Análisis en R para determinar la distribución a la que se ajustan los datos

Consideraciones para evaluar si la distribución se ajusta a los datos:

1. **Pruebas de bondad y ajuste**: El código realiza diferentes pruebas para determinar la distribución bajo la cual se hace la comparación. Entre las pruebas están:
   1. Kolmogorov-Smirnov:
      1. Hipótesis nula (H0): La hipótesis nula establece que no hay diferencia significativa entre la distribución empírica de tus datos y la distribución teórica.
      2. Hipótesis alternativa (H1): La hipótesis alternativa sugiere que existe una diferencia significativa entre la distribución empírica y la distribución teórica.
   2. Chi-Cuadrado: evalúa si hay una diferencia significativa entre la distribución observada de datos y una distribución teórica o esperada.
2. **Interpretación de los p-values**: Si el p-value es pequeño, se rechaza la hipótesis nula y se acepta la alternativa, por lo tanto, la distribución no se ajusta a los datos de la muestra

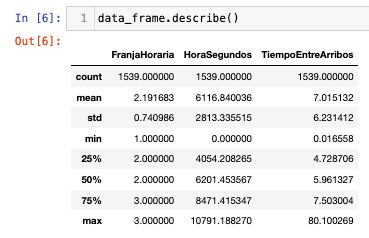
**Contexto**

Los datos representan los tiempos entre arribos de clientes a un punto de servicio. El data set cuenta con las siguientes columnas:

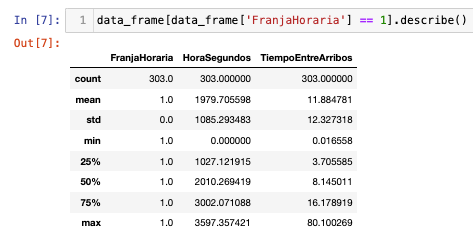
* Franja horaria en que el cliente arriba al punto – Existen 3 franjas horarias
* Hora del arribo en segundos
* Tiempo entre arribos entre cliente y cliente

**Descriptivas**

* Datos totales



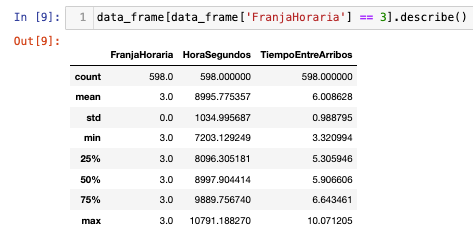
* Datos Franja Horaria 1



* Datos Franja Horaria 2



* Datos Franja Horaria 3



**Evaluación**

Se realizan las pruebas con las siguientes distribuciones:

* Distribución normal: "norm"
* Distribución log-normal: "lnorm"
* Distribución exponencial: "exp"
* Distribución de Poisson: "pois"
* Distribución de Cauchy: "cauchy"
* Distribución gamma: "gamma"
* Distribución logística: "logis"
* Distribución binomial negativa: "nbinom"
* Distribución geométrica: "geom"
* Distribución beta: "beta"
* Distribución Weibull: "weibull"

**Datos Totales:**

Luego de realizar la prueba con todas las distribuciones, se observa que todos los resultados de las pruebas de bondad y ajuste se rechaza la hipótesis nula, por lo tanto, los datos no se ajustan correctamente a ninguna distribución.

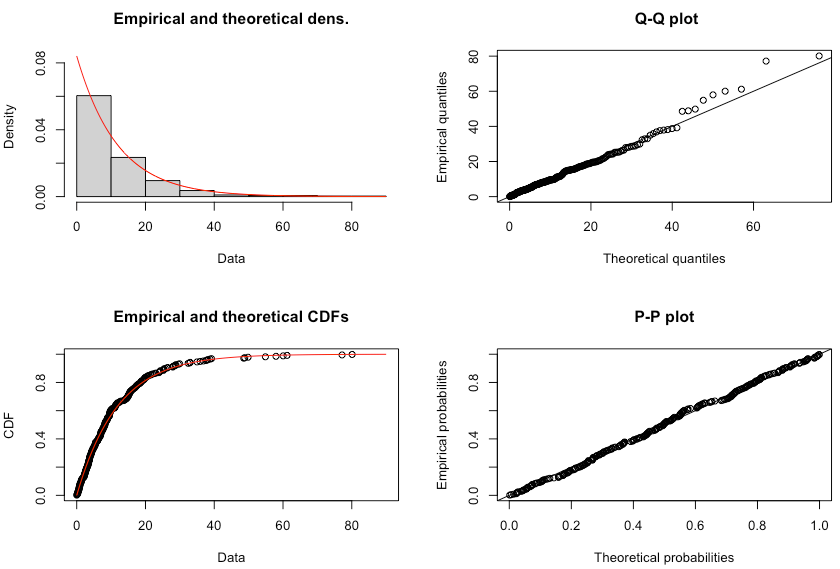
A partir de la evaluación de los datos por cada franja horaria se obtienen las siguientes distribuciones que se ajustan a los datos

1. Datos totales

|  |  |
| --- | --- |
| **Prueba de Bondad y Ajuste** | **Resultado** |
| Kolmogorov-Smirnov | Rechazado |
| Chi-Cuadrado (p-value) | 0 |

1. Datos Franja Horaria 1 - Exponencial

|  |  |
| --- | --- |
| **Prueba de Bondad y Ajuste** | **Resultado** |
| Kolmogorov-Smirnov | No rechazado |
| Chi-Cuadrado (p-value) | 0.3057602 |

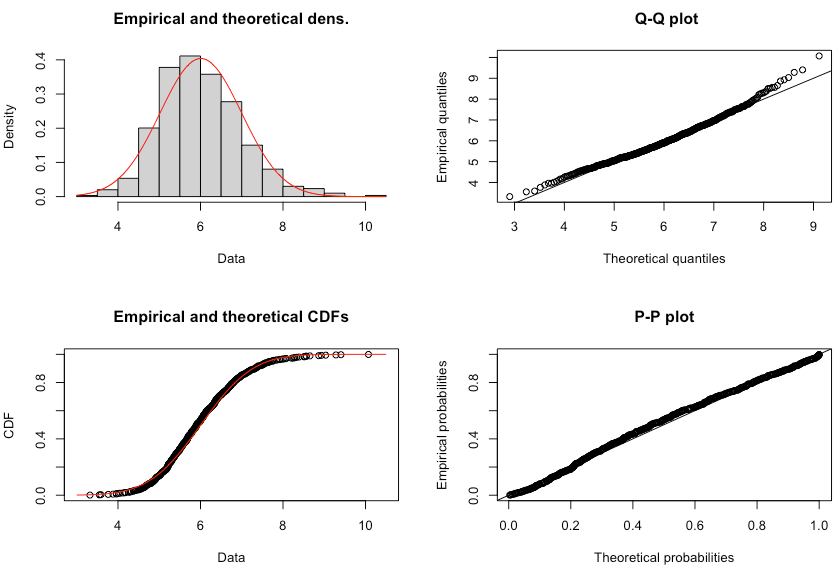


1. Datos Franja Horaria 2 – Ninguna distribución se ajusta a los datos

|  |  |
| --- | --- |
| **Prueba de Bondad y Ajuste** | **Resultado** |
| Kolmogorov-Smirnov | Rechazado |
| Chi-Cuadrado (p-value) | < 0 |

1. Datos Franja Horaria 3 – Distribución Normal

|  |  |
| --- | --- |
| **Prueba de Bondad y Ajuste** | **Resultado** |
| Kolmogorov-Smirnov | No rechazado |
| Chi-Cuadrado (p-value) | 0.05123061 |

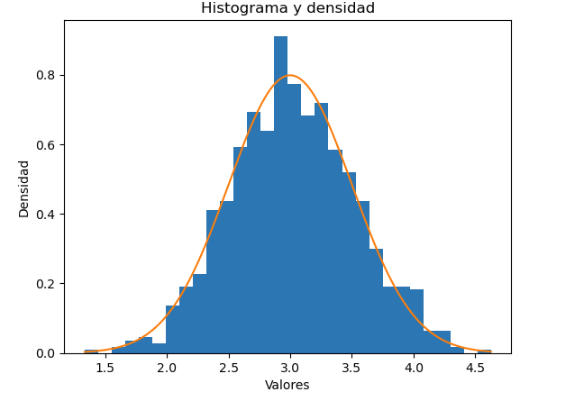


**PUNTO 3**

1. Modificar los parámetros de las distribuciones
2. **Distribución Normal**

Parámetros:

* Media: 3.0
* Varianza: 0.25
* Asimetría: 0.0
* Curtosis: 0.0



Parámetros de la distribución normal ajustada

* Localización: 3.0229927766501343
* Escala: 0.5039009495403807

Momentos de la distribución normal ajustada

* Media: 3.0229927766501343
* Varianza: 0.25391616694769725
* Asimetría: 0.0
* Curtosis: 0.0

