

## Ejercicio 3

### Librerías

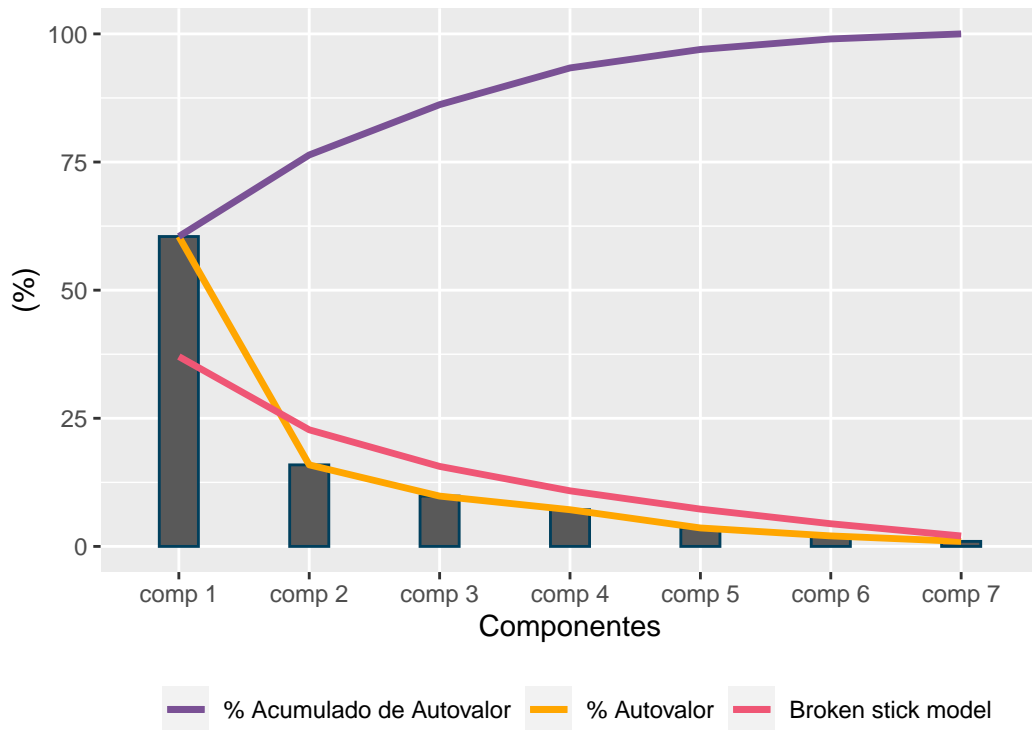
### Datos

En el archivo CUANTITATIVAS se presentan los datos correspondientes a la caracterización agronómica cuantitativa de las 41 variedades de pepino tratadas en el ejercicio anterior. Los caracteres cuantitativos analizados fueron: número de flores femeninas por nudo (FLORES), número de espinas en el ovario (ESPINAS), número de aristas (ARISTAS) y estrías (ESTRIAS) en fruto, longitud de fruto (FRUTO), intensidad de cuello (CUELLO) e intensidad de color de cuello (CCUELLO)

**A.** Aplique sobre estos datos un Análisis de Componentes Principales a partir de matriz de correlaciones.

Para analizar los datos y realizar el Análisis de Componentes Principales, se utilizó la librería “FactoMineR”. Como el objetivo era utilizar la matriz de correlación como base del análisis, se utilizó la opción de escalar los datos dentro del método de la librería. Cabe destacar que el paquete utiliza la noción francesa a la hora de computar las matrices (n, no n-1). A continuación se presentan los resultados:

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	4.23	60.47	60.47
comp 2	1.11	15.91	76.38
comp 3	0.69	9.82	86.20
comp 4	0.50	7.16	93.36
comp 5	0.25	3.60	96.96
comp 6	0.14	2.05	99.01
comp 7	0.07	0.99	100.00



Como puede observarse en la tabla, las dos primeras componentes acumulan el 76.38% del porcentaje de los autovalores. Tal como denota el gráfico, se observa que el criterio de “broken srick model” apunta a que las dos primeras componentes son suficientes para representar la variabilidad de los datos.

**B.** Realice la representación de las variedades en el plano principal, encuentre grupos y caracterícelos

A partir de la representación en dos dimensiones de los datos, se podría establecer que existen visualmente 3 grandes grupos de Pepinos. Los P/F, los A y los H. Esta separación coloca a las variedades Holandesas (H), en el cuadrante superior derecho. Este cuadrante está caracterizado por ser la dirección de crecimiento de las variables Cuello, Estrias, Frutos y Flores. En el cuadrante superior izquierdo se aloca las variedades Pepinillo y Francesas (P y F). Tal cuadrante queda caracterizado por ser la dirección de crecimiento de las variables Espinas, Aristas y Ccuello. Finalmente, en el cuadrante inferior, se sitúan las variedades Alpha-Beta (A).

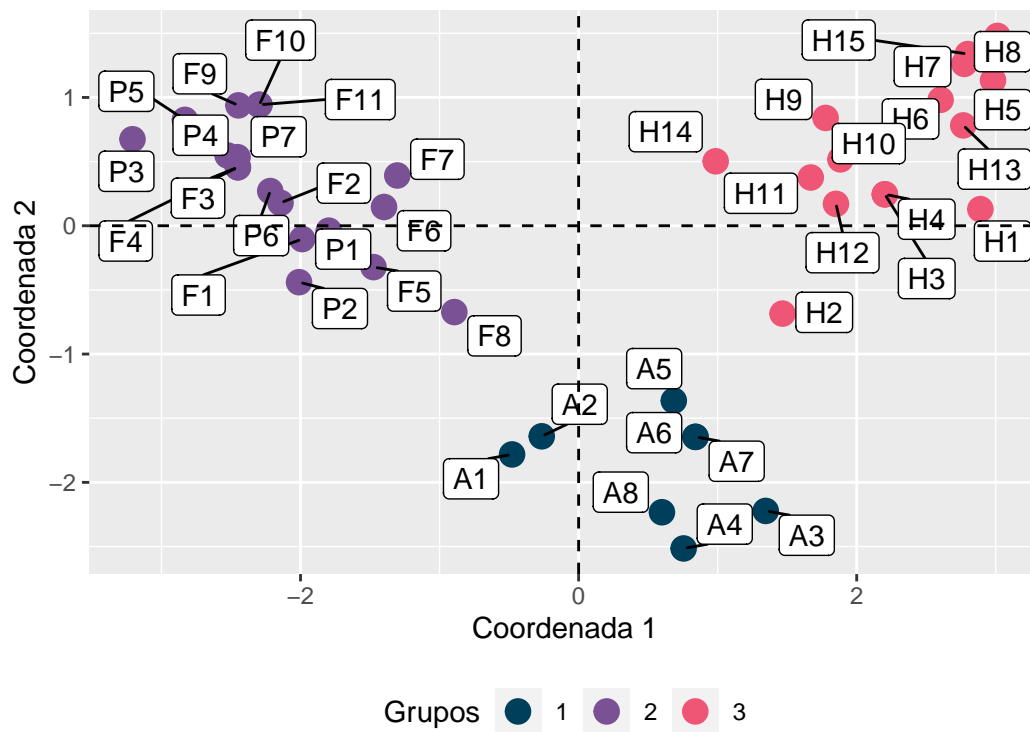


Figure 1: Caracterización cuantitativa de las variedades de pepino.

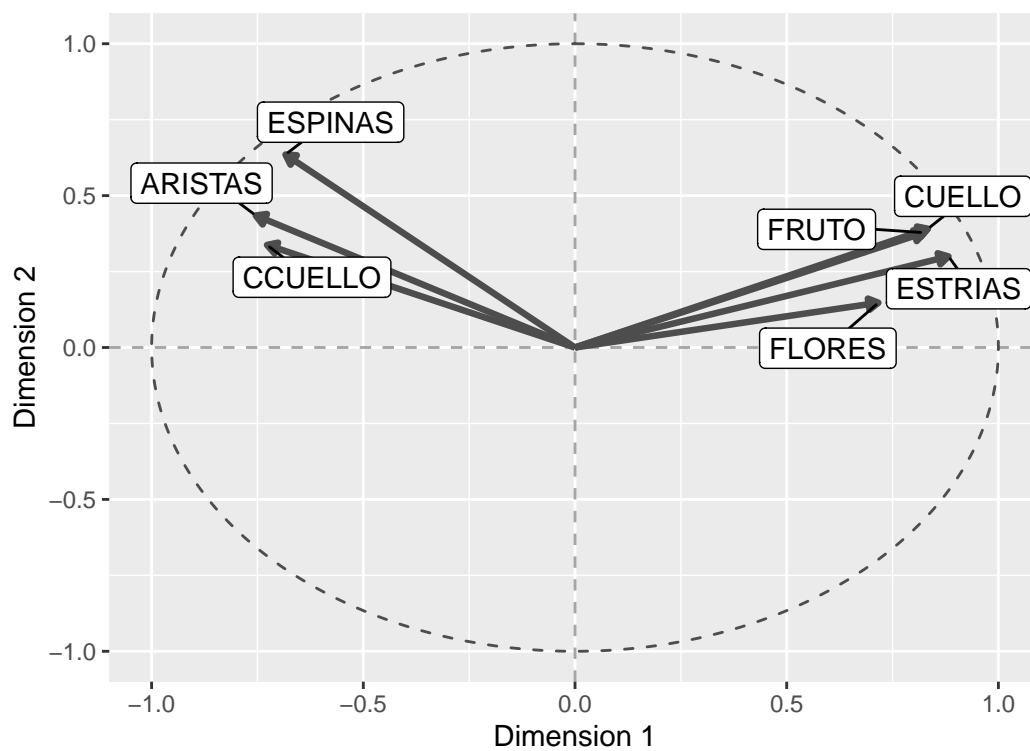


Figure 2: Gradiente de las Variables

C. Evalúe visualmente y a través de la correlación entre matrices la concordancia general entre esta configuración y la hallada en función de los caracteres cualitativos

Visualmente, se observa que la clusterización en base a componentes principales no logra separar las variedades P y F en dos dimensiones. En contra posición, se observa que con dos dimensiones, el análisis basado en las variables cualitativas logrará discriminar con notable exactitud las variedades y hasta colocarlas en cuadrantes opuestos.

Desde un punto de vista cuantitativo, si se calculan las matrices de distancias a partir de la representación en dos dimensiones para tipo de análisis y luego se toma la correlación entre los elementos de la diagonal subinferior, se observa que hay una concordancia del 0.7797612. Esto mismo puede corroborarse en el diagrama de dispersión de puntos.

```
## [1] 0.7797612
```

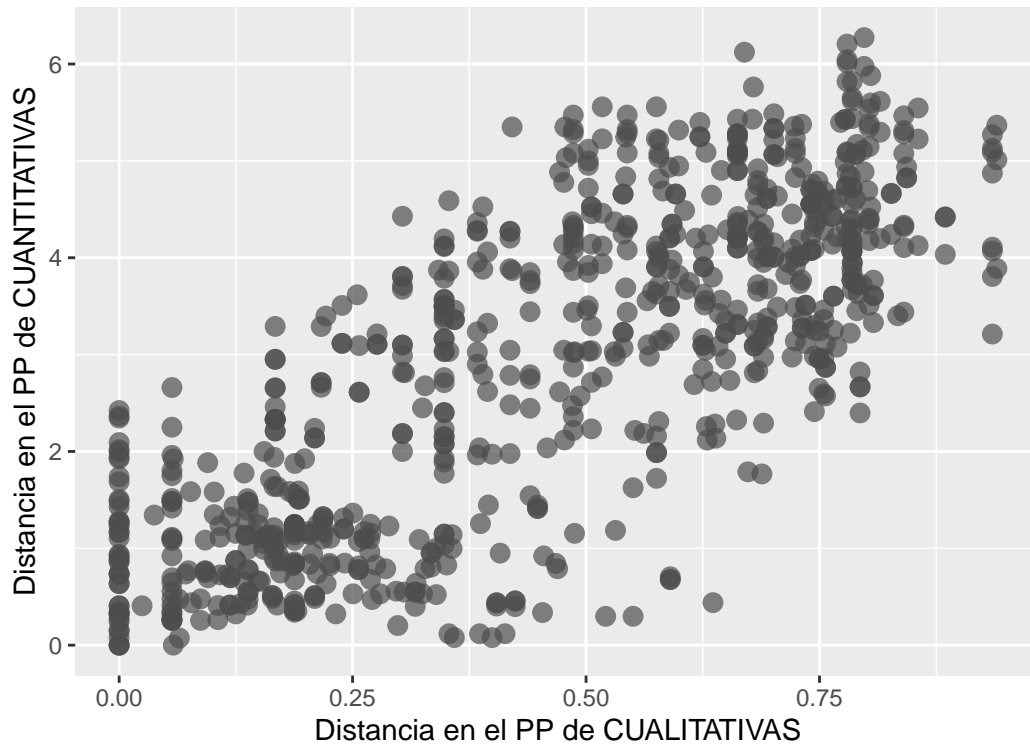


Figure 3: Distancia entre Matrices