\documentclass{article}

\usepackage[letterpaper, margin=1in]{geometry}

\usepackage{caption}

\usepackage{float}

\usepackage{array}

\usepackage{subfig}

\usepackage{graphicx}

\usepackage[spanish]{babel}

\usepackage[latin1]{inputenc}

\title {Análisis Multivariado\\Taller No. 5: ACS de la TC manzanas de Bogotá según localidades y estratos}

\author {Laura Elizabeth Holguín\\Eduardo Contraras Bohorquez\\Diana Catalina Martínez}

\begin{document}

\SweaveOpts{concordance=TRUE}

\maketitle

\section{Comente la repartición de las manzanas según estratos (histograma de la distribución

de las manzanas en los 6 estratos - distribución marginal)}

<<fig=TRUE, echo=FALSE>>==

library(FactoClass)

data("Bogota")

tabb<-plotct(Bogota, "row", col=c("darkgreen","brown","orange","darkblue",

" gray","white","green"), tables=TRUE)

tabb2<-plotct(Bogota[,-1], "row", col=c("brown","orange","darkblue",

"gray","white","green"), tables=TRUE)

@

Según el histograma de perfiles por estrato, en la distribución marginal los estratos 2 y 3 son los que más manzanas tienen, es decir, geográficamente Bogotá tiende a concentrar habitantes de clase media-baja. En contraste, los estratos más altos, 4, 5 y 6, son los que están en menor proporción de manzanas. Las manzanas no estratificadas, por su parte, corresponden aproximadamente al 10\% del total de la información. El estrato con menor proporción de manzanas es el 5.

\section{?`Cómo es la distribución de las manzanas según localidades (distribución marginal)?}

<<fig=TRUE, echo=FALSE>>==

tabbc<-plotct(Bogota, "col", col=colors()[c(1,8,26,32,37,52,57,68,73,74,81,82,84,88,300,635,652,568,640)], tables=TRUE)

tabbc2<-plotct(Bogota[,-1], "col", col=colors()[c(1,8,26,32,37,52,57,68,73,74,81,82,84,88,300,635,652,568,640)], tables=TRUE)#Sin categoría NoSTR

E<-unclass(table(Bogota))

@

Las localidades con mayor proporción de manzanas, de acuerdo al histograma, son Ciudad Bolívar, Suba y Kennedy, respectivamente. En contraste, la Candelaria, Antonio Nariño y Los Mártires son las localidades con menor proporción de manzanas. En general, ninguna localidad sobrepasa el 12\% de dicha proporción.

\section{?`Utilizaría la columna sin estrato como activa en un análisis de correspondencias?;

?`Porque sí? ; ?`Por qué no?}

La columna sin estrato corresponde a manzanas que contienen áreas comunes, servicios sociales, entre otros, los cuales pueden estar relacionados con la estratificación de una zona. Por tanto, esta columna no sería utilizada como activa en un análisis de correspondencias, sino como variable ilustrativa.

\section{Compare la estadística chi2 asociada a la tabla de contingencia con la teórica. ?`Hay asociación entre estratos y localidades?}

<<echo=FALSE, results=hide>>==

B<-as.matrix(Bogota)

F <- B/sum(B) #P(carr\_i, estr\_i)

#Perfil fila

Dn <- diag(rowSums(B)) #Sum per carr

perfilFil <- solve(Dn)%\*%B #P(carr\_i, estr\_i | carr\_i)

#Perfil columna

Dp <- diag(colSums(B)) #Sum per estr

perfilCol <- B%\*%solve(Dp) #P(carr\_i, estr\_i | estr\_i)

Pn<- rowSums(B)/sum(B) #P(carr\_i)

Pp<- colSums(B)/sum(B) #P(estr\_i)

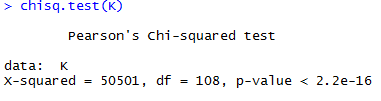
A <- Pn %\*% t(Pp) #P(carr\_i)P(estr\_i)

round(A,2)

Diff <- F-A

round(Diff,2)

@



Se realizó una prueba de Independencia X^2 de entre las variables estrato y localidad para los datos reportados en "Bogotá"; la cual arrojó evidencias de que dichas variables no se comportan de manera independiente rechazando Ho con un p valor de 2.2 exp -16 y un nivel de confianza del 95%.

De igual forma se realizó la comparación entre los estadísticos teóricos (media y varianza) y el estadístico calculado X2 los cuales dieron por resultado X=152,1, x=50500.99 X<x rechazando también Ho.

Estas variables pueden estar asociadas, pues existen factores como preferencia y capacidad adquisitiva que determinan la demanda de vivienda en las diferentes localidades que se caracterizan por su estrato predominante, así una población con nivel adquisitivo alto demanda vivienda estrato 6 y prefiere que sea en una localidad cuyo estrato dominante es el 6.

Una población con capacidad adquisitiva media alta aún puede elegir la zona de vivienda en una localidad específica con dominio de estratos 4 y 5, contrario a la población de capacidad media baja (estratos 2 y 3) que se encuentra ocupando localidades diversas de estrato.

Existe otro conjunto de localidades que están representadas por manzanas de estrato 1 y corresponden a aquellos lugares de fácil acceso para poblaciones con capacidad adquisitiva mínima, es decir cada localidad puede estar caracterizada por la dominancia de un estrato particular.

\section{?`Cuántos ejes retiene para el análisis? ?`Por qué?}

library(ade4)

acs <- dudi.coa(K, scannf=FALSE)

barplot (acs$eig,las=1)

inertia.dudi(acs)

Realizando un análisis de correspondencias simple sobre la tabla de contingencias, se encuentra que los tres primeros ejes acumulan el 85% de la inercia; además, la inercia promedio es de 0.21, estando solo los tres primeros ejes por encima de ésta, por lo tanto retenemos los tres primeros ejes para el análisis (OJO, ESTO SE HIZO INCLUYENDO LOS QUE NO TIENEN ESTRATO).

Si realizamos el análisis descartando la primera columna “Sin estrato” se acumula el 85,5% de la inercia en el tercer eje también, y la inercia promedio es 0,28 estando solo los dos primeros ejes por encima.

k<- K[,-1]

acs <- dudi.coa(k, scannf=FALSE)

barplot (acs$eig,las=1)

inertia.dudi(acs)

\section{Identifique en el primer eje las localidades más contributivas y sus oposiciones (localidades con coordenadas negativo sobre el eje vs. las de signo positivo).}

Para pintar sin estrato como de apoyo

Noestrato<- K[,1]

indicativa<-cor(Noestrato,k)

text ( indicativa , " Sin estrato " ,col = " darkgreen " ,pos =1 , cex =2)

dev. print ( device = xfig , file = "Bogota" )

Las localidades que más contribuyen al eje 1 son Barrios unidos, Suba y Fontibón a la izquierda y opuestos a estos se encuentran Bosa, Usme y Ciudad Bolivar.

\section{Identifique los estratos más contributivos al primer eje y sus oposiciones}

Los estratos 1, 2 y 3 son los que más contribuyen al primer eje siendo evidentemente opuestos 1 y 3, el estrato dos se encuentra representado muy cerca del origen.

\section{Repita 6 para el segundo eje}

Los estratos que más contribuyen al eje 2 se agrupan en la parte positiva de la recta Santafé, San Cristobal, Rafael Uribe, Candelaria, Kennedy, Puente aranda, Antonio Nariño, Los Mártires y Tunjuelito.

\section{Repita 7 para el segundo eje}

Los estratos 123 se encuentran opuestos a los estratos 456 según el segundo eje, el que más contribuye es el estrato 2.

\section{Resuma la comparación de los perfiles de las localidades utilizando el primer plano

Factorial.}

Los perfiles de las localidades en el primer plano factorial se aglutinan en tres grupos principalmente, la localidad que más se diferencia del promedio es Teusaquillo según cálculos de distancia euclidiana, las localidades más relacionadas con el primer eje según porcentaje de contribución son Usme y Ciudad Bolívar cuyas cuadras se caracterizan por ser opuestas a las localidades de Fontibón, Barrios Unidos, Puente Aranda y Antonio Nariño. La localidad que más se acerca al promedio según cálculos de distancia euclidiana es Santafé. Las localidades que más se relacionan con el segundo eje son Chapinero, Teusaquillo y Usaquén. Las demás localidades tienen manzanas diversas que seguramente contienen diferentes estratos en combinaciones variadas en proporción.

\section{Resuma la comparación de los perfiles de los estratos utilizando el primer plano

factorial.}

Se observa que los perfiles de los estratos se distribuyen ordenadamente desde el estrato 1 al 3 sobre el primer eje en el primer plano factorial, siendo opuestos principalmente estrato 1 y el 4. En el segundo eje se oponen los estratos 3 y 6 y en general el estrato más cercano al promedio es el 2 y el más lejano es el 6.

\section{Según el primer plano factorial, ?`cómo es la asociación entre localidades y estratos?}

Según el primer plano factorial existen localidades con estratos dominantes que las caracterizan y existen localidades con estratos mixtos. Se pueden observar principalmente tres grupos de localidades definidos por su estrato: estrato 1; estratos 2 y 3; estratos 4, 5 y 6; y unas localidades dispersas entre el estrato 3 y el 4.

Los estratos más altos agrupan a las localidades de Usaquén, Teusaquillo y Chapinero, los cuales se localizan geográficamente hacia el norte de la ciudad. El estrato 1 se localiza hacia el extremo sur-oriente en las localidades de Usme y Ciudad Bolívar, localidades menos homogéneas en estrato como Barrios Unidos, Fontibón y Suba y la masa de localidades de estrato 2 y 3 se localizan tan aglomeradas que podría decirse que comparten las características de distribución de manzanas en estratos.

\section{?`Hay efecto Guttman? Explique.}

De acuerdo al primer plano factorial, se puede observar que se presenta un efecto Guttman. El primer eje opone los estratos extremos, es decir los estratos altos (4, 5 y 6) contra estratos bajos (1 y 2); el eje 2 opone los estratos medios (2 y 3). Así mismo, se observa que los estratos se ordenan siguiendo la parábola y de forma secuencial, de derecha a izquierda aparecen los estratos de menor a mayor. De forma análoga, el orden de los estratos se traslada a las localidades siguiendo una parábola invertida: Usme y Ciudad Bolívar son las localidades que se encuentran en los estratos más bajos, mientras que Usaquén, Teusaquillo y Chapinero están en los estratos más altos.

\section{?`Hay contraposiciones en el tercer eje que no se observen en el primer plano factorial?

Según lo anterior, ?`vale la pena interpretar el tercer eje?}

En el eje 3 se puede observar que el estrato 4 se contrapone a los estratos 5 y 6, lo cual no era evidente en el primer plano factorial. Sin embargo, este eje no recoge suficiente inercia como para contemplarlo en el análisis.

Así mismo, de acuerdo a las tablas de inercia, la localidad Rafael Uribe no estaría bien representada en el eje tres.

\section{Agregue a los datos una columna de orden de las localidades según el primer plano

factorial. Ordene la TC por esa variable y haga una gráfica que muestre los perfiles

de las localidades así ordenadas y el perfil promedio. No incluya la columna sin

estrato. Resuma la comparación de los perfiles utilizando esta gráfica y el primer

plano factorial.}

Observando los perfiles ordenados de acuerdo al primer eje en la Figura 9, aquellos que cuentan con proporciones muy similares de estratificación, en el primer plano factorial se encuentran muy cercanos espacialmente. En particular, las localidades de Antonio Nariño, Mártires y Puente Aranda tienen una gran proporción del estrato 3, lo cual es similar a lo observado en el primer plano factorial. Así mismo, en las localidades de Usme y Ciudad Bolívar predomina el estrato 1, lo cual es visible en el primer plano factorial, pues éstas se encuentran en el lado positivo del eje 1 y están contiguas.

\section{Utilizando el primer plano factorial proponga una partición de las localidades en

cinco clases.}

De acuerdo a la distribución de las localidades en el primer plano factorial, la partición de las mismas sería así:

\textbf{Grupo 1:} Usme, Ciudad Bolivar

\textbf{Grupo 2:} Kennedy, La Candelaria, Tunjuelito, Rafael Uribe, San Cristóbal, Santafe, Bosa

\textbf{Grupo 3:} Puente Aranda, Antonio Nariño, Los Mártires, Engativá

\textbf{Grupo 4:} Fontibon, Suba, Barrios Unidos

\textbf{Grupo 5:} Usaquén, Teusaquillo, Chapinero

\section{Proponga una partición en cinco clases utilizando la gráfica de perfiles entre localidades.}

\textbf{Grupo 1:} Usme, Ciudad Bolivar

\textbf{Grupo 2:} Bosa, San Cristóbal, Santafe, Rafael Uribe, Tunjuelito, La Candelaria, Kennedy.

\textbf{Grupo 3:} Puente Aranda, Antonio Nariño, Los Mártires, Fontibon,Engativá

\textbf{Grupo 4:} Barrios Unidos, Teusaquillo

\textbf{Grupo 5:} Usaquén, Chapinero, Suba

\section{Proponga un partición ``final" de las localidades en cinco clases y haga una gráfica

de perfiles incluyendo el perfil marginal. Como otra síntesis del análisis comente la

gráfica obtenida.}

De acuerdo a la figura 10, las particiones 1, 2 y 3 coinciden con la ubicación de los tres primeros estratos. La partición 5 está asociada a los estratos más altos (4,5,6), mientras que la partición 4 contiene un poco disímiles pero cercanos espacialmente entre sí.

\section{Construya una TC más pequeña formando algunos grupos de localidades de perfiles

muy parecidos: Usme-Ciudad Bolívar, etc. (las dos filas se colapsan en una sumándolas

en la TC). Realice un ACS de esta tabla, compare los resultados de la tabla

completa. Comente la propiedad de equivalencia distribucional.}

\section{Compruebe ``a mano" (utilizando R) las relaciones de transición. Por ejemplo: calcule

la coordenada, sobre el eje 1, de Usme a partir de las coordenadas de los 6 estratos.

Analice el ejercicio (?`quién atrae a quién y porqué?).}

\section{Describa la distribución geográfica de los habitantes de Bogotá según su nivel socioecon

ómico, utilizando el estrato de la manzana donde vive cada uno como indicador

de ese nivel.}

\end{document}