

Ejercicio 8 Guía 1

Agustín Muñoz González

20/4/2020

Los datos **TITANIC3** del paquete **PASWR2** contienen información sobre los pasajeros del Titanic, incluyendo clase, sexo y si sobrevivieron o no, entre otras características.

En primer lugar debemos cargar la librería **PASWR2** para poder acceder a los datos **TITANIC3**. Además definimos la variable **tit** con los datos de **TITANIC3** y ‘attachamos’ **tit** para que los datos sean más manejables.

```
library('PASWR2')
tit=TITANIC3
attach(tit)
```

a) Determine la proporción de sobrevivientes por clase.

Usaremos la función *mean()* que nos permite calcular la media de la lista que ingresemos. En este caso las listas serán

```
sobrevivientes_1st=tit[pclass=='1st','survived']
sobrevivientes_2nd=tit[pclass=='2nd','survived']
sobrevivientes_3rd=tit[pclass=='3rd','survived']
```

en la que filtramos los datos de cada clase y nos quedamos con la columna de sobrevivientes. Como los datos de sobrevivientes son 1 ó 0, calcular la media es calcular la proporción.

```
mean(sobrevivientes_1st) #prop 1st
```

```
## [1] 0.619195
```

```
mean(sobrevivientes_2nd) #prop 2nd
```

```
## [1] 0.4296029
```

```
mean(sobrevivientes_3rd) #prop 3rd
```

```
## [1] 0.2552891
```

Observamos que la proporción de muertos aumenta a medida que aumenta el número de la clase. En particular, las personas de la primera clase fueron los que más sobrevivieron y los de la tercera clase fueron los que más murieron, proporcionalmente.

b) Calcule la proporción de sobrevivientes por clase y sexo. ¿Quién tuvo una tasa más alta de supervivencia: los varones de 1ra clase o las mujeres de 3ra?

Nuevamente usaremos la función *mean()*, pero en este caso filtramos los datos por clase y sexo.

```
sobrevivientes_1st_male=tit[pclass=='1st' & sex == 'male','survived']
sobrevivientes_1st_female=tit[pclass=='1st' & sex == 'female','survived']

sobrevivientes_2nd_male=tit[pclass=='2nd' & sex == 'male','survived']
sobrevivientes_2nd_female=tit[pclass=='2nd' & sex == 'female','survived']
```

```
sobrevivientes_3rd_male=tit[pclass=='3rd' & sex == 'male', 'survived']
sobrevivientes_3rd_female=tit[pclass=='3rd' & sex == 'female', 'survived']
```

Tenemos las siguientes proporciones de supervivencia.

```
mean(sobrevivientes_1st_male) #prop 1st male
```

```
## [1] 0.3407821
```

```
mean(sobrevivientes_1st_female) #prop 1st female
```

```
## [1] 0.9652778
```

```
mean(sobrevivientes_2nd_male) #prop 2nd male
```

```
## [1] 0.1461988
```

```
mean(sobrevivientes_2nd_female) #prop 2nd female
```

```
## [1] 0.8867925
```

```
mean(sobrevivientes_3rd_male) #prop 3rd male
```

```
## [1] 0.1521298
```

```
mean(sobrevivientes_3rd_female) #prop 3rd female
```

```
## [1] 0.4907407
```

Notamos que las mujeres de la tercera clase tuvieron una tasa más alta de supervivencia que los varones de primera. Más aún, las mujeres de cada clase tuvieron una tasa de supervivencia muy superior a la de los hombres de la misma clase, de casi 3 veces más. Concretamente tenemos

```
mean(sobrevivientes_1st_female)/mean(sobrevivientes_1st_male) #prop 1st female/prop 1st male
```

```
## [1] 2.832536
```

```
mean(sobrevivientes_2nd_female)/mean(sobrevivientes_2nd_male) #prop 2nd female/prop 2nd male
```

```
## [1] 6.06566
```

```
mean(sobrevivientes_3rd_female)/mean(sobrevivientes_3rd_male) #prop 3rd female/prop 3rd male
```

```
## [1] 3.225802
```

c) ¿Cuál era la edad de la mujer más grande que sobrevivió?

Usaremos la función `max()` que devuelve el máximo elemento de la lista que ingresemos. En este caso la lista de entrada será la columna de edades de las mujeres que sobrevivieron. Como en la lista aparecen datos NA debemos removerlos.

```
max(tit[survived == 1 & sex=='female', 'age'], na.rm=T)
```

```
## [1] 76
```

Con lo cual observamos que la mujer más grande que sobrevivió tenía 76 años.

Para terminar ‘des-attachamos’ `tit` y borramos los registros.

```
detach(tit)
rm(list=ls())
```