

# Comparación de AC

Cristian Hernando González Runza  
Yosef Shmuel Guevara Salamanca  
Hermann Betancourth Romero

November 9, 2020

## Casos de dos variables

Para las variables carr y estr realice el ACS, el ACM y el AC de la tabla de Burt respectiva. Verifique que se cumplen cada una de las fórmulas dadas en Lebart, Piron Morineau (2006) para la comparación. Escriba la tabla 5.2 reemplazando las fórmulas por los resultados numéricos del ejercicio. Haga un resumen de la comparación. Responda las siguientes preguntas:

### Preguntas

1. En el ACS de la TC, las coordenadas del centro de gravedad de la nube de los perfiles fila son:

Dado que las coordenadas del centro de gravedad son el marginal de la tabla de frecuencias relativas (F) entre el estrato y la carrera, Las coordenadas del centro de gravedad para los perfiles fila es:

$$gp = [0.402, 0.416, 0.182]',$$

como se puede ver a continuación

	Ebajo	Emedio	Ealto	Suma
Biología	5.2	5.8	3.1	14.2
Estadística	6.5	6.5	1.8	14.8
Farmacia	6.7	8.1	1.6	16.4
Física	6.1	8.1	4.3	18.4
Geología	4.0	2.0	4.0	10.1
Matemáticas	4.7	5.6	1.6	11.9
Química	7.0	5.4	1.8	14.2
Suma	40.2	41.6	18.2	100.0

Figure 1: Tabla de Frecuencias relativas F

2. En el ACS, la nube está soportada en:(Dimensiones)

Al centrar la nube de puntos sobre el centro de gravedad el plano pierde una dimension tal que  $(\min(n,p)-1) = 2$ , es decir que la nube de puntos quedara soportada en R2 para este caso en particular esta conformado por las dimensiones Dim1 Y Dim2 que a la vez generan el primer plano factorial, por ende es posible comporar los perfiles sin perder información.

3. La inercia asociada al ACS es: ¿Qué significa?

La inercia asociada con el ACS es el 0.0656 y es el promedio del cuadrado de la suma de las diferencias entre la matriz frecuencia relativas (F) y la matriz de independencia(A).

El significado de la inercia asociada define que tan dispersos están los datos del centro de coordenadas del plano factorial sobre el cual se encuentran representado, pues la inercia mide las distancia de cada uno de los puntos de la nube al centro del plano, cuando la inercia es baja significa que una gran cantidad de datos se encuentran cercanos al centro por lo tanto cada punto es muy similar y cuando es alto indica una gran grandes diferencias de cada uno de los puntos.

4. En el ACS, el porcentaje de inercia retenido por el primer plano es: (porcentaje)

El porcentaje de inercia retenido por el primer plano factorial es el 100%, el 85.7 % (0.0562) para la dimensión 1 y el 14.3% (0.094) para la dimensión 2, tal como se a continuación.

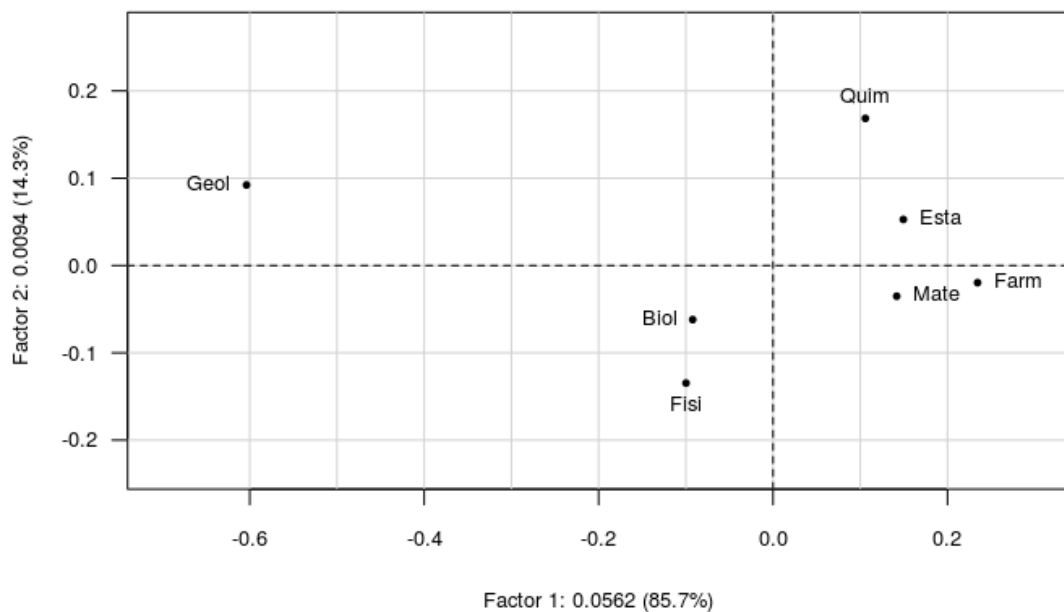


Figure 2: Primer plano Factorial ACS

5. En el ACM equivalente al ACS, la TDC tiene 445 filas(individuos) y 10 columnas(categorías), sabiendo que:.

$$TDC = matriz(n, s)$$

6. En el ACM de la TDC las nubes de “individuos” y de categorías tienen, respectivamente, 21 y 10 puntos, como se ve a continuación.

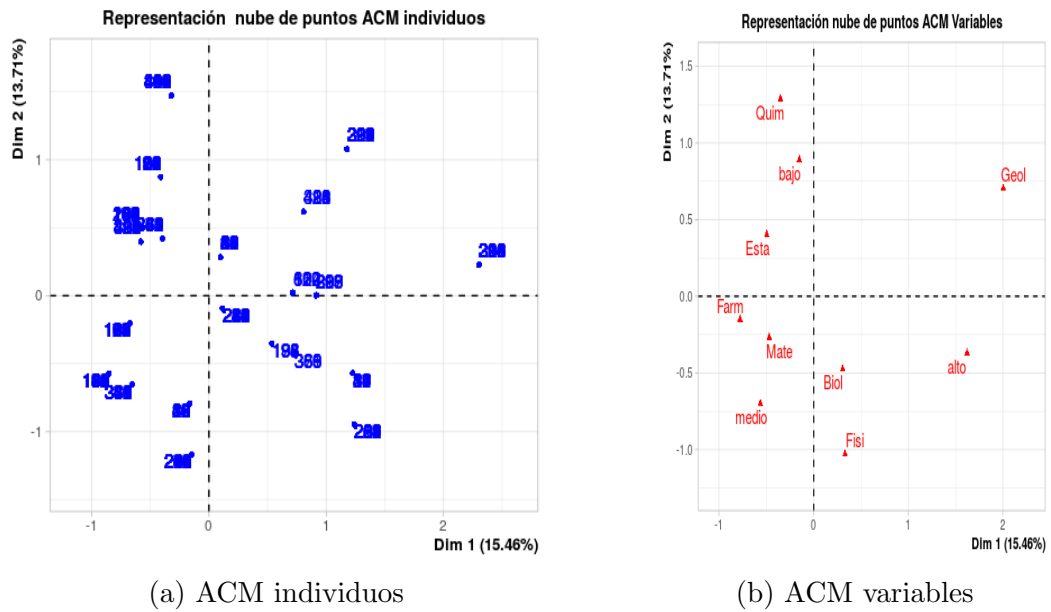


Figure 3: ACM individuos y variables

7. En este ACM las nubes están soportadas en 8 dimensiones.

Esto debido a que la nube de puntos esta soportada por el subespacio generado por la diferencia entre el numero de categorias y el numero de particiones definidas por variables ( $p - s$ ) sabiendo que:

$$p = 10 (Categorias)$$

$$s = 2 (Particiones definidas por las variables)$$

$$p - s = 10 - 2 = 8 (Dimensiones)$$

8. La inercia asociada al ACM es? y su significado estadístico es?

La inercia de la nube de puntos es 4 y está dada por:

$$\frac{p}{s} - 1$$

$$\frac{10}{2} - 1 = 4$$

Esta inercia no tiene ningún sentido estadístico, ya que depende del cociente entre el número de categorías y el número de variables, no de los valores internos de la tabla.

9. La tabla de Burt asociada a la TDC tiene 10 filas y 10 columnas y su total es 890, como se puede ver a continuación :

	carr.Biol	carr.Esta	carr.Farm	carr.Fisi	carr.Geol	carr.Mate	carr.Quim	estr.bajo	estr.medio	estr.alto
carr.Biol	63	0	0	0	0	0	0	23	26	14
carr.Esta	0	66	0	0	0	0	0	29	29	8
carr.Farm	0	0	73	0	0	0	0	30	36	7
carr.Fisi	0	0	0	82	0	0	0	27	36	19
carr.Geol	0	0	0	0	45	0	0	18	9	18
carr.Mate	0	0	0	0	0	53	0	21	25	7
carr.Quim	0	0	0	0	0	0	63	31	24	8
estr.bajo	23	29	30	27	18	21	31	179	0	0
estr.medio	26	29	36	36	9	25	24	0	185	0
estr.alto	14	8	7	19	18	7	8	0	0	81

Figure 4: Tabla de Burt

$$El\ total\ es\ (n \cdot s) = (445 \cdot 2) = 890$$

10. Según el criterio de Benzécri, en el ACM para el histograma del los , se deben tener en cuenta 6 ejes y el primer plano retiene el 79.17 % de la suma de los , como vemos a continuación.

Según el criterio de Benzécri solo deberemos analizar los ejes asociados a los valores asociados a la siguiente inecuación

$$\lambda z > \frac{1}{s}$$

donde:

$$\frac{1}{s} = \frac{1}{2} = 0.5$$

Mediante la tabla de valores propios:

	Valor p	%	% acum
dim 1	0.62	15.46	15.46
dim 2	0.55	13.71	29.17
dim 3	0.50	12.50	41.67
dim 4	0.50	12.50	54.17
dim 5	0.50	12.50	66.67
dim 6	0.50	12.50	79.17
dim 7	0.45	11.29	90.46
dim 8	0.38	9.54	100.00

Table 1: Tabla de Valores propios

Es decir, solo debemos analizar los ejes asociados a valores propios superiores a 0.5, en otras palabras la dimensión 1 (0.62), la dimensión 2 (0.55), la dimensión 3 (0.50), la dimensión 4 (0.50), la dimensión 5 (0.50) y la dimensión 6 (0.50) que retienen el 79.17 % .