

ANÁLISIS DE GRUPO (Cluster Analysis)

Por Tevni Grajales G.

Objetivo de la Técnica:

Esta técnica procura clasificar los objetos o cosas en categorías aún cuando se desconocen las características de los grupos. El objetivo es identificar grupos homogéneos (Clusters).

Criterios Básicos a Tomar en Cuenta

En primer lugar es importante determinar ¿cuáles variables serán usadas como base para agrupar? ¿de qué manera se medirán las distancias entre las cosas observadas? ¿qué criterio usar para combinar en grupo lo que se observa?. Esta etapa es crucial, pues de no considerar una variable importante para el asunto que se trata de resolver, puede llegarse a una resolución muy pobre del problema. Eso hace necesario que se disponga de un conocimiento teórico adecuado, con respecto al asunto que se trata.

Los conceptos fundamentales del análisis son:

- El concepto de distancia el cual determina cuánto se alejan dos objetos entre sí.
- El concepto de similitud que atiende a una medida de cercanía entre los objetos.

Concepto de Similitud y de Distancia

En el caso de la técnica de agrupamiento, debe recordarse que ésta se realiza a partir del concepto de cercanía (similitud) entre los objetos. Existen diferentes definiciones de distancia y similitud. En el ejemplo del anexo E relacionado con un estudio acerca de las esposas de pastores ASD en México, éstas difieren en su preparación académica, edad, número de hijos, etc. Cada persona puede diferir en cierta medida respecto al valor de una variable similar en otra persona. Las diferencias entre las personas y entre las variables necesitan ser combinadas de modo que se pueda disponer de un índice simple o medida de diferencia.

El índice más usado es El Cuadrado de la Distancia Euclidiana. Tiene la desventaja que depende de las unidades de medida usadas para las variables. Otra desventaja consiste en que cuando las medidas están en diferentes escalas, la variable con medidas mayores contribuye más, al computar las distancias, que la variable con medidas menores. La forma recomendada para superar estas desventajas consiste en transformar las medidas de las variables en medidas estandarizadas. (Distancia Mahalanobis D cuadrado)

Métodos para la Formación de Grupos

También existen diversos métodos para formar los grupos. El más común procede a agrupar las cosas u objetos en grupos cada vez más grandes hasta que lleguen a constituir uno sólo. Este es el método de Análisis Jerárquico Aglomerativo.

El proceso se inicia uniendo dos casos (sujetos, cosas) que son las más cercanas o similares entre si y luego se siguen añadiendo otros, paso a paso, según la cercanía o similitud respecto al grupo ya formado. Una vez que se identifica un caso con un grupo, el caso no puede ser transferido a otro grupo, sólo puede ser parte de otro grupo en el momento que su grupo original se integre para formar otro grupo mayor.

En el Anexo E aparecen seis variables que se observaron entre las esposas de pastores ASD de México. Según aparece en la página 85 del anexo, las variables amistad y aconsejar resultan las más cercanas y por consiguiente forman el grupo inicial: luego se añade al mismo grupo en un segundo paso la variable disciplina. En el tercer paso se constituye un nuevo grupo entre las variables indecisa y tensión. Después procede un cuarto paso en el que el grupo inicial que se inició entre amistad y aconsejar, se funde con la variable distorsión (lo que puede indicar que estas variables que constituyen el grupo son las que más se acercan a la forma como idealizan sus respuestas las esposas de pastores). En el quinto paso nuevamente aparece el segundo grupo donde aparece las variables tensión e indecisa, al cual se integra la variable migraña. Así puede seguirse el análisis del proceso según aparece graficado en las páginas 86 y 87 del mismo anexo.

El anexo E también presenta los resultados de otro agrupamiento relacionado con las esposas de los pastores. En este caso se agruparon las señoras según las medidas personales de las variables antes mencionadas. Según la gráfica adjunta, se nota que a la altura del antepenúltimo paso de agrupamiento, las 269 esposas incluidas se agrupaban en tres conglomerados claramente distinguidos.

Estudios posteriores de cada uno de los grupos por separado pueden conducir a caracterizar las personas que constituyen cada uno de los grupos. Según el marco teórico que sustenta el estudio.

Criterios para Asignar los Casos

Existen diversos criterios para decidir si un caso pertenece a un grupo. Estos criterios están basados en una matriz de distancia o similitud. Entre ellos:

La técnica del vecino más cercano: esta técnica combina o integra aquellos que tiene la distancia más pequeña entre sí; entonces se computa la distancia entre el grupo recién formado y cada uno de los otros casos fuera de grupo como la distancia mínima entre un caso individual y un caso perteneciente al grupo. Las distancias entre los casos que no pertenecen al grupo se mantienen sin alterar. En cada paso, la distancia entre dos grupos se define como la distancia que existe entre sus puntos más cercanos.

La técnica del vecino más lejano: la distancia entre dos grupos se define como la distancia entre sus dos puntos más lejanos.

La forma más recomendable para atender los detalles de este tipo de análisis consiste en detenerse a observar los resultados de una aplicación como aparece en el Anexo E. El docente ofrecerá una explicación detallada de cada una de las secciones del mencionado anexo E no se ausente de clase por favor.

Anexos disponibles por solicitud a tevgra@umontemorelos.edu.mx

[Altius](#)

tgrajales

©Tevni Grajales G.