


Laboratorio 3

Entrega: 3 de setiembre de 2020

Instrucciones: Trabaje las siguientes instrucciones en , creando uno o varios scripts. Entregue en la plataforma Mediación Virtual los scripts y genere un archivo Word o L^AT_EX con las respuestas. Puede entregar un archivo HTML usando Markdown, que contenga el script y las salidas.

1. Haremos la discretización de una variable cuantitativa, de dos formas diferentes:

- Cortándola a través de valores predeterminados.
- Cortándola de manera que las clases formadas tengan aproximadamente la misma cantidad de individuos.

a) La primera forma sería, en 3 clases (note la forma de los intervalos, abiertos a la izquierda y cerrado a la derecha):

```
> set.seed(654) # semilla aleatoria
> X <- rnorm(15, mean=0, sd=1) # genera un vector aleatorio normal
> Xcual <- cut(X, breaks=c(min(X)-1e-10, -0.2, 0.2, max(X)))
> Xcual # muestra los intervalos
```

b) Para la segunda forma especificaremos los cortes a través de los cuantiles:

```
> corte <- quantile(X, probs=seq(0,1,length=4))
> Xcual2 <- cut(X, breaks=corte, include.lowest=TRUE)
> table(Xcual2)
```

Obsérvese el efecto de la instrucción `include.lowest`.

2. Considere la tabla de datos de las Notas Escolares y la métrica identidad $\mathbf{M} = \mathbf{I}_5$.

- Calcule el centro de gravedad de la nube de puntos.
- Calcule la inercia respecto a Lucía, Andrés y Sonia.
- Calcule la inercia respecto al centro de gravedad, llamada simplemente *inercia*.
- Centre las variables.
- Recalcule el centro de gravedad y la inercia. ¿Hay algún cambio?
- Calcule la inercia con la métrica de los inversos de las varianzas $\mathbf{M} = \mathbf{D}_s = \text{diag}(1/\sigma^2)$.
- Calcule la inercia con la métrica de Mahalanobis $\mathbf{M} = \mathbf{V}^{-1}$ (inversa de la matriz de varianzas-covarianzas).

3. Considere de nuevo la tabla de datos de las Notas Escolares.

- Centre y estandarice las variables.
- Calcule la inercia con la métrica identidad. ¿Qué puede decir sobre este resultado?
- Encuentre la variable más correlacionada con todas las demás.

4. Retome los datos de las Notas Escolares, con las variables centradas y estandarizadas. Ahora consideramos que las materias tienen distintas ponderaciones: Matemática vale 4 puntos, Español 4, Ciencias 3, Historia 3, y Educación Física 1. Construya una métrica que refleje esta situación y calcule la inercia con esa métrica.