Este método pretende dividir la muestra en 2^n segmentos, siendo n el número de variables explicativas.

Por ejemplo, imaginemos que ante la pregunta “¿Usted baila?” obtuviésemos las siguientes respuestas:

Sexo

Sí

%

No

%

Hombre

10

20,00%

40

80,00%

Mujer

30

60,00%

20

40,00%

Edad

Sí

%

No

%

Jóvenes (<30)

14

43,75%

18

56,25%

Adultos (de 30 a 50)

14

31,82%

30

68,18%

Mayores (>50)

12

50,00%

12

50,00%

El número de variables explicativas es 2: el sexo y la edad. Por lo tanto, el número de segmentos se tiene que reducir a 4 (2^2)

Pues bien, teniendo en cuenta que la media global de encuestados que bailan es del 40% (han contestado “Sí” 40 encuestados de 100), formaremos dos nuevos segmentos para la variable edad:

- Segmento de jóvenes y mayores: ya que sus respectivas medias (43,75% y 50,00%), superan la media global (40%).
- Segmento de adultos: ya que su media (31,82%) es inferior a la media global (40%).

De esta forma, los datos de los cuatro segmentos definitivos serían:

Sexo

Sí

%

No

%

Hombre

10

20,00%

40

80,00%

Mujer

30

60,00%

20

40,00%

Edad

Sí

%

No

%

Jóvenes y mayores

26

46,43%

30

53,57%

Adultos

14

31,82%

30

68,18%

Llegados a este punto, definimos el poder discriminante (D) de una variable explicativa, como la diferencia (en valor absoluto) entre el valor observado o real y su valor teórico (aplicando la media global de la muestra).

Cuanto mayor sea D, mayor será la capacidad explicativa de la variable. Si D fuese cero, la variable no tendría relación con la característica explicada.

En nuestro caso, para un estado igual a “Sí”, el poder discriminante de las dos variables se obtendría a partir de la siguiente fórmula:

D = | Sí (real) - Sí (teórico) |

Haciendo cálculos…..

Sexo

Sí (real)

Sí (teórico)

D = I Sí - Sí (teórico) I

Hombre

10

20

10

Mujer

30

20

10

Edad

Sí (real)

Sí (teórico)

D = I Sí - Sí (teórico) I

Jóvenes y mayores

26

22

4

Adultos

14

18

4

Así, por ejemplo, el dato real es que el número de hombres que han contestado “Sí” es de 10 (de los 50 entrevistados). Si esos 50 hombres hubieran contestado según la media global de toda la muestra (40%), un total de 20 hubieran dicho “Sí”. Este último, es el dato teórico.

Como podemos observar, el poder discriminante de la variable sexo es superior al de la variable edad, por lo que deberíamos segmentar en base al sexo.

Para ampliar vuestros conocimientos sobre este método, echad un vistazo al siguiente documento:

Analisis multivariable