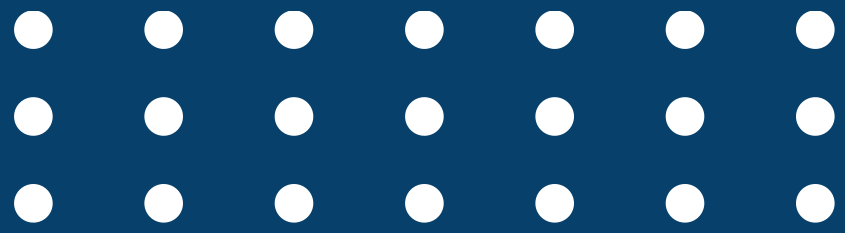


# Modelado de Tiempos de Espera con Distribución de Erlang

Por Mario Pérez





# Objetivo

Modelar tiempos de procesamiento de solicitudes en un centro de atención al cliente con Distribución de Erlang

# Dataset chickwts

Este dataset contiene información sobre el peso de los pollos en diferentes tipos de alimentación, que se interpretarán como tiempos de crecimiento.



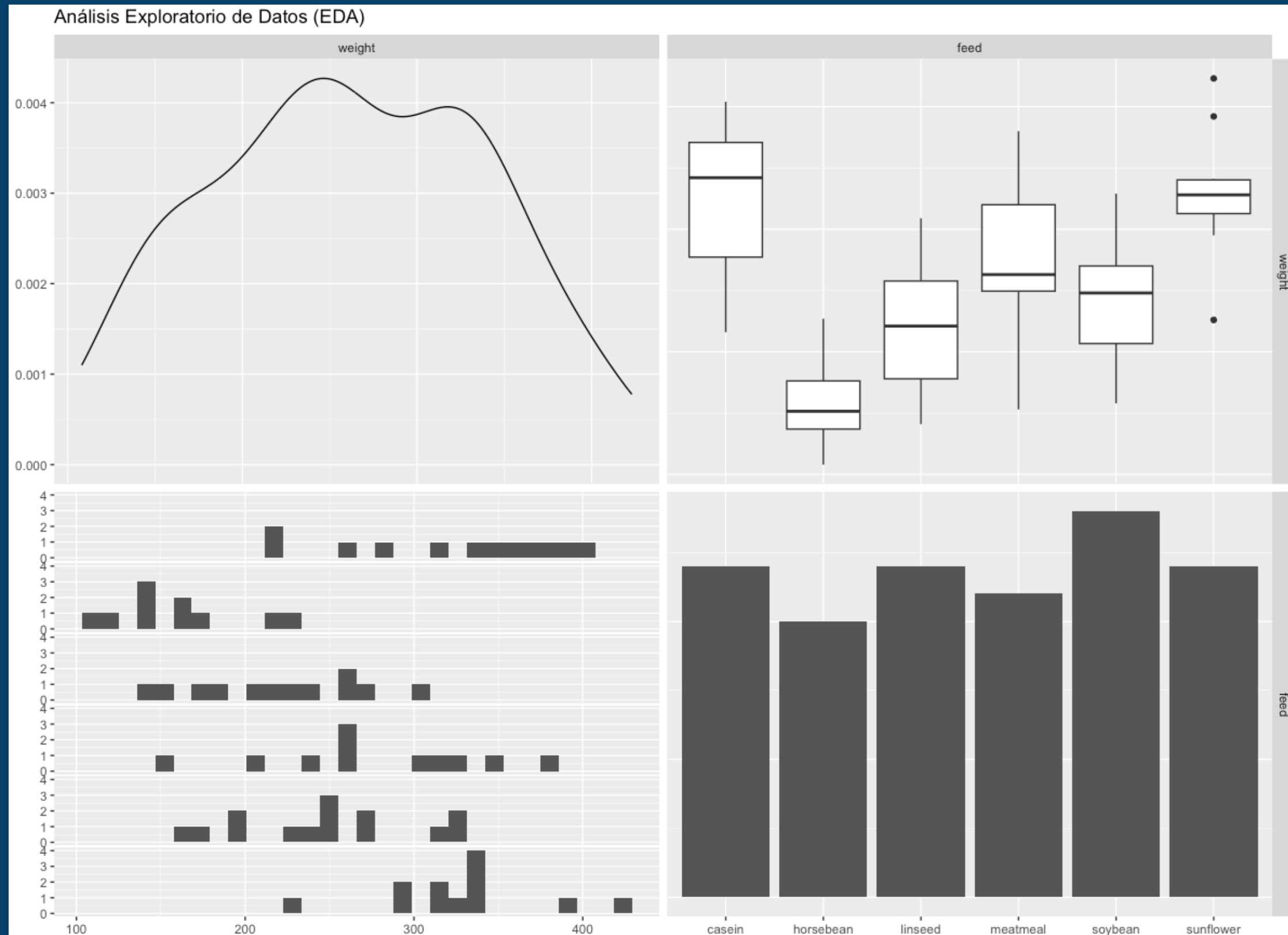
# Datos

71 observaciones con dos variables:

- weight: peso de los pollos en gramos. Simulara los tiempos de espera en un centro de atención al cliente
- feed: tipo de alimentación proporcionada a los pollos. Se interpretará como diferentes tipos de solicitudes en el contexto del centro de atención.

	weight	feed
1	179	horsebean
2	160	horsebean
3	136	horsebean
4	227	horsebean
5	217	horsebean
6	168	horsebean
7	108	horsebean
8	124	horsebean
9	143	horsebean
10	140	horsebean
11	309	linseed
12	229	linseed
13	181	linseed
14	141	linseed
15	260	linseed
16	203	linseed
17	148	linseed
18	169	linseed
19	213	linseed
20	257	linseed

# Análisis Exploratorio de Datos

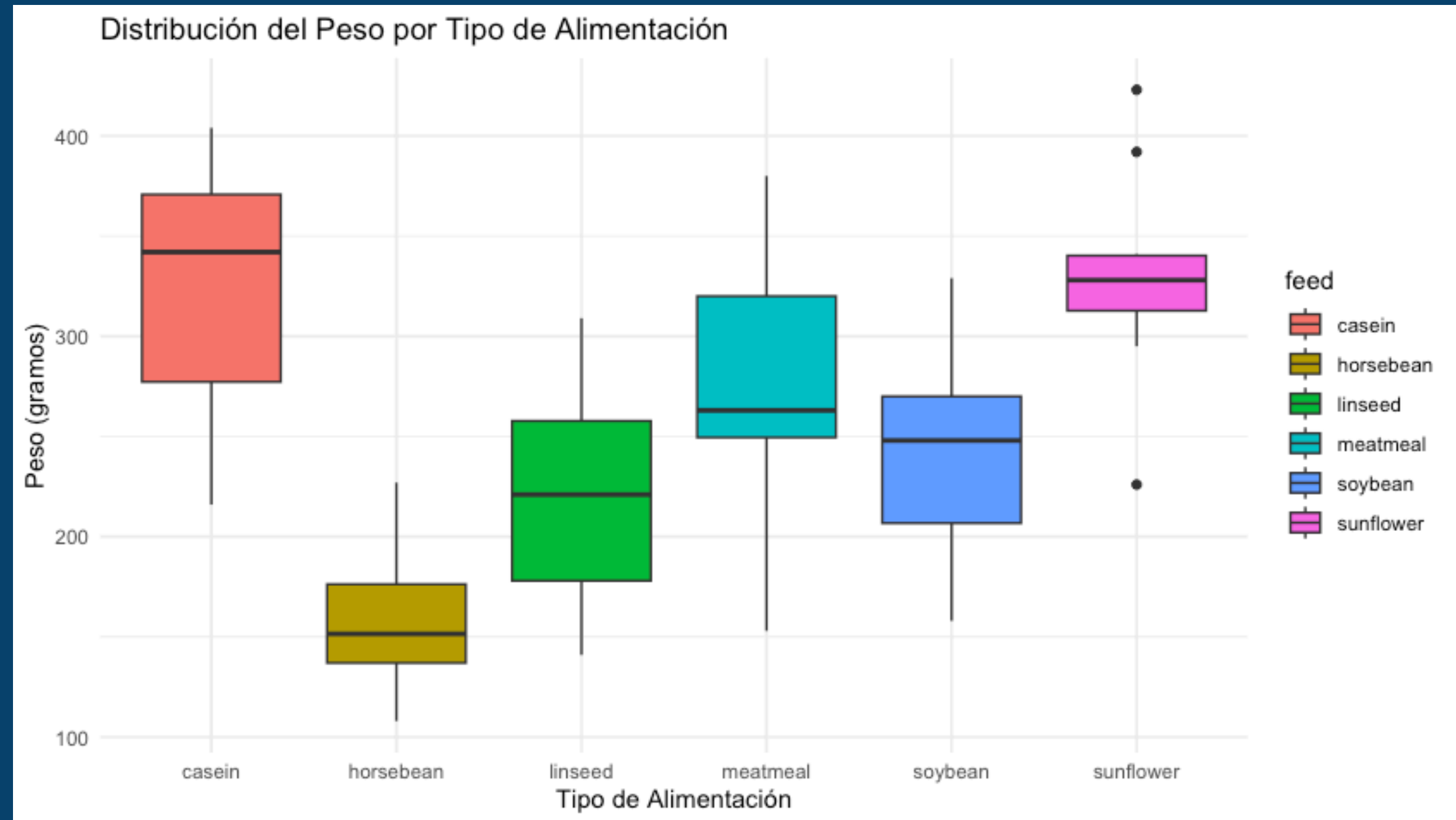


Se observa una distribución bimodal, esto quiere decir que es un tipo de distribución de datos que tiene dos picos distintos.

Podemos observar una gran variabilidad en los pesos de los pollos dependiendo el tipo de alimento.

# Boxplot

- Los tipos de alimento **caseína** y **semilla de girasol** muestran las medianas más altas.
- La dieta con **haba** muestra que los pollos con esta dieta tienden a ser mas livianos y con menos variabilidad en peso
- Podemos observar como hay mucha variabilidad entre los diferentes grupos, lo que nos dice que el tipo de alimentacion afecta significativamente el peso de los pollos

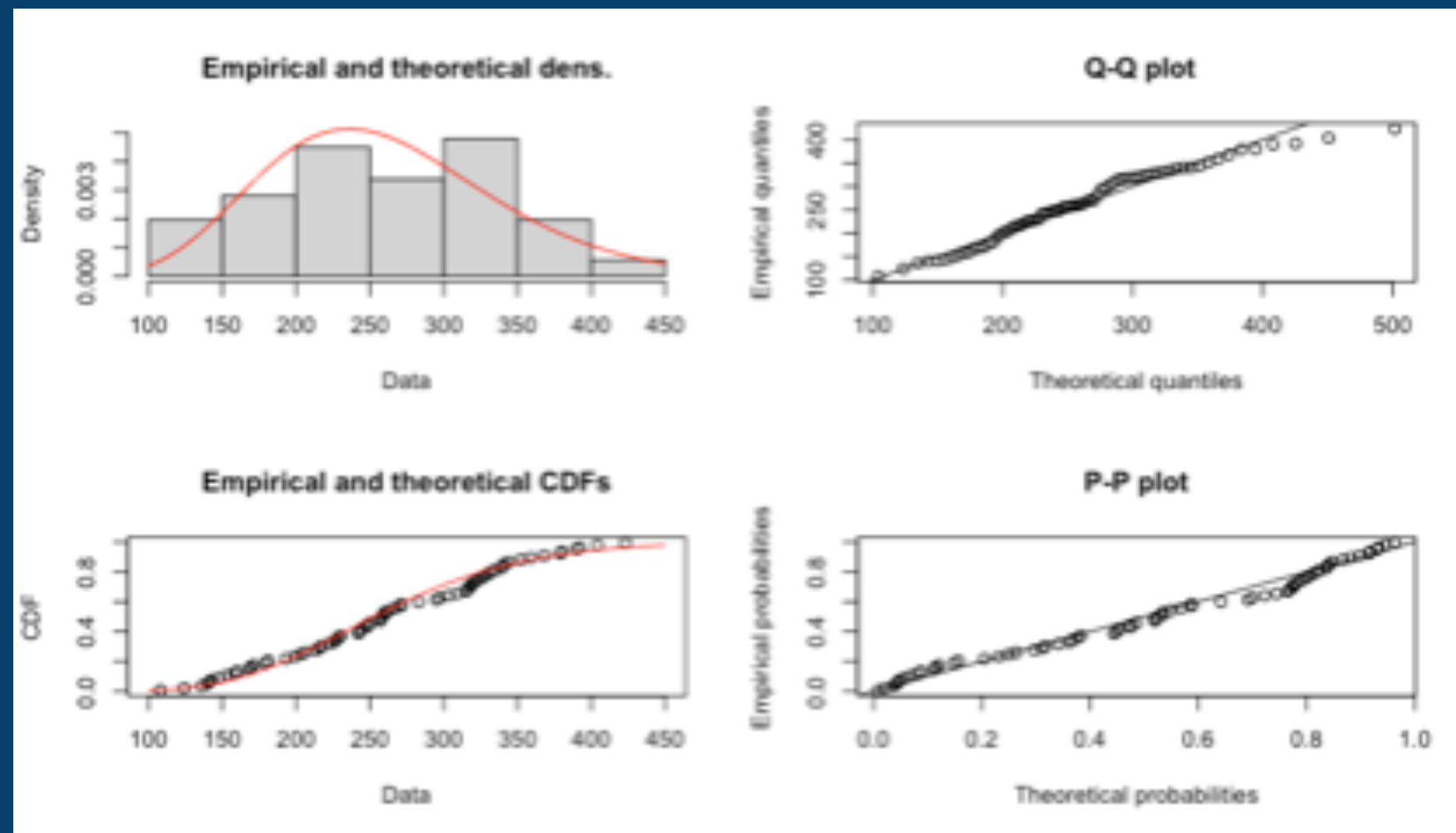


# Ajuste a la distribución de Erlang

## Densidad Empírica y Teórica

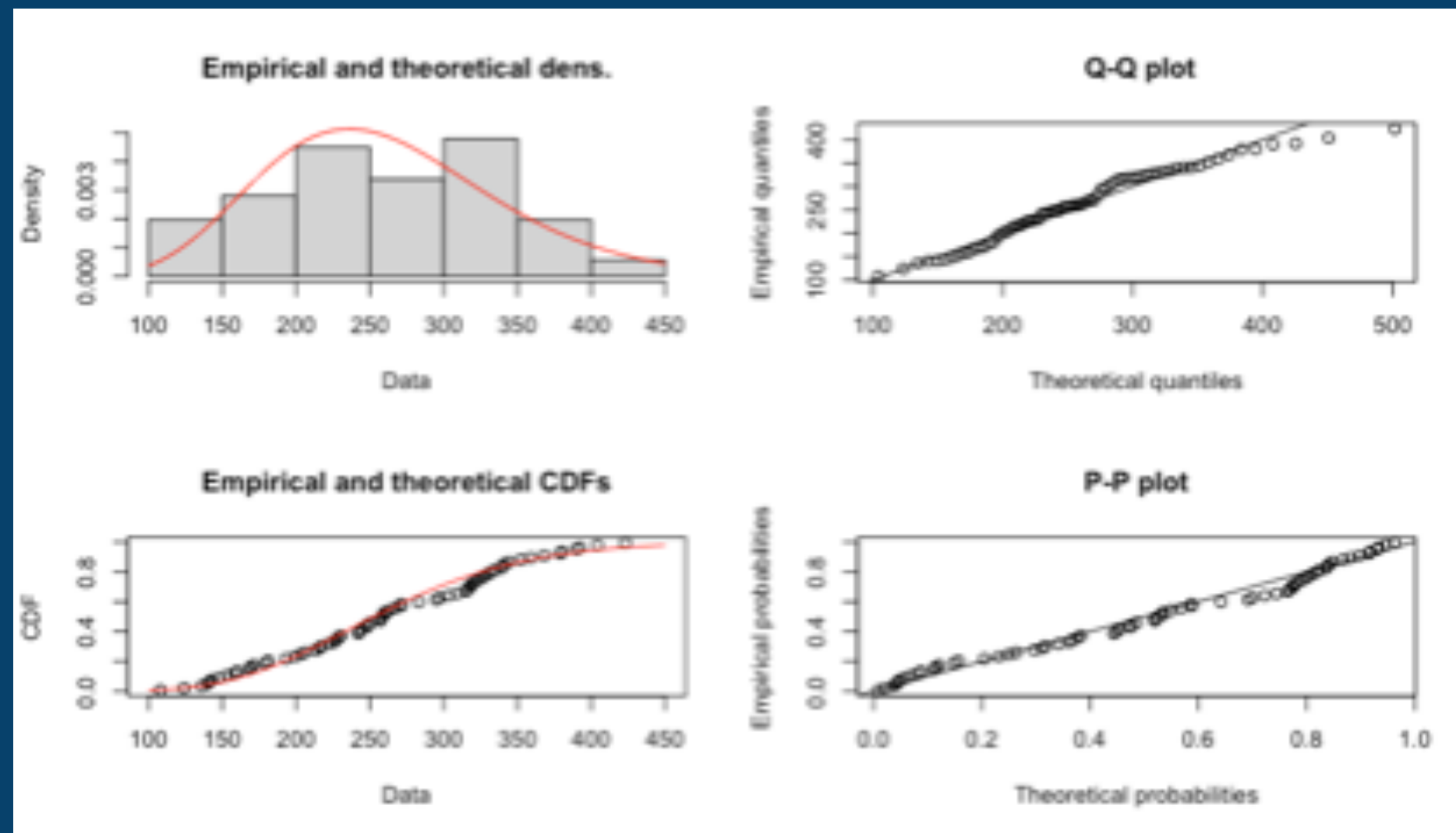
Observamos la comparación de la densidad observada del peso con la densidad teórica de la Distribución de Erlang.

La distribución empírica se ajusta bien, lo que nos dice que Erlang es una buena aproximación para los valores más comunes, sin embargo, si observamos los extremos, encontramos que hay discrepancias.



# Ajuste a la distribución de Erlang

## CDF's



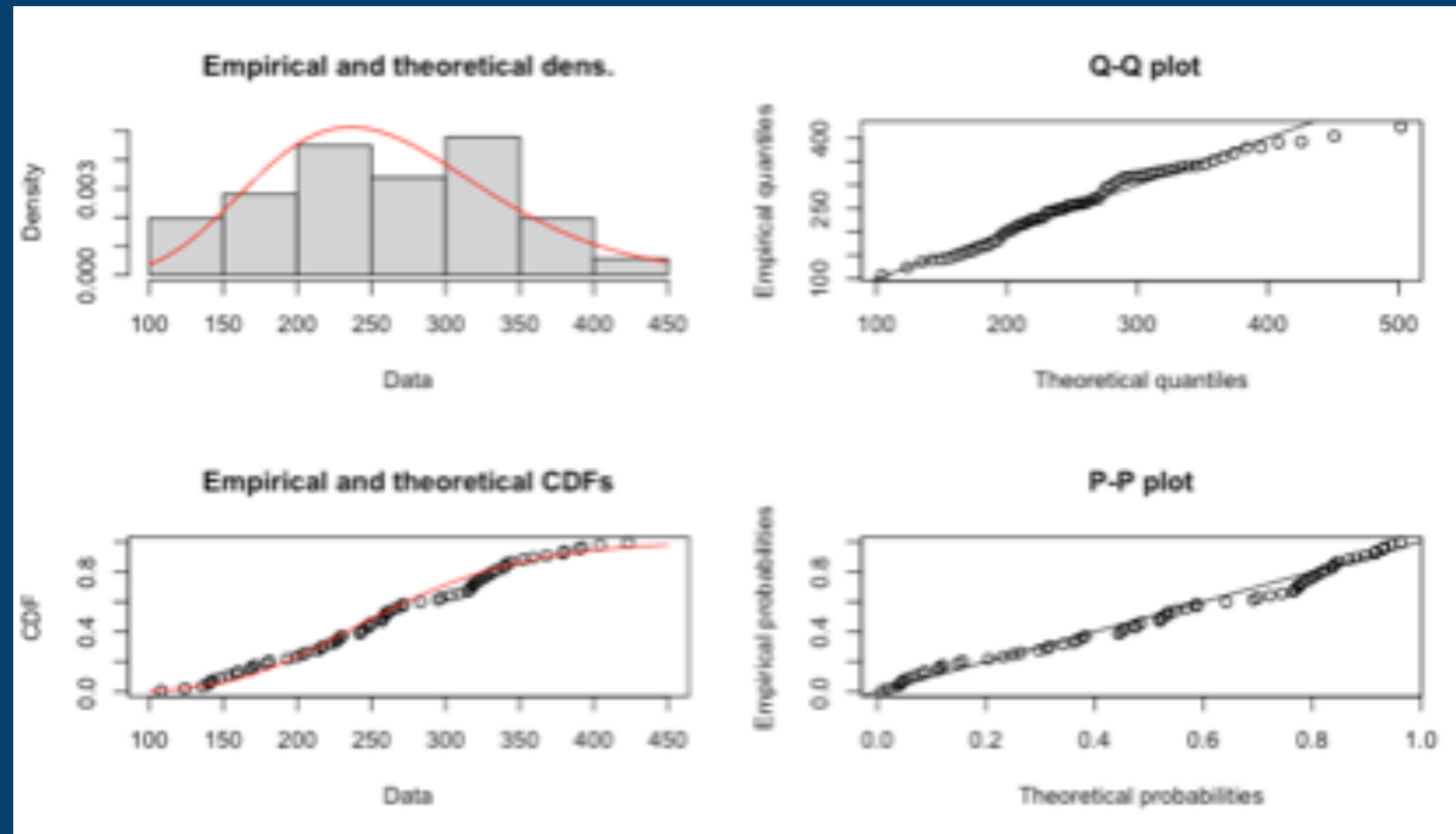
Observamos como la curva teórica sigue muy bien la tendencia que tienen la curva empírica, sin embargo, se observan ligeras desviaciones en los extremos, sobre todo en los valores más altos.

Esto nos dice que Erlang modela muy bien el comportamiento general pero, no captura completamente la variabilidad en los pesos extremos



# Ajuste a la distribución de Erlang

## P-P Plot

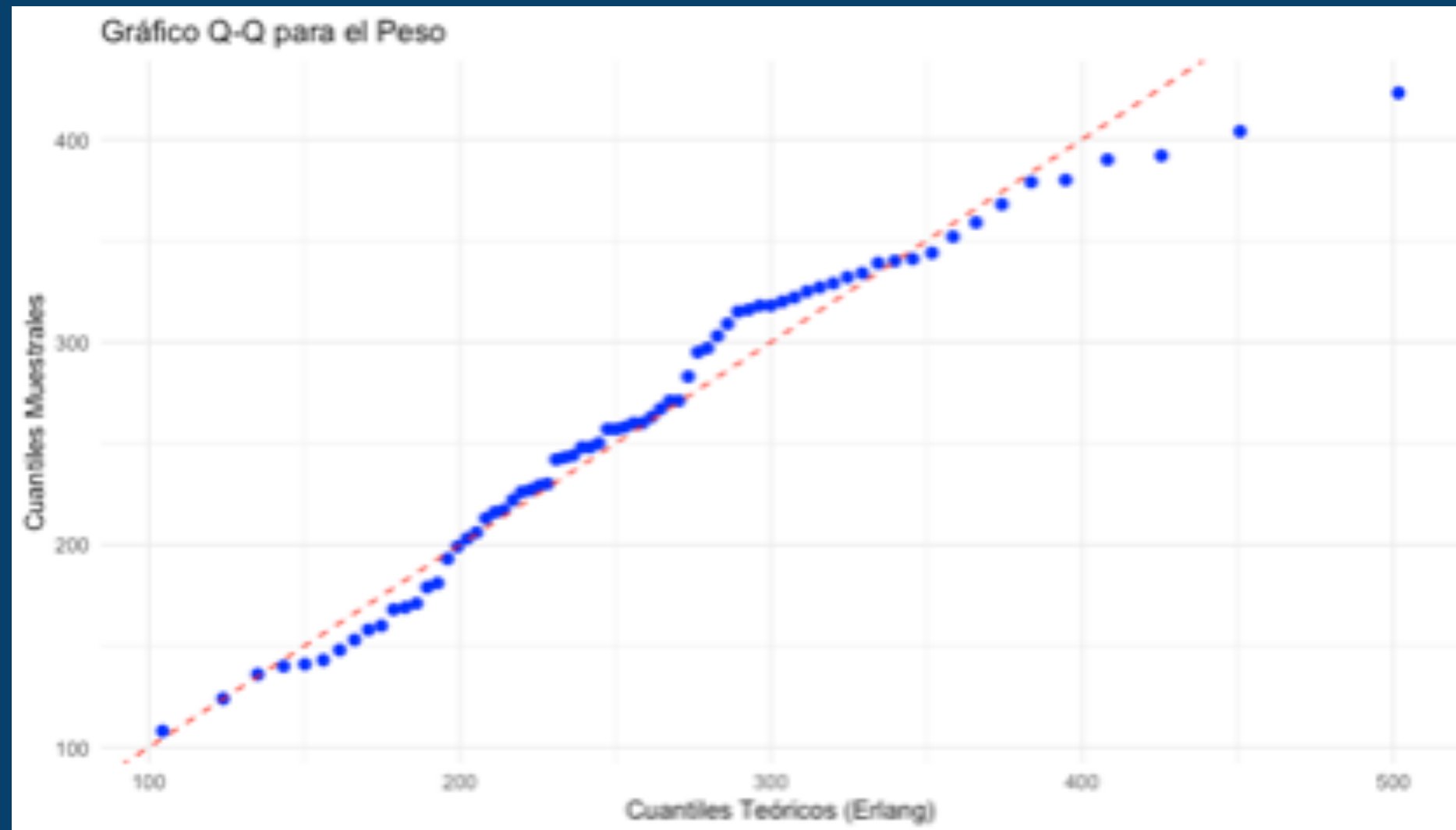


Se observa un patrón cercano a la diagonal, con algunas desviaciones, lo que confirma lo que se menciona con anterioridad.



# Ajuste a la distribución de Erlang

## Q-Q Plot



Compara los cuantiles de los datos con los de la distribución de Erlang.

Podemos observar como la alineación con la distribución de Erlang es bastante buena en la parte central de la distribución, sin embargo, en los extremos vemos como se desvían de la línea.

En pesos mayores, los valores observados son inferiores a los predichos por Erlang, lo que nos dice que el modelo está sobreestimando la probabilidad de pesos altos.

# Calcular probabilidades

```
49 # Calcular la probabilidad de que un proceso tarde más de X minutos
50 X <- 300 # Umbral de peso (interpretado como tiempo de espera)
51 probabilidad <- 1 - pgamma(X, shape = erlang_fit$estimate["shape"], rate = erlang_fit$estimate["rate"])
52 cat("La probabilidad de que un proceso tarde más de", X, "gramos (minutos) es:", probabilidad, "\n")
53
54 # Punto 5: Conclusión y reflexión
55 cat("\nConclusión:\n")
56 cat("El modelo de Erlang se ajusta a los datos con los siguientes parámetros:\n")
57 cat(" - Shape (forma):", erlang_fit$estimate["shape"], "\n")
58 cat(" - Rate (tasa):", erlang_fit$estimate["rate"], "\n")
59 cat("La probabilidad calculada sugiere que", round(probabilidad * 100, 2), "% de los procesos tardan más de", X, "gramos (minutos).\n")
60
```

La probabilidad de que un proceso tarde más de 300 gramos (minutos) es: 0.2883564

```
> cat("\nConclusión:\n")
```

Conclusión:

```
> cat("El modelo de Erlang se ajusta a los datos con los siguientes parámetros:\n")
```

El modelo de Erlang se ajusta a los datos con los siguientes parámetros:

```
> cat(" - Shape (forma):", erlang_fit$estimate["shape"], "\n")
```

- Shape (forma): 10.38868

```
> cat(" - Rate (tasa):", erlang_fit$estimate["rate"], "\n")
```

- Rate (tasa): 0.03975894

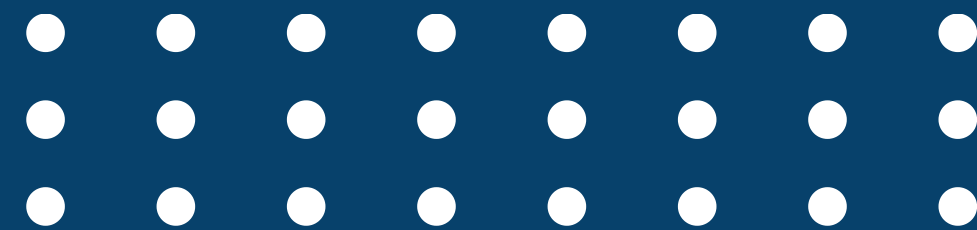
```
> cat("La probabilidad calculada sugiere que", round(probabilidad * 100, 2), "% de los procesos tardan más de", X, "gramos (minutos).\n")
```

La probabilidad calculada sugiere que 28.84 % de los procesos tardan más de 300 gramos (minutos).

# ¿Por qué la distribución de Erlang es útil para modelar tiempos de espera? ¿Qué características la hacen adecuada para este tipo de problemas?

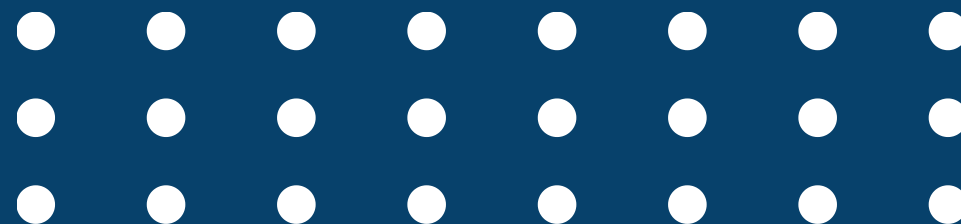
Principalmente por las propiedades de esta distribución, esta es una generalización de la exponencial y se usa comúnmente para modelar tiempos de espera o procesos de colas donde los eventos ocurren en una secuencia de etapas.

Cómo observamos, los datos se ajustan bien a la distribución, por lo que las predicciones serán más precisas.



# Limitaciones del modelo: ¿Cuáles son las limitaciones de usar el dataset chickwts para simular tiempos de espera en un centro de atención al cliente?

- Los datos no muestran tiempos de llegada o de servicio
- Es únicamente una analogía estadística ya que el peso de los pollos no tiene relación alguna con tiempos de espera
- El dataset únicamente se centra en la relación que hay entre alimentación y crecimiento
- Nos da únicamente valores enteros



# Mejoras Futuras: Otras Distribuciones para Modelar Tiempos de Espera

- Distribución Exponencial: modela el tiempo entre llegadas de clientes
- Distribucion Gamma: permite valores no enteros, haciendolo más flexible para modelar tiempos de espera

