

Análisis multivariante sobre la vivienda

Housing multivariate analysis

Elsy Garnica Olmos*

Resumen

En este trabajo se describen algunos resultados, acerca de las características de la vivienda, de la Encuesta de Presupuestos Familiares, efectuada en Mérida, Venezuela, en 1986. Los datos se observan bajo la óptica multivariante del Análisis Factorial de Correspondencias Múltiples. En la investigación, se efectúan “retratos” de la posición que toman las numerosas características, y se analizan las interrelaciones entre todas las variables bajo estudio, lográndose destacar dos perfiles: “bienestar” y “vivienda”. El ingreso familiar se toma, a efectos del análisis, como una variable ilustrativa, observándose cómo los recursos monetarios inciden en gran parte, sobre el “bienestar” familiar y, en menor grado, sobre la estructura física de la vivienda.

1. 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 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.

* Universidad de Los Andes, Instituto de Investigaciones Económicas y Sociales

En esta investigación se centra el interés sólo en los resultados de la EPF realizada en la ciudad de Mérida, en 1986 (Vielma, 1992), acerca de las características de la vivienda: tipo, ambientes, servicios, equipamiento y condiciones de habitabilidad. Y, con fines ilustrativos, se toma el monto total de ingreso familiar. El perfil de la vivienda de la comunidad emeritense es uno de los resultados finales del trabajo.

Se emplea una técnica multivariante, el Análisis de Correspondencias Múltiples, para analizar el gran número de variables categóricas de la EPF y para no obviar información importante acerca de las interrelaciones entre todas ellas. Además, la variable Ingreso, ilustra el resultado al describir su relación con el perfil de la vivienda.

2. Hipótesis y objetivos

Los hallazgos de este trabajo se limitan al contexto de la observación. No se establecen hipótesis cuantitativas ni modelos preconcebidos acerca de la estructura y equipamiento de la vivienda del merideño, ni siquiera se realizan inferencias estadísticas proyectadas en el tiempo porque la estructura de las características de una vivienda tipo, no permanece inmutable en el tiempo. Sólo se refleja la realidad que plasman, para 1986, las 487 familias observadas.

Por lo antes descrito, no se pretende ser riguroso en cuanto a la generalización de los resultados. La hipótesis general del trabajo es: el ingreso familiar cumple un papel importante dentro del perfil y equipamiento de la vivienda. El resultado puede ser la base de estudios inferenciales posteriores en el campo de los presupuestos familiares.

No debe dejarse de lado el hecho de que no todas las familias entrevistadas expresan la verdad acerca de sus recursos monetarios, sean éstos ordinarios o extraordinarios, bien sea porque no quieren hacerlo, o bien, porque no recuerdan el monto con exactitud. Aunque este comportamiento es conocido, el ingreso recabado en la EPF continúa tomándose como la mejor aproximación al ingreso familiar verdadero.

Son dos los objetivos del estudio:

1. Determinar las características de la vivienda en Mérida, destacando la interrelación que existe entre las variables de esta naturaleza.
2. Describir el papel que juega el ingreso en este contexto.

La técnica multivariante aplicada a los datos relacionados a la vivienda y al ingreso, logra cumplir con los dos objetivos.

3. La encuesta

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

El tamaño de la muestra, con un diseño de muestreo estratificado, con afijación proporcional y polietápico, abarcó un número de 487 familias. Los estratos se definieron 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 etapas del muestreo fueron: primero, las cuadras y segundo, las familias.

Además de obtener en la encuesta, datos sobre la vivienda, se recabó información acerca de los diferentes tipos de gasto que realizaba la familia: gastos en alimentos, bebidas y tabaco, gastos en vestido y calzado, gastos del hogar y gastos diversos. El análisis multivariante de la información cuantitativa de la EPF se publicará en un próximo número de *Economía*.

Por otra parte, el dato de ingreso se obtuvo al precisar el monto percibido por la familia el mes anterior. Otros ingresos se referían a los aguinaldos, bonos, prestaciones, etc.; las familias debían recordar el monto percibido, por estos conceptos, durante el año anterior.

Los datos parciales (en su mayoría cualitativos) de la EPF realizada en Mérida, durante el año 1986, constituye la fuente de información del presente estudio.

4. Metodología estadística

El análisis estadístico multivariante utilizado en la presente investigación es el Análisis Factorial de Correspondencias Múltiples (AFCM). Esta técnica es esencialmente descriptiva. La información contenida en la tabla de datos se presenta, en forma resumida y gráfica, lo menos deformada posible. La situación es similar a la del fotógrafo que observa un paisaje y trata de sacarle el mejor provecho con su cámara, tomando la foto desde el “mejor” ángulo. El método procesa, inicialmente, una matriz de datos con la finalidad de formar la estructura e interrelaciones entre las características bajo consideración.

El AFCM es una técnica estadística, que fue estructurada, inicialmente, para el uso de un grupo de investigación, dirigidos por el profesor J.P. Benzecri, a principios de los años 60; más tarde, la técnica fue actualizada por su equipo de investigación de la Universidad de París VI. Su práctica se hizo popular a partir de los años 80, con la llegada de las micro computadoras.

El objetivo esencial del método AFCM 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. Cuando en el análisis resultan importantes los siguientes ejes (tercero, cuarto, etc.) se debe tener cuidado en la interpretación y explicación de los mismos porque puede complicar, en lugar de simplificar, el análisis descriptivo.

La técnica está desarrollado en obras tales como las de De Dervin (1991), Benzecri (1984), Lagarde (1983), Benzecri (1980) y Rennes (1979). El lector interesado en la esencia del método de Análisis Factorial de Correspondencias Múltiples (AFCM) debe consultar la bibliografía propuesta. A continuación se describen sólo las líneas generales de la técnica AFCM:

1. *Tabla de datos*: los datos a ser analizados pueden ser cualitativos (de variables nominales u ordinales), o cuantitativos provenientes de variables discretas y/o de variables continuas (éstas últimas clasificadas en categorías). El usuario de esta técnica debe, en

- primer lugar, presentar su tabla inicial de la forma $N \times Q$, donde N son las líneas o individuos y Q , las variables del análisis. Luego, el programa del AFCM, convierte esa tabla inicial en otra de forma disyuntiva completa de N columnas y J filas, donde J es el número de categorías totales correspondiente a las Q variables del análisis.
2. *Tabla de Burt*: contiene todas las frecuencias cruzadas de cada dos categorías. Antes de analizar los resultados del AFCM se aconseja observar esta Tabla de Burt ($J \times J$). Esta Tabla pone en evidencia la información inconsistente o extraña.
 3. *Distancia II-cuadrado* (χ^2): es la medida de la distancia entre dos puntos proyectados en un sistema de coordenadas. Anteriormente se había mencionado que el objetivo del AFCM era la reducción de la información a pocas dimensiones, para observar en un gráfico, la estructura de los puntos proyectados en el mismo. Pero, ¿qué se observa en este plano formado por los dos primeros ejes y en donde hay una nube de puntos? Se observan y, dado el caso, se interpretan, las proximidades entre esos puntos, medidas por la distancia (χ^2), sobre sus perfiles; líneas y columnas de la Tabla de Datos juegan un mismo papel en el AFCM¹.

Cuando, en el plano, dos puntos-individuos están cercanos, puede decirse que esos dos individuos revelan un comportamiento semejante, en el grupo de variables con mayor contribución en la formación de los ejes del plano. Lo mismo vale para dos puntos-categorías cercanos: esas dos categorías tienen un comportamiento semejante para el grupo de individuos de mayor peso en el primer plano.

Pero, cuando un punto-categoría está cercano a un punto-individuo, la interpretación no es tan simple, porque debe tenerse en cuenta los valores de la inercia y de los valores propios².

Los valores propios (V_j) están comprendidos entre 0 y 1; un valor propio cercano a 1, destaca diferencias entre los perfiles de los puntos; un valor propio cercano a 0, no lo hace. Cada valor propio representa la parte de la variación, de la nube de puntos, que es explicada por el eje en consideración. Por lo general, los valores propios son más cercanos de 0 que de 1, entonces es preferible

utilizar un indicador de la disparidad entre los puntos: $\sqrt{(V_j)}$; mientras mayor sea la disparidad, más real es la interpretación de las proximidades en el eje respectivo. Cuando los valores propios de los dos primeros ejes son muy pequeños (mucho menores que uno) no puede concluirse que la proyección de ese punto-individuo esté explicado por el punto-categoría adyacente; el punto-individuo puede estar ubicado cerca del punto-categoría, sin que exista una relación verdadera. Si las cifras de los valores propios de los ejes involucrados en el plano, son cercanas a uno (1), entonces puede hacerse aquella generalización: el individuo está parcialmente explicado por la categoría cercana.

4. *Gráficos*: Se mencionó anteriormente que puede haber más de dos ejes importantes en el AFCM³. Estos mismos ejes pueden combinarse con los dos primeros, formando planos para dar más información del fenómeno en estudio, pero la simplificación debe ser el criterio a seguir a la hora de explicar y “resumir” información. De aquí en adelante se hará alusión sólo al primer plano, formado por los ejes 1 y 2, para explicar la forma de interpretación de los puntos proyectados. En el gráfico compuesto por el plano principal de los dos primeros ejes, se proyectan los puntos-categorías. En este plano se analizan las cercanías entre las categorías de una variable con las categorías de las otras. Existen varios aspectos que deben ser tomados en consideración cuando se analizan los gráficos:
 - a) Cuando las diversas categorías de una misma variable presentan un “orden” natural; entonces tiene sentido “unir” las diversas categorías, por medio de trazos continuos, y en este caso, se estudia su recorrido;
 - b) Cuando la nube de puntos categorías puede ser reducida a una sola dimensión, la imagen que se obtiene en el plano principal, presenta una forma parabólica: es el efecto “Guttman”⁴.
 - c) Cuando en el primer plano se observan puntos-individuos aislados o “raros”, entonces es necesario verificar los datos originales y la Tabla de Burt, descrita anteriormente; puede ser que este comportamiento se deba a un error de codificación, o bien, a un comportamiento real del individuo observado.

- d) Cuando se observa la formación de varias nubes de puntos-individuos se debe decidir entre hacer estudios separados de AFCM en esas sub-poblaciones o seguir con la exploración de los datos.
 - e) Un resultado gráfico interesante de analizar es el plano donde se proyectan, conjuntamente, los puntos-categorías y los puntos-individuos. Debido a que las columnas y filas de la matriz de datos juegan un rol simétrico, la representación simultánea de estas dos nubes de puntos, es posible.
5. *Variables y categorías:* Para lograr una buena representación de los datos, el número de modalidades o categorías de cada variable debe ser parecido. Las diferentes modalidades, también deben tener, en la medida de lo posible, frecuencias similares para no distorsionar la información gráfica. Según una regla empírica, el efectivo de una modalidad debe ser mayor que $\sqrt{(3N)}$ (C. Dervin, 1991). Cuando una modalidad tiene una pequeña frecuencia, el punto respectivo queda proyectado en un extremo del gráfico, distorsionando así, toda la nube de puntos. Lo mismo pasa cuando existen variables con diferentes números de clases; puede ser que las variables dicotómicas tengan un recorrido más reducido debido al gran contingente en cada una de sus clases, al contrario de otras variables, que repartirán, con menores frecuencias, el total de individuos dentro de sus modalidades, dispersando sus proyecciones en el plano.
6. *Indicadores del AFCM:* Para determinar cuáles categorías (y, por supuesto, sus variables respectivas) son las más importantes en la formación de los ejes principales, se analizan los indicadores: coseno cuadrado y contribución relativa. El *coseno cuadrado* permite describir la calidad de la representación, de cada una de las categorías, sobre el eje en consideración. Cuanto más grande sea el valor de este indicador, mejor será la calidad de la representación del punto-categoría respectivo. Por otra parte, la *contribución relativa* de cada uno de los puntos a la inercia explicada por el eje en consideración, es otro indicador que especifica las categorías que más contribuyen a la formación

del eje (las contribuciones relativas son sus pesos marginales). La suma de las contribuciones relativas de todas las categorías activas es igual a 100.

El nombramiento de los ejes es un paso relativamente subjetivo porque tiene que ver con la experiencia y el conocimiento del investigador. En la denominación de los ejes debe observarse las interrelaciones entre las categorías, los recorridos de las variables, la calidad de la representación (cosenos cuadrados) y las contribuciones relativas. Las interrelaciones existentes entre las variables sólo las podrá explicar y sustentar el experto en el asunto.

Además, en la nominación de cada eje prevalece lo que el investigador crea importante destacar. Se trata, en la medida de lo posible, de ajustar esta denominación a la realidad, con base en los valores de los dos indicadores señalados.

7. *Variables e individuos ilustrativos*: Son elementos que no intervienen en el cálculo de los ejes principales, pero que son proyectados en los gráficos con la finalidad de ayudar en la interpretación de los ejes. En la mayoría de los casos, es conveniente incorporar variables y/o individuos ilustrativos o suplementarios, después de realizar un AFCM preliminar.

Una variable ilustrativa debe estar asociada a la definición de los ejes principales. Un individuo ilustrativo, por ejemplo, puede tratarse de un caso hipotético con valores extremos en algunas variables importantes en el plano principal. Cualquiera de estos puntos ilustrativos (variables, individuos o ambos), puede ayudar al investigador a definir el eje en consideración.

5. Variables utilizadas

Se realizan numerosos análisis exploratorios para depurar la gran masa de información aportada por familias de la ciudad de Mérida en el año 1986. Veintidós (22) fue el total de variables estudiadas, entre categóricas y numéricas, medidas, originalmente, sobre 487 encuestas. Luego, después de eliminar aquellas con datos dudosos, la información

se reduce a 476 encuestas. La lista de las variables originales del Estudio de Presupuestos Familiares, que se toma en cuenta para la presente, investigación, se detalla a continuación:

Vivienda:

1. Tipo de vivienda: “rancho”, “casa de vecindad”, “apartamento” y “casa o quinta”.
2. Número de habitaciones totales: 2, 3 ó 4, 5 ó 6, 7 u 8, 9 ó 10 y 11 ó más.
3. Cocina separada: “no tiene” y “si tiene”.
4. Comedor separado: “no tiene” y “si tiene”.

Servicios:

5. Agua: “fuente común” y “acueducto”.
6. Ducha: “no tiene”, “uso común” y “uso exclusivo”.
7. Water: “no tiene”, “uso común” y “uso exclusivo”.

Tenencia actual de la vivienda:

8. Forma de posesión: “cedida”, “alquilada”, “hipotecada”, “propia pagándose” y “propia”.

Equipamiento de la vivienda:

9. Cocina: “de kerosene”, “a gas” y “eléctrica”.
10. Calentador de agua: “no tiene” y “si tiene”.
11. Lavadora: “no tiene” y “si tiene”.
12. Nevera: “no tiene” y “si tiene”.
13. Radio: “no tiene” y “si tiene”.
14. Reproductor o tocadiscos: “no tiene” y “si tiene”.
15. Televisor a color: “no tiene” y “si tiene”.
16. Vehículos: “no tiene” y “si tiene”.
17. Motocicleta: “no tiene” y “si tiene”.
18. Bicicleta: “no tiene” y “si tiene”.

Condiciones de habitabilidad:

19. Número de habitaciones para la familia: 1 ó 2, 3 ó 4, 5 ó 6, 7 u 8, 9 ó 10 y 11 o más.

Variables demográficas:

20. Número de personas en el hogar: 1 ó 2, 3 ó 4, 5 ó 6, 7 u 8, 9 ó 10 y 11 o más.
21. Número de empleados domésticos: 1, 2 y 3 sirvientes.

Variables ilustrativas:

22. Ingreso: aunque en la encuesta se desglosó la fuente de ingresos en variedad de categorías, para efectos de este análisis exploratorio se requería sólo el monto total, medido en bolívars mensuales por persona⁵. Esta variable se clasificó en cinco categorías uniformes: la primera categoría, con los menores ingresos y la quinta categoría, con los mayores⁶.
23. Sirv/nf (número de sirvientes / número de miembros familiares), este indicador puede corroborar la interpretación de los ejes. Las categorías, a excepción de la clase cero, tienen números similares de individuos: Clase 0 (valor de cero, con el mayor contingente de individuos), Clase 1 (con valores entre 0,10 a 0,14), Clase 2 (con valores entre 0,14 a 0,20), Clase 3 (con valores entre 0,20 a 0,25) y Clase 4 (con valores entre 0,25 a 1,00).

6. Resultados

El proceso estadístico se realiza sobre la información, en su mayoría, categórica que aportan 476 familias en veintiún (21) variables activas y dos (2) variables ilustrativas; este conjunto de variables producen 63 categorías activas y 10 categorías ilustrativas. El análisis, netamente descriptivo, logra poner en relieve los dos objetivos del estudio: por un lado, las principales características de la vivienda en Mérida y las interrelaciones entre ellas y, por el otro, descripción del rol que juega el ingreso en este contexto.

Anteriormente se había mencionado que el AFCM servía para detectar información incongruente o extraña, y eso sucedió, en los gráficos resultantes de un AFCM inicial; se observaron varios puntos-individuos “extraños”, no congruentes con el conjunto de datos. Luego en, la Tabla de Burt (frecuencias cruzadas de cada dos categorías), se determinaron algunas inconsistencias. Un ejemplo patético de falla en el asentamiento de datos, en la encuesta que nos ocupa, es el de una familia que no cuenta en su vivienda con un comedor separado, calentador, reproductor, TV a color, ni automóvil y con la clasificación de ingreso más baja, tenga, a su vez, 8 sirvientes; evidentemente, en este caso, hubo error de información. Al analizar cuidadosamente la Tabla de Burt y los gráficos iniciales, se detectan algunas incongruencias de esta naturaleza. De 487 familias originalmente observadas, once fueron los casos eliminados en el preliminar AFCM, quedando un total de 476 familias o individuos a estudiar.

Del análisis de las 476 familias, resultan 62 valores propios⁷, los tres primeros con 9%, 7% y 5% de la inercia total. La explicación de los resultados se centra en los dos primeros valores propios, con un 16% de la inercia total. El tercer eje no se describe porque no aclara el panorama.

En los resultados también influyen dos hechos: a) el número de modalidades no es igual para todas las variables (es decir, sus pesos no son equivalentes) y b) existen modalidades raras en el sentido de que sus frecuencias son muy pequeñas⁸.

6.1. Indicadores del AFCM

Debe recordarse que el AFCM es netamente descriptivo, y en esta investigación, la proyección de la nube de puntos-categorías sobre el plano principal, es utilizado para observar las interrelaciones entre el conjunto de puntos-categorías que engloba el aspecto de vivienda del merideño. Los valores que toman los cosenos cuadrados y las contribuciones relativas son decisivos al definir o interpretar los ejes.

6.1.1. Cosenos cuadrados

Los cosenos cuadrados permiten describir la calidad de la representación, sobre el eje en consideración, de cada una de las categorías. Cuanto más grande sea este indicador, mejor será la calidad de la representación del punto-categoría respectivo.

Los valores de los cosenos cuadrados, que más se destacan, en la formación del primer eje del AFCM están dentro de un rango de 0,37 y 0,28. Estas categorías (“tiene” y “no tiene”) corresponden a las variables: televisión, automóvil, lavadora y calentador (véase en el anexo, Cuadro A-1).

Con respecto al segundo eje del AFCM, los valores de los cosenos cuadrados más importantes, con rango entre 0,44 y 0,17, (véase en el anexo, Cuadro A-2), conciernen a las categorías de las variables: tipo de vivienda (“apartamento” y “casa”), water y ducha (sólo las clasificaciones de “uso común” y “uso exclusivo”), forma de posesión de la vivienda (sólo la condición “propia pagándose”), número de habitaciones (las clases 3 ó 4, 7 u 8 y 9 ó 10) y número de habitaciones para la familia (las categorías 3 ó 4 y 7 u 8).

En la figura 1 sólo se proyectan los puntos-categorías importantes en el primer eje. En la figura 2, aquellas importantes en el segundo eje.

6.1.2. Contribuciones relativas

Las contribuciones relativas de los puntos a la inercia explicada por el eje en consideración, es otro indicador que describe cuáles son las categorías que más contribuyen a la formación del eje (las contribuciones relativas son sus pesos marginales).

Las contribuciones relativas más importantes (entre 9,9% y 7,5%) en la formación del eje 1, son las de las variables: televisión a color, automóvil, número de habitaciones para la familia, lavadora y calentador (véase en el Anexo, Cuadro A-3).

Las contribuciones relativas agregadas (por variable, no por categorías)⁹ sobresalientes (entre 20,9% y 10,3%) en la formación del eje 2, son las correspondientes a las variables: tipo de vivienda, número total de habitaciones, número de habitaciones para la familia, water, ducha y tenencia de la vivienda (véase en el Anexo, Cuadro A-4).

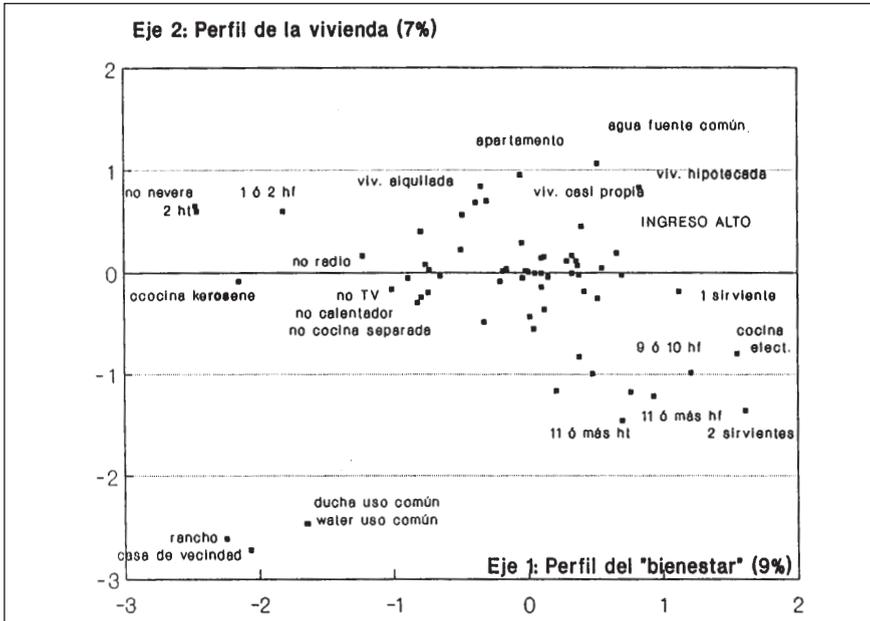


Figura 1. Clases activas extremas e ingreso (v.i)

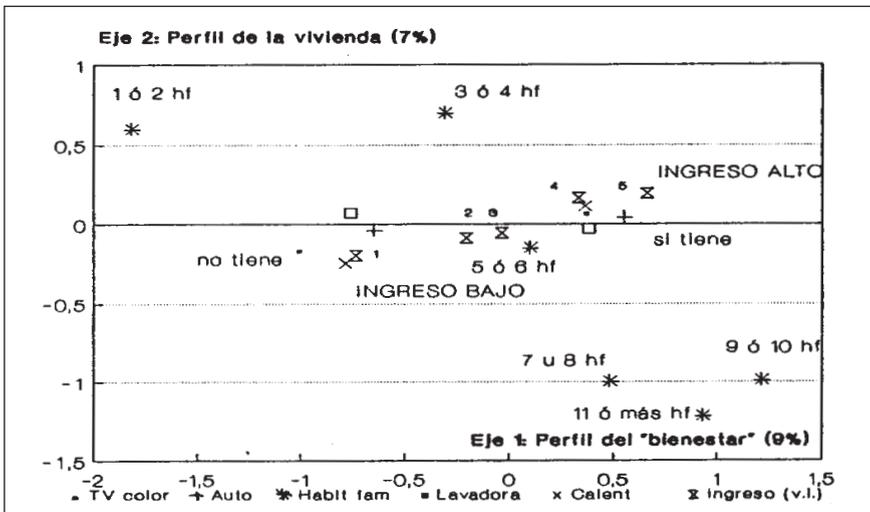


Figura 2. Variables importantes eje 1 del AFCM.

6.2. Análisis de los ejes principales

El nombramiento de los ejes es un paso relativamente subjetivo. La interpretación de cada uno de los ejes se basa en la experiencia y conocimiento del investigador; también, en lo que él crea importante destacar dentro del análisis estadístico descriptivo, con base en los valores de los dos indicadores señalados del AFCM: los cosenos al cuadrado y las contribuciones relativas.

La figura 1 muestra la proyección de algunas categorías en el primer plano principal, algunos gráficos ofrecidos resultan “alargados” debido a limitaciones del paquete computacional utilizado; no obstante, ellos sirven de ilustración para describir la estructura de los puntos proyectados.

6.2.1. Análisis del primer eje (9% de la variación total): “perfil del bienestar”

En la figura 2 puede observarse como el ingreso, variable ilustrativa, se mueve desde el tercer cuadrante (ingreso más bajo) hasta el primer cuadrante (ingreso más elevado). 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. Este es un resultado esperado.

La variable número de habitaciones para la familia (hf) presenta una mayor dispersión dentro del gráfico debido a un número mayor de categorías. Los individuos del análisis se distribuyen en seis (6) de ellas, en contraposición de las dos (2) clases, en la mayoría del resto de las variables (dicotómicas) importantes en la formación de este primer eje. Esta variable (hf) indica el espacio físico destinado a la familia, pero no, el grado de hacinamiento de los miembros familiares puesto que la variable (hf) no indica el número de habitaciones por miembro familiar, sino número de habitaciones para todos ellos. El comportamiento de (hf) dentro del gráfico, es también esperado, pues una familia con vivienda bien equipada dispondrá, por lo general, de mayor desahogo en espacio físico para sus miembros.

Este primer eje podría designarse perfil del “bienestar”¹⁰. Se le asigna esta denominación porque las categorías que tienen la mejor representación dentro de este primer eje se refieren a las variables de

equipamiento de la vivienda: televisor a color, automóvil, lavadora y calentador. Son estas mismas variables, más el número de habitaciones para la familia (hf), las que tienen las contribuciones relativas más importantes. Esta última variable nombrada presenta una gran contribución relativa, pero sus clases (seis en total) no tienen cosenos cuadrados altos como para situarse dentro de las categorías mejor representadas.

6.2.2. Análisis del segundo eje (7% de la variación total): "perfil de la vivienda"

En este segundo eje el ingreso tiene poca variabilidad visible. Existen algunas variables activas que se observan muy dispersas en cuanto al segundo eje de la figura 3. En este plano sólo se proyectan las categorías de las variables más importantes del segundo eje principal.

La primera variable esencial es Tipo de vivienda¹¹. En la figura 3, las categorías más bajas ("rancho" y "casa de vecindad") se ubican en el tercer cuadrante; la condición "apartamento", en el primero y, finalmente, la más alta clasificación, "casa" (o quinta), se sitúa en el cuarto cuadrante, cerca del origen. Dos de sus categorías ("apartamento" y "casa") tienen las mejores representaciones (máximos cosenos cuadrados) y la máxima contribución relativa al Eje 2.

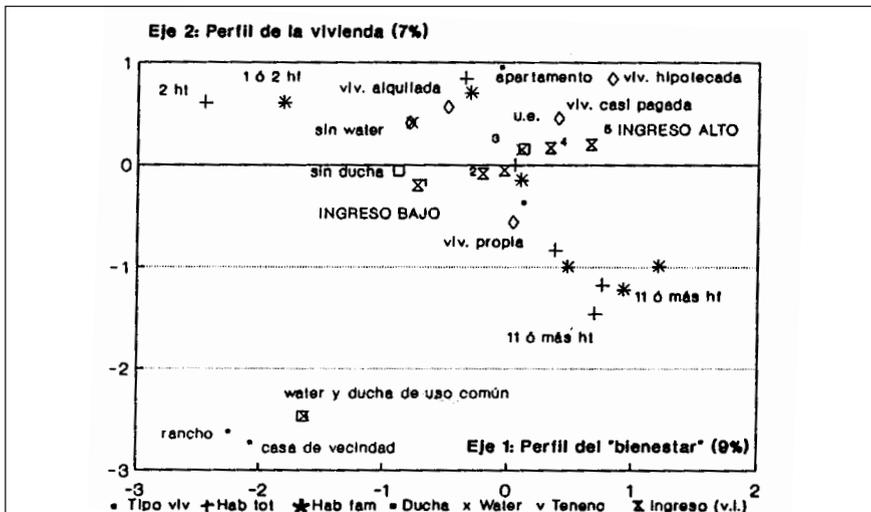


Figura 3. Variables importantes eje 2 del AFCM.

Las dos siguientes variables importantes, de acuerdo a sus contribuciones relativas a la formación del segundo eje, son: Número de habitaciones totales (hab) y Número de habitaciones para la familia (hf). Cada una de estas variables está clasificada en seis (6) categorías y algunas de ellas son muy importantes en cuanto a la representación en el segundo eje. Estas dos variables (hab y hf) se mueven de manera similar, iniciando su recorrido en el segundo cuadrante con la condición de “una o dos habitaciones”, hasta la última clasificación, “11 o más habitaciones”, ubicadas en el cuarto cuadrante.

Los servicios más importantes en este segundo eje, water y ducha, presentan similitud de posición dentro del gráfico. Inician su movimiento con la categoría más baja, “no tiene”, ubicadas cerca del origen, a pesar de las bajas frecuencias (de 4 y 12 respectivamente), pasan a la segunda jerarquía, “uso común”, a un extremo del tercer cuadrante y cerca de los puntos-categorías “rancho” y “casa de vecindad” (de la variable Tipo de vivienda), y finalmente llegan, de nuevo, a ubicarse cerca del origen, con su tercera y última categoría, “uso exclusivo”, cerca del punto-categoría “casa” o quinta (de la variable tipo de vivienda). El tipo de vivienda y su baño, perfila la estructura física de la vivienda.

La tenencia de la vivienda, otra de las variables importantes en la formación del segundo eje, tanto en calidad de representación como en contribuciones relativas, proyecta sus puntos-categorías muy cerca del origen. La categoría “hipotecada”, se observa un poco alejada del origen, quizás debido a su baja frecuencia (sólo 9 casos). La variable Tenencia de la vivienda presenta sus clases “cedida” y “alquilada” en el segundo cuadrante (cerca de los puntos-categorías “no tiene ducha” y “no tiene wáter”); las dos siguientes, “hipotecada” y “propia pagándose”, están en el primer cuadrante, y esta última, específicamente, cerca de los puntos-categorías “uso exclusivo de la ducha” y “uso exclusivo del water”. Termina el recorrido con la categoría “propia” (ubicada cerca de la categoría “casa” de la variable Tipo de vivienda).

Este segundo eje podría denominarse perfil de la vivienda, porque las variables más importantes, en la formación del mismo, se refieren a la vivienda, su tamaño y sus servicios.

6.3. Variables ilustrativas

En el análisis se introducen dos variables ilustrativas cuyos puntos-categorías se proyectan en la figura 4. La primera, el Ingreso, con cinco (5) categorías uniformes. El recorrido que toma esta variable, tal como puede observarse en los gráficos 2 y 3, está mejor relacionado con el Eje 1 (perfil del "bienestar") que con el Eje 2 (perfil de la vivienda). La otra variable ilustrativa es el indicador sirv/nf (número de sirvientes/número de miembros familiares), con un rango entre cero y uno; el valor cero indica ausencia de sirvientes, el valor uno manifiesta que el número de sirvientes es idéntico al número de miembros de la familia. Este indicador se mueve ampliamente sobre los dos ejes principales, quiere decir que probablemente esté relacionado con ambos ejes.

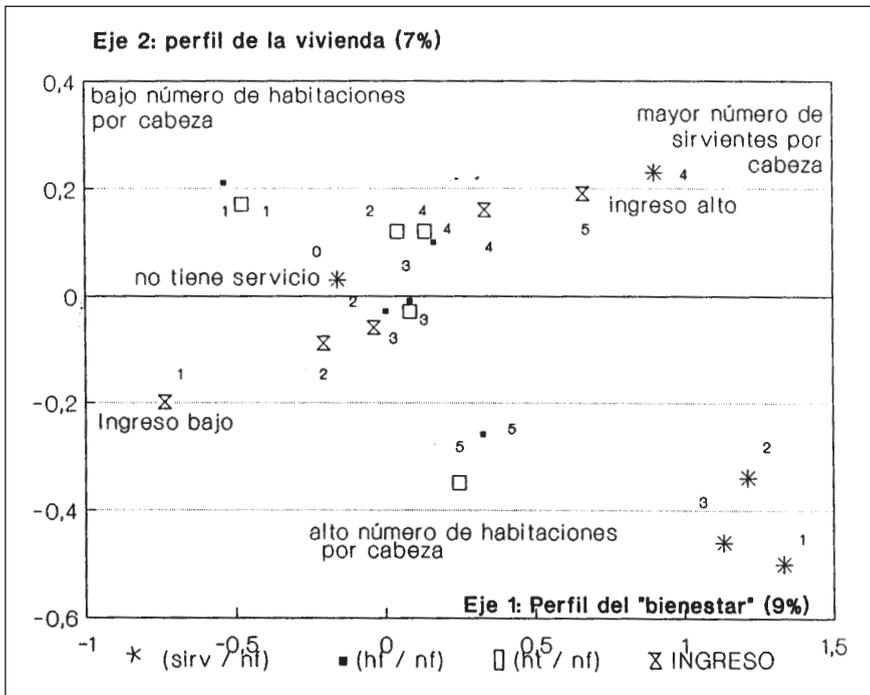


Figura 4. Variables ilustrativas del AFCM.

Del análisis de las proximidades (véanse los puntos-categorías de la figura 4) entre las dos variables ilustrativas se deduce que puede existir una relación entre el valor más alto del indicador de *sirv/nf* y el nivel más alto de ingreso, porque la proyección de la clase 4 (clase superior de *sirv/nf*) está cerca de la proyección del nivel de ingresos más elevado (clase 5). El número de sirvientes tiene mucho que ver con los niveles de ingreso de la familia, más aún cuando los ingresos son altos. ¿El número de sirvientes por persona, está asociado al “bienestar” de la familia y al perfil de la vivienda? Definitivamente sí.

Las clases centrales de *sirv/nf* (1, 2 y 3) se proyectan a un extremo del plano, en la esquina inferior derecha de la figura 4, indicando una relación de estas tres categorías con el mayor “bienestar” y bajo perfil de la estructura física de la vivienda. Mientras que las clases extremas (0 y 4) se ubican en la parte superior del plano, con perfiles diferentes del “bienestar” (bajo y alto, respectivamente), pero ambas clases con un alto perfil en vivienda.

Una vivienda con todos los servicios básicos (alto perfil de vivienda) de water, ducha, comedor separado, cocina separada, calentador, etc., puede no disponer de “ningún sirviente, por cabeza” (clase 0) o de un alto valor en este indicador (clase 4), ambas posiciones extremas.

La situación de alto perfil en la estructura física de la vivienda asociada con gran número de sirvientes es explicable, pero ¿por qué una familia con alto perfil de su vivienda no dispone de servicio doméstico? Varias pueden ser las razones, pero la proximidad de la categoría “1” (bajos ingresos) con la categoría “0” (ningún empleado doméstico) indica que la limitación es precisamente la monetaria (véase el segundo cuadrante de la figura 3).

Otro aspecto es el de los puntos-categorías de las clases “1”, “2” y “3” de este indicador *sirv/nf*, proyectados hacia un “alto perfil del bienestar” (extremo derecho del primer eje) y “bajo perfil en la estructura física de la vivienda” (zona inferior del segundo eje). ¿Cuál es la explicación en el contexto de los datos? Lo que refleja la proyección de puntos-categorías, y sus proximidades, es que el número de empleados domésticos está asociado al “bienestar” personal ¿Por qué se afirma? porque las clases

referidas a un número alto de habitaciones (viviendas grandes) están muy próximas a las categorías “1”, “2” y “3” del indicador *serv/nf*. Véase el cuarto cuadrante de la figura 4 y recuerde la posición que tomaban las categorías altas de las variables Número total de habitaciones (*hab*) y Número de habitaciones para la familia (*hf*) en la figura 3. Es decir, el empleado doméstico está presente, por lo general, cuando la vivienda es grande, aunque no necesariamente este tipo de vivienda cuente con buenos servicios (de water y ducha, por ejemplo).

6.4. Nube de puntos individuos

En el AFCM se utiliza la representación simultánea de los puntos-categorías y de los puntos-individuos ya que las líneas y las columnas juegan un rol simétrico. La mayor parte de la información se encuentra en el plano principal, formado por los ejes 1 y 2. La proyección de los puntos-individuos en el plano principal, conjuntamente con la proyección de los puntos-categorías, sirve de ilustración al análisis.

Cuando dos puntos-individuos están cercanos, puede decirse que esos dos individuos revelan un comportamiento semejante con respecto al grupo de variables importantes en el primer plano.

En la figura 5 se observan esas proximidades. Por ejemplo, los puntos-individuos extremos proyectados en el segundo cuadrante del plano, 046, 339 y 381, tienen un comportamiento similar en variables importantes: Número de habitaciones para la familia (*nf*) y Número de habitaciones totales (*hab*) iguales, “no tienen” los principales equipos del hogar, pero tienen “uso exclusivo” de la Ducha y del Water, aunque ellos pertenezcan a diferentes niveles de ingresos. Lo mismo puede decirse de los puntos-individuos proyectados en el tercer cuadrante, 212 y 213: no tienen los principales equipos del hogar, tienen ducha y water de uso común, el mismo nivel de ingresos, pero diferentes valores en *nf* y *hab*. Comportamiento parecido también se observa en los puntos-individuos 215 y 219. En el Anexo (Cuadro A-5) se listan los valores de algunas variables importantes para algunos de los individuos proyectados a los extremos del plano principal.

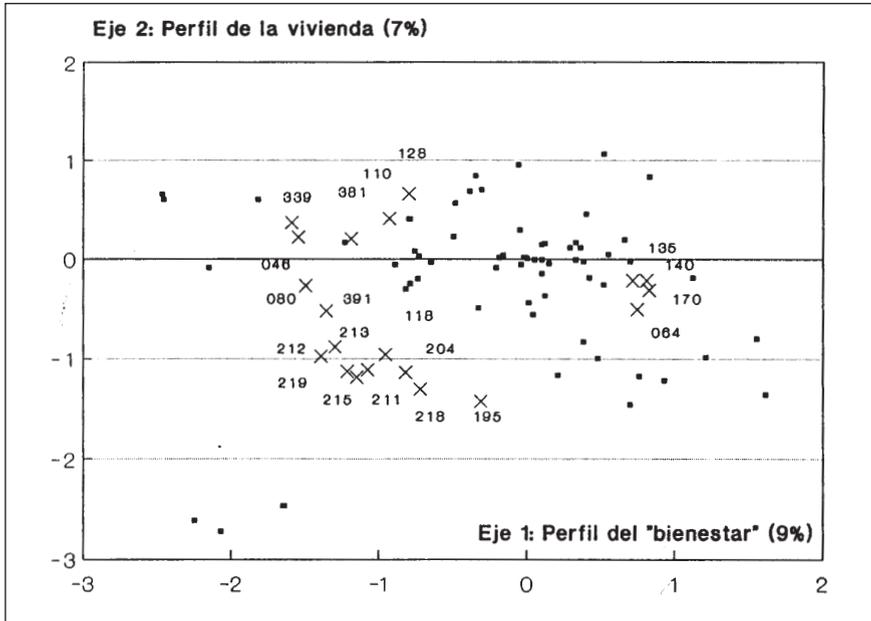


Figura 5. Clases (•) e individuos extremos (X).

Las proximidades no reflejan exactamente la similitud porque el porcentaje de inercia recogido por el primer plano del AFCM es bajo (16%). Los índices de disparidad, para los dos primeros ejes, $\sqrt{V1} = 0,42$ y $\sqrt{V2} = 0,36$, son bajos. Debe recordarse que cuanto más altos sean estos valores mejor será la interpretación de las proximidades de los puntos.

7. Conclusiones

Los resultados de la aplicación del método de Análisis Factorial de Correspondencias Múltiples son meramente descriptivos y referidos sólo a los resultados parciales de la Encuesta de Presupuestos Familiares efectuada en Mérida en el año 1986. Con el AFCM, en la descripción de las variables cualitativas acerca de las características de la vivienda, se obtienen las siguientes conclusiones:

- Un perfil del “bienestar” –o confort– donde intervienen variables como equipamiento de la vivienda: tv, automóvil, lavadora, calentador, cocina, reproductor, etc., en el primer eje.
- Un perfil de la vivienda, donde intervienen variables como el tipo de vivienda, Número de habitaciones, Tenencia de la vivienda, water y ducha. Estas dos últimas se refieren a los servicios de la misma, en el segundo eje.
- La variable ilustrativa, Ingreso, clasificada en cinco (5) categorías uniformes se proyecta bien en el eje 1 y con una visual débil asociación al eje 2. Es decir, el ingreso, como dato expresado por los encuestados, está más asociado con el perfil del “bienestar” personal, que con el perfil de la vivienda.
- La vivienda de la comunidad emeritense, vista a través de los datos que proporcionaron 467 familias en 1986, se caracteriza en primer lugar, por el equipamiento (tv, lavadora, calentador, cocina, reproductor principalmente) y si hay automóvil. En segundo plano, la familia toma en cuenta el tipo de vivienda (rancho, casa de vecindad, apartamento o quinta), su número de habitaciones, la forma de posesión de la misma y los servicios de water y de ducha. Es decir, la familia emeritense requiere de su vivienda, primero, la comodidad (con equipamiento) y, luego, se preocupa por la estructura física y forma de posesión de la misma.
- Un resultado esperado, aunque no previsto dentro de los objetivos del estudio, es la asociación existente entre dos categorías de las variables ilustrativas. Al más alto ingreso, corresponde un valor alto de la razón: número de sirvientes por miembro de familia. Sólo si la familia tiene un gran número de empleados domésticos, por miembro familiar, puede asegurarse que posee altos ingresos. Esto no quiere decir lo contrario. Es decir, el hecho de que una familia posea altos ingresos, no implica que deba tener un gran número de empleados domésticos.

En Venezuela, en estudios acerca de la Encuesta de Presupuestos Familiares, no se ha analizado, en profundidad, la forma cómo la familia, percibe su vivienda. Los estudios se centran en el aspecto físico

de la misma, y restan importancia al interés que tiene la familia en el equipamiento de su vivienda. Quizás este estudio local abra las puertas a equipos multidisciplinarios compuestos por analistas de datos, sociólogos, psicólogos, ingenieros de sistemas, estadísticos y economistas, entre otros, para estudiar la vivienda como lugar común del núcleo familiar: su estructura física y las “comodidades” que éste ofrece.

8. Notas

- 1 La distancia χ^2 entre dos puntos (i e i') se calcula de la siguiente forma: $d_2(i,i') = \sum (x_{ij} / x_i - x_{i'j} / x_{i'})^2 / x_{.j}$, donde, x_i es la suma de la línea i , sobre todas las columnas; $x_{i'}$ es la suma de la línea i' , sobre todas las columnas y $x_{.j}$ es la suma de la columna j , sobre todas las líneas.
- 2 *Inercia y valores propios*: La inercia total de un estudio de AFCM es $(J-Q)/Q$ y el número de valores propios (V_j) del AFCM, excluyendo el trivial, es $\inf(N,J)-1$. Los valores propios, reflejan la cantidad de información explicada por los ejes (éstos calculados a partir de los vectores propios). Para las demostraciones matemáticas del cálculo de los valores y vectores propios el lector puede referirse a J.P. Benzécri *et al.* (1980).
- 3 Aunque no hay normas precisas para decidir cuántos ejes escoger en el AFCM, una simple regla empírica revela que, una vez ordenados los valores propios, en forma descendente, se toman los primeros valores propios, antes de cambiar bruscamente hacia abajo, el porcentaje de variación explicada. Otros investigadores escogen todos los valores propios mayores a la relación: $[1 / (\inf(N,J) - 1)]$, valor que conllevaría cada uno de los ejes si la repartición de la inercia es igual para todos los ejes. En este punto, vuelve a recordarse que el objetivo del AFCM es la simplificación. Por lo general, en la práctica se escogen los dos primeros valores propios (máximo, los tres primeros), aunque el porcentaje de variación sea pequeño, pues lo que interesa es simplificar la información.
- 4 El efecto Guttman se obtiene cuando la nube de puntos tiene forma de parábola. Cuando las categorías de las variables pueden “unirse” por medio de trazos continuos, es de mucha utilidad observar si en el plano se presenta este efecto. Teóricamente es el efecto de veredictos extremos

- y contradictorios, el centro de gravedad está en el interior de la parábola. Si la concavidad está dirigida al primer eje, por ejemplo, entonces el segundo eje se convertirá en una escala de opiniones “fuertes” (extremas y contradictorias), en la parte inferior, y de opiniones medias o suaves en la parte superior del segundo eje. Para suavizar este efecto puede ser recomendable trabajar con variables transformadas (logaritmos, inversas, etc.).
- 5 Se realizó esta simplificación, en primer lugar, para tener una idea de la distribución por cabeza y, en segundo lugar, para reducir los montos.
 - 6 El Ingreso fue agrupado en las siguientes cinco (5) clases uniformes: Clase 1: 115,78 hasta 940,22 bolívares/miembro/mes (b/m/m); Clase 2: ingresos mayores a 940,22 hasta 1.365,18 b/m/m; Clase 3: ingresos mayores a 1.365,18 hasta 1.860,00 b/m/m; Clase 4: ingresos mayores a 1.860,00 hasta 2.862,50 b/m/m; Clase 5: ingresos mayores a 2.862,50 hasta 15.000 b/m/m (valor máximo observado).
 - 7 El número de valores propios de este análisis (excluyendo el trivial) es: $\inf(N,J)-1 = 62$ y la inercia total es $(J-Q)/Q = (63 - 21) / 21 = 2$. El primer valor propio es igual a 0,18 (con un 9% de la inercia total); el segundo, es igual a 0,13 (con un 7% de inercia total); el tercero, es igual a 0,11 (con un 5% de la inercia total). (Cada uno de los tres ejes principales conlleva, de porcentaje de inercia, más del doble del 1,6% aproximadamente que, en igualdad de condiciones contendría cada uno de los ejes. En este estudio, pueden escogerse, para su análisis todos aquellos ejes con valores propios mayores a la relación $[1 / (\inf(N,J) - 1)] = 0,016129$.
 - 8 La regla empírica del número de efectivos para cualquier categoría (el número de efectivos debe ser mayor que $\sqrt{(3N)} = \sqrt{(3*476)} = 38$) no se cumple en muchos de los casos, pues se toman las variables cualitativas originales, sin arreglos previos del número de efectivos dentro de cada clase. El número de variables cualitativas que cumplían estrictamente la regla empírica era mínimo y el AFCM efectuado con esas pocas variables perdía objetividad.
 - 9 Se denominan “agregados”, porque se suman las contribuciones relativas de todas las categorías involucradas.
 - 10 Aunque el “bienestar” es un concepto que tiene una connotación, más compleja, en esta investigación se limita a la parte de equipamiento que ayuda al “comfort”.

- 11 En la figura 3 se observa la ubicación de las categorías del Tipo de vivienda, sólo tiene un (1) caso en la clase más baja, “rancho”, y doce (12) casos en la siguiente categoría “casa de vecindad”, ello hace que los puntos respectivos se ubiquen muy lejos del centro.

9. Referencias

- Altimir, Oscar (1974). “Un archivo de datos sobre distribución del ingreso procedentes de encuestas de hogares en países latinoamericanos” Segunda conferencia Latinoamericana de la Asociación Internacional para la Investigación del Ingreso y la Riqueza en Río de Janeiro del 9-12 de Enero de 1974, Santiago de Chile: CEPAL.
- Benzécri F. y otros (1980). *L'Analyse des Données. II: L'Analyse des Correspondances*. París: Dunod.
- Benzécri, F. et collaborateurs (1984). *Pratique de L'Analyse des Données. 1: Analyse des Correspondances Exposé Élémentaire*, 2e. ed. París: Dunod.
- CEPAL (1986). “Encuestas de ingresos y gastos, conceptos y métodos en la experiencia latinoamericana” Cuadernos de la CEPAL. Santiago de Chile: Publicación de las Naciones Unidas.
- De Lagarde, Jean (1983). *Initiation à l'analyse des données*. París: Dunod.
- Dervin, C. (1991). *¿Cómo interpretar los resultados de un Análisis Factorial de Correspondencias?* A. Tomassone y J. Jhonson-Alvarez (traductores). París: Institut Technique des Céréales et des Fourrages.
- Ferguson, C. E. (1976). *Teoría microeconómica*. Bogotá: Fondo de Cultura Económica.
- Institut Technique des Céréales et des Fourrages (1988). *Manual del STAT-ITCF*, versión 4. París: ITCF.
- Lebart L., A. Morineau y N. Tabard (1977). *Techniques de la description statistique, methods et logiciels pour l'analyse des grands tableaux*. París: Bordas.
- Naciones Unidas / Oficina Internacional del Trabajo / Organización de las Naciones Unidas para la Agricultura y la Alimentación / Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura / Organización Mundial de la Salud (1964): “Manual de encuestas sobre hogares, guía práctica para investigación del nivel de vida (texto

- provisional)”, *Estudio de métodos*, serie F, No. 10. Nueva York: Oficina de Estadística de las Naciones Unidas, Departamento de Asuntos Económicos y Sociales.
- Rennes, A. (1979). *L'Analyse des Données*. París: Dunod.
- República de Chile, Ministerio de Economía, Fomento y Reconstrucción, Instituto Nacional de Estadísticas (s.f). “Índices de precios al consumidor”. Santiago de Chile: Talleres del Instituto de Estadísticas. (Mimeografía)
- República de Venezuela, Presidencia de la República, Oficina Central de Estadística e Informática (1986). Situación habitacional en Venezuela, XI Censo General de Población y Vivienda. Caracas: Taller Gráfico de la OCEI.
- Tinto Arandes, Jaime (1988). “Funciones de consumo en la ciudad de Mérida, influencia de factores socio-demográficos” (mimeografía), informe final de pasantías, Escuela de Economía, Facultad de Economía y Ciencias Sociales; tutores; Héctor Mata Brito y Rafael Solórzano. Mérida: Facultad de Ciencias Económicas y Sociales, Universidad de Los Andes.
- Vielma, Gilberto (Coord.) / Corporación de Los Andes / Universidad de los Andes (1992). “Estudio costo de vida de las principales ciudades de la Región de Los Andes”. Mérida: Gerencia de Planificación y Coordinación de CORPOANDES - Instituto de Investigaciones Económicas y Sociales, Facultad de Economía y Ciencias Sociales, Universidad de Los Andes (mimeografía).

10. Anexos

Cuadro A-1. Categorías mejor representadas en el eje 1 del AFCM

Categoría	Coseno Cuadrado con el Eje 1
0TV (No Tiene TV a Color)	0,371
1TV (Si Tiene TV a Color)	0,370
0AU (No Tiene Auto)	0,352
1AU (Si Tiene Auto)	0,351
0LA (No Tiene Lavadora)	0,291
1LA (Si Tiene Lavadora)	0,290
0CA (No Tiene Calentador)	0,283
1CA (Si Tiene Calentador)	0,282

Cuadro A-2. Categorías mejor representadas en el eje 2 del AFCM

Categoría	Coseno Cuadrado con el Eje 2
2VI (apartamento)	0,440
3VI (casa o quinta)	0,244
1WC (water de uso común)	0,353
2WC(water de uso exclusivo)	0,294
1DU (ducha de uso común)	0,335
2DU (ducha de uso exclusivo)	0,241
4TE(vivienda propia pagando)	0,283
2HT (3 ó 4 habit. totales)	0,283
4HT (7 u 8 habit. totales)	0,282
5HT (9 ó 10 habit. totales)	0,117
2HF (3 ó 4 habit. para miembros fam.)	0,273
4HT (7 u 8 habit. para miembros fam.)	0,167

Cuadro A-3. Contribuciones relativas del AFCM, en orden descendente

Variable	CR. en Eje 1
TV/Color	9,9
Auto	9,4
N-Hab- Fam (HF)	8,0
Lavadora	7,8
Calentador	7,5
N-Tot-Habit (HAB)	6,8
Cocina/Separ	6,3
Reproductor	6,3
Comed/Separ	5,4
Ducha	4,9
Sirvientes	4,7
WC	4,3
Bicicleta	3,5
Tipo vivienda	3,4
Radio	3,4
Tenencia/Vivienda	3,3
Nevera	3,1
Tipo/Cocina	1,1
Número	0,7
Moto	0,3
Agua	0,0

Cuadro A-4. Contribuciones relativas del AFCM, en orden descendente

Variable	CR En Eje 2
Tipo Vivienda	20,9
N-Tot-Habit (HAB)	18,8
N-Hab-Fam (HF)	15,0
WC	12,6
Ducha	12,0
Tenencia/Viv.	10,3
Número	4,9
Comed/Separ	1,5
Cocina/Separ	1,1
Calentador	1,0
Sirvientes	0,5
TV/Color	0,4
Nevera	0,3
Agua	0,2
Tipo/Cocina	0,2
Lavadora	0,1
Radio	0,1
Auto	0,1
Moto	0,1
Bicicleta	0,0
Reproductor	0,0

Cuadro A-5. Individuos extremos y sus valores en algunas variables del AFCM

Individuos proyectados a la derecha del gráfico										
Indiv	Ivc	Auto	Lav	Cal	Ducha	Wc	Nf	Hab	Tenencia	Nivel
064	SI	SI	SI	SI	U.E.	U.E.	4	4	Propia	5
135	SI	SI	SI	SI	U.E.	U.E.	4	4	Prop.Pag	5
140	SI	SI	SI	SI	U.E.	U.E.	5	5	Hipotecad	5
170	SI	SI	SI	SI	U.E.	U.E.	5	5	Prop.Pag	4
Individuos proyectados a la izquierda del gráfico										
046	NO	NO	NO	NO	U.E.	U.E.	1	1	Alquilada	1
080	NO	NO	NO	NO	U.C.	U.C.	1	1	Alquilada	1
110	NO	NO	SI	NO	U.E.	U.E.	2	2	Alquilada	2
118	SI	SI	NO	NO	U.E.	U.E.	2	2	Propia	2
128	NO	NO	NO	NO	U.E.	U.E.	2	2	Alquilada	4
195	SI	SI	SI	SI	U.C.	U.C.	4	6	Propia	1
204	NO	NO	NO	NO	U.C.	U.C.	4	6	Propia	3
211	NO	NO	NO	SI	U.C.	U.E.	3	5	Propia	5
212	NO	NO	NO	NO	U.C.	U.C.	3	3	Alquilada	1
213	NO	NO	NO	NO	U.C.	U.C.	2	2	Propia	1
215	NO	NO	NO	NO	U.C.	U.C.	2	5	Propia	2
218	SI	SI	NO	SI	U.C.	U.C.	4	4	Propia	3
219	NO	NO	NO	NO	U.C.	U.C.	3	3	Propia	3
339	NO	NO	NO	NO	U.E.	U.E.	1	1	Alquilada	5

* Notas: U.C. Uso Com N; U.E.: Uso Exclusivo; NF (Habitaciones para la fía) y Hab (Habitaciones totales):

1:1 2,2 = 3 4,3 = 5 6,4 = 7 u 8, 5:9 10, 6 =11 o M S; Nivel de ingresos 1: Bajo 2: Medio Bajo 3 : Medio 4: Medio-Alto, 5: Alto.