la experiencia y conocimientos del investigador. Interpretaciones diferentes del mismo fenómeno (utilizando la misma matriz de datos) pueden surgir de investigadores diferentes. Pero se trata, en la medida de lo posible, de ajustar esta denominación a la realidad. El estadístico ofrece los resultados, pero el especialista en el tópico o materia interpreta esos resultados, según sus conocimientos o experiencia en el campo.

El gráfico de la nube de individuos, proyectada en el primer plano, es útil en la interpretación de los resultados del ACP. Es posible ver cuáles son los individuos que contribuyen más a la formación de los ejes: por lo general, corresponden a los puntos que se encuentran más alejados del centro.

El hecho de que en el gráfico de la nube de individuos se observen algunos puntos raros, o situados en un extremo del plano principal, debe llamar a la reflexión. En este caso esos individuos pueden ser sospechosos y debe investigarse el comportamiento de sus medidas o datos; puede tratarse de individuos con una conducta algo extraña, pero consistente, en todas sus variables, o bien, ser errores en la información. Cuando existen datos considerados raros, es preciso sacar estos casos del ACP y dejarlos como ilustrativos.

También puede presentarse el caso de varias sub-nubes de individuos, separadas perfectamente entre sí; en esta situación, el ACP ha logrado destacar la formación de grupos naturales, de acuerdo a las variables principales del análisis; este puede ser un resultado digno de estudios posteriores (de ACP en cada sub-grupo, de Análisis Factorial Discriminante, de un *Cluster Analysis*, de MANOVA o de cualquiera otra técnica estadística de análisis de grupos).

#### **4.5 INERCIA Y VALORES PROPIOS**

El ACP transforma las  $p$  variables cuantitativas iniciales en  $p$  nuevas variables cuantitativas no correlacionadas, denominadas **componentes principales**. El paso matemático para obtener los  $p$  componentes principales se llama **diagonalización** y proporciona los vectores y valores propios  $V_j$  (para  $j = 1, 2, \dots, p$ ). Las demostraciones matemáticas pueden estudiarse en Morrison (1976). Mardia y otros (1979) y Anderson (1984).

Los valores propios representan las varianzas de los datos sobre los ejes principales correspondientes. Cada valor propio puede expresarse en valor relativo, es decir, en porcentaje de varianza explicada sobre la varianza total.

La inercia total de un estudio de ACP, basado en la matriz de correlaciones ( $\mathbf{R}$ ), es la suma de todos los valores propios de esta matriz ( $\sum V_j = p$ ). En un ACP, basado en la matriz de varianzas-covarianzas ( $\mathbf{S}$ ), la inercia total es la suma de los valores propios de esta matriz o bien, la traza de  $\mathbf{S}$ ; en notación matricial:  $\sum V_j = \text{tr}(\mathbf{S})$ , es decir, la suma de las varianzas de las  $p$  variables originales.

En los estudios de ACP, bien sean estos basados en la matriz de correlaciones o en la matriz de varianzas-covarianzas, no existen normas estrictas sobre cuántos componentes principales deben retenerse. Existen investigaciones orientadas hacia métodos de descarte de variables en el ACP, como por ejemplo el de Jolliffe (1972).

El criterio del investigador se impone. Existen algunas indicaciones y cualquiera de ellas es válida. Así, el investigador puede escoger entre:

- a) conservar aquellos componentes con valores propios superiores a 0,8, ó bien,
- b) retener los componentes cuyos valores propios representen un buen porcentaje de varianza explicada, o tal vez,
- c) una vez ordenados los valores propios en forma descendente, escoger los primeros, antes de cambiar bruscamente hacia abajo, el porcentaje de variación explicada; otro criterio válido es,

- d) escoger aquellos valores propios que resulten superiores al promedio general de ellos. En este caso, si se utiliza la matriz de correlaciones, entonces se escogen los valores propios superiores a uno (1).

Por lo general, en un estudio de ACP cuyo objetivo principal sea el de explicar o describir la estructura fundamental de la masa de datos, se escogen los dos o tres primeros valores propios. Si el porcentaje de variación es muy pequeño, el ACP no ha logrado su cometido cual es el de resumir la masa de datos en uno o dos planos principales. Si el porcentaje de variación es mínimo, es decir,  $V_j \approx (1/p)*100$ , el ACP no resume información porque todas las variables están incorrelacionadas y, por consiguiente, todas son importantes en la explicación del fenómeno; en este caso se recomienda no usar el ACP. Lo que interesa es simplificar la información en los planos principales, formados por los primeros ejes.

#### **4.6 ESTRUCTURA DE LAS VARIABLES**

La estructura de las variables está reflejada en la correlación entre las variables y sus componentes. El ACP del paquete estadístico STATITCF (ITCF, 1988) de la Escuela Francesa, forma, automáticamente, el círculo de correlaciones en los planos principales. Las direcciones de esas variables se construyen uniendo, el centro u origen del plano, con los diferentes puntos-variables.

La importancia de cada variable, en un componente, está reflejada en los indicadores que se enuncian en la siguiente Sección.

#### **4.7 INDICADORES DEL ACP**

Antes de realizar un ACP, es conveniente observar las estadísticas elementales e histogramas de cada variable, con la finalidad de detectar errores de medida o fenómenos raros, asimetría, curtosis, bimodalidad o múltiples modas. Luego, debe organizarse una matriz de datos “limpios” de errores de asentamiento o de valores dudosos.

Existen ciertos indicadores que ayudan a evaluar la información contenida en los resultados de un ACP y realizar una adecuada interpretación de los mismos, ellos son:

- 1) **Los valores propios  $V_j$**  o porcentajes de variación explicados por cada componente principal. Algunas pautas acerca de los valores propios ya fueron señaladas en la Sección 4.5.
- 2) **El coeficiente de correlación  $r$** , entre la variable y el componente. El coeficiente de correlación al cuadrado ( $r^2$ ), es un indicador de la calidad de la representación de esa variable en el componente: es la parte de variación de la variable que es explicada por el componente. Para determinar cuáles son las variables mejor representadas en el plano principal, se pueden sumar los  $r^2$  de esas variables, en los dos primeros ejes, y determinar cuáles son las que están bien, mediana o mediocrementemente representadas en ese plano principal.

- 3) La estructura de las variables en los planos retenidos. **Las variables más importantes se encuentran proyectadas cerca del círculo de correlaciones.** Las menos importantes, lo están cerca del origen del plano; estas últimas variables no aportan nada especial en la interpretación de los ejes en consideración.
- 4) **La repartición de los individuos en los planos principales.** Con respecto a la nube de individuos, algunas consideraciones fueron descritas en el último párrafo de la Sección 4.4.

#### **4.8 VARIABLES E INDIVIDUOS ILUSTRATIVOS**

Este tipo de estudio permite la introducción de variables o individuos ilustrativos, que no aportan nada a la formación de los componentes del análisis, pero su representación gráfica puede ser de gran utilidad en la denominación de los ejes.

El ACP permite la utilización de elementos ilustrativos que pueden ayudar a la interpretación de los resultados. Por ejemplo, una variable ilustrativa de un estudio de ACP puede ser una característica que esté muy relacionada con una definición o denominación tentativa de uno de los primeros componentes principales. Por otra parte, un individuo ilustrativo, puede ser un individuo con valores extremos en algunas variables importantes en el plano principal. Cualquiera de estos puntos ilustrativos, variables o individuos, ayudaría al investigador a definir el eje en consideración.

### **5 RESULTADOS**

La información cualitativa de la Encuesta de Presupuestos Familiares (características de la vivienda), fue procesada mediante el Análisis de Correspondencias Múltiples y los resultados se publicaron en la edición anterior de **Economía** (Garnica, 1995). Ahora, se describe la información cuantitativa (gastos en subgrupos de la canasta familiar) de la EPF, mediante el Análisis de Componentes Principales. En entregas posteriores se publicarán:

(a) los resultados de los análisis de correlación y regresión múltiple, utilizados con la finalidad de estimar el ingreso familiar y (b) un análisis multivariante sobre el gasto en el grupo de alimentos. Todos estos análisis tienen como fuente de datos la mencionada EPF.

En la actual investigación se escogió el ACP para analizar los cuarenta subgrupos del gasto familiar y el **ingreso** (véase Sección 3, Variables Estudiadas), de la Encuesta de Presupuestos Familiares realizada en la ciudad de Mérida, en 1986. Este análisis permite detectar las interrelaciones entre esas cuarenta variables, y además, la relación de todas ellas con el **ingreso** (como dato aportado por los encuestados), característica importante en el estudio, puesto que el conocimiento del papel que juega éste, dentro de la estructura del gasto, puede ser útil en futuras investigaciones. La información, que originalmente viene dada en forma familiar y anual, se transformó en personal y mensual, para facilitar el manejo de las cifras, con valores pequeños y para describir, al mismo tiempo, la información por cabeza.

Las 40 variables de gasto se distribuyen en los cuatro grupos de la canasta familiar: **Grupo 1**: alimentos, bebidas y tabaco, **Grupo 2**: vestido y calzado, **Grupo 3**: hogar y **Grupo 4**: diversos. La primera letra de la variable gasto, identifica al grupo al cual ella pertenece (Véase la Sección 3, Variables Estudiadas)

La tabla de datos está formada por 41 columnas o variables (40 activas referentes al gasto y una ilustrativa, el **ingreso**) y 476 filas o encuestas. El ACP transforma las  $p = 40$  variables activas iniciales, en 40 variables nuevas no correlacionadas denominadas "componentes principales" y, precisamente, son los primeros componentes los que aportan la máxima información sobre la composición del gasto.

Como todas las variables tienen las mismas unidades de medida (bolívaes mensuales, por persona), se efectúa el primer ACP con los datos centrados, empleando la matriz de varianzas-covarianzas, ya que interesa observar cuáles son las variables con mayor dispersión y su relación con las demás. Luego, se realiza un segundo ACP con los datos centrados y reducidos, utilizando la

matriz de correlaciones, por la sospecha de que pueda existir otro tipo de información que produzcan los datos estandarizados.

### **5.1 PRIMER ACP: CON LA MATRIZ DE VARIANZAS-COVARIANZAS**

En este primer ACP, con una inercia total de  $\text{Tr}(S) = 403.113,29$ , los tres primeros componentes recogen un 70% de la variación, pero para la descripción de los resultados generales, se considera suficiente el 60% de la información que aportan los dos primeros ejes o componentes (40% en el primer componente y 20%, en el segundo). La descripción del tercer componente (con un 10% de la variación total) se ofrece al final de la sección, como complemento informativo.

El gráfico del plano principal (ejes 1 y 2) permite describir, en forma resumida, la masa de datos. En éste es importante observar la posición que toman los puntos variables y su interrelación. Los puntos-variables más alejados del origen (es decir, más cercanos al círculo) se refieren a las variables con mayores contribuciones a la formación de los ejes. El coeficiente de determinación,  $r^2$ , es un indicador útil que mide la importancia de cada variable en el ACP y es la correlación, al cuadrado, entre la variable y el eje en discusión.

#### **5.1.1 INTERPRETACIÓN DEL GRÁFICO PRINCIPAL (EJES 1 Y 2)**

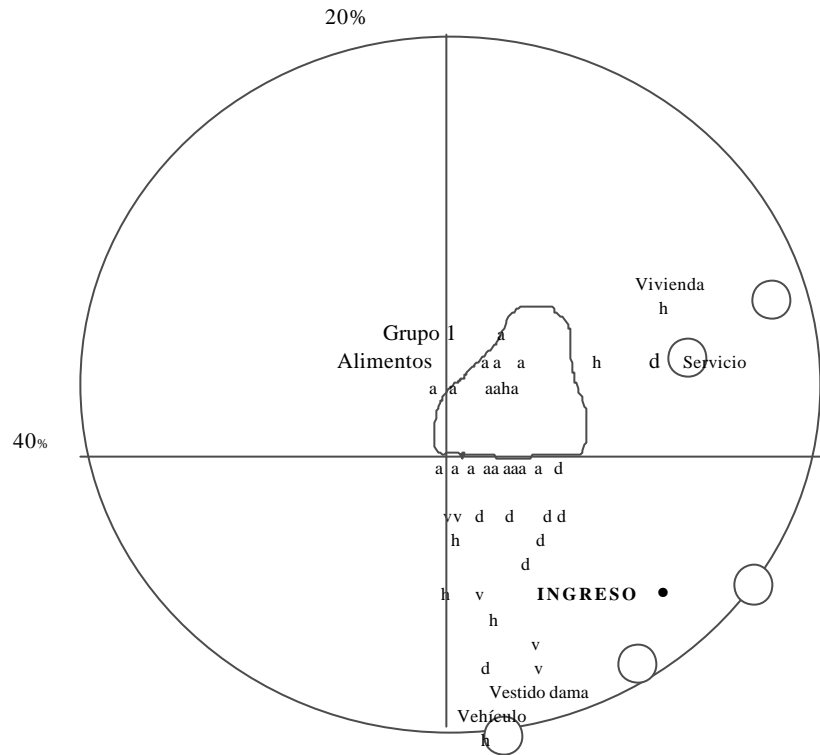
En el plano principal, Gráfico 1, se proyectan las 40 variables activas que se refieren al gasto y el **ingreso**, como variable ilustrativa, dentro de un círculo de correlaciones. La nube marcada en el Gráfico 1 se refiere a los gastos en los rubros **alimenticios**, señalados con la letras “**a**”. Un poco más alejado del origen, se encuentra la nube de puntos de gastos **diversos**, denotados por letras “**d**”. Además, de la proyección de la variable **ingreso**, el punto oscuro hacia la derecha del plano, se destacan cuatro puntos cerrados dentro de círculos para destacarlos, dos de ellos referidos a los gastos del **hogar**, identificados con “**h**”, uno del

Grupo Diversos, denotado con la letra “**d**”, y otro, de gastos en **vestido** y calzado, señalado con la letra “**v**”.

¿Por qué la nube de puntos perteneciente a los rubros alimenticios, está ubicada cerca del origen? Porque son gastos básicos de la canasta familiar. Una vez cubierta la necesidad alimenticia, la familia dirige el resto de su presupuesto hacia otros aspectos. En otras palabras, sea alto o bajo el ingreso, las personas cubren primero el aspecto primario: alimentos. Este resultado concuerda con la ley empírica del alemán Christiam Lorents Ernst Engel, propuesta en 1857, cuyo enunciado parcial dice así: “...el porcentaje del gasto destinado a alimentos, en relación al gasto total, tiende a bajar a medida que el ingreso aumenta” (Ferguson, 1976). El gasto en rubros



*Análisis de componentes principales en los presupuestos familiares*



**LEYENDA:**  
 a Grupo 1: Alimentos  
 v Grupo 2: Vestido y Calzado  
 h Grupo 3: Hogar  
 d Grupo 4: Diversos  
 0 INGRESO, Variable ilustrativa

**GRÁFICO 1**  
**COMPONENTES PRINCIPALES CON VARIANZAS-COVARIANZAS,**  
**PRIMER PLANO PRINCIPAL**

del Grupo 1: **Alimentos**, tiene la variación más pequeña de la canasta familiar.

Otra observación importante: dentro de esta nube de puntos de gasto en rubros alimenticios se halla un gasto de hogar (el punto “**h**”) que se refiere a **H**diversos, el cual se compone de gasto del hogar dirigido a útiles de limpieza y de cocina, tales como detergentes, limpiadores, ceras, servilletas, bolsas, vasos, esponjas, bombillos y otros útiles de limpieza; su proyección está dentro de la nube de puntos de alimentos, significando con esto que esta variable está asociada, de alguna forma, a los rubros alimenticios; el gasto efectuado en el subgrupo **H**diversos pareciera ser prioritario también.

¿Por qué la nube de puntos (“**d**”) de los rubros del Grupo 4, **D**iversos, se sitúa más alejado del origen? Porque el gasto efectuado en rubros del Grupo 4, presenta gran dispersión. Quiere decir que no son gastos básicos, algunas familias, supuestamente las que perciben mayores ingresos, gastan más, otras familias destinan montos moderados, y otras, simplemente no distraen absolutamente nada de su presupuesto en los rubros de este grupo. En general, los gastos en este grupo son los que originan ciertas diferencias entre los presupuestos familiares.

En este Gráfico 1, además de la proyección del **ingreso**, se observa, como se mencionó anteriormente, cuatro puntos encerrados en círculos, hacia el extremo derecho del plano. El primero (“**h**”), situado en el primer cuadrante se refiere a **H**vivienda (pago de alquiler, agua, teléfono y aseo urbano), el gasto más importante del eje 1 del ACP. El segundo (“**d**”), ubicado también en el primer cuadrante, es **D**servicio (pago a servicio doméstico y hoteles). El tercer punto-variable (“**v**”), situado hacia el centro del cuarto cuadrante, es **V**dama (gasto en abrigos, chaquetas, suéteres, pantalones, faldas, conjuntos, blusas, fondos y medio fondos, brassiers, bloomers, dormilonas, medias, vestidos confeccionados, trajes de baño, carteras, zapatos, sandalias, pantuflas, telas sin confeccionar y reparación de calzado), importante, tanto en el primero, como en el segundo eje del plano principal del ACP. El cuarto (“**h**”), situado en el extremo inferior del plano, **H**vehículos (gasto en compra de automóvil o bicicleta).

En la nominación del primer eje (con un 40% de la variación total) es importante el gasto en vivienda, pago a servicio doméstico (y gasto en hoteles) y la variable ilustrativa **Ingreso**. En

la nominación del segundo eje (con un 20% de la variación total) es preponderante la participación del gasto en vehículo. El gasto en vestido dama es importante en la formación de ambos ejes. Las variable **ingreso** está más asociada con el primer eje que con el segundo.

La evaluación de la posición de esos puntos y su coeficiente de determinación ( $r^2$ ) es muy útil en la denominación de los ejes.

- **Primer componente: Hogar (40% de la variación total)**

El primer componente principal (Eje 1 del Gráfico 1) tiene una variable importante en su formación: **H**vivienda, punto situado hacia el extremo derecho; esta variable se refiere al gasto en alquiler (92% del gasto en vivienda), teléfono, agua y aseo urbano, con un coeficiente  $r^2 = 0,95$ . Hay otras variables que le siguen en orden de importancia: **D**servicio, que se refiere al gasto en servicio doméstico (90% del gasto en servicios) y hoteles (el 10% restante), con un coeficiente  $r^2 = 0,36$ ; **V**dama, es el gasto en ropa y calzado de damas, con un coeficiente  $r^2 = 0,23$ ; **H**combustible (punto “h” situado a la izquierda de servicio) que es el pago en electricidad (56% del gasto en combustible), gas (con un 37%), fósforos, kerosene y velas con un coeficiente  $r^2 = 0,22$ .

El primer eje puede denominarse **Hogar** porque son aspectos relacionados directamente con éste, y el gasto en vestido de damas es otro aspecto resaltante, porque se percibe que es la mujer la que, por lo general, administra los gastos de su hogar y, por consiguiente, destina parte importante de ese presupuesto familiar, en su vestido y calzado; adicionalmente, las “cuentas secretas” de las mujeres casadas es un punto de discusión y de investigación al igual que las de los esposos (El Nacional, 14/9/95). Por otra parte, el ingreso, variable ilustrativa, está asociada a este primer eje.

- **Segundo componente: Aspecto Personal (20% de la variación total)**

El segundo componente principal (Eje 2 del Gráfico 1) también tiene una variable básica en su constitución: **H**vehículo, punto (“h”)situado en un extremo inferior del plano; esta variable se refiere al gasto en adquisición de automóviles (98% del gasto en

vehículo) y bicicletas para niños, con un coeficiente  $r^2 = 0,88$ . Le siguen en orden de importancia otras variables tales como: **D**diversiones (punto “**d**” anexo al punto marcado “**v**” de vestido para dama), con un coeficiente  $r^2 = 0,18$ , que agrupa el gasto en los viajes de recreo (54% del gasto en diversiones), cine y teatro (con un 14%), juguetes, juegos, diversiones, apuestas de caballos, loterías, películas y revelados; **V**dama, que se refiere al gasto en ropa y calzado de damas, con un coeficiente  $r^2 = 0,16$  y **V**bebé (punto “**v**”, cerca de la proyección de la “**d**” de diversiones), con el gasto en ropa y calzado de bebés, con  $r^2 = 0,14$ . Este segundo componente puede denominarse **Aspecto Personal**, debido a que está asociado al disfrute, comodidad o apariencia personal.

• **Tercer componente: Equipamiento (10% de la variación total)**

Aunque no se ofrecen los gráficos correspondientes (planos 1-3 y 2-3), se describe brevemente el análisis del tercer componente: una sola variable del **Grupo Hogar**, resalta en su formación, con un coeficiente  $r^2 = 0,76$  (las demás variables se proyectan cerca del origen). Esta variable es **Hequipo**, cuyos rubros constituyentes son los siguientes: gastos en compra de equipo de sonido, televisor, lavadora, secadora, nevera, juego de recibo, juegos de dormitorio, juego de comedor, ayudante de cocina, cocina, máquina de coser, aspiradora, pulidora, betamax, VHS, planchas, licuadoras, relojes de pared, utensilios de cocina, cunas, lámparas y accesorios, discos y cassettes, aparatos telefónicos, maletas, maletines, reparación de los equipos del hogar.

La variable ilustrativa **ingreso**, tiene muy poca importancia en este tercer componente (con una  $r^2 = 0,03$ ). El subgrupo **Hequipo**, preponderante en la formación del tercer eje, no se relaciona directamente con el **ingreso**. En el estudio de Garnica (1995) se analizan los datos cualitativos de la misma encuesta utilizando el Análisis Factorial de Correspondencias Múltiples (AFCM) y en una de las conclusiones de entonces se declaraba: “El ingreso más bajo está relacionado con la falta de bienes de equipamiento de la vivienda y el ingreso más alto, con la posesión de éstos” ; pareciese que ese hallazgo contradijera los resultados de ahora, pero no es así, en el AFCM se estudiaron los datos de **posesión** de equipos del hogar, mientras que en este estudio de ACP se

analizan los **gastos** efectuados en la compra de esos equipos. En la encuesta EPF, el periodo de referencia de los datos de gasto en la compra de equipos del hogar, es el año anterior, en cambio la pregunta acerca de la posesión de los mismos tienen un periodo de referencia de la misma fecha cuando se realizó la encuesta. Puede hacerse la suposición de que la **tenencia** de equipos, probablemente adquiridos años atrás, no está directamente relacionada con el **gasto** actual (para el momento de la encuesta), en equipos del hogar.

### 5.1.2 GRÁFICO DE PUNTOS INDIVIDUOS DEL ACP CON LA MATRIZ DE VARIANZAS-COVARIANZAS

La nube de puntos-individuos puede observarse en el Gráfico 2. Éste es un gráfico ilustrativo que ayuda a la interpretación de los ejes del ACP. Algunos de los individuos (cuyos números se encierran en un círculo) se listan en la Tabla 1, con los valores de las variables más importantes de este primer plano.

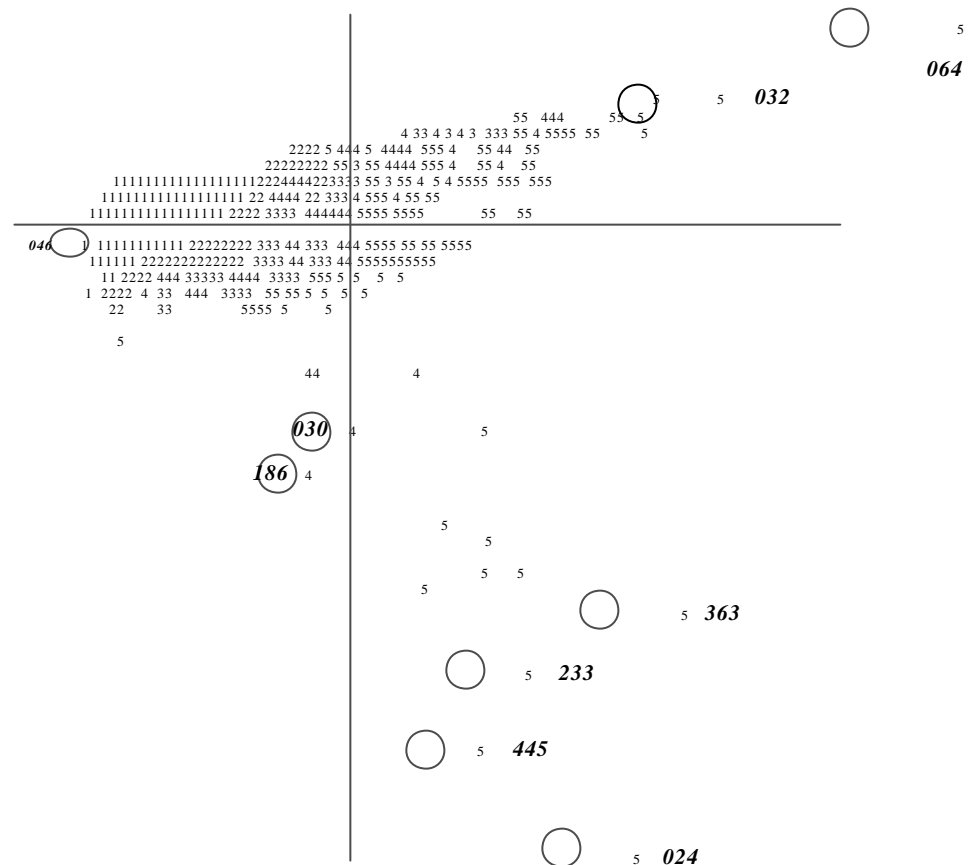
El individuo identificado con el número 064 es el que tiene el ingreso más alto y el gasto más elevado, en **Hvivienda**, de los 476 individuos. Por otra parte, el individuo identificado con el número 024 es el que tiene el más alto gasto en **Hvehículo** y en **Vdamas** de todo el grupo de individuos. Los individuos más dispersos dentro del plano principal (primer y cuarto cuadrante) tienen los niveles más altos de ingresos. Puede observarse, en este mismo gráfico, la alta concentración de puntos-individuos cerca del origen, (a la izquierda del primer eje); estos puntos se refieren a individuos con el nivel más bajo de ingresos.

**TABLA 1  
PUNTOS-INDIVIDUOS MAS RESALTANTES DEL ACP CON  
VARIANZAS-COVARIANZAS**

Individuo	Nivel Ingreso	GASTO PERSONAL MENSUAL (Bolívares)		
		Hvivienda	Hvehículo	Vdamas
024	5	792	3.500	1.160
030	4	269	833	78
032	5	2.030	0	162
046	1	63	0	31
064	5	4.467	0	0

186	4	134	833	61
233	5	691	1.833	732
363	5	1.819	1.938	535
445	5	493	2.500	107
<b>Media:</b>		<b>75</b>	<b>600</b>	<b>129</b>
<b>Mediana</b>		<b>426</b>	<b>0</b>	<b>57</b>
<b>Desv.Típ:</b>		<b>375</b>	<b>278</b>	<b>129</b>
<b>Máximo:</b>		<b>4.467</b>	<b>3.500</b>	<b>1.160</b>

Fuente: EPF, Mérida, 1986. Cálculos propios.



**LEYENDA** de escala de ingresos:

1: Bajo; 2: medio-bajo; 3: medio; 4: medio-alto y 5: alto

**GRÁFICO 2 NUBE DE PUNTOS-INDIVIDUOS  
ACP CON VARIANZAS-COVARIANZAS  
PRIMER PLANO PRINCIPAL**

## 5.2 SEGUNDO ACP: CON LA MATRIZ DE CORRELACIONES

La inercia total de este estudio de ACP es  $p = 40$ . Aunque el plano principal recoge sólo el 28% de la variación total (19% en el primer componente y 9%, en el segundo), al contrario del 60% en el análisis previo, es interesante observar, en el plano principal, el comportamiento de las 40 variables gasto y de la variable ilustrativa **ingreso**, cuando todas ellas se estandarizan.

En el Gráfico 3 se detectan claramente dos nubes de puntos variables: una pertenece al grupo **Alimentos** (cuyas proyecciones, en su mayoría, se ubican en la parte positiva del Eje 2) y a la otra al grupo **Diversos** (cuyas proyecciones se sitúan en la parte negativa del Eje 2). La proyección de la variable ilustrativa **ingreso** se sitúa a un extremo del gráfico. Otro punto variable se destaca: el subgrupo **Vdama**.

### 5.2.1 INTERPRETACIÓN DEL GRÁFICO PRINCIPAL (EJES 1 Y 2)

En el Gráfico 3 se percibe cómo el ACP con la matriz de correlaciones, ofrece otro tipo de información. En el primer cuadrante del plano principal del Gráfico 3, se observa la mayoría de productos alimenticios, esta nube se enmarca para resaltar los rubros "a", del Grupo 1. En el cuarto cuadrante se proyectan los renglones del Grupo 4 (Diversos), también la nube de estos elementos (señalados con "d"). se enmarcan para resaltarlos. La variable ilustrativa **ingreso** se proyecta cerca del círculo de correlaciones, en el cuarto cuadrante del plano.

El Gráfico 4 muestra, con más detalle, la nube de puntos variables de este segundo ACP. Se identifican todos los puntos variables a objeto de iniciar la interpretación de este plano.

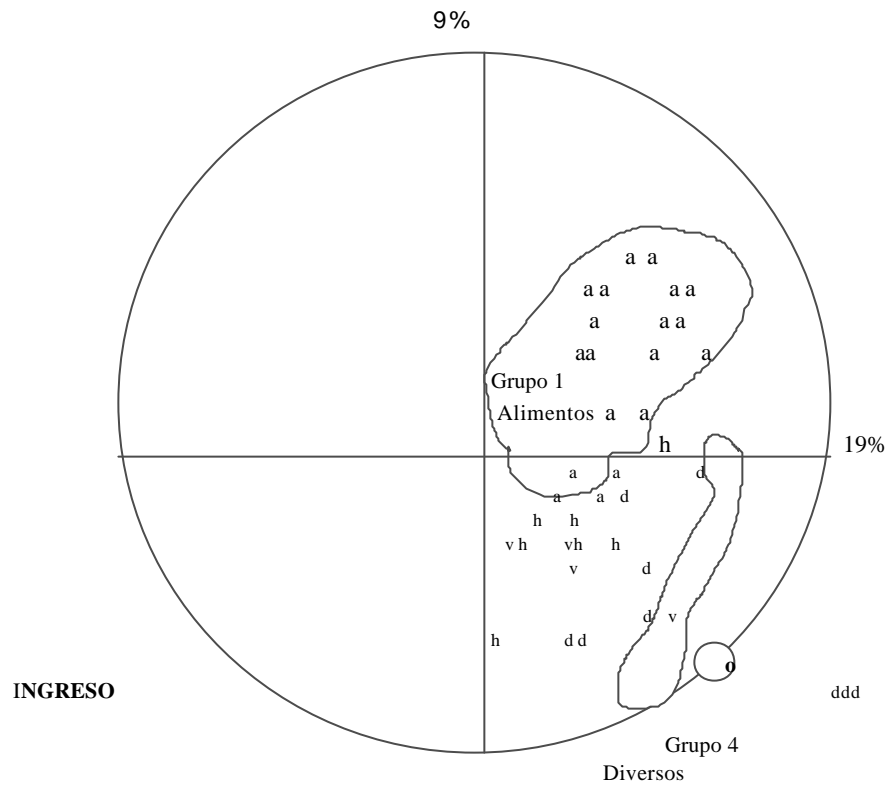
- **Primer eje: Hábitos de Consumo**

La proyección de los estos puntos en el primer eje, refleja, de alguna forma, los **hábitos de consumo** de los 476 individuos observados. En la parte izquierda del plano, cerca del origen



*Análisis de componentes principales en los presupuestos familiares*

(véase el Gráfico 4), se sitúan las variables con valor mediano igual a cero,



**LEYENDA:**  
 a Grupo 1: Alimentos  
 v Grupo 2: Vestido y Calzado  
 h Grupo 3: Hogar  
 d Grupo 4: Diversos  
 0 INGRESO, Variable ilustrativa

**GRÁFICO 3**  
**COMPONENTES PRINCIPALES CON CORRELACIONES**  
**PRIMER PLANO PRINCIPAL**

quiere decir que, por lo menos el 50% de los individuos no consume ese tipo de rubros de la canasta familiar. Las otras

variables proyectadas, más o menos, de la mitad de eje 1, hacia afuera, presentan algún valor mediano. Debe destacarse que la única variable que presenta un gasto en la totalidad de los individuos, es **H**vivienda (gasto en alquiler, agua, teléfono y aseo urbano). Este Eje 1 pudiera ser examinado por un analista de mercado porque revela, en cierto sentido, **los gustos o poder de compra** del grupo observado. La variable ilustrativa **ingreso** está relacionada a este Eje 1.

Los gastos que definen hábitos de consumo en la población merideña, definidos en el primer eje son, en orden de importancia: productos lácteos, cereales, cuidado personal, aceite, frutas, vestido para damas, alimentos preparados, carnes, educación, alimentos varios, enseres, servicios básicos (vivienda), diversos, azúcar, transporte, hortalizas, médico, diversión, servicio, vestido para hombres, huevos, impuestos, refrescos, aves, combustible, otros en diversos, seguros, vestido para bebé, raíces, legumbres, alimentos para niños, vestido para niñas, pescado, equipos, alimentos fuera del hogar, cigarros, vehículo materiales de construcción y ferretería, vestido para niño y bebidas.

- **Segundo eje: Escala de necesidades**

En el Gráfico 4 se observa, con todo detalle, la proyección de cada una de las 40 variables del gasto con respecto al segundo eje.

Este segundo eje del ACP, realizado con la matriz de correlaciones, podría definir una **escala de necesidades**. Se inicia la escala con los gastos básicos que cubren, en su mayoría, las necesidades alimenticias, luego, si la persona dispone de ingresos suficientes, el gasto se dirige hacia los demás rubros, en orden de prioridades, hasta llegar a las diversiones, impuestos (a mayores ingresos, mayores impuestos) y otros.

La escala de necesidades se enuncia en la Tabla 2. En esta tabla se observa claramente el rango de prioridades que la familia merideña otorga a cada uno de los rubros. Los quince (15) primeros elementos de esta lista, por ejemplo, se refieren a productos alimenticios, el elemento número 16 destaca la importancia, dentro de un presupuesto familiar, de **H**diversos, el cual engloba productos de limpieza del hogar, algunos accesorios

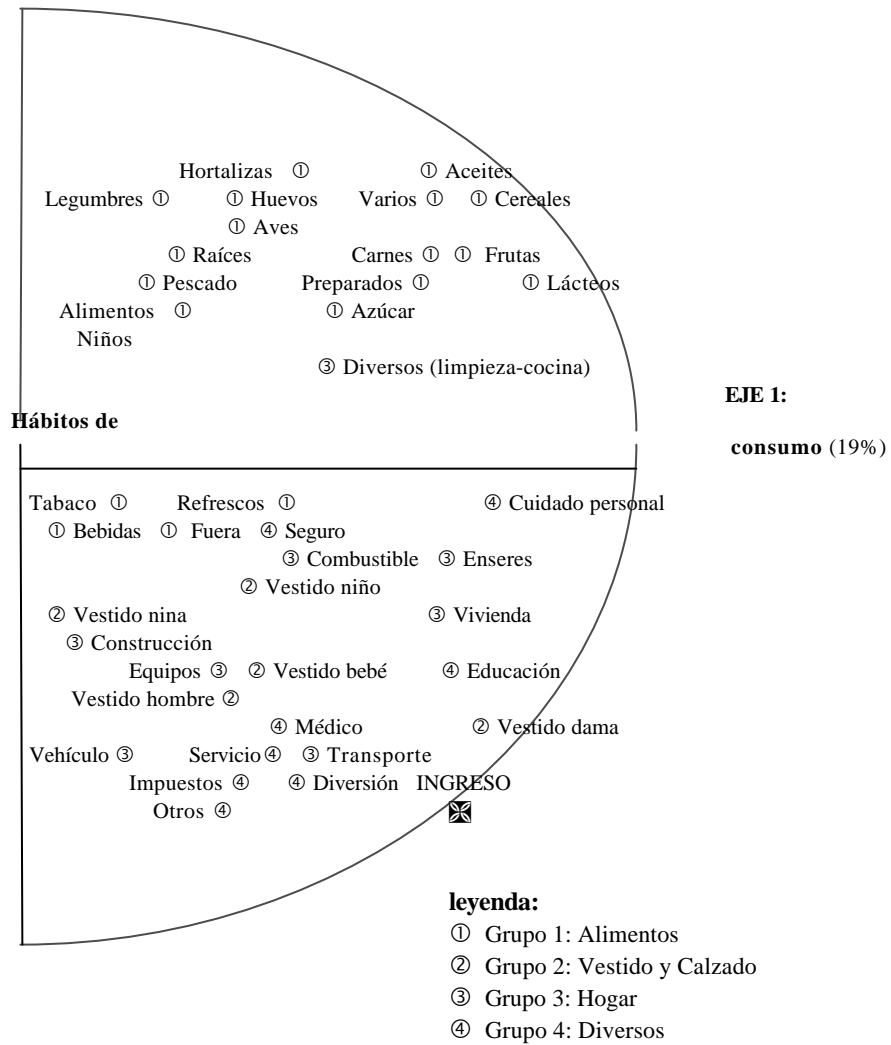
de cocina y mercerías (detergente, jabón para lavar, limpiadores, bombillos, cera, cloro, coletas, mopas, útiles de limpieza, insecticidas, papel parafinado, papel metálico, servilletas de papel, bolsas plásticas, bolsas de colar café, hilos, botones, agujas, vasos, esponjas, enjuagues, suavizantes y teteros). El rubro número

**TABLA 2**  
**ESCALA DE NECESIDADES PROYECTADAS EN**  
**EL EJE 2 DEL SEGUNDO ACP**  
**(CON CORRELACIONES)**

<b>Grado de Prioridad</b>	<b>Subgrupo</b>	<b>Importancia (Proyección)</b>
1	Ahortalizas	0,41
2	Aaceites	0,40
3	Acereales	0,40
4	Avarios	0,39
5	Alegumbres	0,38
6	Ahuevos	0,35
7	Aaves	0,32
8	Afrutas	0,29
9	Acarnes	0,29
10	Araíces	0,28
11	Apescado	0,24
12	Alácteos	0,24
13	Apreparados	0,24
14	Aazúcar	0,17
15	Aniños	0,15
16	Hdiversos	0,10
17	Dcuidperson	- 0,01
18	Atabaco	- 0,05
19	Arefrescos	- 0,05
20	Afuera	- 0,08
21	Abebidas	- 0,10
22	Dseguro	- 0,10
23	Hcombustible	- 0,14
24	Vniño	- 0,17
25	Henseres	- 0,17
26	Hvivienda	- 0,18
27	Vniña	- 0,20
28	Hconstrucc	- 0,21
29	Hequipos	- 0,22
30	Vbebé	- 0,24
31	Deducación	- 0,26
32	Vhombre	- 0,28
33	Dmédico	- 0,34
34	Vdama	- 0,37
35	Hvehículo	- 0,40
36	Dtransporte	- 0,40
37	Dservicio	- 0,42
38	Ddiversión	- 0,47
39	Dimpuesto	- 0,47
40	Dotros	- 0,50

Fuente: ACP, Matriz de correlaciones de la EPF, Mérida, 1986.

**EJE 2: Escala de necesidades (9%)**



**GRÁFICO 4**  
**COMPONENTES PRINCIPALES CON CORRELACIONES**  
**DETALLES DEL PRIMER PLANO PRINCIPAL**

17 en la lista es el gasto en cuidado personal (**D**cuidado), en este subgrupo están los cosméticos, champú, peluquería, barbería, cepillos dentales, dentífricos, hojillas de afeitarse, jabones de baño, papel higiénico, pasta y jabón de afeitarse, talco, loción para niños, toallas sanitarias, etc.

Después de ellos, continúan, con los rangos 18, 19, 20 y 21, cuatro renglones del Grupo Alimentos, ellos son: **A**tabaco, **A**refresco, **A**fuerahogar y **A**bebidas, que, al parecer, no son tan importantes como los 15 primeros renglones alimenticios. Luego, con menor prioridad están los gastos en seguros, combustible, vestido para niños, etc., hasta llegar a los cuatro últimos elementos referidos al Grupo 4, **D**iversos. Los gastos en transporte, servicio, diversión, impuesto y otros, son los menos prioritarios, para la familia merideña.

El último subgrupo de esta escala, **D**otros, se refiere a: gastos funerarios, regalos, intereses pagados, flores, matas, adornos, prendas, etc. Para más información sobre los rubros que conforman cada uno de los subgrupos de la canasta familiar, véase el Anexo.

### **5.2.2 GRÁFICO DE PUNTOS-INDIVIDUOS DEL ACP CON CORRELACIONES**

El Gráfico 5 refleja la proyección de la nube de individuos en el plano principal del ACP con correlaciones. Los puntos-individuos con el nivel más alto de ingresos se encuentran dispersos hacia el lado derecho del plano, entre el primer y cuarto cuadrante (los individuos proyectados en el cuarto cuadrante tienen las mayores cifras de ingreso); los del nivel más bajo de ingresos, se sitúan hacia el lado izquierdo del plano, cerca del punto de origen. En los demás niveles de ingresos (medio-bajo, medio y medio-alto) se observa que no existe una división clara entre los puntos porque se mezclan hacia el punto de origen.

Los gastos en los diferentes rubros de la canasta familiar determina la posición de los puntos-individuos. En la Tabla 3 se

describen los gastos de algunas variables de los individuos situados en un extremo del plano.



**GRÁFICO 5**  
**PROYECCIÓN DE PUNTOS-INDIVIDUOS DEL ACP**  
**CON CORRELACIONES, PLANO PRINCIPAL**

Puede agregarse que la variable **ingreso** está explicada, de cierta manera, por la proyección de los puntos-individuos dentro del plano, aunque sea ésta una variable ilustrativa, los individuos con los ingresos más altos son los que están situados cerca de la proyección del punto-**ingreso**, en el plano. Pero no es sólo el **ingreso**, también el gasto en las últimas variables de la escala de necesidades, descrita anteriormente, son esenciales en la proyección de los puntos-individuos en ese cuarto cuadrante.

En la Tabla 3, se listan algunos de los individuos cuyas proyecciones resultaron extremas. Por ejemplo, el individuo 064 tiene el máximo ingreso del estudio y el máximo gasto en **Dservicio**. El individuo 135 es el tercero de mayores ingresos y el que tiene el máximo gasto en **Dotros** (última variable en la escala de necesidades) y valores muy altos en gastos de **Deducación**, **Dtransporte** y **Dmédico**. El individuo 024 es el segundo de mayores ingresos y el que tiene el máximo gasto en **Ddiversión**.

**TABLA 3**  
**PUNTOS-INDIVIDUOS MÁS RESALTANTES DEL ACP**  
**CON CORRELACIONES**

<b>GASTO PERSONAL MENSUAL EN BOLÍVARES</b>													
Ind.	Nivel	Dotros	Dimpuesto	Ddiver-Siones	Dservicio	Deduca-Cion	Alac-Teos	Deuida-Do Per	Ace-Real	Aacei-Te	Avarios	Vdama	
012	5	13	0	23	0	156	256	494*	213	46	172	202	
014	5	1	0	100	0	17	132	199	212	108	146	0	
024	5	181	0	928*	0	209	65	204	36	0	4	1.160*	
063	5	40	92	683	600	22	148	24	58	17	0	594	
064	5*	154	406	0	1.500*	299	103	145	115	61	186	0	
090	5	415	61	163	200	143	0	159	0	0	0	61	
099	4	10	24	2	0	60	189	151	146	35	71	57	
107	1*	0	0	0	0	0	0	0	19	5	1	6	
135	5	1110*	491	303	1.117	305	93	179	83	36	61	74	
138	5	373	579*	96	0	157	209	77	10	65	14	37	
140	5	493	376	131	0	7	124	74	108	33	0	173	
189	3	23	60	23	3	106	329	38	132	74	142	132	
233	5	18	207	387	20	40	117	108	41	10	50	732	
284	5	54	129	217	231	87	414*	247	123	102	145	496	
314	5	0	0	71	0	489*	282	194	205	40	210*	194	
342	5	11	0	13	0	83	136	82	284*	151*	43	575	
357	5	470	190	181	260	116	49	21	16	0	0	111	
363	5	23	13	74	250	20	209	185	191	87	43	535	
368	5	178	20	247	267	90	71	135	186	0	0	424	

388	2	0	0	0	0	14	192	45	146	40	133	23
468	5	117	73	400	195	65	143	249	107	46	91	358
<b>Media:</b>	<b>25</b>	<b>18</b>	<b>44</b>	<b>34</b>	<b>37</b>	<b>66</b>	<b>55</b>	<b>51</b>	<b>16</b>	<b>32</b>	<b>100</b>	<b>100</b>
<b>Mediana:</b>	<b>2</b>	<b>1</b>	<b>11</b>	<b>0</b>	<b>22</b>	<b>89</b>	<b>104</b>	<b>77</b>	<b>114</b>	<b>110</b>	<b>129</b>	<b>129</b>
<b>Desv.Típ:</b>	<b>79</b>	<b>55</b>	<b>92</b>	<b>115</b>	<b>52</b>	<b>59</b>	<b>58</b>	<b>39</b>	<b>18</b>	<b>35</b>	<b>129</b>	<b>129</b>
<b>*Máximo:</b>	<b>1.110</b>	<b>579</b>	<b>928</b>	<b>1.500</b>	<b>489</b>	<b>414</b>	<b>494</b>	<b>284</b>	<b>151</b>	<b>210</b>	<b>1.160</b>	<b>1.160</b>

Fuente: EPF, Mérida, 1986.

En el primer cuadrante del plano principal (Gráfico 5) están ubicados puntos-individuos con valores máximos en las variables importantes de ese cuadrante. Por ejemplo, el individuo identificado con el 284 tiene el gasto más alto en Alácteos de toda la muestra. El individuo 314 tiene los valores máximos del gasto en Avarios y en Afrutas. El individuo 342 los tiene en Aceites y Acereal. El individuo 189, por ejemplo, tiene el segundo gasto más alto (después del 284) en Alácteos y, en diferencia con los demás descritos, tiene un nivel medio de ingreso. Por otra parte, también puede destacarse el individuo 107, con el nivel de ingreso más bajo y, por supuesto, con gastos mínimos o muy bajos en los subgrupos de la canasta familiar.

Todos estos individuos con gastos extremos en algún subgrupo importante del ACP, ilustran este análisis estadístico.

## 6 CONCLUSIONES

Ambos análisis, ACP con matriz de varianzas-covarianzas y ACP con matriz de correlaciones, resultaron ser complementarios en esta investigación. Sin embargo, el primero de ellos es el más importante puesto que recoge un 60% de la variación, en el primer plano, contrapuesto al segundo análisis, con sólo un 28% de la variación explicada en el primer plano.

Los aspectos más importantes en los presupuestos familiares, en orden de prioridades, son los siguientes:

1. Los gastos para mantenimiento del **hogar**, como son los desembolsos causados por alquiler, teléfono, agua, aseo urbano, servicio doméstico, ropa y calzado de damas, electricidad y gas. El gasto en ropa y calzado de dama se entiende, en el contexto de la investigación, como el gasto realizado por el ama de casa,

como administradora principal del presupuesto del hogar. El ingreso, como variable ilustrativa del ACP, está relacionado, directamente, con este primer componente (a mayores ingresos, mayores gastos en los renglones mencionados).

2. Los gastos en **aspecto personal**, en el sentido de que no van dirigidos a los servicios básicos de la vivienda, sino más bien a los miembros que conforman la familia. Estos son los desembolsos realizados para la adquisición de vehículo y para los viajes de recreo, principalmente. En este segundo componente, de nuevo, es importante el gasto en vestido y calzado para dama y, como aspecto adicional, el gasto en ropa para bebé. El ingreso se encuentra bastante asociado a los gastos en vestido y calzado para dama (sus proyecciones están muy cercanas), pero esta relación se observa débil respecto a los otros subgrupos de este segundo componente.
3. Los gastos en **equipamiento** de la vivienda. Se refieren, principalmente, a la compra de electro-domésticos que ayudan a obtener la comodidad del ama de casa, principalmente. Además de los electro-domésticos, en este tercer componente, se encuentran los gastos para adquisición de muebles y otros accesorios de la vivienda. El ingreso está débilmente asociado a este componente.
4. El gasto en el grupo de alimentos se mantiene más estable que el gasto en los demás grupos de la canasta familiar. Sea cual sea el ingreso familiar, se cubren primero, las necesidades alimenticias y el resto del ingreso se destina a otros grupos. Es decir, el incremento de los ingresos no parece causar un incremento, igual o proporcional, en los gastos en alimentos.
5. El Análisis de Componentes Principales realizado con la matriz de correlaciones muestra otros aspectos de la investigación: los **hábitos de consumo** (Sección 5.2.1, primer eje) y la **escala de necesidades** (Tabla 2) de la familia merideña.

## **BIBLIOGRAFÍA**

- Ahamad, B. (1967): "An Analysis of Crimes by the Method of Principal Components". *Appl. Statist.*, **16**, 17-35.
- Anderson, T.W. (1984): **An Introduction to Multivariate Statistical Analysis**. 2d. ed. New York. Jhon Wiley & Sons.
- Altimir, Oscar (1974): "Un archivo de datos sobre distribución del ingreso procedentes de encuestas de hogares en países latinoamericanos". mimeografía presentada en el Segundo Congreso Latinoamericano de la Asociación Internacional para la Investigación del Ingreso y la Riqueza, Río de Janeiro, 9 al 12 de enero de 1974. Chile. Comisión Económica para la América Latina.
- Batista Foguet, J.M. y María del Rosario Martínez Arias (1989): **Análisis multivariante, análisis en componentes principales**. Barcelona, España. Editorial Hispano Europea, S.A.
- Comisión Económica para América Latina y el Caribe, CEPAL, (1986): "Encuesta de ingresos y gastos, conceptos y métodos en la experiencia latinoamericana". Resultados del Seminario Regional sobre Métodos de Encuestas de Ingresos y Gastos del Hogar realizado en Bogotá del 5 al 8 de junio de 1984 y auspiciado conjuntamente por el Departamento Administrativo Nacional de Estadísticas de Colombia (DANE) y la CEPAL. **Cuadernos de la CEPAL**, LC/G. 1386, Nov. 1986.
- Díaz, Magali (1995) "Estudio comparativo del rendimiento estudiantil en FACES, usuarios de comedor y becarios". Mimeografía. Mérida, Venezuela. Instituto de Investigaciones Económicas y Sociales, FACES, ULA.
- El Nacional** (1995): "Las mujeres casadas mantienen la tradición: una cuenta propia". Jueves 14 de septiembre de 1995, p. Economía/7. Columna Tendencias.
- Ferguson, C.E. (1976): **Teoría Microeconómica**. Bogotá. Fondo de Cultura Económica.

- Garnica O., Elsy (1995): "Análisis multivariante sobre la vivienda". **Economía, 10**, 21-52.
- \_\_\_\_\_ (1993): "Análisis multivariante de los presupuestos familiares". Mimeografía. Mérida. Instituto de Investigaciones Económicas y Sociales, FACES, ULA.
- \_\_\_\_\_ (1992): Índice de precios al consumidor, Región de Los Andes, Informe 1992. Mérida: Instituto de Investigaciones Económicas y Sociales, FACES, ULA.
- González, Pilar, Amelia D. de Pascual, Enrique Torres y Elsy Garnica (1994): "Una aplicación del análisis de componentes principales en el área educativa". **Economía, 9**.
- Hotelling, H. (1933): "Analysis of a Complex of Statistical Variables into Principal Components". **J. Educ. Psychol., 24**:417:441 and 498:520.
- Jeffers, J.N.R. (1967): "Two Case Studies in the Application of Principal Component Analysis". **Appl. Statist., 16**, 225-236.
- Johnson, Richard and Dean W. Wichern (1982): **Applied Multivariate Statistical Analysis**. Englewood Cliffs, N.J. Prentice-Hall, Inc.
- Jolliffe, I.T. (1972): "Discarding Variables in a Principal Components Analysis. I: Artificial Data". **Appl. Statist., 21**, 160-173.
- Mardia, K.V., J.T. Kent and J.M. Bibby (1979): **Multivariate Analysis**. London. Academic Press, Inc.
- Márquez, Víctor (1989): "Apuntes sobre análisis multivariante, Volumen 1". Mimeografía. Mérida, Venezuela. Instituto de Estadística Aplicada, FACES, ULA (Publicación No. 89-01).
- Morrison F., Donald (1976): **Multivariate Statistical Methods**. 2d. ed. Tokyo. McGraw-Hill Kogakusha, LTD.
- Naciones Unidas (1964): "Manual de encuestas sobre hogares, guía práctica para investigación del nivel de vida (texto

provisional)”. Mimeografía resultado de un trabajo conjunto de las Naciones Unidas, la Oficina Internacional del Trabajo, la Organización de las Naciones Unidas para la Agricultura y la Alimentación, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura y la Organización Mundial de la Salud. Nueva York: Naciones Unidas.

Pachano A., Felipe A. (1995): “Construcción de patrones de consumo de alimentos: análisis exploratorio en caso Mérida”. Mérida, Venezuela. Instituto de Estadística Aplicada y Computación, Facultad de Ciencias Económicas y Sociales, Universidad de Los Andes. Tesis de grado para optar el título de Magister Scientiae en Estadística Aplicada (mimeografía). Profesora asesora : Elsy Garnica Olmos.

Pearson, Karl (1901): “On Lines and Planes of closed Fit to System of Point in Space”, **Phil. Mag.**, **6**: 559-572.

Pla, L. (1986): **Análisis multivariado: método de componentes principales**. Coro, Falcón. Venezuela. Departamento de Producción Vegetal, Área de Ciencias del Agro y del Mar, Universidad Nacional Experimental Francisco de Miranda / Secretaría General de la Organización de los Estados Americanos, Programa Regional de Desarrollo Científico y Tecnológico, Washington D.C.

ITCF, Institut Technique des Céréales et des Fourrages (1988): **STATITCF**, paquete estadístico computacional. 4ta versión.

Tinto Arandes, Jaime (1988): “Funciones de consumo en la ciudad de Mérida, influencia de factores sociodemográficos”. Mimeografía. Informe final de pasantías, Escuela de Economía, Facultad de Ciencias Económicas y Sociales, estudio realizado bajo la tutoría de Héctor Mata Brito y Rafael Solórzano. Mérida, Venezuela. FACES, ULA.

Tinto Arandes, Jaime: (1992): “Aproximación al verdadero índice de costo de vida (marco teórico)”. Mimeografía. Mérida, Venezuela. Instituto de Investigaciones Económicas y Sociales, FACES, ULA.

Vielma, Gilberto (Coord.) / Corporación de Los Andes / Universidad de Los Andes (1992): **Estudio de costo de vida de las principales ciudades de la Región de Los Andes**. Mérida, Venezuela. Gerencia de Planificación y Coordinación de CORPOANDES / Instituto de Investigaciones Económicas y Sociales, Facultad de Ciencias Económicas y Sociales (FACES), Universidad de Los Andes.

Walker, Monica (1967): "Some Critical Comments on 'An Analysis of Crimes by the Method of Principal Components' by B. Ahamad". Miscellanea in **Appl. Statist.** 16, 36-50.

## ANEXO

**VARIABLES DEL GASTO FAMILIAR** (40 variables cuantitativas del gasto). La primera letra de la variable identifica al grupo al cual ella pertenece. Las variables del gasto se listan a continuación:

• **Grupo 1 (alimento, bebidas y tabaco):**

1. **Acereales** (arepas, arroz, avena, galletas, harinas, pan, etc.),
2. **Araíces** (papas, apio, ocumo, yuca),
3. **Aleguminosas** (arvejas, caraotas, frijoles, garbanzos, lentejas),
4. **Aazúcar** (azúcar, bombones, caramelos, gelatinas, chocolates, flan, frutas en almíbar, mermeladas, miel de abejas, papelón, tortas),
5. **Ahortalizas** (tomates, cebollas, zanahorias, acelga, ajo, auyama, berenjena, cebollín, coliflor, espinaca, lechuga, pepino, pimentón, remolacha, repollo, vainitas, ají dulce, apio española)
6. **Afrutas** (aceituna, alcaparra, aguacate, cambures, duraznos, fresas, guanábanas, lechosas, limones, mandarinas, melón, moras, naranja, parchita, plátano, piña, uvas),
7. **Acarne** (carne de vacuno, hígado, cochino),
8. **Aaves** (carne de gallina y carne de pollo),
9. **Apescado** (pescado fresco, pescado salado, mariscos y crustáceos),
10. **Apreparados** (jamones cocidos, atún, sardinas, carne en lata, chorizos, diablitos, jamonadas, morcillas, salchichas, salchichones, tocineta, chuletas ahumadas),
11. **Ahuevos** (huevos de gallina),
12. **Alácteos** (leche en polvo, queso blanco, leche pasteurizada, cuajada, leche condensada, queso amarillo, queso parmesano, natilla y suero),
13. **Aaceites** (aceites de maíz, aceite de ajonjolí, mantequilla, margarina, manteca vegetal)
14. **Avarios** (café molido, salsa de tomate, mayonesa, alimentos para animales, chocolate y cacao, especies condimentos, sal, salsa inglesa, sopas deshidratadas, vinagre, salsas preparadas para espagueti, café instantáneo)
15. **Aniño** (cereales preparados, compotas, leches especiales, maizina o maicena),
16. **Arefresco** (refrescos gaseosos, jugos de frutas naturales, jugos de frutas enlatados, cepillados y helados, agua mineral, malta, refrescos no gaseosos, chicha, café, té),
17. **Abebidas** (cerveza, ron, vino, ponche crema, anís, brandy, ginebra y aguardiente),
18. **Atabaco** (cigarrillos rubios, cigarrillos negros y tabacos),
19. **Afuerahogar** (comidas fuera del hogar, bebidas alcohólicas y no alcohólicas tomadas fuera del hogar).



• **Grupo 2 (vestido y calzado):**

20. **V**hombre (fluxes, trajes, sacos, chaquetas, pantalones, camisas, corbatas, calzoncillos, calcetines, camisetas, cinturones, carteras, pañuelos, pijamas, zapatos, botas, pantuflas, sandalias, reparación de calzado),
21. **V**dama (abrigos, chaquetas, suéteres, pantalones, faldas, conjuntos, blusas o cotas, fondo y medio fondos, brassiers, bloomers, carteras, dormilonas, medias, vestidos confeccionados, trajes de baño, zapatos, sandalias, pantuflas, telas sin confeccionar y reparación de calzado),
22. **V**niños (fluxes, chaquetas, camisas, pantalones, suéteres y jerseys, uniformes, conjuntos, calcetines, calzoncillos, zapatos, botas y pijamas),
23. **V**niñas (chaquetas, abrigos, suéteres, jerseys, vestidos confeccionados, faldas, pantalones, blusas, camisas franelas, uniformes, pijamas, dormilonas, calcetines, bloomers, zapatos, sandalias, botas y pantuflas).
24. **V**bebés (abrigos, vestidos, pantalones, conjuntos, camisas, camisetas, franelas, bloomers, pantaleticas, bragas, escaarpines, gorros, fajitas, medias, monos, panteleticas plásticas, pañales de tela, pañales desechables, delantales, baberos, zapatos, botas, mantas, canastilla).

• **Grupo 3 (hogar):**

25. **H**vivienda (alquiler, agua, teléfono y aseo urbano),
26. **H**construcción-Ferretería (brocha, pintura, rodillo, cemento, cal, yeso, bloques, ladrillos, herramientas generales, cables, tomacorrientes, lijas, mecate, clavos, manguera),
27. **H**combustible-Alumbrado (electricidad, gas, velas, kerosene y fósforos),
28. **H**enseres (cortinas, sábanas, cubrecamas, cobijas, colchonetas, colchones, almohadas, alfombras, paños de cocina, toallas, telas sin confeccionar para la vivienda)
29. **H**equipos (televisor, equipo de sonido, ayudante de cocina, cocina, nevera, lavadora, secadora, máquina de coser, aspiradora, pulidora, betamax, VHS, planchas, licuadoras, relojes de pared, utensilios de cocina, juego de comedor, juego de dormitorio, juego de dormitorio matrimonial, juego de dormitorio de soltero, juego de recibo, cunas, lámparas y accesorios, discos y cassettes, aparatos telefónicos, maletas, maletines, reparación de equipos del hogar),
30. **H**vehículos (automóvil, bicicleta),
31. **H**diversos (detergente, jabón para lavar, limpiadores, bombillos, ceras, cloro, coletas, mopas, útiles de limpieza, insecticidas, papel parafinado y metálico, servilletas de papel, bolsas plásticas, bolsas de colar café, hilos, botones, agujas, vasos, esponjas, enjuagues, suavizantes, teteros),

• **Grupo 4 (diversos):**

32. **D**médico (medicinas, gastos de hospitalización, radiografías, visitas médicas, dentista, exámenes de laboratorio),
33. **D**cuidadop (cosméticos, champo, peluquería, barbería, cepillo para dientes, dentífricos, hojillas de afeitar, jabones de baño, papel higiénico, pasta y jabón de afeitar, talco, loción para niños, toallas sanitarias),
34. **D**educación (periódicos, revistas, libros de lectura, matrícula de inscripción, mensualidad por pago enseñanza, textos de enseñanza, útiles de enseñanza, maletines, instrumentos),
35. **D**diversiones (viajes de recreo, cine, teatro, juguetes, juego, películas, revelado, diversiones en general, apuestas de caballos, loterías y similares),
36. **D**transporte (pasajes en transporte urbano, bus extraurbano, avión, porpuesto extraurbano, escolar; lavado del vehículo, engrase del vehículo, estacionamiento, gasolina, lubricantes, repuestos, reparación),
37. **D**servicio (servicio doméstico, hoteles),
38. **D**seguros (de responsabilidad civil, personales, vivienda, casco de automóvil),
39. **D**impuestos (matriculación de auto, derecho de frente, municipal de auto),
40. **D**otros (intereses, regalos a personas, gastos funerarios, cuotas políticas, artículos de deporte, adornos para la casa, reparación de prendas, compra de prendas, flores, matas, adornos).