

# A4 - Regresión Poisson

Diego Alberto Baños Lopez  
A01275100

2023-10-13

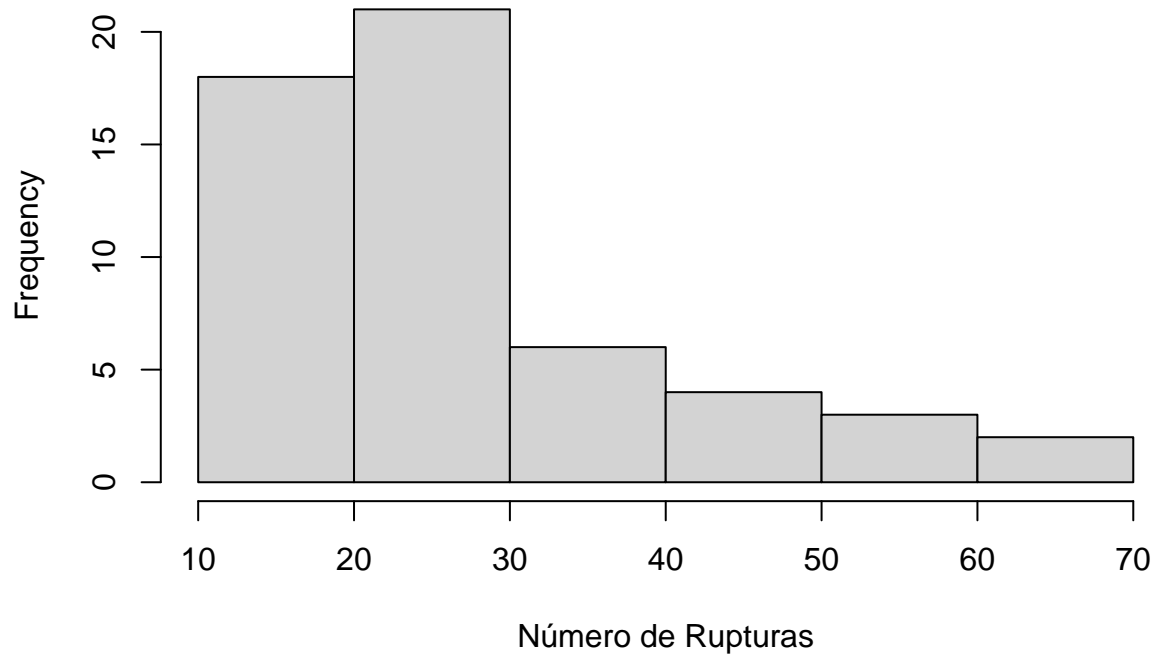
```
library(datasets)
data <- warpbreaks
head(data, 10)
```

```
##      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

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

## Histograma de Rupturas



### Media y varianza

```
mean_breaks <- mean(data$breaks)
var_breaks <- var(data$breaks)
print(mean_breaks)
```

```
## [1] 28.14815
```

```
print(var_breaks)
```

```
## [1] 174.2041
```

### Modelo Poisson

```
poisson_model <- glm(breaks ~ wool + tension,
  data = 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
```

## Interpretación del Modelo Poisson

En cuanto a este modelo los coeficientes estimados indican cómo las variables wool y tension afectan el número de rupturas de urdimbre:

- Tipo de lana (wool): La lana del tipo B está asociada con una disminución en el número de rupturas, con un coeficiente estimado de -0.206 ( $p < 0.001$ ). Esto sugiere que el tipo de lana B podría ser más resistente a las rupturas en comparación con la lana A (que es la categoría de referencia).
- Nivel de tensión (tension): Los niveles de tensión media y alta están asociados con disminuciones en el número de rupturas, con coeficientes estimados de -0.321 ( $p < 0.001$ ) y -0.518 ( $p < 0.001$ ), respectivamente. Esto indica que aumentar la tensión puede resultar en menos rupturas, aunque esto podría ser contraintuitivo y requeriría una investigación adicional.

La desviación residual es de 210.39 con 50 grados de libertad, lo que sugiere que hay una dispersión que no es capturada por el modelo. La desviación nula es de 297.37, lo que indica la desviación del modelo sin predictoras frente al modelo con predictoras.

## Modelo Cuasi Poisson

```
poisson_model2 <- glm(breaks ~ wool + tension,
  data = data, family = quasipoisson(link = "log")
)
summary(poisson_model2)
```

```
##
## 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
```

## Interpretación del Modelo Cuasi Poisson

El modelo Cuasi Poisson también revela una relación similar entre las variables predictoras y el número de rupturas, aunque con errores estándar más grandes debido a la dispersión adicional considerada:

- Tipo de lana (wool): La lana del tipo B sigue estando asociada con una disminución en el número de rupturas, aunque ahora el p-valor es 0.059, lo que sugiere que esta relación es menos significativa en el modelo Cuasi Poisson.
- Nivel de tensión (tension): Los niveles de tensión media y alta siguen estando asociados con disminuciones en el número de rupturas, con p-valores significativos.

El parámetro de dispersión para la familia quasipoisson es 4.2615, lo que indica una dispersión mayor que la suposición de Poisson. Esta dispersión adicional es una indicación de la variabilidad extra no capturada por el modelo Poisson, lo que justifica la consideración del modelo Cuasi Poisson.

## Evaluación de la Dispersión

```
res_dev <- poisson_model$deviance
df <- poisson_model$df.residual
print(res_dev)
```

```
## [1] 210.3919
```

```
print(df)
```

```
## [1] 50
```

```
res_dev <- poisson_model2$deviance
df <- poisson_model2$df.residual
print(res_dev)
```

```
## [1] 210.3919
```

```
print(df)
```

```
## [1] 50
```

La desviación residual y los grados de libertad son los mismos en ambos modelos, esto sugiere que hay una dispersión no capturada por los modelos

## Evaluación de la Desviación Nula Excesiva

```
null_dev <- poisson_model$null.deviance
print(null_dev)
```

```
## [1] 297.3722
```

```
null_dev <- poisson_model2$null.deviance
print(null_dev)
```

```
## [1] 297.3722
```