

# Prueba de Kolmogorov - Smirnov



# Recordatorio

- ▶ Prueba de bondad de ajuste para variables absolutamente continuas
- ▶ Dada una muestra aleatoria  $x_1, \dots, x_n$  se busca saber si los datos vienen de  $F_0(X)$
- ▶ La prueba utiliza la distribución empírica

$$F_n(x) = \begin{cases} 0 & \text{si } x < x_{(1)} \\ \frac{i}{n} & \text{si } x_{(i)} \leq x < x_{(i+1)} \\ 1 & \text{si } x > x_{(n)} \end{cases}$$

- ▶ Si  $H_0 : F(x) = F_0(x)$  es cierta, ent.  $\sup_{x \in \mathbb{R}} | F_n(x) - F(x) | \rightarrow 0$  c.s. (Glivenko-Cantelli)

# Recordatorio

- ▶ La idea entonces es considerar como estadístico a

$$D_n = \sup_X | F_n(x) - F(x) |$$

- ▶ Se calcula como  $D_n = \max(D_n^+, D_n^-)$  donde

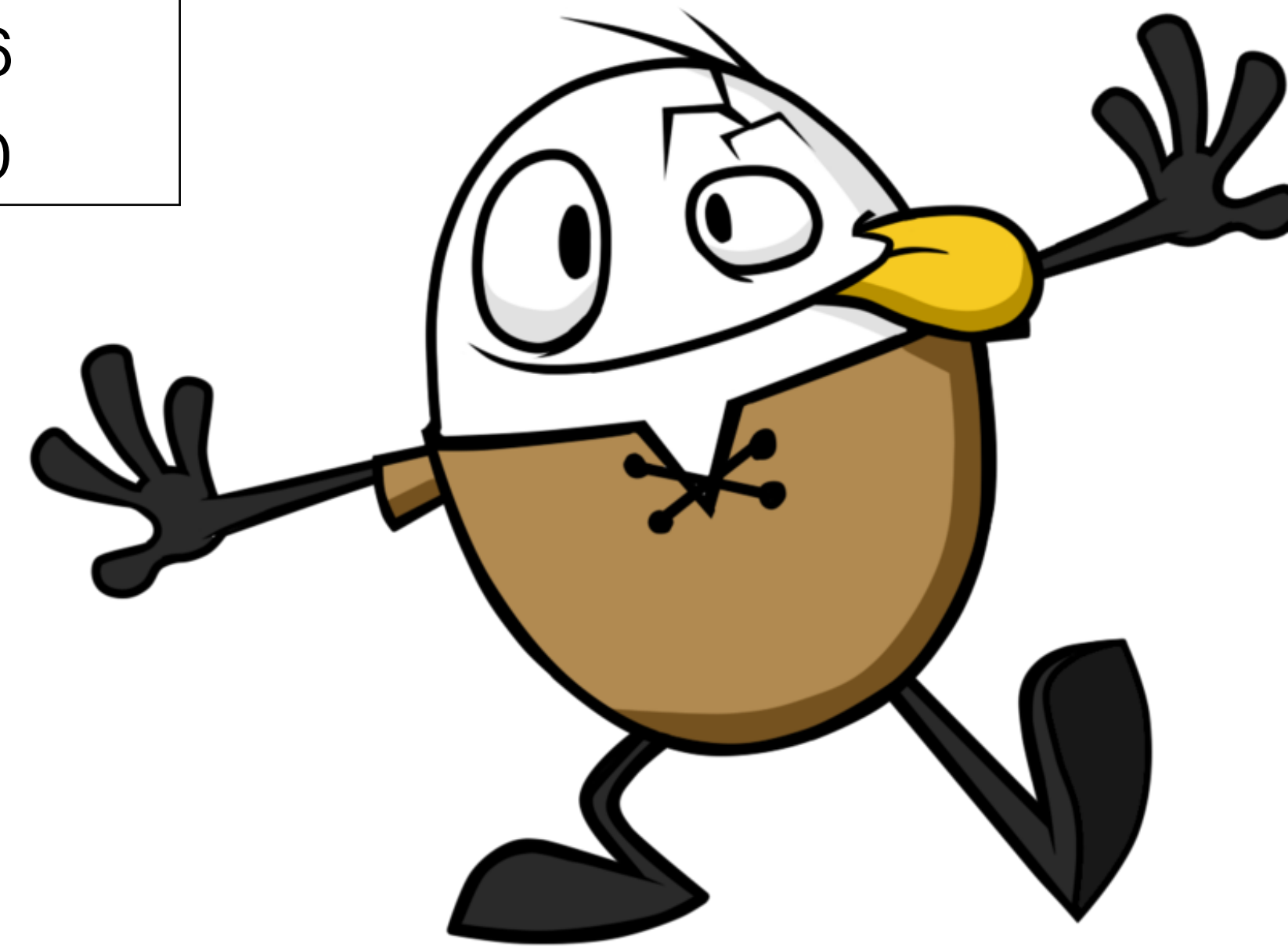
- $D_n^+ = \max \left( \frac{i}{n} - F(x_{(i)}) \right)$

- $D_n^- = \max \left( F(x_{(i)}) - \frac{i-1}{n} \right)$

# Huevos de gallina

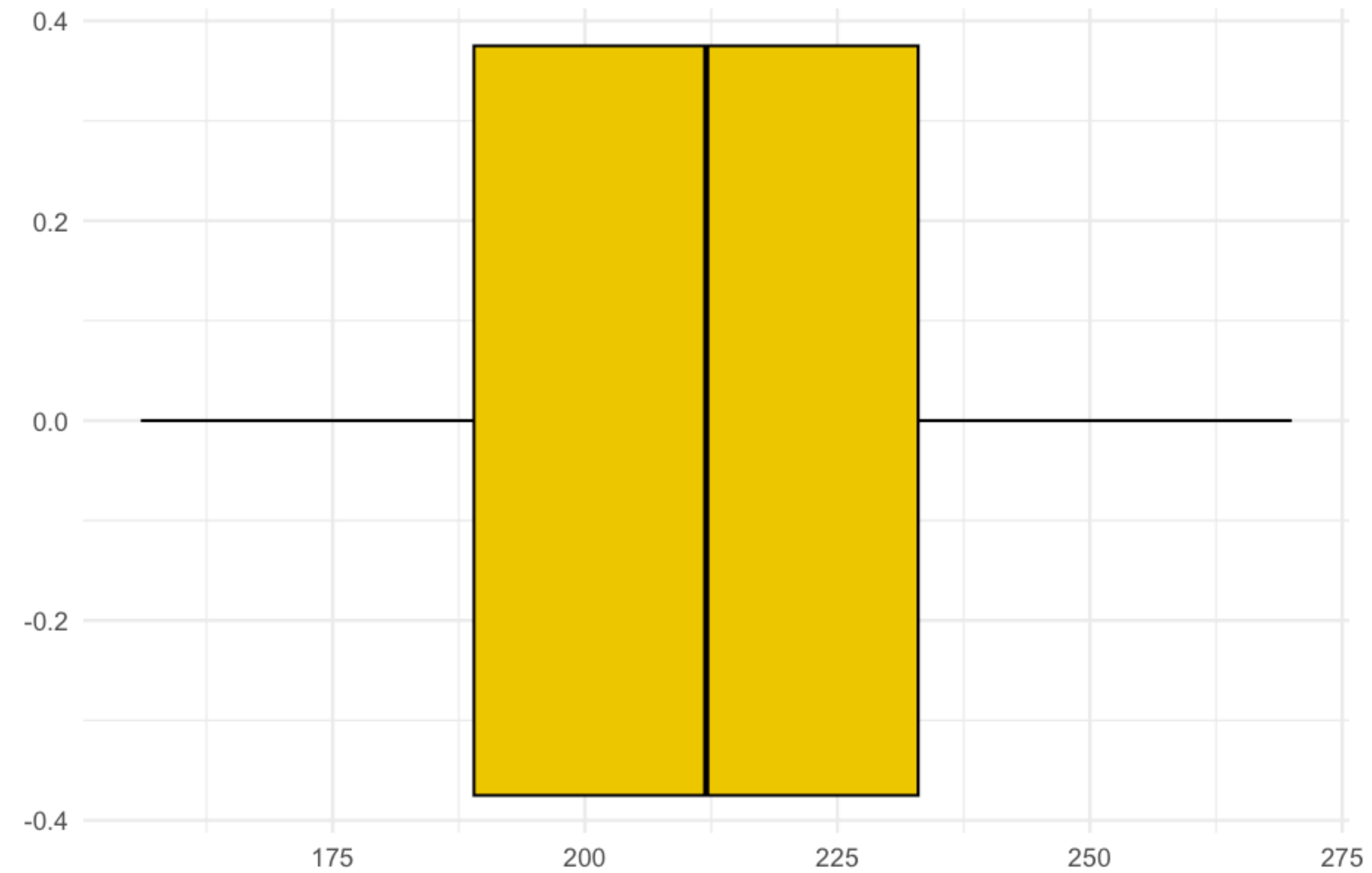
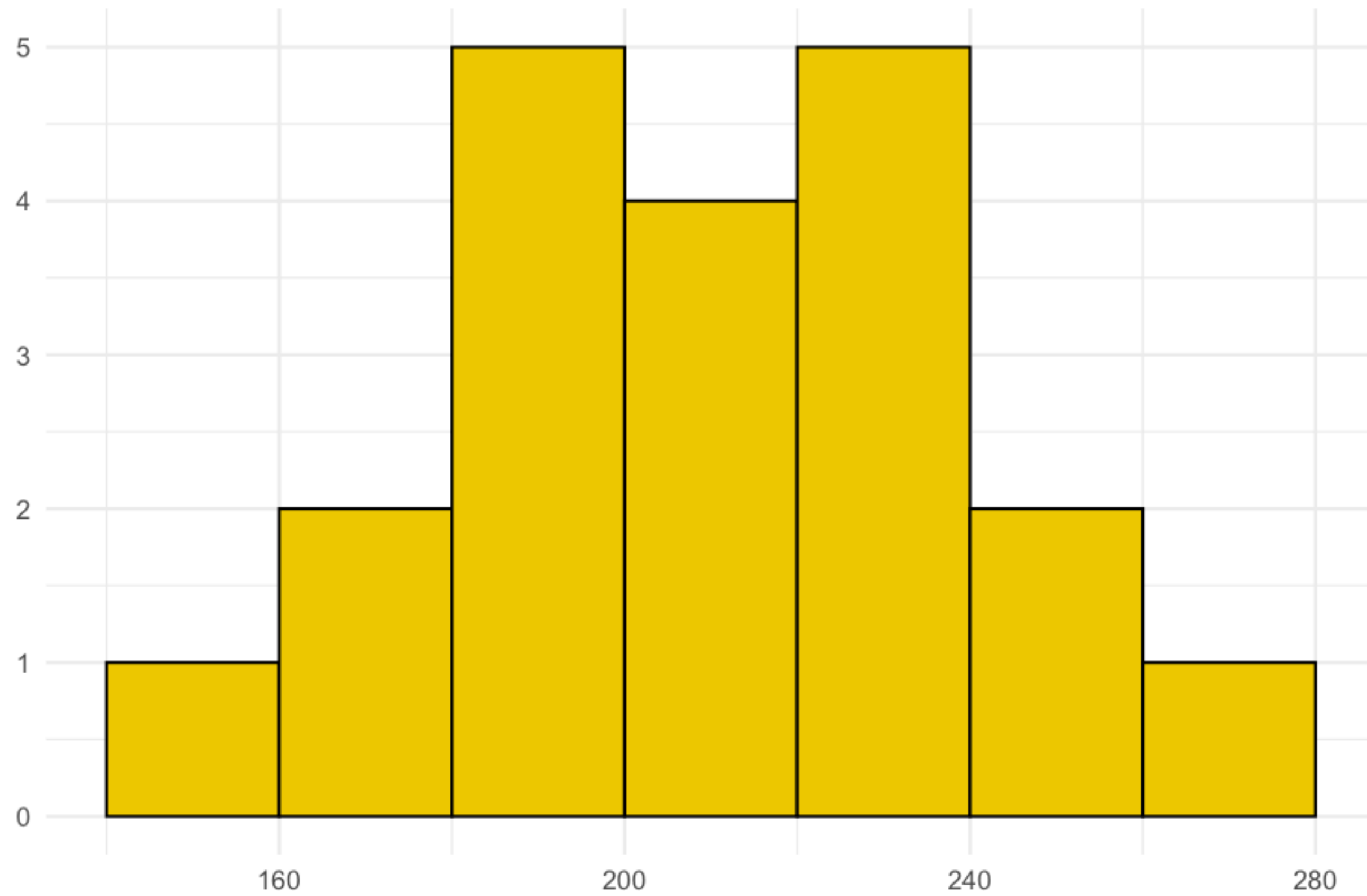
- Se busca saber si el peso de 20 huevos de gallina siguen una distribución normal de media  $\mu = 200$  y varianza  $\sigma^2 = 1225$

156	162	168	182
186	190	193	196
202	210	214	220
226	230	232	236
238	242	246	270



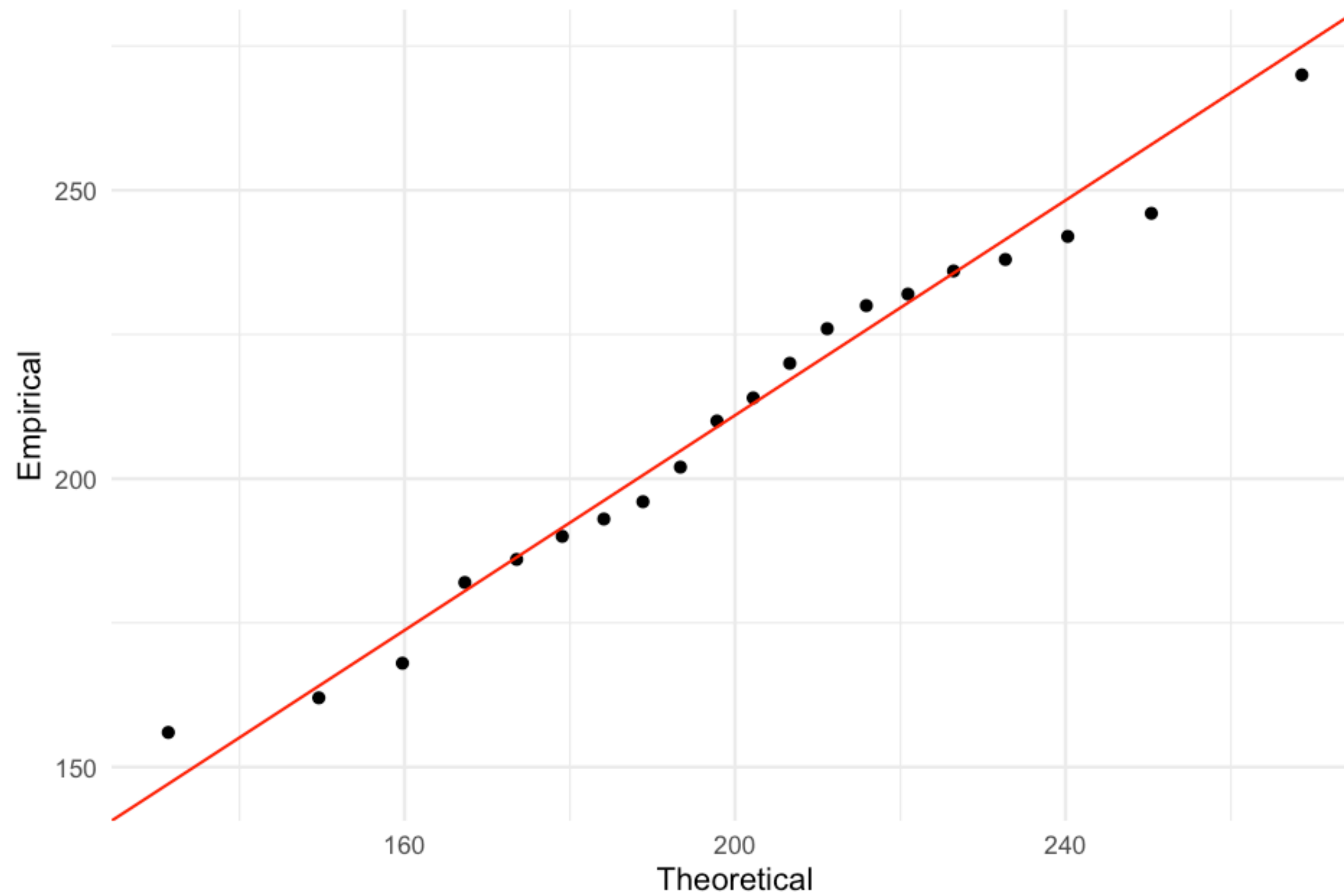
# Huevos de gallina

- ▶ Histograma y boxplot



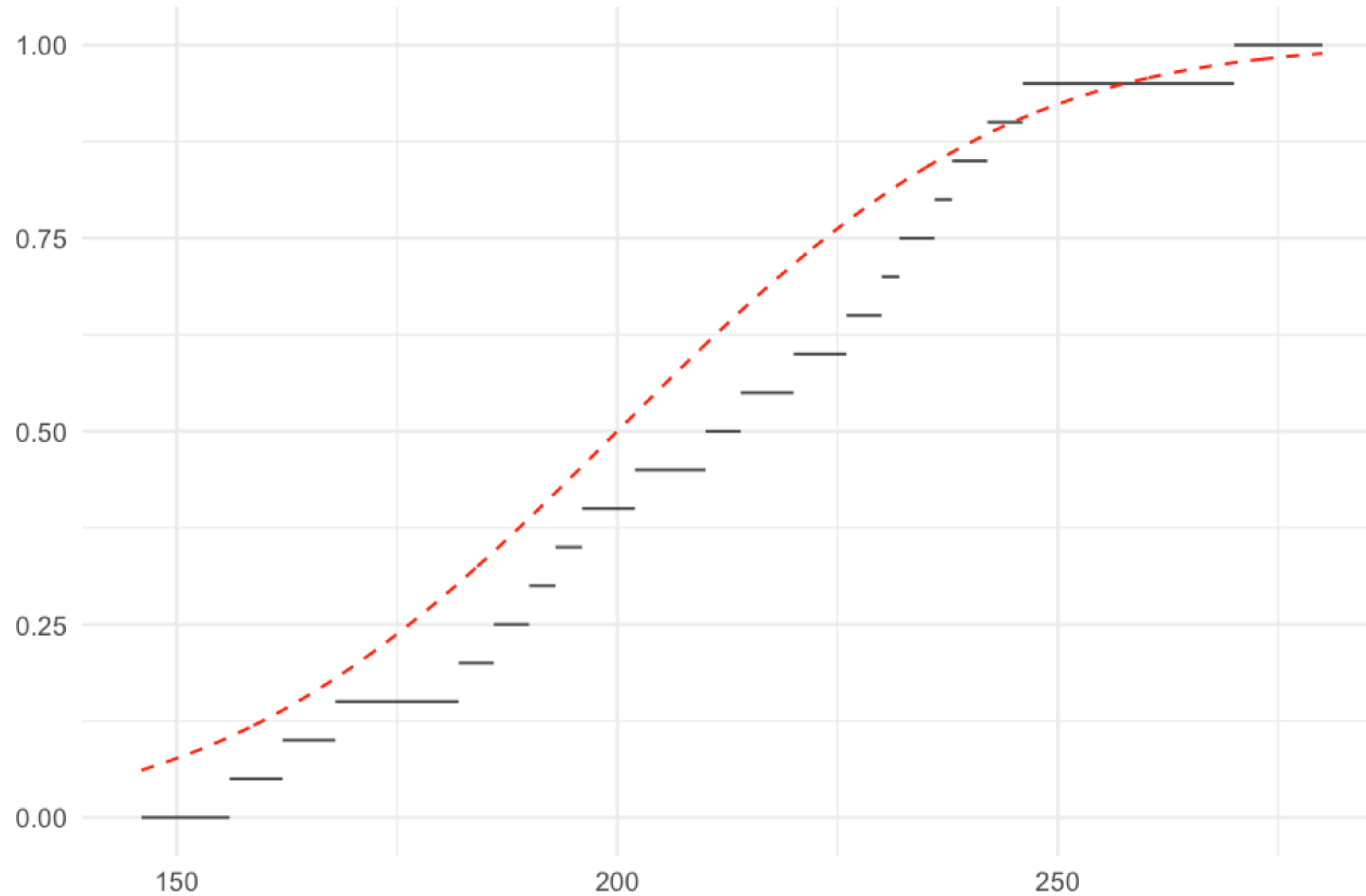
# Huevos de gallina

- qqplot



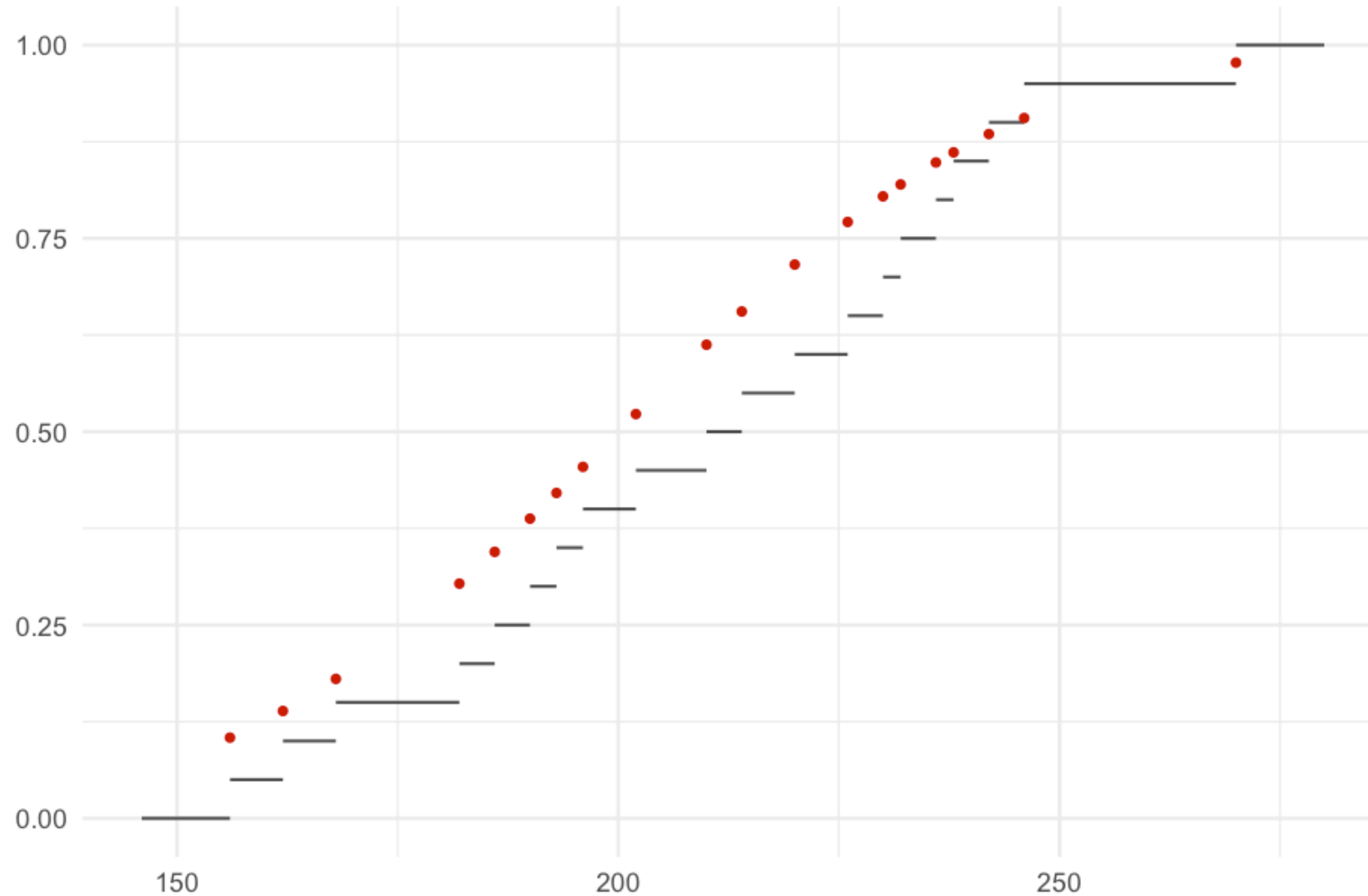
# Huevos de gallina

- ▶ Función empírica vs teórica



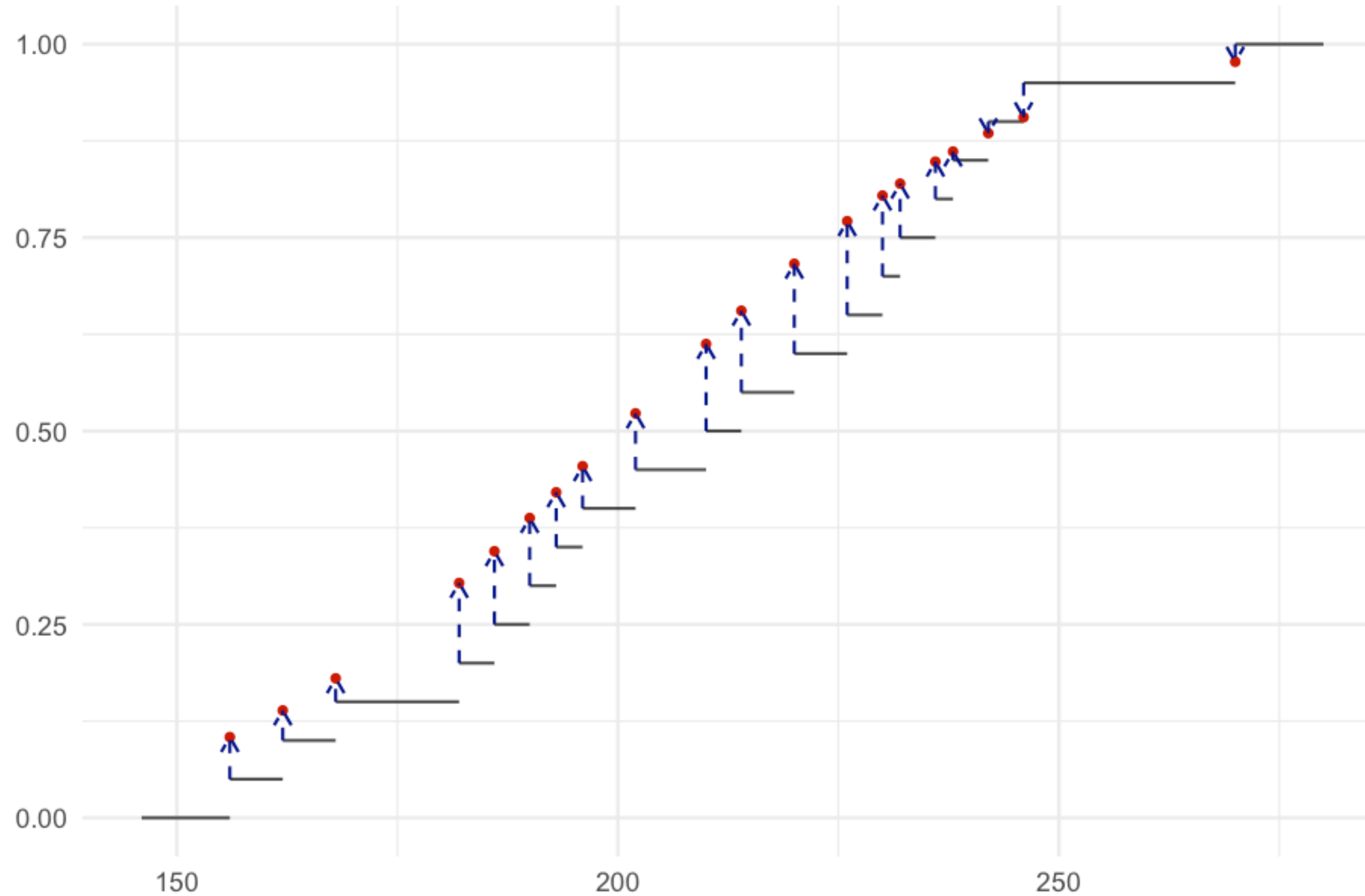
# Huevos de gallina

- ▶ Para la prueba nos fijamos en la función evaluada en la muestra (ordenada)



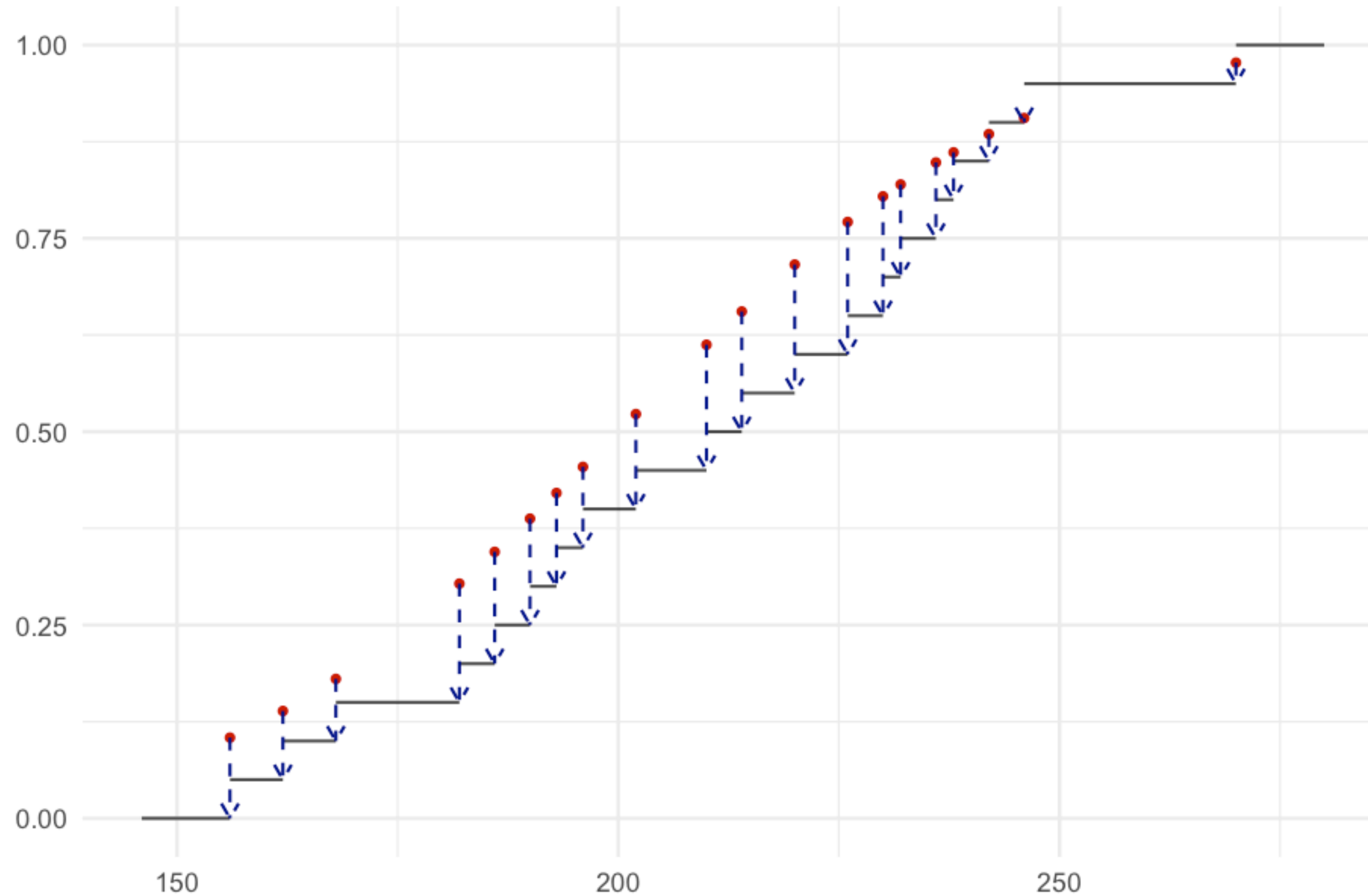
# Huevos de gallina

- ▶ Calcular  $D_n^+$



# Huevos de gallina

- ▶ Calcular  $D_n^-$



# Huevos de gallina

$i$	$x_{(i)}$	$F(x_{(i)})$	$i/n - F(x_{(i)})$	$F(x_{(i)}) - (i-1)/n$
1	156	0.1043	-0.0543	0.1043
2	162	0.1388	-0.0388	0.0888
3	168	0.1802	-0.0302	0.0802
4	182	0.3035	-0.1035	0.1535
5	186	0.3445	-0.0945	0.1445
6	190	0.3875	-0.0875	0.1375
7	190	0.3875	-0.0375	0.0875
8	196	0.4545	-0.0545	0.1045
9	202	0.5227	-0.0727	0.1227
10	210	0.6124	-0.1124	0.1624
11	214	0.6554	-0.1054	0.1554
12	220	0.7161	-0.1161	0.1661
13	226	0.7712	-0.1212	0.1712
14	230	0.8043	-0.1043	0.1543
15	230	0.8043	-0.0543	0.1043
16	236	0.8481	-0.0481	0.0981
17	236	0.8481	0.0018	0.0481
18	242	0.8849	0.0150	0.0349
19	246	0.9056	0.0443	0.0056
20	270	0.9772	0.0227	0.0272

▸  $D_n^+ = 0.0443$

▸  $D_n^- = 0.1712$

▸  $D_n = 0.1712$

# Huevos de gallina

- ▶ En R: `ks.test(weights, 'pnorm', 200, sqrt(1225))`
  - Prueba exacta de una muestra de Kolmogorov-Smirnov
  - $D = 0.17122$
  - p-value = 0.5433
  
- ▶ No rechazamos  $H_0$