

# A4-Regresión Poisson

Rodolfo Sandoval A01720253

2023-10-12

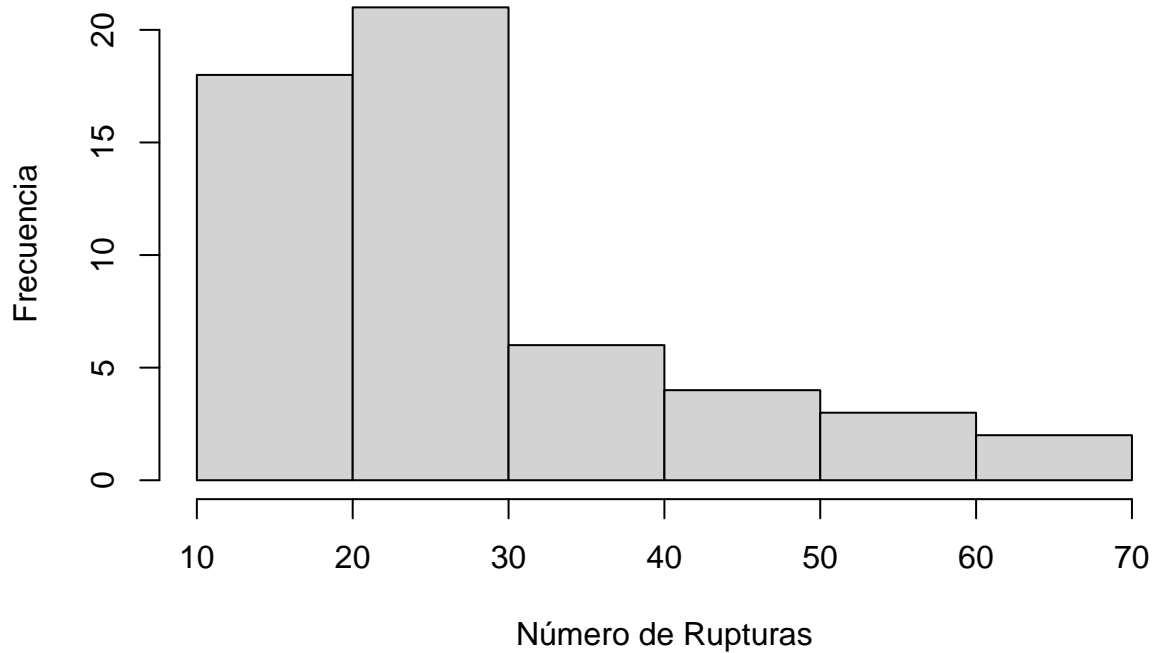
## R Markdown

```
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
```

```
hist(data$breaks, main = "Histograma de Rupturas", xlab = "Número de Rupturas", ylab = "Frecuencia")
```

## Histograma de Rupturas



```
# Calcular 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
```

```
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)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.6871  -1.6503  -0.4269   1.1902   4.2616
##
```

```
## 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
```

*# Valores p*

```
summary(poisson.model)$coefficients[, "Pr(>|z|)"]
```

```
## (Intercept)          woolB          tensionM          tensionH
## 0.000000e+00 6.489775e-05 9.728642e-08 5.209021e-16
```

*# Desviacion residual*

```
residual_deviance <- poisson.model$deviance
```

*# Grados de libertad*

```
df_residual <- poisson.model$df.residual
```

*# Ajustar un modelo cuasi Poisson*

```
quasipoisson_model <- glm(breaks ~ wool + tension, data, family = quasipoisson(link = "log"))
summary(quasipoisson_model)
```

```
##
## Call:
## glm(formula = breaks ~ wool + tension, family = quasipoisson(link = "log"),
##      data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.6871  -1.6503  -0.4269   1.1902   4.2616
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.69196    0.09374  39.384 < 2e-16 ***
## woolB       -0.20599    0.10646  -1.935 0.058673 .
## tensionM    -0.32132    0.12441  -2.583 0.012775 *
## tensionH    -0.51849    0.13203  -3.927 0.000264 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 4.261537)
##
```

```
##      Null deviance: 297.37  on 53  degrees of freedom
## Residual deviance: 210.39  on 50  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```
# Comparar los coeficientes
coef(poisson.model)
```

```
## (Intercept)      woolB      tensionM      tensionH
##   3.6919631  -0.2059884  -0.3213204  -0.5184885
```

```
coef(quasipoisson_model)
```

```
## (Intercept)      woolB      tensionM      tensionH
##   3.6919631  -0.2059884  -0.3213204  -0.5184885
```

Interpretacion y conclusiones de ambos modelos

Con el modelo de Cuasi Poisson, obtenemos que la media del numero de rupturas es de 28.14815 y la varianza del numero de rupturas es de 174.2041. Esto nos indica la dispersion de los datos. Estos resultados tambien indican que la varianza es significativamente mayor y sugiere sobredispersión. Tambien se puede concluir que todos los valores P son significativos para ambos modelos, es decir que todos los coeficientes tienen un alto impacto ya que los valores P son bajos. Con de acuerdo a los modelos, se puede interpretar que el modelo Poisson asume que la varianza de los errores es igual a la media. El modelo de Poisson proporciona la desviacion Residual. El cual evalua el ajuste del modelo. En este caso, la desviacion residual fue menor que la desviacion nula e interpreta un buen ajuste de modelo. El modelo de Cuasi Poisson permite que la varianza sea mayor a la media. Estos fueron los resultados que se plasmaron anteriormente donde se establecia que la media del numero de rupturas es de 28.14815 y la varianza del numero de rupturas es de 174.2041. Donde se indica la sobredispersión. Finalmente, podemos concluir que el modelo cuasi Poisson se utiliza cuando se sospecha que la varianza real es mayor que la media, lo que se conoce como sobredispersión. Este modelo permite estimar una varianza mayor que 1 y se adapta mejor a datos con este tipo de variabilidad. Y el modelo Poisson asume equidispersión, donde la varianza es igual a la media, y es apropiado cuando esta suposición se cumple. Ambos modelos tienen desviaciones residuales que son mayores que los grados de libertad, lo que sugiere que no hay una dispersion excesiva. Con esto podemos concluir que ambos modelos son adecuados para los datos analizados.