

A4-Regresión Poisson

A01275465 Carol Arrieta Moreno

2023-10-13

Trabajaremos con el paquete dataset, que incluye la base de datos warpbreaks, que contiene datos del hilo (yarn) para identificar cuáles variables predictoras afectan la ruptura de urdimbre.

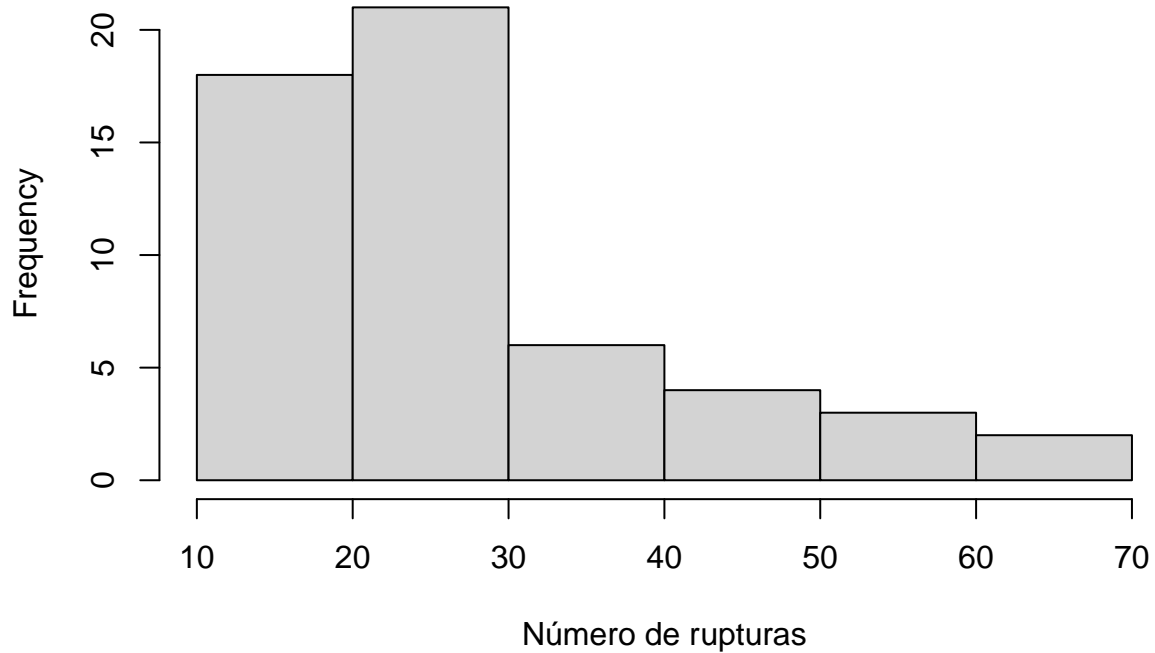
```
data<-warpbreaks  
head(data,10)
```

```
##      breaks wool tension  
## 1         26    A      L  
## 2         30    A      L  
## 3         54    A      L  
## 4         25    A      L  
## 5         70    A      L  
## 6         52    A      L  
## 7         51    A      L  
## 8         26    A      L  
## 9         67    A      L  
## 10        18    A      M
```

Histograma del número de rupturas

```
hist(data$breaks, main = "Histograma del número de rupturas", xlab = "Número de rupturas")
```

Histograma del número de rupturas



Obtén la media y la varianza

```
media <- mean(data$breaks)
varianza <- var(data$breaks)
cat("Media:", media, "\n")
```

```
## Media: 28.14815
```

```
cat("Varianza:", varianza, "\n")
```

```
## Varianza: 174.2041
```

Ajusta el modelo de regresión Poisson.

```
poisson.model <- glm(breaks ~ wool + tension, data, family = poisson(link = "log"))
summary(poisson.model)
```

```
##
## Call:
## glm(formula = breaks ~ wool + tension, family = poisson(link = "log"),
##      data = data)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
```

```

## (Intercept)  3.69196    0.04541  81.302  < 2e-16 ***
## woolB       -0.20599    0.05157  -3.994  6.49e-05 ***
## tensionM    -0.32132    0.06027  -5.332  9.73e-08 ***
## tensionH    -0.51849    0.06396  -8.107  5.21e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##    Null deviance: 297.37  on 53  degrees of freedom
## Residual deviance: 210.39  on 50  degrees of freedom
## AIC: 493.06
##
## Number of Fisher Scoring iterations: 4

```

Conclusión

El modelo que se generó es bueno, el modelo ajustado mejora sobre el modelo residual significativamente, las variables predictoras tienen p-valores muy bajos, el AIC es de 493.06 lo que indica que el modelo tiene buen equilibrio entre ajuste y complejidad.