

PW03

Equipo 22

1/29/2021

Postwork Sesión 3

Objetivo

- Realizar descarga de archivos desde internet
- Generar nuevos data frames
- Visualizar probabilidades estimadas con la ayuda de gráficas

Requisitos

1. R, RStudio
2. Haber realizado el prework y seguir el curso de los ejemplos de la sesión
3. Curiosidad por investigar nuevos tópicos y funciones de R

Desarrollo Ahora graficaremos probabilidades (estimadas) marginales y conjuntas para el número de goles que anotan en un partido el equipo de casa o el equipo visitante.

```
library(ggplot2)
data <- read.csv("temporadas.csv")
```

1. Con el último data frame obtenido en el postwork de la sesión 2, elabora tablas de frecuencias relativas para estimar las siguientes probabilidades: *La probabilidad (marginal) de que el equipo que juega en casa anote x goles (x=0,1,2,)

```
casa <- table(data$FTHG)
casa_pr_mar <- prop.table(casa)
casa_pr_mar
```

```
##
##           0           1           2           3           4           5
## 0.232456140 0.327192982 0.266666667 0.112280702 0.035087719 0.019298246
##           6           7           8
## 0.005263158 0.000877193 0.000877193
```

*La probabilidad (marginal) de que el equipo que juega como visitante anote y goles (y=0,1,2,)

```
visita <- table(data$HTAG)
visita_pr_mar <- prop.table(visita)
visita_pr_mar
```

```
##
##           0           1           2           3           4           5
## 0.619298246 0.292105263 0.077192982 0.009649123 0.000877193 0.000877193
```

*La probabilidad (conjunta) de que el equipo que juega en casa anote x goles y el equipo que juega como visitante anote y goles (x=0,1,2,, y=0,1,2,)

```
conjunta <- table(data$FTHG,data$HTAG)
conjunta_pr_con <- prop.table(conjunta)
conjunta_pr_con
```

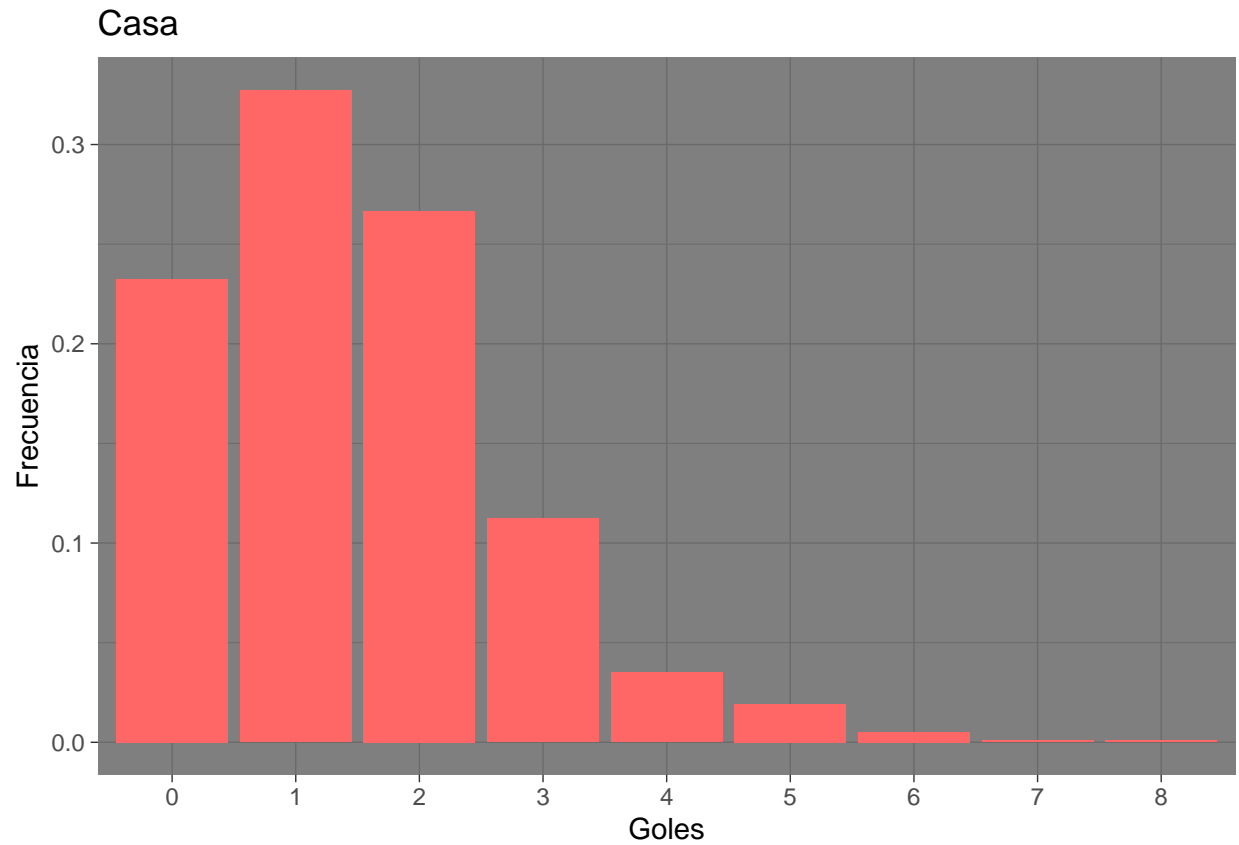
```
##
##           0           1           2           3           4           5
## 0 0.141228070 0.071052632 0.014912281 0.005263158 0.000000000 0.000000000
## 1 0.214035088 0.086842105 0.024561404 0.001754386 0.000000000 0.000000000
## 2 0.156140351 0.082456140 0.024561404 0.001754386 0.000877193 0.000877193
## 3 0.068421053 0.034210526 0.008771930 0.000877193 0.000000000 0.000000000
## 4 0.022807018 0.008771930 0.003508772 0.000000000 0.000000000 0.000000000
## 5 0.013157895 0.006140351 0.000000000 0.000000000 0.000000000 0.000000000
## 6 0.003508772 0.001754386 0.000000000 0.000000000 0.000000000 0.000000000
## 7 0.000000000 0.000877193 0.000000000 0.000000000 0.000000000 0.000000000
## 8 0.000000000 0.000000000 0.000877193 0.000000000 0.000000000 0.000000000
```

2. Realiza lo siguiente:

- Un gráfico de barras para las probabilidades marginales estimadas del número de goles que anota el equipo de casa

```
casa <- as.data.frame(casa_pr_mar)

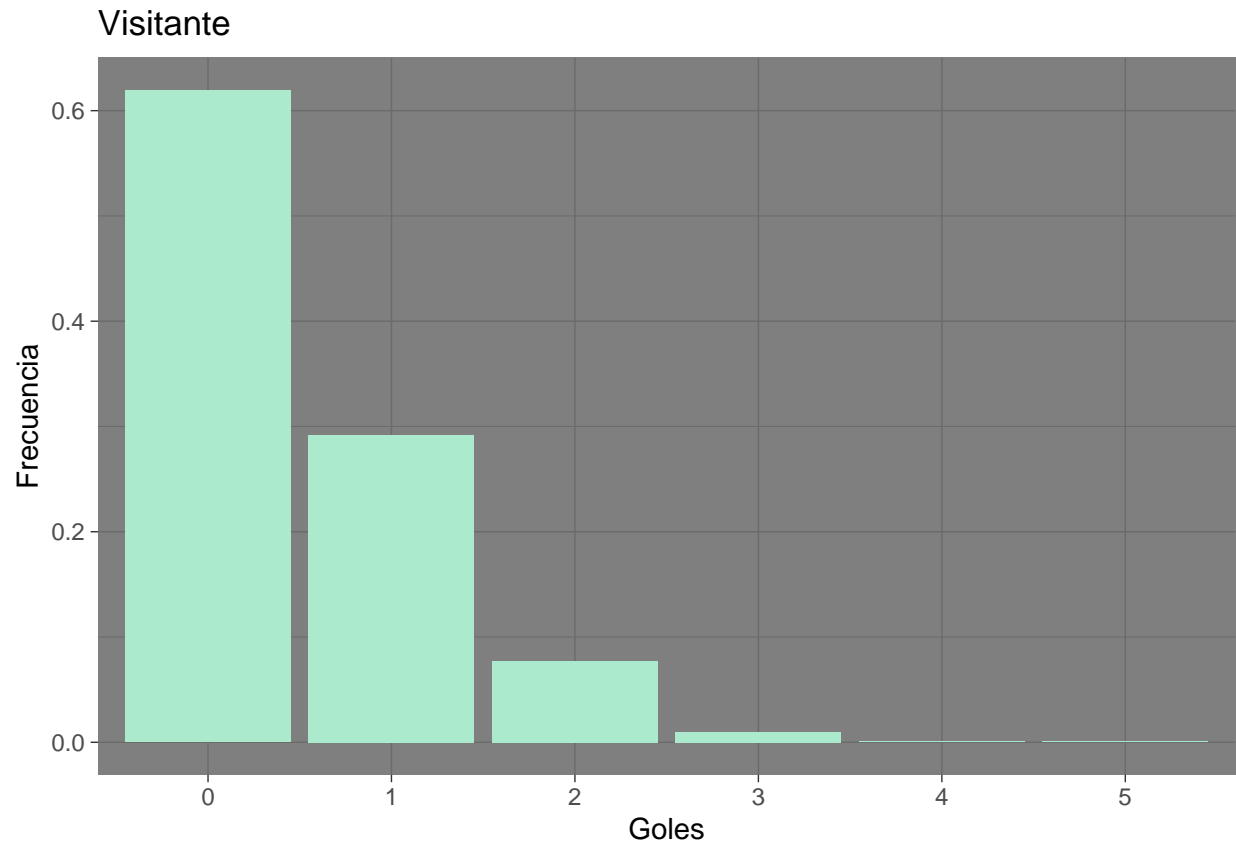
ggplot(data=casa, aes(x=Var1, y=Freq)) +
  geom_bar(fill = "#FF6666", stat="identity") +
  ggtitle("Casa") +
  ylab("Frecuencia") +
  xlab("Goles") +
  theme_dark()
```



- Un gráfico de barras para las probabilidades marginales estimadas del número de goles que anota el equipo visitante.

```
visitante <- as.data.frame(visita_pr_mar)

ggplot(data=visitante, aes(x=Var1, y=Freq)) +
  geom_bar(fill = "#ACEACE", stat="identity") +
  ggtitle("Visitante") +
  ylab("Frecuencia") +
  xlab("Goles") +
  theme_dark()
```



- Un HeatMap para las probabilidades conjuntas estimadas de los números de goles que anotan el equipo de casa y el equipo visitante en un partido.

```
prob_conjunta <- as.data.frame(conjunta_pr_con)

ggplot(data=prob_conjunta, aes(x=Var1, y=Var2, fill = Freq)) +
  geom_tile() + scale_fill_gradient(low = "white", high = "steelblue") +
  ggtitle("Probabilidad conjunta") +
  ylab("Visitante") +
  xlab("Casa") +
  theme_light()
```

