ANÁLISIS

1. Análisis estadístico descriptivo de las precipitaciones históricas máximas mensuales de un estado

A. Descarga la base de datos de precipitaciones máximas históricas mensuales de todos los estados de la república de la siguiente liga: precipitaciones mensuales. Esta base de datos se construyó con información de los resúmenes mensuales de lluvia y temperatura de CONAGUA Selecciona un estado que sea diferente a los del resto de tu equipo.

```
# Cargar librerías necesarias
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(tidyr)
library(moments) # Para el cálculo de sesgo y curtosis

# Cargar los datos desde el archivo
datos <- read.table("precipitaciones_maximas_mensuales.txt", header = TRUE, sep = "\t")
datos$Lluvia <- as.numeric(as.character(datos$Lluvia))
```

B. Elabora una gráfica de las precipitaciones máximas mensuales por año para tu estado. Para ello deberás calcular la precipitación mensual máxima de cada año y graficarla.

C. Analiza los datos de precipitaciones máximas mensuaels del estado seleccionado.

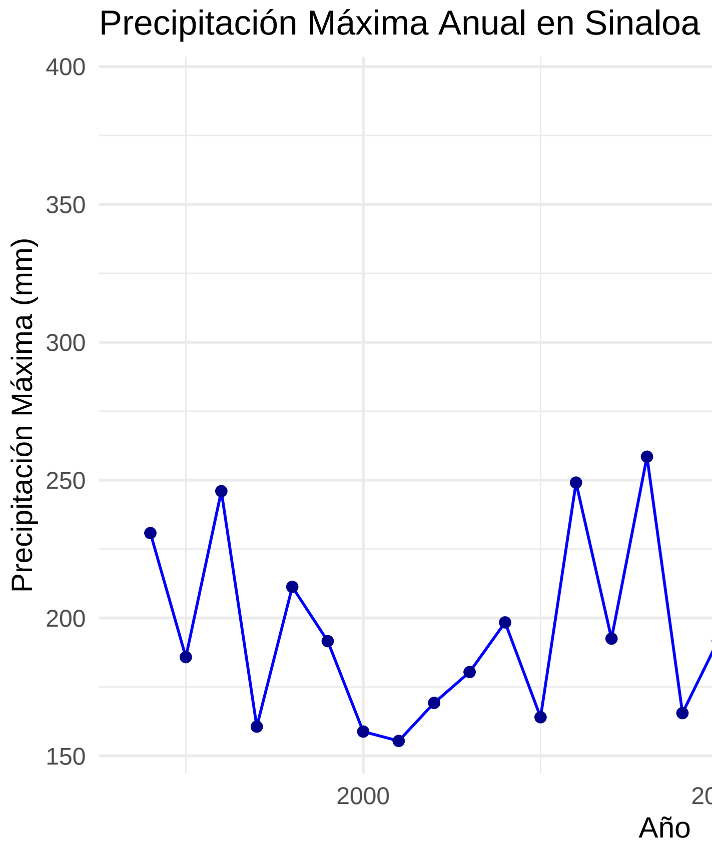
Calcula las medidas de centralización y variación de las precipitaciones máximas mensuales

Realiza gráficos que te sirvan para describir la distribución de las lluviar máximas mensuales: histograma y boxplot

```
# Seleccionar el estado de Sinaloa
estado <- "Sinaloa"
datos_estado <- datos %>% filter(Estado == estado)
```

```r
# Calcular la precipitación máxima mensual por año
datos_max_anuales <- datos_estado %>%
  group_by(Anio) %>%
  filter(Lluvia == max (Lluvia, na.rm = TRUE))
#datos_max_anuales[datos_max_anuales == 0] <- 0.001
# Graficar precipitación máxima mensual por año
ggplot(datos_max_anuales, aes(x = Anio, y = Lluvia)) +
  geom_line(color = "blue") +
  geom_point(color = "darkblue") +
  labs(title = paste("Precipitación Máxima Anual en", estado),
       x = "Año", y = "Precipitación Máxima (mm)") +
  theme_minimal()
```

Describe el comportamiento de la distribución: centralización, sesgo, variación, ..os de precip-



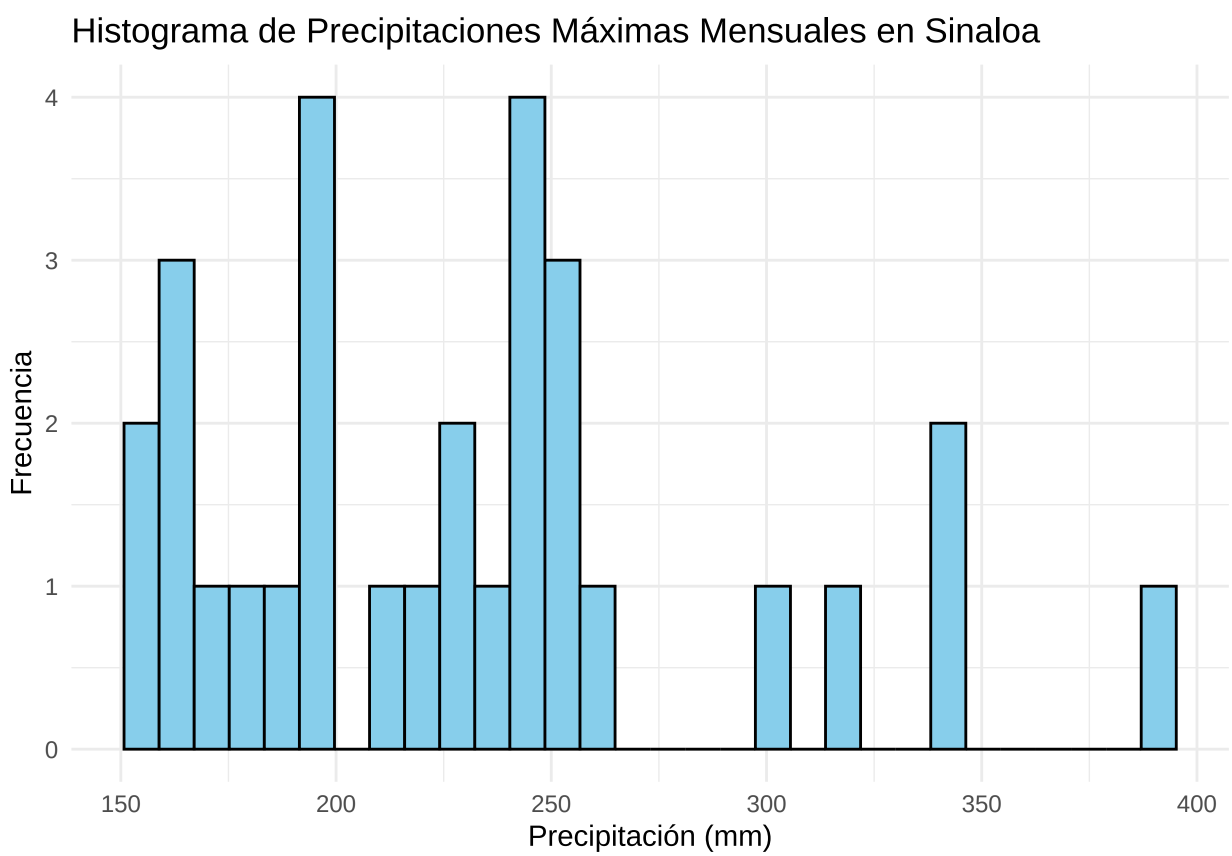
itaciones máximas mensuaels del estado seleccionado.

```r
# Análisis Estadístico Descriptivo
media <- mean(datos_max_anuales$Lluvia, na.rm = TRUE)
mediana <- median(datos_max_anuales$Lluvia, na.rm = TRUE)
desviacion <- sd(datos_max_anuales$Lluvia, na.rm = TRUE)
sesgo <- skewness(datos_max_anuales$Lluvia, na.rm = TRUE)
curtosis <- kurtosis(datos_max_anuales$Lluvia, na.rm = TRUE)
# Crear un data frame con las estadísticas
summary_estadistico <- data.frame(
  Metrica = c("Media", "Mediana", "Desviación Estándar", "Sesgo", "Curtosis"),
  Valor = c(media, mediana, desviacion, sesgo, curtosis)
)
```

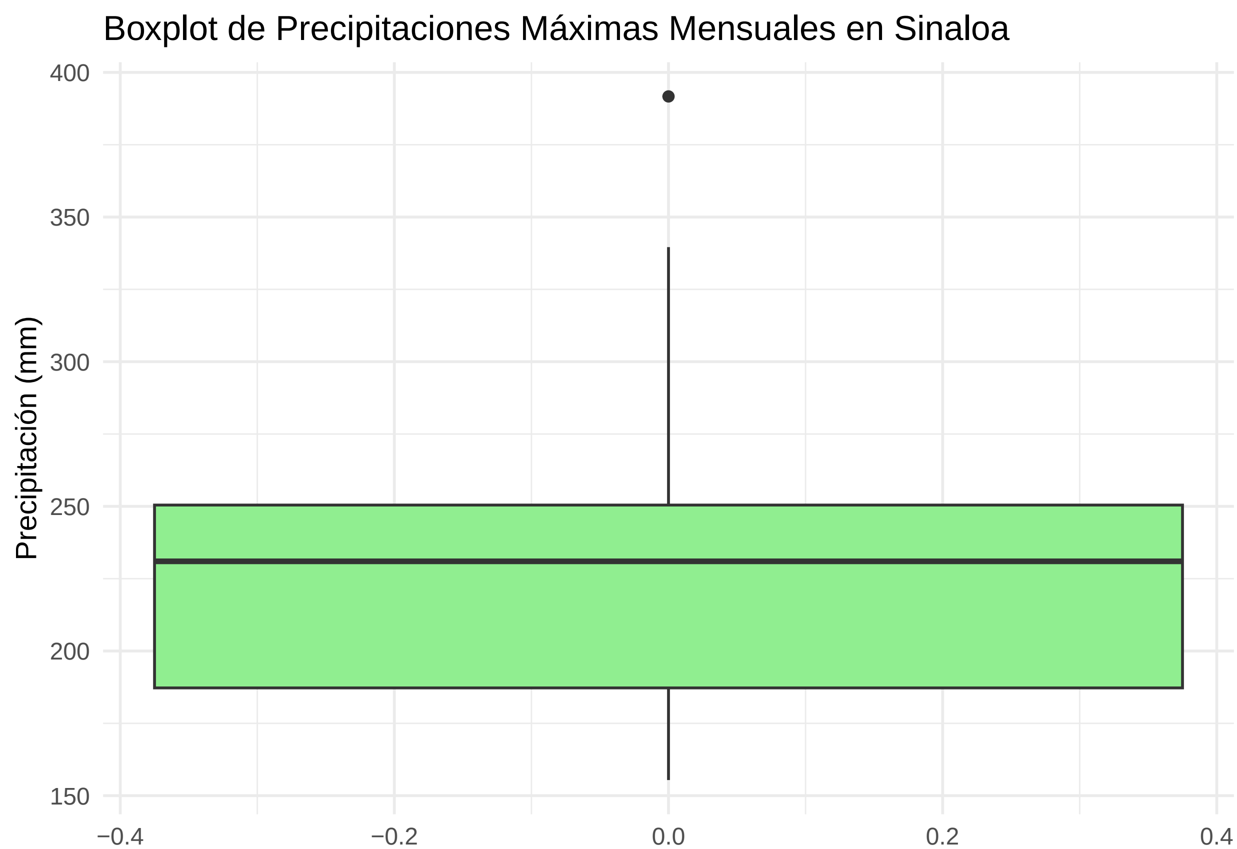
```r
# Mostrar estadísticas descriptivas
print(summary_estadistico)
```

```
##               Metrica       Valor
## 1               Media 230.9266667
## 2             Mediana 231.0000000
## 3 Desviación Estándar  59.8781172
## 4               Sesgo   0.8803158
## 5            Curtosis   3.2961284
```

```r
# Histograma de la distribución de precipitaciones máximas mensuales
ggplot(datos_max_anuales, aes(x = Lluvia)) +
  geom_histogram(color = "black", fill = "skyblue", bins = 30) +
  labs(title = paste("Histograma de Precipitaciones Máximas Mensuales en", estado),
       x = "Precipitación (mm)", y = "Frecuencia") +
  theme_minimal()
```




```r
# Boxplot para observar variación y posibles valores atípicos
ggplot(datos_max_anuales, aes(y = Lluvia)) +
  geom_boxplot(fill = "lightgreen") +
  labs(title = paste("Boxplot de Precipitaciones Máximas Mensuales en", estado),
       y = "Precipitación (mm)") +
  theme_minimal()
```



D.¿Qué puedes concluir observando la gráfica de los máximos mensuales anuales para tu Estado? ¿Observas alguna tendencia? ¿Puedes concluir que cada determinado número de años la cantidad de precipitación sube o baja? ¿Para qué nos sirve analizar este tipo de gráficas

Análisis del comportamiento de la distribución de precipitación máxima mensual en Sinaloa:

```
Centralización:
    La media de la precipitación máxima mensual en Sinaloa es de 230.93 mm, y la mediana es de 231 mm. E

Sesgo:
    La distribución presenta un sesgo positivo leve (0.88), lo cual sugiere una ligera inclinación hacia

Variación:
    La desviación estándar de 59.88 mm refleja una variabilidad moderada en las precipitaciones máximas

Curtosis:
    La curtosis de 3.29 indica que la distribución tiene colas ligeramente más largas que una distribuci
```

Conclusiones sobre las gráficas de los máximos mensuales anuales para Sinaloa:

```
Tendencias Observadas:
    En la gráfica de la precipitación máxima anual, se observa un incremento gradual con algunos picos n

Ciclos o Fluctuaciones:
    No se observa un ciclo constante, pero sí una tendencia ascendente en la variabilidad a medida que n
```

Analizar la tendencia de los máximos de precipitación anuales y mensuales es fundamental para poder frenar las consecuencias o daños de eventos extremos, diseñar infraestructura resiliente y tomar decisiones informadas en la planificación de recursos y gestión de riesgos. Esto contribuye a una respuesta más efectiva y a una mejor preparación para los desafíos climáticos que pudiese llegar a enfrentar Sinaloa.

2. Análisis de Frecuencias Método Gráfico

A. En el data frame de los datos de precipitación máxima se agrega una columna con los datos de lluvias máximas ordenados de mayor a menor.

B. Se agrega una columna con el número de orden que tiene asignado cada precipitación máxima. A ese número se le llama "rank" (rango en español) y se simboliza por m

C. Se calcula la probabilidad de excedencia o de ocurrencia de acuerdo con Weibull, donde el numerador es el número de orden (m) o "rank" y el denominador es la suma del total de datos (N) y 1:

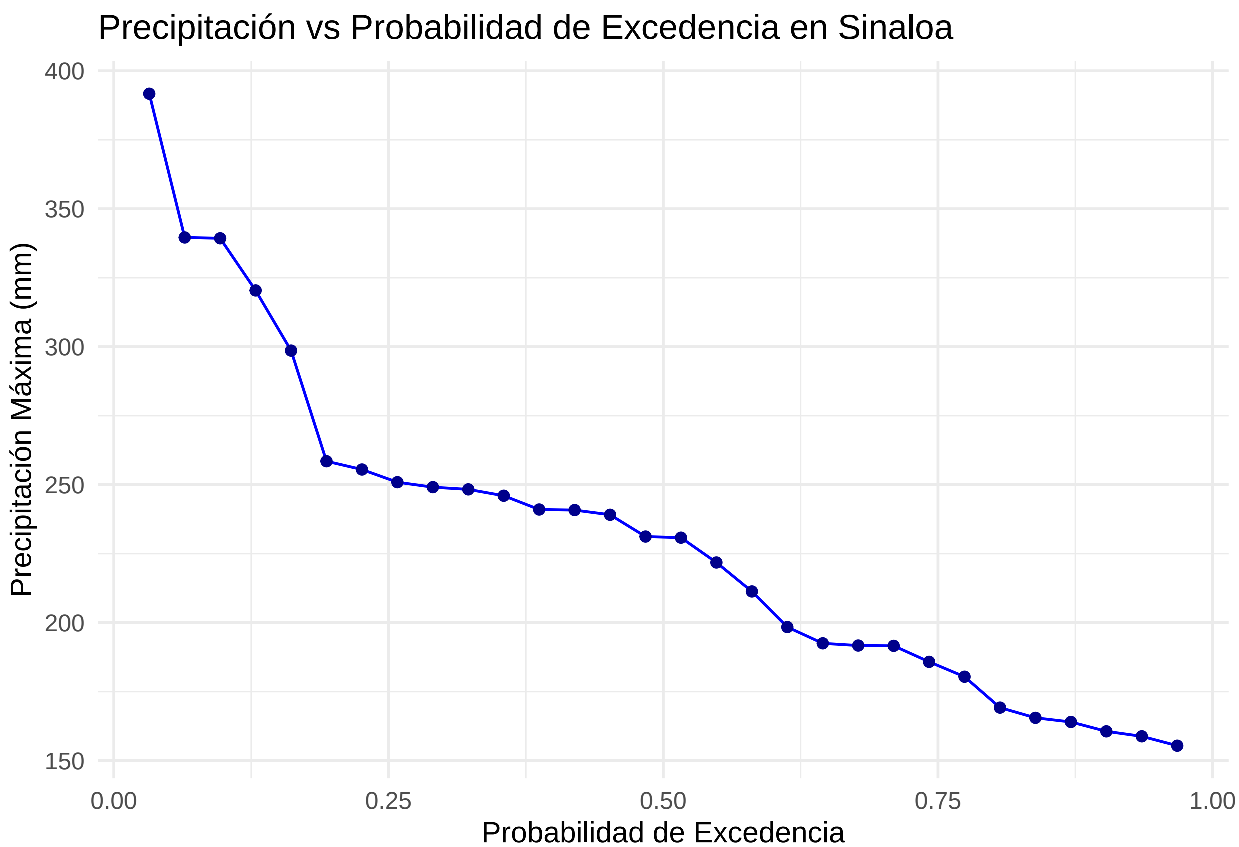
D. Se calcula la probabilidad de no excedencia para cada precipitación (complemento de la probabilidad de excedencia):

E. Se calcula el periodo de retorno como el inverso de la probabilidad de excedencia:

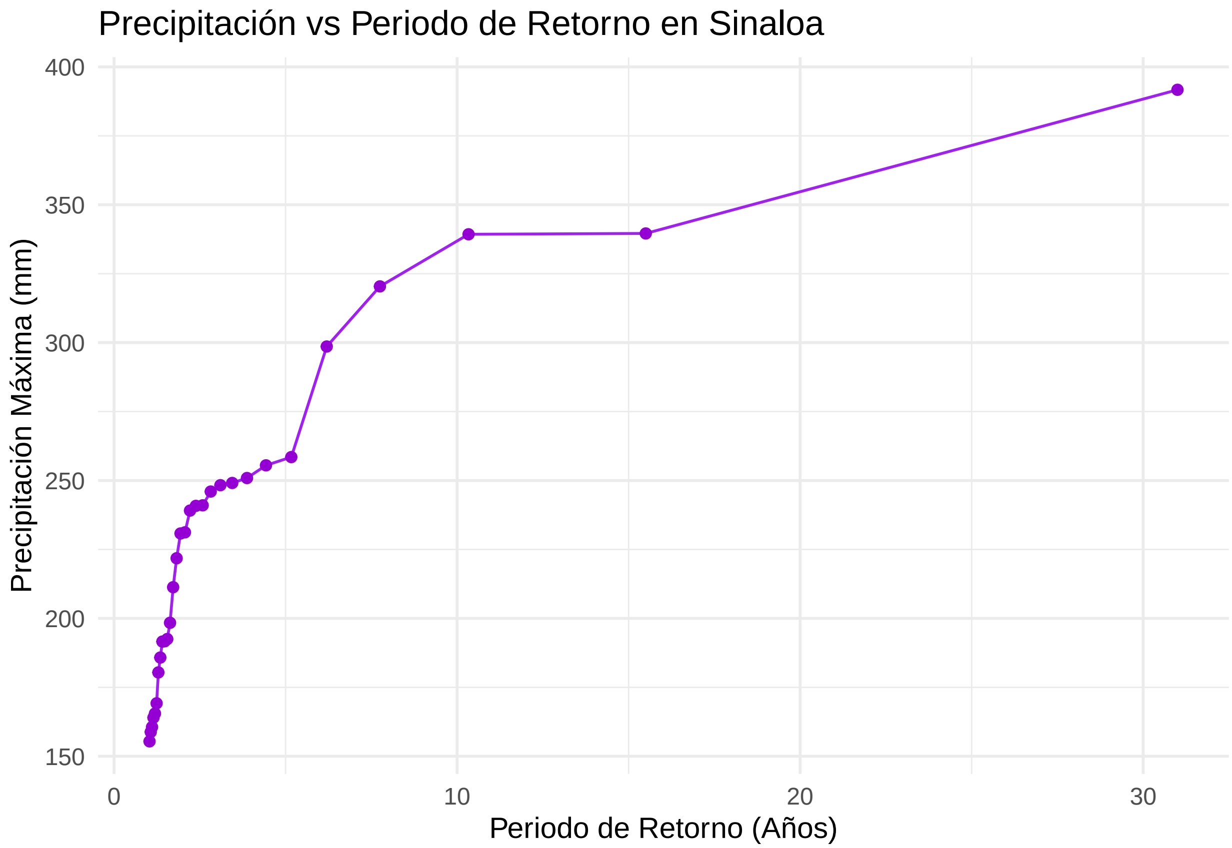
```r
# Ordenar las precipitaciones de mayor a menor
datos_max_anuales<- datos_max_anuales[order(-datos_max_anuales$Lluvia),]
# Crear columna de rango
datos_max_anuales$Rango <- 1:nrow(datos_max_anuales)

# Calcular probabilidad de excedencia usando la fórmula de Weibull
N <- nrow(datos_max_anuales)  # Número total de datos
datos_max_anuales$P_exc <- datos_max_anuales$Rango / (N + 1)

# Calcular la probabilidad de no excedencia
datos_max_anuales$P_no_exc <- 1 - datos_max_anuales$P_exc
# Calcular el periodo de retorno
datos_max_anuales$P_ret <- 1 / datos_max_anuales$P_exc
head(datos_max_anuales)
```

```
## # A tibble: 6 x 8
## # Groups:   Anio [6]
##    Anio Mes   Estado  Lluvia Rango  P_exc P_no_exc P_ret
##   <int> <chr> <chr>    <dbl> <int>  <dbl>    <dbl> <dbl>
## 1  2021 Ago   Sinaloa   392.     1 0.0323    0.968 31
## 2  2019 Ago   Sinaloa   340.     2 0.0645    0.935 15.5
## 3  2022 Ago   Sinaloa   339.     3 0.0968    0.903 10.3
## 4  2018 Sep   Sinaloa   320.     4 0.129     0.871  7.75
## 5  2013 Sep   Sinaloa   299.     5 0.161     0.839  6.2
## 6  2008 Ago   Sinaloa   258.     6 0.194     0.806  5.17
```

```r
library(ggplot2)

ggplot(datos_max_anuales, aes(x = P_exc, y = Lluvia)) +
  geom_line(color = "blue") +
  geom_point(color = "darkblue") +
  labs(title = paste("Precipitación vs Probabilidad de Excedencia en", estado),
       x = "Probabilidad de Excedencia",
       y = "Precipitación Máxima (mm)") +
  theme_minimal()
```




```r
ggplot(datos_max_anuales, aes(x = P_ret, y = Lluvia)) +
  geom_line(color = "purple") +
  geom_point(color = "darkviolet") +
  labs(title = paste("Precipitación vs Periodo de Retorno en", estado),
       x = "Periodo de Retorno (Años)",
       y = "Precipitación Máxima (mm)") +
  theme_minimal()
```



F. Describe las gráficas obtenidas. ¿Qué significa la probabilidad de excedencia? ¿Qué significa el periodo de retorno? ¿Por qué es importante en hidrología? ¿Qué valores son deseables en la probabilidad de excedencia para una precipitación de diseño de una obra? Puedes consultar el siguiente video de apoyo:

Descripción de las Gráficas

```
Precipitación vs Probabilidad de Excedencia:
    La primera gráfica (Probabilidad de Excedencia) muestra cómo las precipitaciones máximas disminuyen

Precipitación vs Periodo de Retorno:
    La segunda gráfica (Periodo de Retorno) muestra cómo las precipitaciones extremas están asociadas c
```

Entonces, la probabilidad de excedencia se refiere a la probabilidad de que un valor específico de precipitación máxima sea superado en un año dado. Es crucial en hidrología, pues permite estimar la frecuencia con la que se pueden esperar eventos extremos de precipitación. Una probabilidad de excedencia baja indica que el evento no es frecuente y por ende, se espera que ocurra con menor frecuencia.

Ahora bien, el periodo de retorno es el inverso de la probabilidad de excedencia y representa el intervalo promedio de tiempo (en años) en el que se espera que un evento específico, como una precipitación máxima dada, ocurra o sea excedido al menos una vez. Un periodo de retorno alto implica que el evento es poco común y se presenta con menor frecuencia.

Ambos son esenciales para el diseño y la planificación de infraestructuras, la evaluación del riesgo de inundaciones y el desarrollo de estrategias de mitigación, la gestión del agua, y la toma de decisiones para reducir el impacto de eventos climáticos severos. Normalmente, se suelen elegir valores de probabilidad de excedencia bajos, correspondientes a eventos no frecuentes (con periodos de retorno altos), para garantizar la seguridad de la infraestructura en condiciones extremas. Por ejemplo, un diseño de infraestructura crítica podría considerar una probabilidad de excedencia de 1% (o un periodo de retorno de 100 años), lo que implica que la infraestructura debería soportar eventos de esta magnitud que se esperaría ocurran, en promedio, una vez cada 100 años.

Estas métricas ayudan a ingenieros y planificadores a decidir el nivel de seguridad y resistencia necesario para la infraestructura en función de los riesgos climáticos y el costo de construcción.

3 . Análisis de Frecuencias Método Analítico

El método analítico consiste en asumir que los datos pueden ser ajustados a través de una función de densidad de probabilidades (FDP) conocida la cual nos servirá para modelar y pronosticar precipitaciones y periodos de retorno. Para ello es necesario probar varias distribuciones y emplear pruebas de bondad de ajuste para ver decidir cuál distribución es la que mejor se ajusta. Para nuestro análisis verificaremos el ajuste de las precipitaciones máximas mensuales a diferentes distribuciones.

A. Ajuste a una Distribución Normal. Hay dos maneras de determinar si un conjunto de datos tiene una distribución normal, una visual y la otra mediante una prueba de bondad de ajuste.

Contruye el histograma de la función de densidad empírica de los datos y sobrepon una distribución normal que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. De manera visual, ¿te parece que los datos se ajustan bien a una distribución normal? Explica. ¿Cuántos parámetros tiene la distribución Normal? ¿Cuáles son? ¿Por qué los parámetros se calculan de la forma en cómo se hace en el código?

Construye la gráfica qqplot. De manera visual, ¿Los datos siguen una distribución normal de acuerdo con la Q-Qplot?

Compara las distribuciones de probabilidad acumuladas (ojiva) empíricos y teóricos de la distribución normal que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. Explica qué son datos empíricos y datos teóricos. ¿Se parecen las distribuciones de probabilidad acumuladas?

Utiliza dos pruebas de bondad de ajuste: Shapiro-Wilks y Kolmogorov-smirnov (KS). ¿Qué información nos dan las pruebas? ¿Cuáles son los valores de los estadísticos? ¿Cuál es el p-value de las pruebas? ¿Se aceptan o se rechazan las hipótesis nulas? ¿Podemos concluir que los datos de las precipitaciones máximas mensuales son normales? ¿Por qué?

No te olvides de las hipótesis planteadas: H0: Los datos provienen de una distribución normal H1: Los datos no provienen de una distribución normal

Ajuste a la distribución normal

H_0 : Los datos provienen de una distribución normal H_1: Los datos no provienen de una distribución normal

```r
# Cargar las librerías necesarias
library(tseries)   # Para la prueba de Jarque-Bera
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

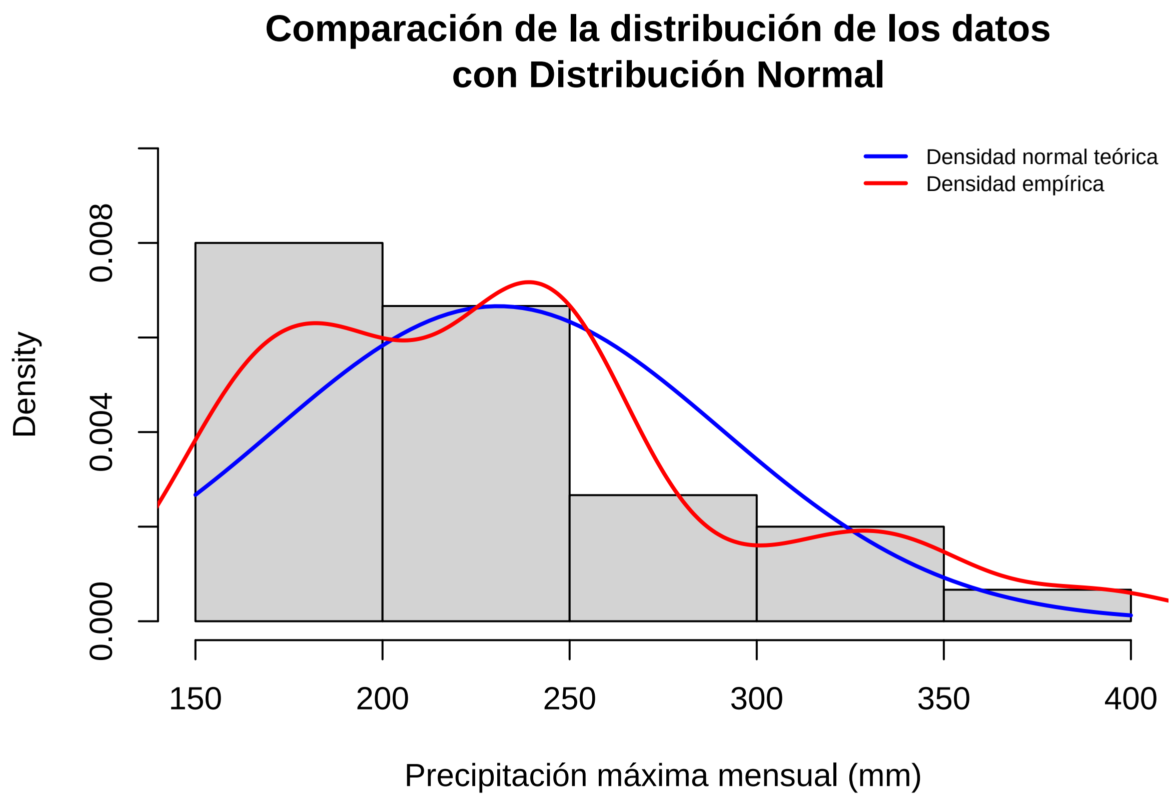
```r
library(nortest)   # Para la prueba de Anderson-Darling
library(car)       # Para la gráfica Q-Q Plot
```

```
## Loading required package: carData
```

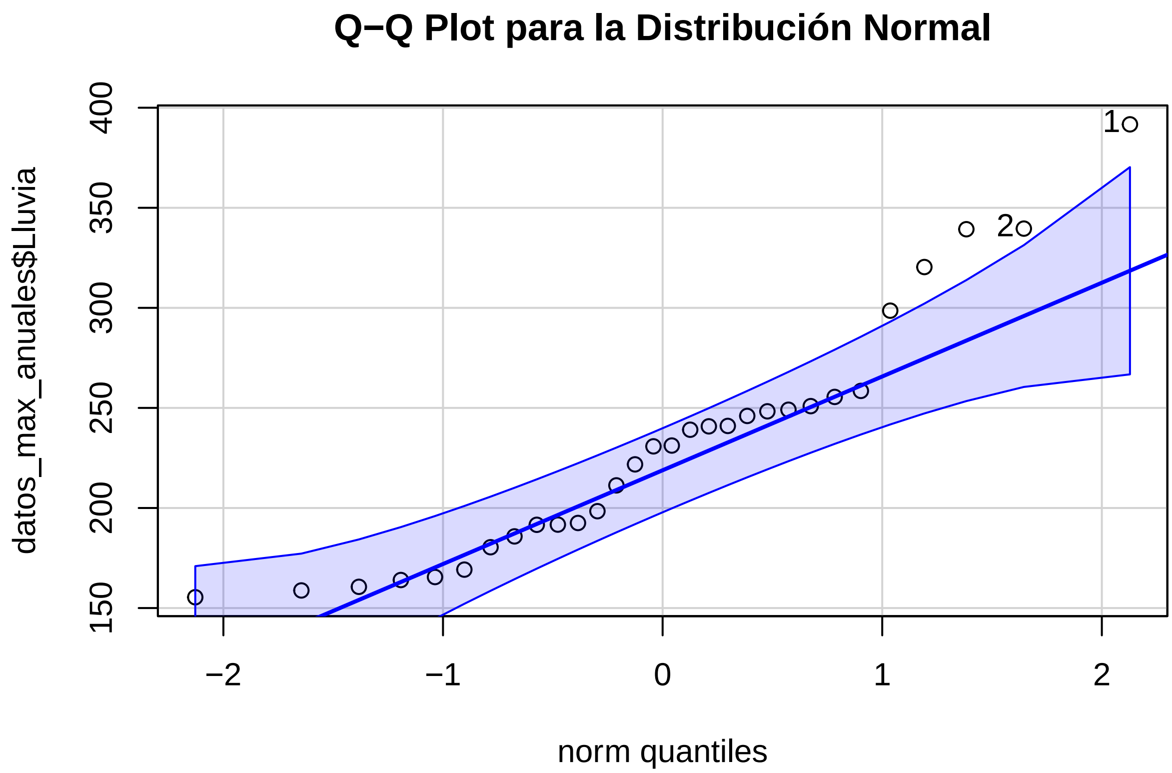
```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```r
# Calcular media y desviación estándar
mean_val <- mean(datos_max_anuales$Lluvia, na.rm = TRUE)
sd_val <- sd(datos_max_anuales$Lluvia, na.rm = TRUE)

# Histograma de densidad empírica con distribución normal teórica
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Normal")
curve(dnorm(x, mean=mean_val, sd=sd_val), add=TRUE, col="blue", lwd=2) # Densidad teórica normal
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2) # Densidad empírica
legend("topright", col=c("blue", "red"), legend=c("Densidad normal teórica", "Densidad empírica"), lwd=
```

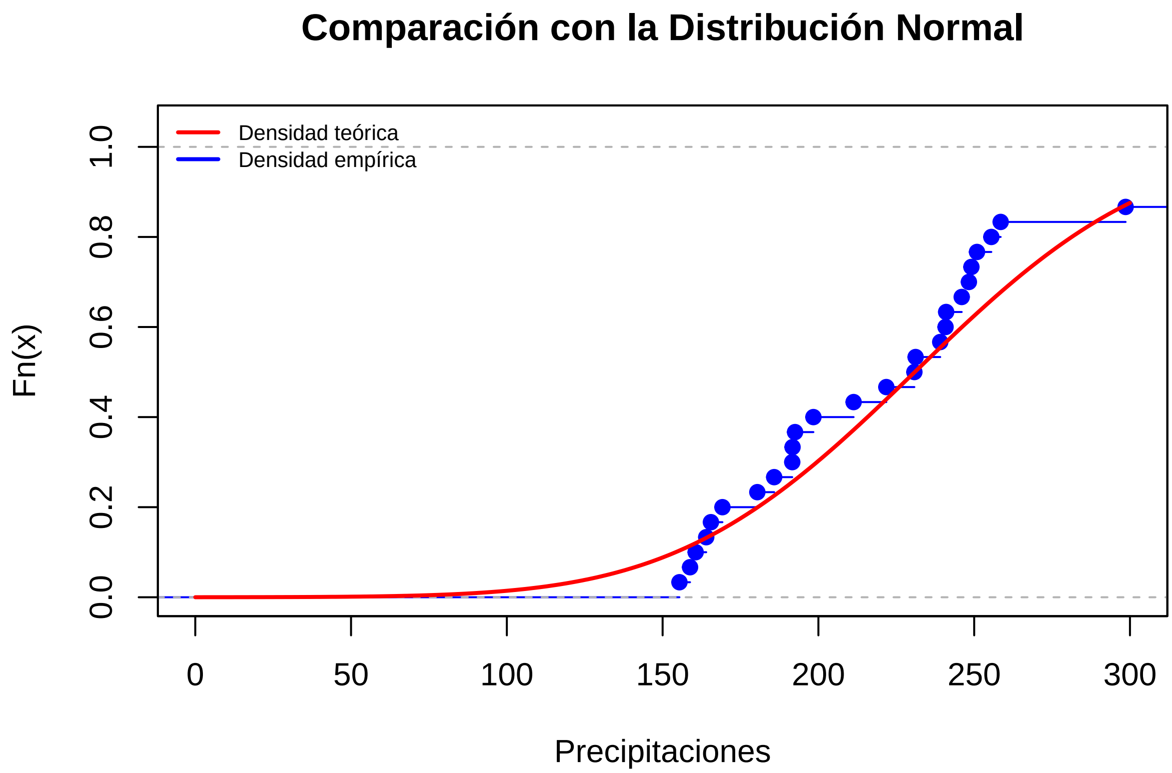



```
# Gráfica Q-Q Plot
qqPlot(datos_max_anuales$Lluvia, main="Q-Q Plot para la Distribución Normal")
```




```
## [1] 1 2
```

```
# Probabilidad acumulada empírica vs teórica
norm_teorica <- pnorm(0:300, mean=mean_val, sd=sd_val)
plot(ecdf(datos_max_anuales$Lluvia), main="Comparación con la Distribución Normal", xlab="Precipitacione
par(new=TRUE)
plot(0:300, norm_teorica, type="l", main="", xlab="", ylab="", col="red", lwd=2, ylim=c(0, 1.05), xaxt=
legend("topleft", col=c("red", "blue"), legend=c("Densidad teórica", "Densidad empírica"), lwd=2, bty="
```




```r
# Prueba de Anderson-Darling
# Prueba de Anderson-Darling
ad_test <- ad.test(datos_max_anuales$Lluvia)

# Prueba de Jarque-Bera
jb_test <- jarque.bera.test(datos_max_anuales$Lluvia)

# Resultados
ad_test
```

```
##
##  Anderson-Darling normality test
##
## data:  datos_max_anuales$Lluvia
## A = 0.78127, p-value = 0.03773
```

```r
jb_test
```

```
##
##  Jarque Bera Test
##
## data:  datos_max_anuales$Lluvia
## X-squared = 3.9844, df = 2, p-value = 0.1364
```

En el histograma, vemos la densidad empírica de nuestros datos representada por las barras (en gris) y la curva de densidad normal teórica en azul. La curva de densidad empírica (en rojo) no sigue exactamente la forma de la curva normal, especialmente en la cola derecha, donde encontramos mayor variabilidad de la esperada bajo una distribución normal. Esto nos sugiere que nuestros datos pueden no ajustarse perfectamente a una distribución normal. 2. Gráfica Q-Q Plot

En la Q-Q Plot, comparamos los cuantiles de nuestros datos empíricos con los cuantiles teóricos de una distribución normal. Idealmente, los puntos deberían alinearse sobre la línea azul si nuestros datos fueran normales. Sin embargo, notamos que se desvían, especialmente en los valores extremos. Esta desviación indica que nuestros datos tienen colas más gruesas o variabilidad en los extremos, algo no característico de una distribución normal. 3. Ojiva Comparativa (Distribución Acumulada Empírica vs Teórica)

En la gráfica de distribución acumulada, observamos que la línea azul (ojiva empírica) no coincide completamente con la curva roja de la distribución normal teórica. Las diferencias entre ambas curvas, en especial en los extremos, nos sugieren que nuestros datos no se ajustan perfectamente a una distribución normal. Esto es consistente con lo que vimos en el histograma y en la Q-Q plot.

Prueba Anderson-Darling: Obtuvimos un estadístico de 0.78127 y un p-valor de 0.03773. Como el p-valor es menor a 0.05, podemos rechazar la hipótesis nula de normalidad a un nivel de significancia del 5%. Esto indica que nuestros datos probablemente no siguen una distribución normal.

Prueba Jarque-Bera: Obtuvimos un p-valor de 0.1364 y un estadístico de 3.9844. Como el p-valor es mayor a 0.05, no tenemos suficiente evidencia para rechazar la hipótesis de normalidad. Esta prueba es menos sensible a las variaciones en los extremos, lo cual podría explicar la discrepancia con el resultado del test de Anderson-Darling.

Conclusión

En conjunto, tanto las pruebas visuales (histograma, Q-Q plot y ojiva) como la prueba de Anderson-Darling sugieren que nuestros datos de precipitaciones máximas mensuales no siguen completamente una distribución normal, debido a las variaciones en las colas. Aunque la prueba de Jarque-Bera no rechazó la normalidad, este test es menos sensible a problemas en las colas.

La evidencia sugiere que los datos no son completamente normales, especialmente en los valores extremos. Entonces, el usar una distribución normal no es la mejor opción.

B.Ajuste a una Distribución Log-Normal. Realiza un análisis visual y otro con pruebas de bondad de ajuste para determinar qué tan certera es la Distribución Log-normal para ajustar los datos.

Contruye el histograma de la función de densidad empírica de los datos y sobrepon una distribución Log-normal que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. De manera visual, ¿te parece que los datos se ajustan bien a una distribución Log-normal? Explica.

Compara las distribuciones de probabilidad acumuladas (ojiva) empíricos y teóricos de la distribución teórica que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. Explica qué son datos empíricos y datos teóricos. ¿Se parecen las distribuciones de probabilidad acumuladas?

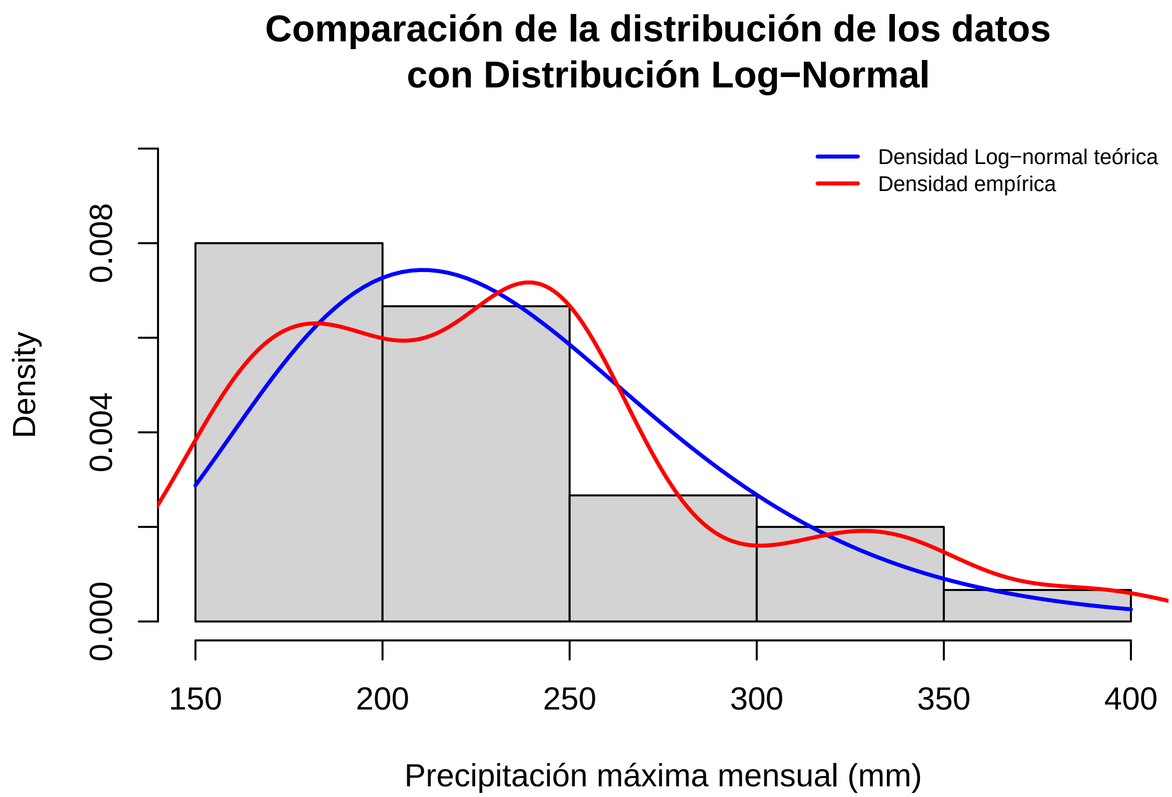
Haz la prueba KS para determinar si los datos se ajustan a una Log-normal

¿Qué información nos da la prueba KS para una Log-normal? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el valor del estadístico? ¿Cuál es el p-value de la prueba? ¿Se acepta o se rechaza la hipótesis nula? ¿Podemos concluir que los datos de las precipitaciones máximas mensuales siguen una distribución log-normal? ¿Por qué?

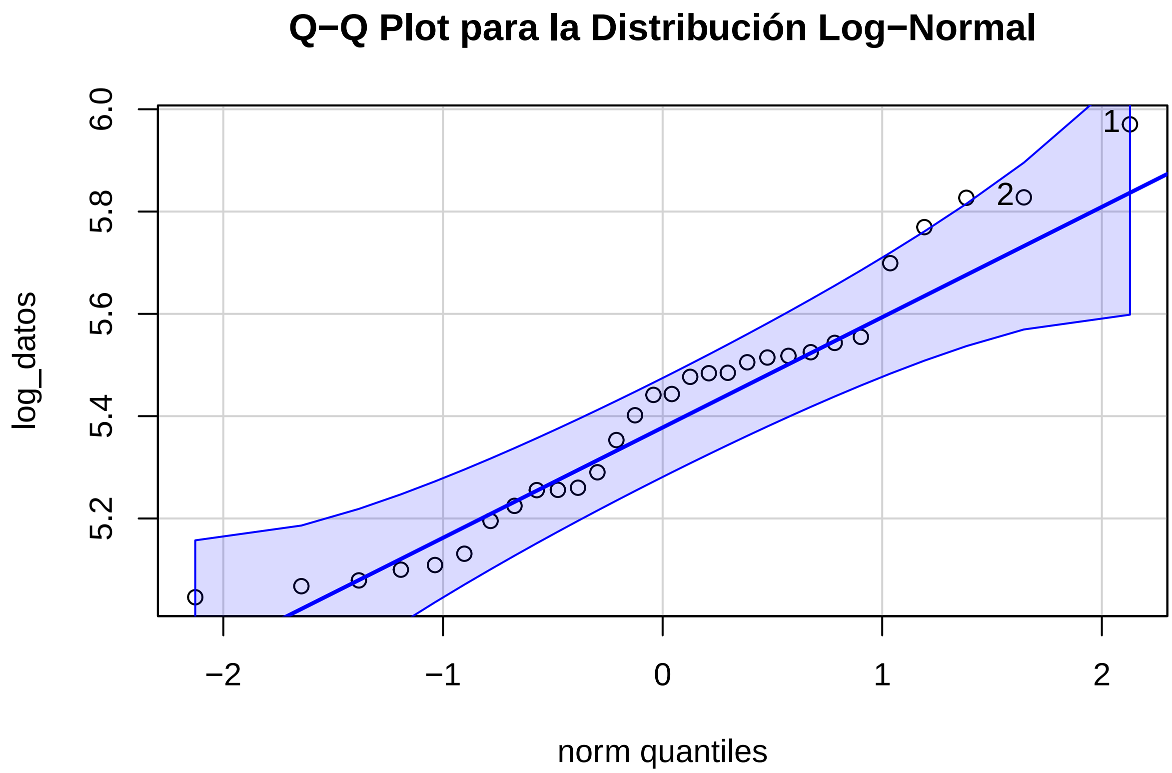
¿Cuántos parámetros tiene la distribución Log-normal? ¿Cuáles son? ¿Por qué los parámetros se calculan de la forma en cómo se hace en el código? Sigue paso a paso el método de momentos para demostrar que los parámetros están bien calculados.

Ajuste a la distribución log-normal

```
# Transformación logarítmica
log_datos <- log(datos_max_anuales$Lluvia)

# Calcular media y desviación estándar de los datos transformados (logaritmo)
meanlog_val <- mean(log_datos, na.rm = TRUE)
sdlog_val <- sd(log_datos, na.rm = TRUE)

# Histograma con densidad log-normal teórica superpuesta
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Log-Normal")
curve(dlnorm(x, meanlog=meanlog_val, sdlog=sdlog_val), add=TRUE, col="blue", lwd=2)
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2)
legend("topright", col=c("blue", "red"), legend=c("Densidad Log-normal teórica", "Densidad empírica"), 
```

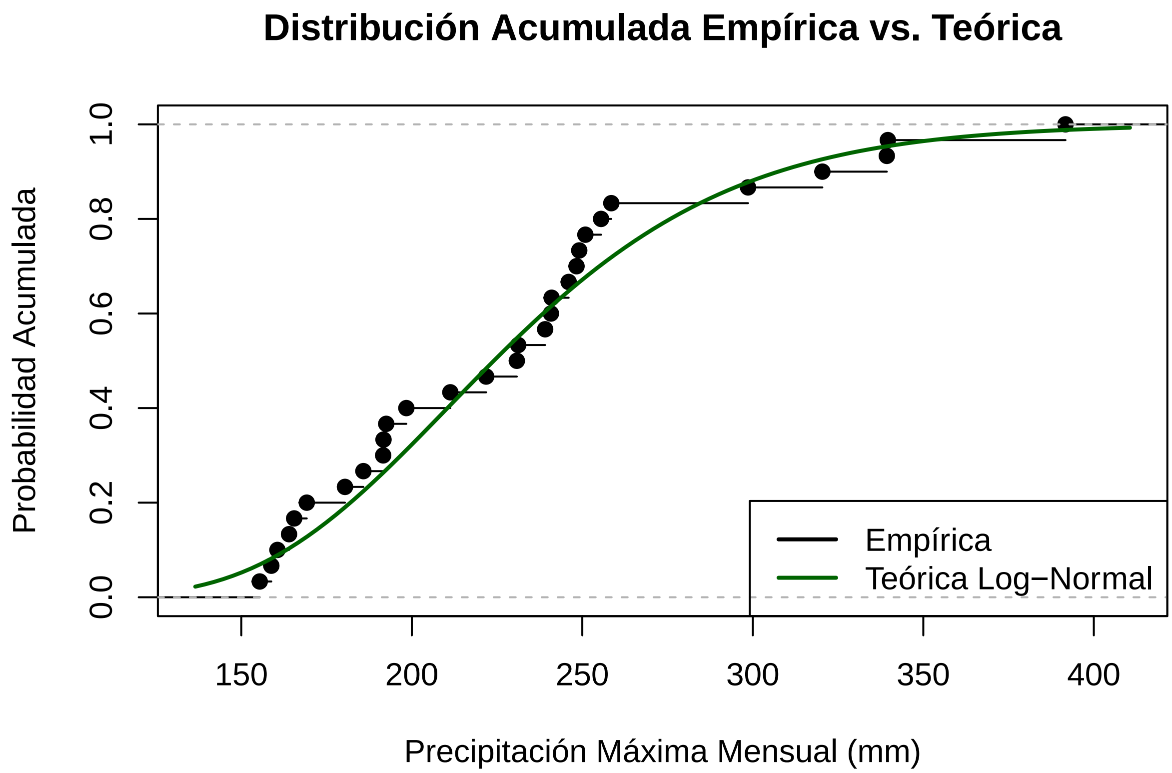



```
# Q-Q Plot para log-normalidad
qqPlot(log_datos, main="Q-Q Plot para la Distribución Log-Normal")
```




```
## [1] 1 2
```

```
# Distribución acumulada empírica
ecdf_empirica <- ecdf(datos_max_anuales$Lluvia)
plot(ecdf_empirica, main = "Distribución Acumulada Empírica vs. Teórica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad Acumulada")

# Distribución acumulada teórica log-normal
curve(plnorm(x, meanlog = meanlog_val, sdlog = sdlog_val),
      add = TRUE, col = "darkgreen", lwd = 2)
legend("bottomright", legend = c("Empírica", "Teórica Log-Normal"),
       col = c("black", "darkgreen"), lwd = 2)
```




```r
# Prueba de Kolmogorov-Smirnov para log-normal
# Añadir un pequeño ruido a los datos
ks_test <- ks.test(datos_max_anuales$Lluvia, "plnorm", meanlog = meanlog_val, sdlog = sdlog_val)

# Resultados
ks_test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datos_max_anuales$Lluvia
## D = 0.1145, p-value = 0.7849
## alternative hypothesis: two-sided
```

Construcción del histograma y ajuste visual:

```
En el histograma de la función de densidad empírica, observamos que la curva log-normal (en azul) se aj
```

Comparación de distribuciones acumuladas (ojiva):

```
La ojiva empírica representa la distribución acumulada observada de los datos, mientras que la ojiva te
Visualmente, ambas distribuciones (empírica y teórica) se parecen bastante, especialmente en el rango c
```

Prueba KS para la distribución log-normal:

```
La prueba de Kolmogorov-Smirnov (KS) se realizó para evaluar el ajuste de los datos a la distribución lo
Información de la prueba KS: La prueba KS compara la distribución empírica de los datos con la distribuc
Estadístico de prueba: Obtenemos un valor de D=0.1145.
p-value: El p-value es 0.7849.
Interpretación: Con un p-value alto (mayor a 0.05), no tenemos suficiente evidencia para rechazar la hip
```

Distribución Log-Normal: Parámetros y Cálculo

La distribución log-normal tiene dos parámetros fundamentales:

- **Media logarítmica (μ)**: es la media de los logaritmos naturales de los datos.
- **Desviación estándar logarítmica (σ)**: es la desviación estándar de los logaritmos naturales de los datos.

Estos parámetros permiten modelar una variable que solo toma valores positivos y presenta asimetría hacia la derecha, como suele ocurrir en fenómenos hidrológicos.

¿Por qué se calculan los parámetros de esta manera?

Para una variable X con distribución log-normal, al tomar el logaritmo natural de X, obtenemos una nueva variable $Y = \ln(X)$, que sigue una **distribución normal** con media μ y desviación estándar σ. Por lo tanto, al calcular la media y la desviación estándar de los logaritmos de los datos, encontramos los parámetros μ y σ que describen la distribución log-normal de X.

Método de Momentos para calcular μ y σ

El método de momentos se basa en igualar los momentos teóricos de la distribución con los momentos muestrales. Los primeros dos momentos teóricos para una variable log-normal X con media logarítmica μ y desviación estándar logarítmica σ son:

1. **Media teórica de X**:
$$E(X) = e^{\mu + \frac{\sigma^2}{2}}$$
2. **Varianza teórica de X**:
$$\mathrm{Var}(X) = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$$

Para estimar μ y σ con base en los datos, tomamos los logaritmos naturales de los valores de X y llamamos a estos valores $Y = \ln(X)$. Luego, calculamos los momentos muestrales (media y varianza) de Y:

- **Cálculo de μ**:
$$\mu = \mathrm{mean}(\ln(X))$$
- **Cálculo de σ**:
$$\sigma = \mathrm{sd}(\ln(X))$$

Aquí, μ es la media muestral de los logaritmos de los datos originales X, y σ es la desviación estándar muestral de esos logaritmos.

Demostración Paso a Paso con el Método de Momentos

1. **Paso 1**: Transformamos los datos X aplicando el logaritmo natural para obtener $Y = \ln(X)$.
2. **Paso 2**: Calculamos la **media muestral** de Y, que corresponde a μ:

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\ln(X_i)$$

3. **Paso 3**: Calculamos la **desviación estándar muestral** de Y, que corresponde a σ:

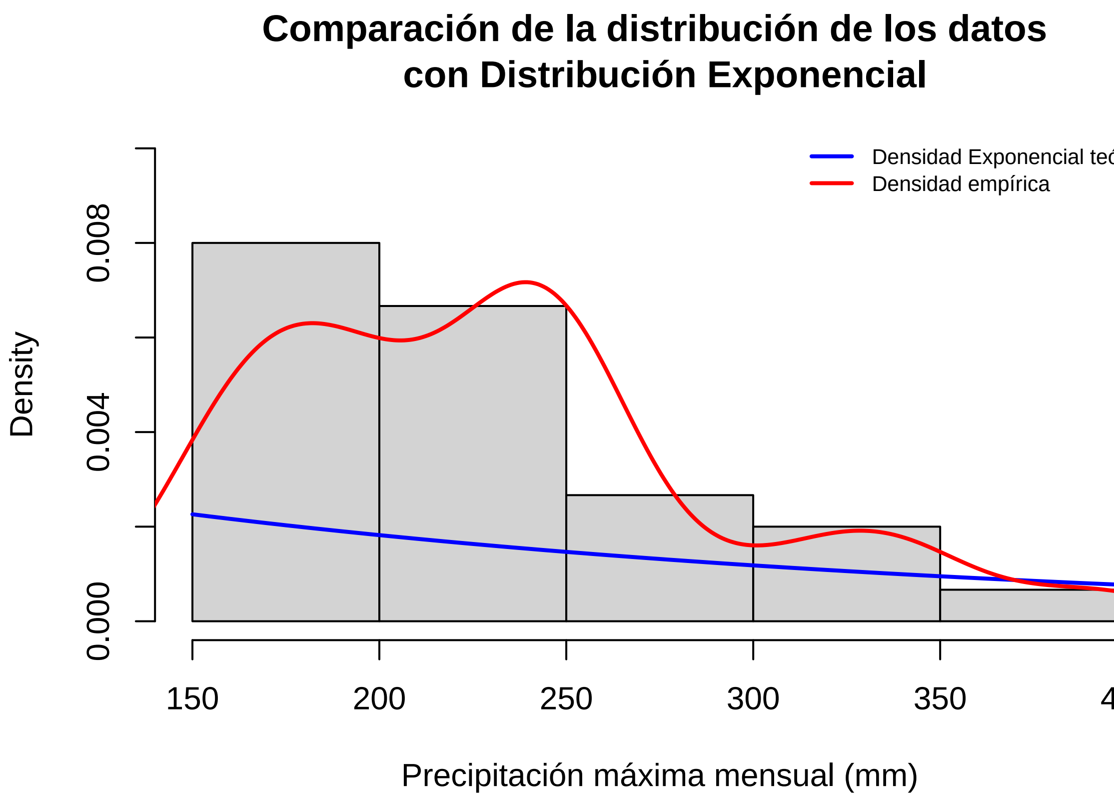
$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\ln(X_i)-\mu)^2}$$

Con estos pasos, encontramos los valores de μ y σ que mejor describen los datos bajo una distribución log-normal. Estos parámetros son los que utilizamos para trazar la curva log-normal teórica y realizar las pruebas de ajuste, asegurando que el modelo representa correctamente la estructura de nuestros datos de precipitaciones.

Ajuste a una Distribución Exponencial. Realiza un análisis visual y otro con pruebas de bondad de ajuste para determinar qué tan certera es la Distribución Exponencial para ajustar los datos.

```r
rate_val <- 1 / mean(datos_max_anuales$Lluvia, na.rm = TRUE)

# Histograma con densidad exponencial teórica superpuesta
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Exponencial")
curve(dexp(x, rate=rate_val), add=TRUE, col="blue", lwd=2)
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2)
legend("topright", col=c("blue", "red"), legend=c("Densidad Exponencial teórica", "Densidad empírica"),
```

Contruye el histograma de la función de densidad empírica de los datos y sobrepon una distribución Exponencial que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. De manera visual, ¿te parece que los datos se ajustan bien a una distribución



Exponencial? Explica.

```
rate_exp = rate_val
```

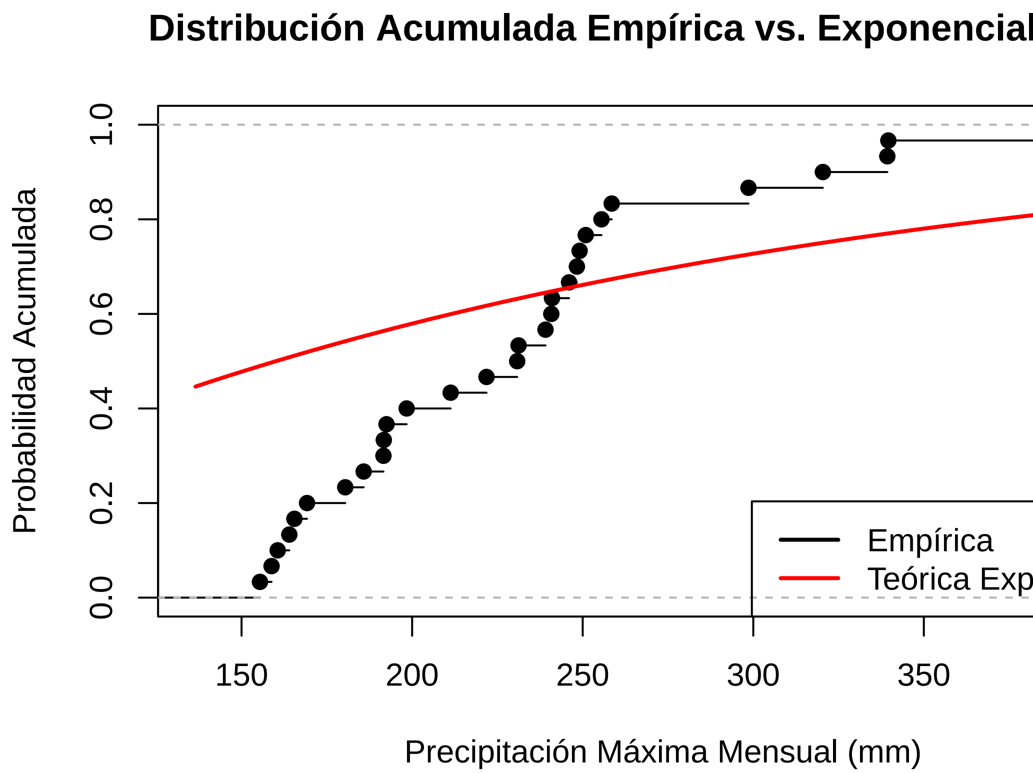
En el histograma de la función de densidad empírica de los datos, hemos sobrepuesto una curva de distribución exponencial teórica en azul. A simple vista, notamos que la distribución exponencial no se ajusta bien a los datos. Esto se observa en cómo la curva exponencial subestima los valores de densidad en la mayor parte del rango de datos, especialmente en los valores más bajos, y no logra capturar la forma general de la distribución empírica (en rojo), que es más dispersa y presenta una mayor variabilidad.

La distribución exponencial suele ser adecuada para datos que representan el tiempo entre eventos independientes que ocurren a una tasa constante, y presenta una fuerte asimetría con una sola cola decreciente. Sin embargo, en este caso, la forma de los datos de precipitación es más compleja, con mayor dispersión y variabilidad en los valores altos, lo cual no se ajusta bien al modelo exponencial.

```
# Distribución acumulada empírica
ecdf_empirica <- ecdf(datos_max_anuales$Lluvia)
plot(ecdf_empirica, main = "Distribución Acumulada Empírica vs. Exponencial Teórica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad Acumulada")

# Distribución acumulada teórica exponencial
curve(pexp(x, rate = rate_val), add = TRUE, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Teórica Exponencial"), col = c("black", "red"), lwd = 2)
```

Compara las distribuciones de probabilidad acumuladas (ojiva) empíricos y teóricos de la distribución teórica que se esperaría que tuvieran los datos con los parámetros calculados por los

mismos datos. Explica qué son datos empíricos y datos teóricos. ¿Se parecen las distribuciones



de probabilidad acumuladas?

Comparación de las Distribuciones

Al comparar ambas curvas, observamos que la distribución acumulada teórica (en rojo) se desvía considerablemente de la empírica. La curva teórica subestima la probabilidad acumulada en casi todo el rango de los datos, especialmente en la parte inferior y superior, donde las precipitaciones son más bajas o más altas. Esto sugiere que la distribución exponencial no ajusta bien nuestros datos de precipitaciones.

Haz la prueba KS para determinar si los datos se ajustan a una Exponencial ¿Qué información nos da la prueba KS para una Exponencial? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el p-value de la prueba? ¿Se acepta o se rechaza la hipótesis nula? ¿Podemos concluir que los datos de las precipitaciones máximas mensuales siguen una distribución exponencial? ¿Por qué? ¿Cuántos parámetros tiene la distribución Gamma? ¿Cuáles son? ¿Por qué los parámetros se calculan de la forma en cómo se hace en el código? Sigue paso a paso el método de momentos para demostrar que los parámetros están bien calculados.

```
# Prueba de Kolmogorov-Smirnov para distribución exponencial
ks_test <- ks.test(datos_max_anuales$Lluvia, "pexp", rate = rate_val)
# Resultados
print(ks_test)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datos_max_anuales$Lluvia
## D = 0.48979, p-value = 3.686e-07
## alternative hypothesis: two-sided
```

Resultados de la Prueba KS

```
Estadístico de prueba: El valor del estadístico D es 0.48979. Este valor mide la diferencia máxima entr

p-value: El p-value obtenido es 3.686e-07, que es un valor muy pequeño (prácticamente cercano a cero).
```

Interpretación

```
Hipótesis nula: La hipótesis nula de la prueba KS es que los datos siguen una distribución exponencial.
Decisión: Dado que el p-value es extremadamente bajo (menor que 0.05), rechazamos la hipótesis nula. Es
```

Conclusión

Podemos concluir que los datos de las precipitaciones máximas mensuales no siguen una distribución exponencial. Esto se debe a que la prueba KS muestra una gran discrepancia entre las distribuciones acumuladas empírica y teórica, respaldada por el bajo p-value, que indica una diferencia significativa entre los datos y el modelo exponencial. ## Parámetro de la Distribución Exponencial

La **distribución exponencial** tiene un solo parámetro, que es:

- **Lambda (λ)**: este es el parámetro de tasa de la distribución exponencial y representa la frecuencia de ocurrencia de un evento por unidad de tiempo o espacio. La media de una distribución exponencial es igual a $\frac{1}{\lambda}$, lo que hace que λ sea el inverso de la media.

Cálculo del Parámetro Lambda (λ)

Para una distribución exponencial, si tenemos una muestra de datos, calculamos λ a partir de la media muestral. La razón es que la media teórica de una variable aleatoria exponencial X con parámetro λ es:

$$E(X) = \frac{1}{\lambda}$$

Dado que $\lambda = \frac{1}{E(X)}$, podemos estimar λ utilizando el inverso de la media muestral. En el código, calculamos λ como:

$$\hat{\lambda} = \frac{1}{\text{mean}(X)}$$

donde X representa los valores de la muestra.

Método de Momentos para Calcular λ

El **método de momentos** nos permite estimar los parámetros de una distribución igualando los momentos teóricos con los momentos de la muestra.

1. **Paso 1**: Para una distribución exponencial, el primer momento (media teórica) es $E(X) = \frac{1}{\lambda}$.
2. **Paso 2**: Calculamos la media muestral $\bar{X}$ de nuestros datos, que es el estimador del primer momento.
3. **Paso 3**: Igualamos el momento teórico con el momento muestral:
$$\bar{X} = \frac{1}{\lambda}$$
4. **Paso 4**: Despejamos λ para obtener la estimación del parámetro de tasa:
$$\lambda = \frac{1}{\bar{X}}$$

Interpretación

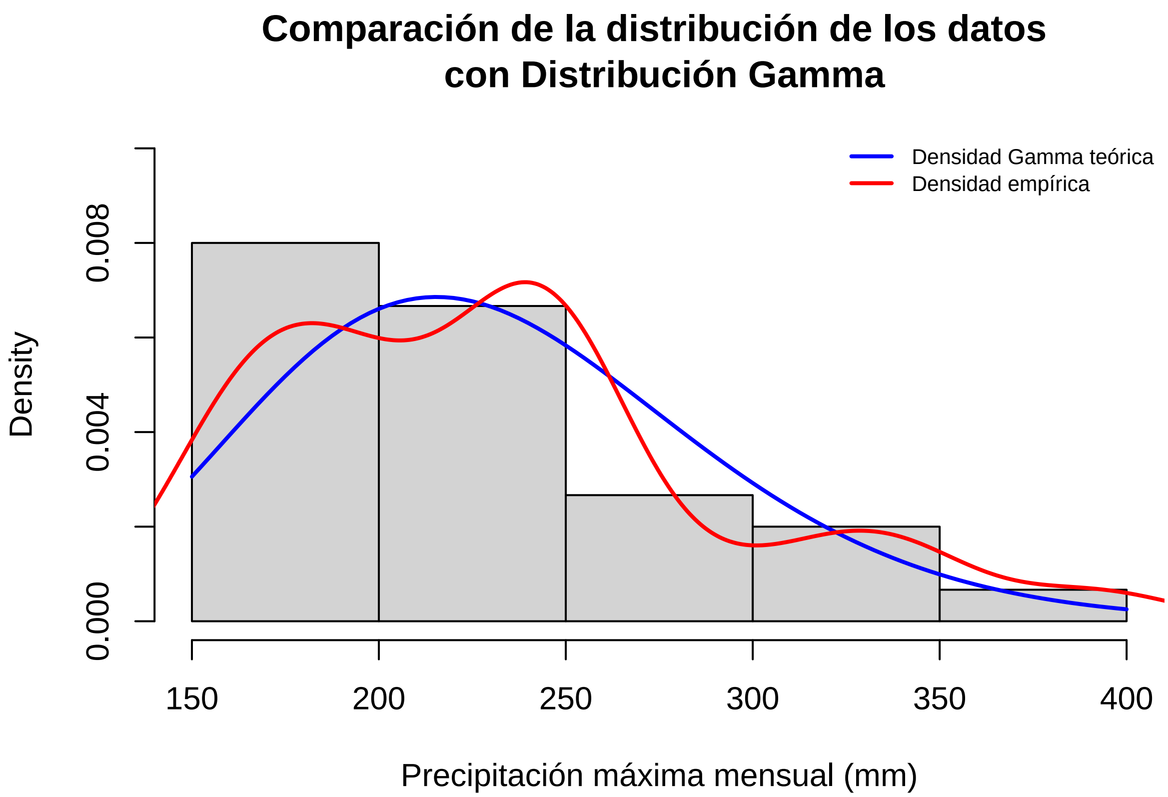
Este método de momentos nos muestra que podemos estimar λ simplemente tomando el inverso de la media muestral. En este caso, calcular λ de esta manera es consistente con la definición de la media en una distribución exponencial y asegura que el parámetro refleja la frecuencia de ocurrencia observada en los datos.

Por lo tanto, el parámetro λ está bien calculado siguiendo este procedimiento, proporcionando una estimación directa y sencilla de la tasa de eventos en nuestros datos.

D. Ajuste a una Distribución Gamma. Realiza un análisis visual y otro con pruebas de bondad de ajuste para determinar qué tan certera es la Distribución Gamma para ajustar los datos.

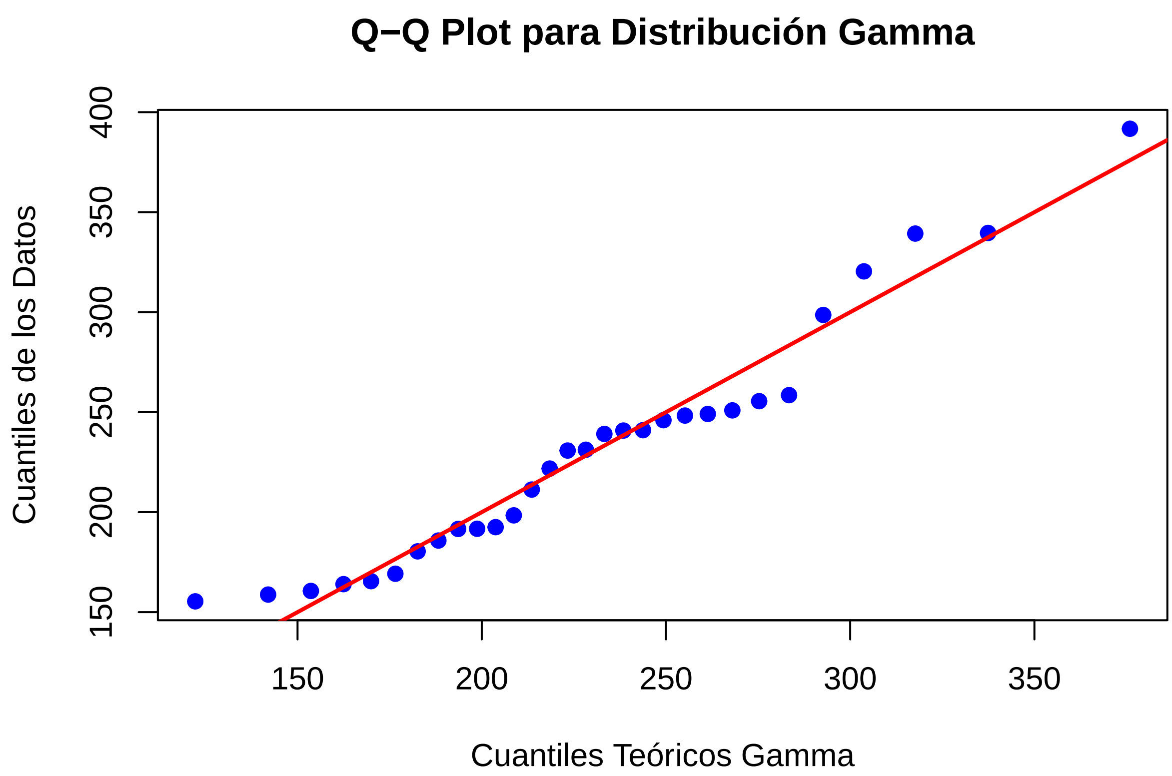
```
#Parámetros para la distribución Gamma (método de momentos)
shape_val <- mean(datos_max_anuales$Lluvia, na.rm = TRUE)^2 / var(datos_max_anuales$Lluvia, na.rm = TRU
scale_val <- var(datos_max_anuales$Lluvia, na.rm = TRUE)  / mean(datos_max_anuales$Lluvia, na.rm = TRUE

# Histograma con densidad gamma teórica superpuesta
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Gamma")
curve(dgamma(x, shape=shape_val, rate=1/scale_val), add=TRUE, col="blue", lwd=2)
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2)
legend("topright", col=c("blue", "red"), legend=c("Densidad Gamma teórica", "Densidad empírica"), lwd=2
```

Contruye el histograma de la función de densidad empírica de los datos y sobrepon una distribución Gamma que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. De manera visual, ¿te parece que los datos se ajustan bien a una distribución



Gamma? Explica.

```
# Generar el Q-Q Plot para la Distribución Gamma
qqplot(qgamma(ppoints(length(datos_max_anuales$Lluvia)), shape = shape_val, scale = scale_val),
       sort(datos_max_anuales$Lluvia),
       main = "Q-Q Plot para Distribución Gamma",
       xlab = "Cuantiles Teóricos Gamma",
       ylab = "Cuantiles de los Datos",
       pch = 19, col = "blue")

# Agregar la línea de identidad (para comparación visual)
abline(0, 1, col = "red", lwd = 2)
```




```
shape_gamma = shape_val
scale_gamma = scale_val
```

Histograma y Densidad Teórica Gamma

En el histograma, sobrepusimos la densidad teórica de la distribución Gamma (curva azul) junto a la densidad empírica (curva roja). Observamos que la distribución Gamma logra capturar de manera razonable la asimetría de los datos y la variabilidad en la cola derecha. La densidad teórica sigue en gran medida la forma de la densidad empírica, aunque con algunas pequeñas desviaciones en ciertos intervalos, especialmente en las fluctuaciones de la densidad empírica alrededor de los 250 mm de precipitación. Gráfico Q-Q Plot para Distribución Gamma

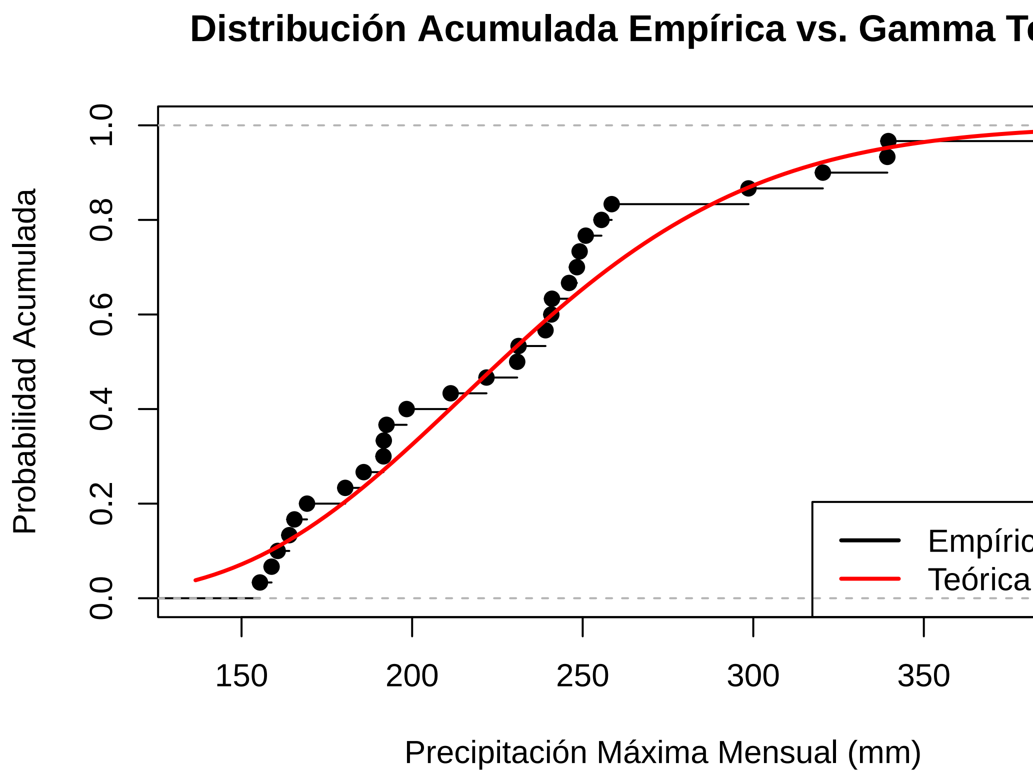
En el Q-Q Plot, comparamos los cuantiles teóricos de la distribución Gamma con los cuantiles empíricos de los datos. La mayoría de los puntos se alinean bien sobre la línea de referencia, lo que sugiere que la distribución Gamma se ajusta bien a los datos. Hay ligeras desviaciones en los valores extremos, pero en general, el Q-Q Plot respalda que los datos siguen una distribución Gamma. Conclusión

Tanto el histograma con la densidad teórica como el Q-Q Plot indican que la distribución Gamma es una buena opción para modelar los datos de precipitaciones máximas mensuales, capturando tanto la asimetría como la dispersión observada en los datos. Aunque no es un ajuste perfecto, es una aproximación razonable que representa bien la estructura general de los datos.

```
# Distribución acumulada empírica
ecdf_empirica <- ecdf(datos_max_anuales$Lluvia)
```

```r
plot(ecdf_empirica, main = "Distribución Acumulada Empírica vs. Gamma Teórica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad Acumulada")

# Distribución acumulada teórica Gamma
curve(pgamma(x, shape = shape_val, scale = scale_val), add = TRUE, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Teórica Gamma"), col = c("black", "red"), lwd = 2)
```

Compara las distribuciones de probabilidad acumuladas (ojiva) empíricos y teóricos de la distribución teórica que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. Explica qué son datos empíricos y datos teóricos. ¿Se parecen las distribuciones de probabilidad acumuladas?



Datos empíricos: La distribución empírica representa la probabilidad acumulada observada directamente en los datos de precipitaciones máximas mensuales. Es una acumulación de los datos reales sin ningún supuesto de modelo.

Datos teóricos: La distribución teórica representa la probabilidad acumulada que esperaríamos bajo el supuesto de que los datos siguen una distribución Gamma. Esta curva teórica se calcula utilizando los parámetros de la distribución Gamma obtenidos a partir de los datos.

Comparación de las Distribuciones

Podemos ver que las distribuciones acumuladas empírica y teórica son bastante similares en la mayor parte del rango de valores. La curva teórica Gamma sigue de cerca la forma de la distribución empírica, capturando bien la tendencia de acumulación. Hay pequeñas diferencias en ciertos puntos, especialmente en los valores extremos, pero en general, ambas distribuciones están bastante alineadas. Conclusión

Las distribuciones de probabilidad acumuladas empírica y teórica se parecen bastante, lo que sugiere que la distribución Gamma es un modelo adecuado para los datos de precipitaciones máximas mensuales. La similitud entre ambas curvas indica que el modelo teórico Gamma puede representar bien la estructura acumulativa observada en los datos. #### Haz la prueba KS para determinar si los datos se ajustan a una Gamma

```
# Prueba de Kolmogorov-Smirnov para distribución Gamma
ks_test <- ks.test(datos_max_anuales$Lluvia, "pgamma", shape = shape_val, scale = scale_val)

# Resultados
print(ks_test)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datos_max_anuales$Lluvia
## D = 0.13171, p-value = 0.6282
## alternative hypothesis: two-sided
```

¿Qué información nos da la prueba KS para una distribución Gamma? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el p-value de la prueba? ¿Se acepta o se rechaza la hipótesis nula? ¿Podemos concluir que los datos de las precipitaciones máximas mensuales siguen una distribución Gamma? ¿Por qué? ¿Cuántos parámetros tiene la distribución Gamma? ¿Cuáles son? ¿Por qué los parámetros se calculan de la forma en cómo se hace en el código? Sigue paso a paso el método de momentos para demostrar que los parámetros están bien calculados. La prueba de Kolmogorov-Smirnov (KS) compara la distribución empírica de nuestros datos con la distribución teórica Gamma para evaluar si los datos siguen dicha distribución.

- **Estadístico de prueba**: El valor del estadístico D es **0.13171**. Este estadístico mide la diferencia máxima entre la función de distribución acumulada empírica y la teórica.

- **p-value**: El p-value obtenido es **0.6282**, un valor considerablemente alto.

Interpretación

- **Hipótesis nula**: La hipótesis nula de la prueba KS es que los datos siguen una distribución Gamma.
- **Decisión**: Dado que el p-value es significativamente mayor que 0.05, no rechazamos la hipótesis nula. Esto sugiere que los datos no presentan diferencias significativas con la distribución Gamma, lo que respalda la adecuación del modelo.

Conclusión

Podemos concluir que los datos de las precipitaciones máximas mensuales **siguen una distribución Gamma** con un buen nivel de ajuste, dado que la prueba KS no rechaza la hipótesis de ajuste a la distribución Gamma.

Parámetros de la Distribución Gamma

La distribución Gamma tiene **dos parámetros**:

1. **Forma (k)**: Describe la forma de la distribución.
2. **Tasa (θ)**: Representa la escala de la distribución, que controla la dispersión de los valores.

Cálculo de los Parámetros: Método de Momentos

El **método de momentos** es una técnica que estima los parámetros igualando los momentos teóricos con los momentos de la muestra.

Paso a Paso para Calcular k y θ

1. **Media teórica de X** para la distribución Gamma:

$$E(X) = k\theta$$

2. **Varianza teórica de X**:

$$\text{Var}(X) = k\theta^2$$

3. **Paso 1**: Calcular la media muestral $\bar{X}$ y la varianza muestral s^2 de los datos.
4. **Paso 2**: Igualar la media muestral con la media teórica para obtener el primer parámetro:

$$\bar{X} = k\theta$$

5. **Paso 3**: Igualar la varianza muestral con la varianza teórica para obtener el segundo parámetro:

$$s^2 = k\theta^2$$

6. **Despejar los parámetros** k y θ:

   - Resolviendo estas ecuaciones, obtenemos:

$$k = \frac{\bar{X}^2}{s^2}$$

$$\theta = \frac{s^2}{\bar{X}}$$

Este método de momentos nos asegura que los parámetros k y θ se calculan correctamente y representan de manera adecuada la estructura de los datos bajo una distribución Gamma.

Ajuste a una Distribución Weibull. Realiza un análisis visual y otro con pruebas de bondad de ajuste para determinar qué tan certera es la Distribución Weibull para ajustar los datos.

El cálculo de los parámetros a partir de los datos es un poco más difícil en la distribución Weibull de lo que fue en las anteriores distribuciones, así que recurriremos a que R los estime con el comando fitdistr. Úsalo para estimar los parámetros de la Weibull.

```
# Cargar librerías necesarias
library(fitdistrplus)
```

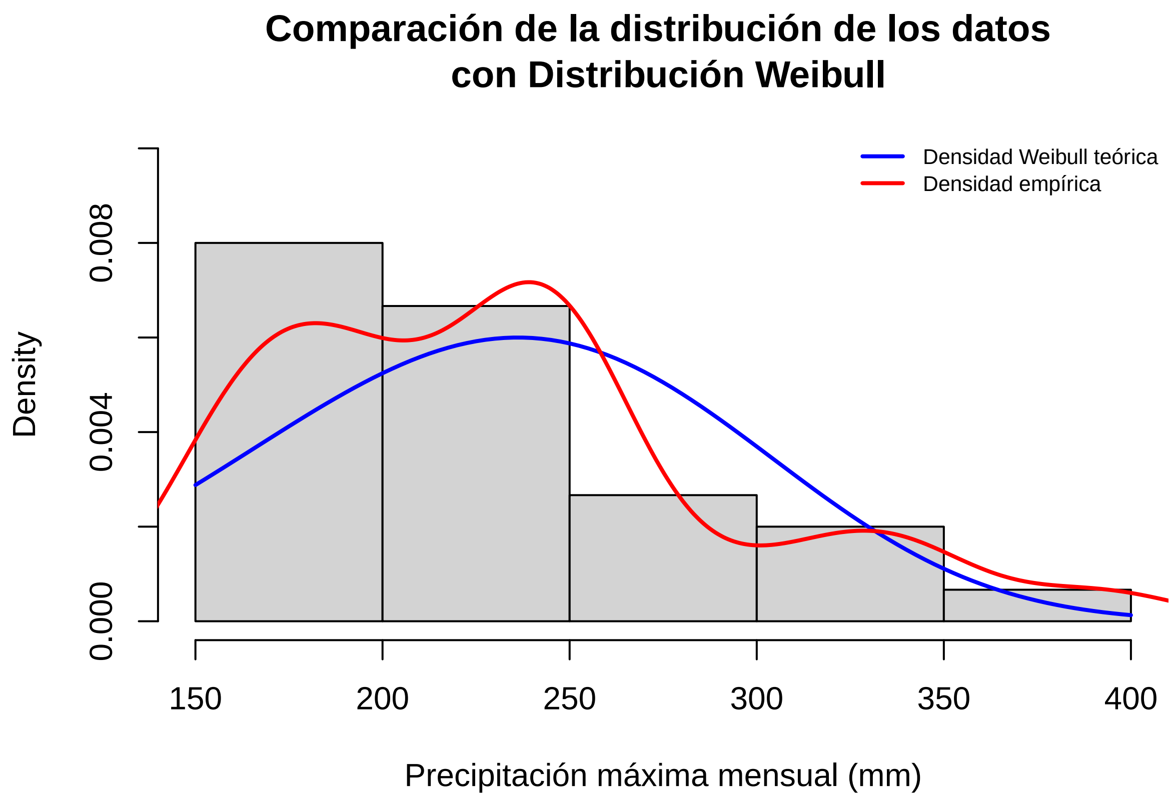
Contruye el histograma de la función de densidad empírica de los datos y sobrepon una distribución Weibull que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. De manera visual, ¿te parece que los datos se ajustan bien a una distribución Weibull? Explica.

```
## Loading required package: MASS

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select

## Loading required package: survival
```

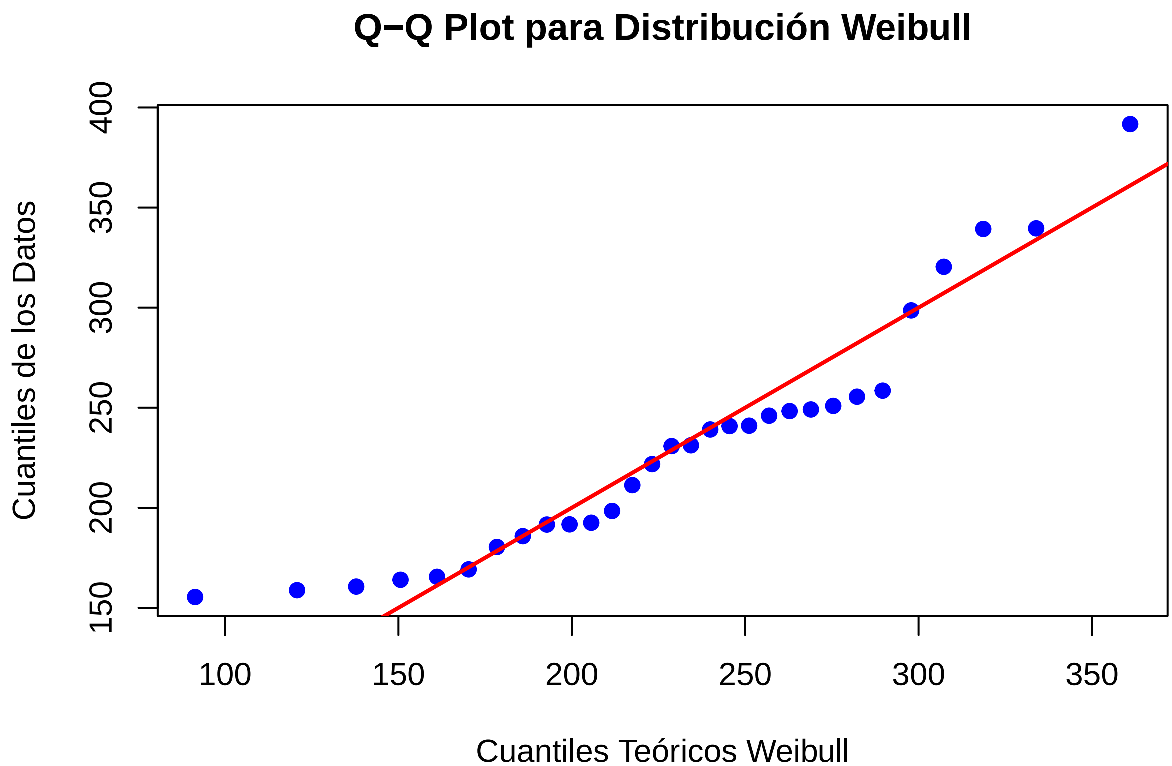
```
library(MASS)
library(ggplot2)
library(car)

# Estimar los parámetros de la distribución Weibull usando fitdistr
params_weibull <- fitdistr(datos_max_anuales$Lluvia, "weibull")
shape_weibull <- params_weibull$estimate[1]  # Parámetro de forma
scale_weibull <- params_weibull$estimate[2]  # Parámetro de escala

# Mostrar los parámetros estimados
print(params_weibull)
```

```
##       shape         scale
##     3.9995785   253.7913239
##   (  0.5266939) ( 12.3059996)
```

```
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Weibull")
curve(dweibull(x, shape=shape_weibull, scale=scale_weibull), add=TRUE, col="blue", lwd=2)
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2)
legend("topright", col=c("blue", "red"), legend=c("Densidad Weibull teórica", "Densidad empírica"), lwd=
```




```r
# Generar el Q-Q Plot para la Distribución Weibull
qqplot(qweibull(ppoints(length(datos_max_anuales$Lluvia)), shape = shape_weibull, scale = scale_weibull),
       sort(datos_max_anuales$Lluvia),
       main = "Q-Q Plot para Distribución Weibull",
       xlab = "Cuantiles Teóricos Weibull",
       ylab = "Cuantiles de los Datos",
       pch = 19, col = "blue")

# Agregar la línea de identidad (para comparación visual)
abline(0, 1, col = "red", lwd = 2)
```



Análisis de Ajuste de la Distribución Weibull

En este análisis, evaluamos si los datos se ajustan bien a una distribución Weibull usando estimaciones de parámetros y herramientas gráficas. 1. Estimación de Parámetros

Usamos el comando fitdistr de R para estimar los parámetros de la distribución Weibull:

```
Shape (forma): 3.9996
Scale (escala): 253.79
```

Estos valores definen la forma y la dispersión de la curva teórica Weibull que usaremos en el análisis visual. 2. Análisis Visual con el Histograma y Densidad Teórica

En el histograma, la densidad teórica de la distribución Weibull (curva azul) se superpone a la densidad empírica (curva roja). Visualmente, la curva teórica sigue la forma general de los datos, aunque hay algunas diferencias menores en ciertos intervalos. 3. Q-Q Plot

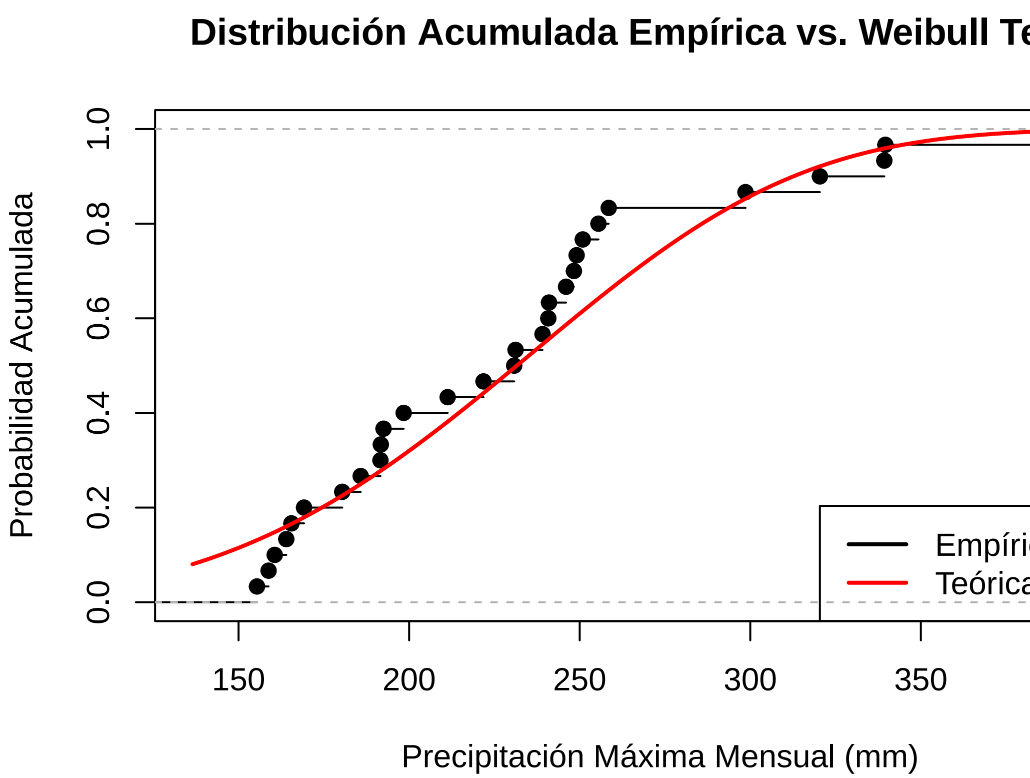
El Q-Q Plot nos ayuda a comparar los cuantiles teóricos de la Weibull con los cuantiles empíricos de los datos. En el gráfico, los puntos se alinean razonablemente bien con la línea de referencia roja, lo que sugiere que la distribución Weibull es un ajuste adecuado para estos datos. Conclusión Visual

De manera visual, el ajuste de la distribución Weibull a los datos parece razonable, ya que tanto el histograma como el Q-Q Plot muestran que la curva teórica sigue la estructura de los datos empíricos.

```r
# Estimación de parámetros usando fitdistr para precisión
library(MASS)
```

```r
params_weibull <- fitdistr(datos_max_anuales$Lluvia, "weibull")
shape_weibull <- params_weibull$estimate["shape"]
scale_weibull <- params_weibull$estimate["scale"]

# Distribución acumulada empírica
ecdf_empirica <- ecdf(datos_max_anuales$Lluvia)
plot(ecdf_empirica, main = "Distribución Acumulada Empírica vs. Weibull Teórica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad Acumulada")

# Distribución acumulada teórica Weibull
curve(pweibull(x, shape = shape_weibull, scale = scale_weibull), add = TRUE, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Teórica Weibull"), col = c("black", "red"), lwd = 2)
```

Compara las distribuciones de probabilidad acumuladas (ojiva) empíricos y teóricos de la distribución teórica que se esperaría que tuvieran los datos con los parámetros calculados por los mismos datos. Explica qué son datos empíricos y datos teóricos. ¿Se parecen las distribuciones



de probabilidad acumuladas?

Definición de Datos Empíricos y Teóricos

```
Datos empíricos: Son los valores acumulados directamente observados en los datos de la muestra. En este

Datos teóricos: Representan la probabilidad acumulada que esperaríamos si los datos siguieran una distri
```

Comparación de las Distribuciones Acumuladas

En el gráfico, podemos ver que las distribuciones acumuladas empírica y teórica son bastante similares. La curva roja, que representa la distribución teórica Weibull, sigue de cerca la forma de la distribución empírica en negro. Aunque hay pequeñas discrepancias en algunos intervalos, en general, la distribución teórica se ajusta bien a la acumulación de los datos observados.

Haz la prueba KS para determinar si los datos se ajustan a una Weibull ¿Qué información nos da la prueba KS para una Weibull? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el p-value de la prueba? ¿Se acepta o se rechaza la hipótesis nula? ¿Podemos concluir que los datos de las precipitaciones máximas mensuales siguen una distribución Weibull? ¿Por qué? ¿Cuántos parámetros tiene la distribución Weibull? ¿Cuáles son? Explica por qué es más compleja la estimación de los parámetros a partir de los datos en esta distribución a diferencia de las distribuciones anteriores.

```r
# Definir la función de distribución acumulada Weibull con los parámetros estimados
pweibull_custom <- function(x) pweibull(x, shape = shape_weibull, scale = scale_weibull)

# Realizar la prueba KS
ks_test_weibull <- ks.test(datos_max_anuales$Lluvia, pweibull_custom)

# Mostrar resultados de la prueba KS
print(ks_test_weibull)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datos_max_anuales$Lluvia
## D = 0.17419, p-value = 0.2878
## alternative hypothesis: two-sided
```

Prueba de Kolmogorov-Smirnov (KS) para la Distribución Weibull

La prueba de Kolmogorov-Smirnov (KS) nos permite comparar la distribución empírica de nuestros datos con la distribución teórica Weibull. La hipótesis nula de esta prueba es que los datos siguen la distribución Weibull. Resultados de la Prueba KS

Estadístico de prueba D: 0.17419 p-value: 0.2878

Interpretación

```
Hipótesis nula: La hipótesis nula establece que los datos de precipitaciones máximas mensuales siguen un

Decisión: Dado que el p-value (0.2878) es mayor que 0.05, no rechazamos la hipótesis nula. Esto indica q
```

Conclusión

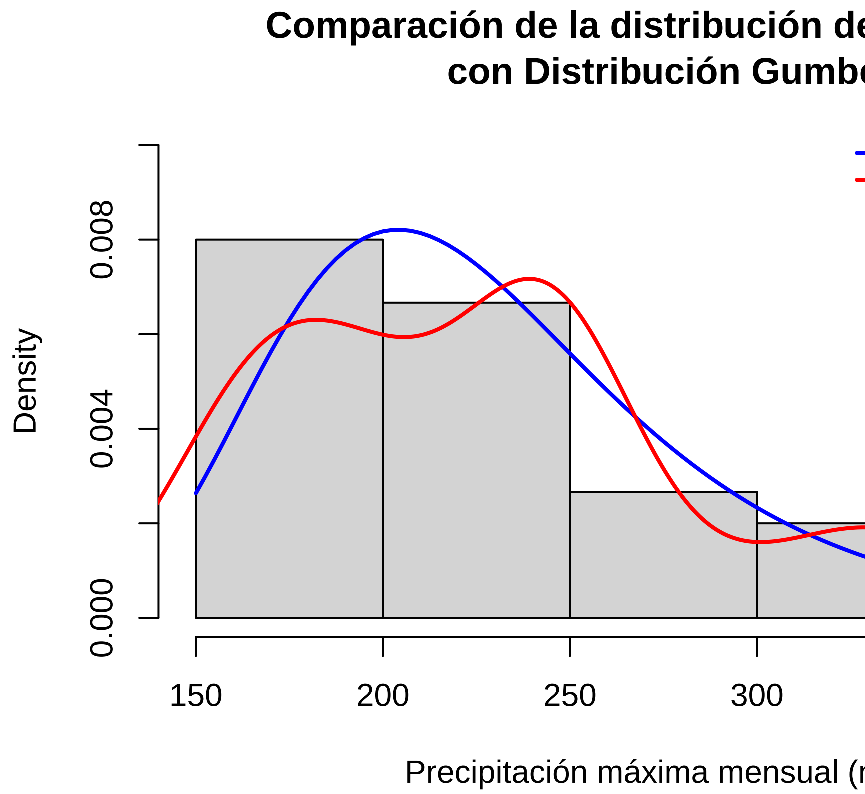
Con base en el valor del estadístico de prueba D y el p-value, podemos concluir que la distribución Weibull es un ajuste razonable para los datos de precipitaciones máximas mensuales. La prueba KS no encuentra diferencias significativas entre la distribución empírica y la teórica de Weibull, lo que respalda la adecuación de este modelo para representar los datos.

Ajuste a una Distribución Gumbel. Realiza un análisis visual y otro con pruebas de bondad de ajuste para determinar qué tan certera es la Distribución Gumbel para ajustar los datos.

Para probar si los datos de precipitación máxima se ajustan a una distribución Gumbel, se necesita definir las funciones de densidad de acuerdo con la función Gumbel. Créalas con las fórmulas de la Distribución Gumbel.

```r
### 6. Distribución Gumbel
library(fitdistrplus)
# Definir funciones de densidad y distribución acumulada para Gumbel
dgumbel <- function(x, a, b) 1/b * exp((a - x)/b) * exp(-exp((a - x)/b))
pgumbel <- function(q, a, b) exp(-exp((a - q)/b))
qgumbel <- function(p, a, b) {
  a - b * log(-log(p))
}
# Ajuste de Gumbel
gumbel_fit <- fitdist(datos_max_anuales$Lluvia, "gumbel", start = list(a=1, b=1))

# Histograma con densidad Gumbel teórica superpuesta
hist(datos_max_anuales$Lluvia, xlab="Precipitación máxima mensual (mm)", freq=FALSE, ylim=c(0, 0.01),
     main="Comparación de la distribución de los datos \n con Distribución Gumbel")
curve(dgumbel(x, a=gumbel_fit$estimate[1], b=gumbel_fit$estimate[2]), add=TRUE, col="blue", lwd=2)
lines(density(datos_max_anuales$Lluvia), col="red", lwd=2)
legend("topright", col=c("blue", "red"), legend=c("Densidad Gumbel teórica", "Densidad empírica"), lwd=2
```

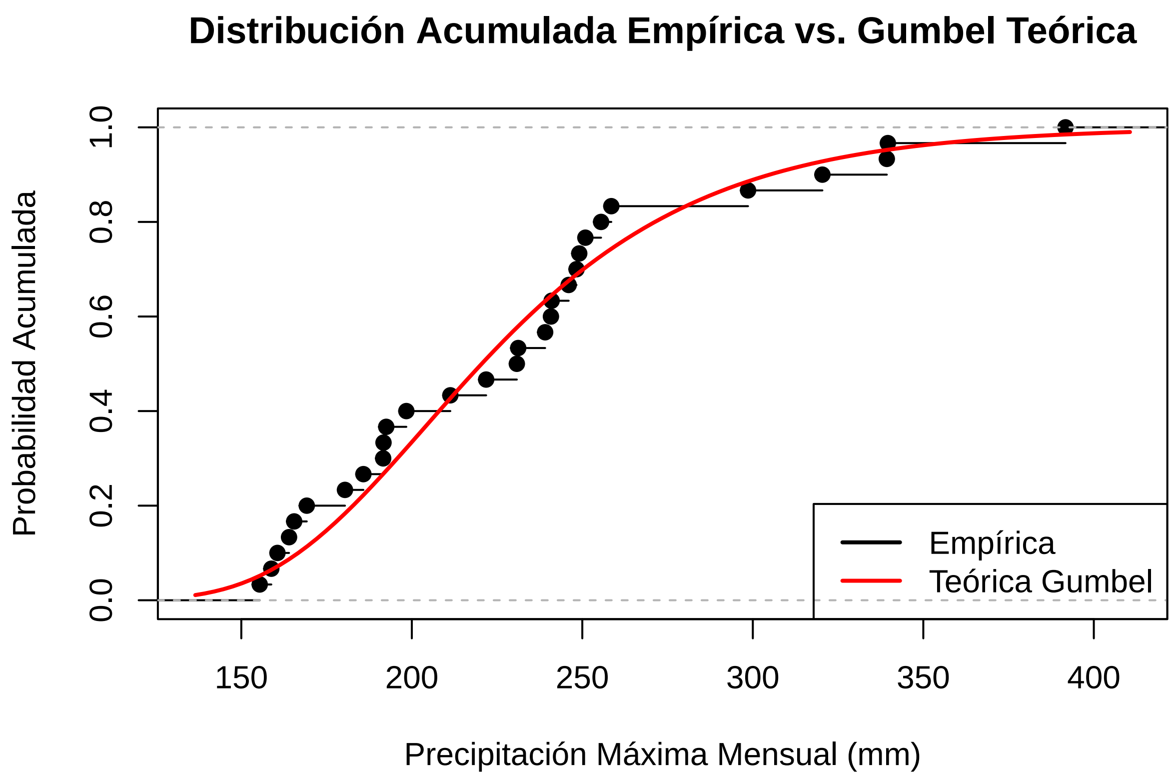
Para estimar los parámetros y hacer el ajuste de la Distribución Gumbel con la biblioteca "fitdistrplus". Haz las gráficas de histograma de densidad empírica y teórica, la probabilidad de



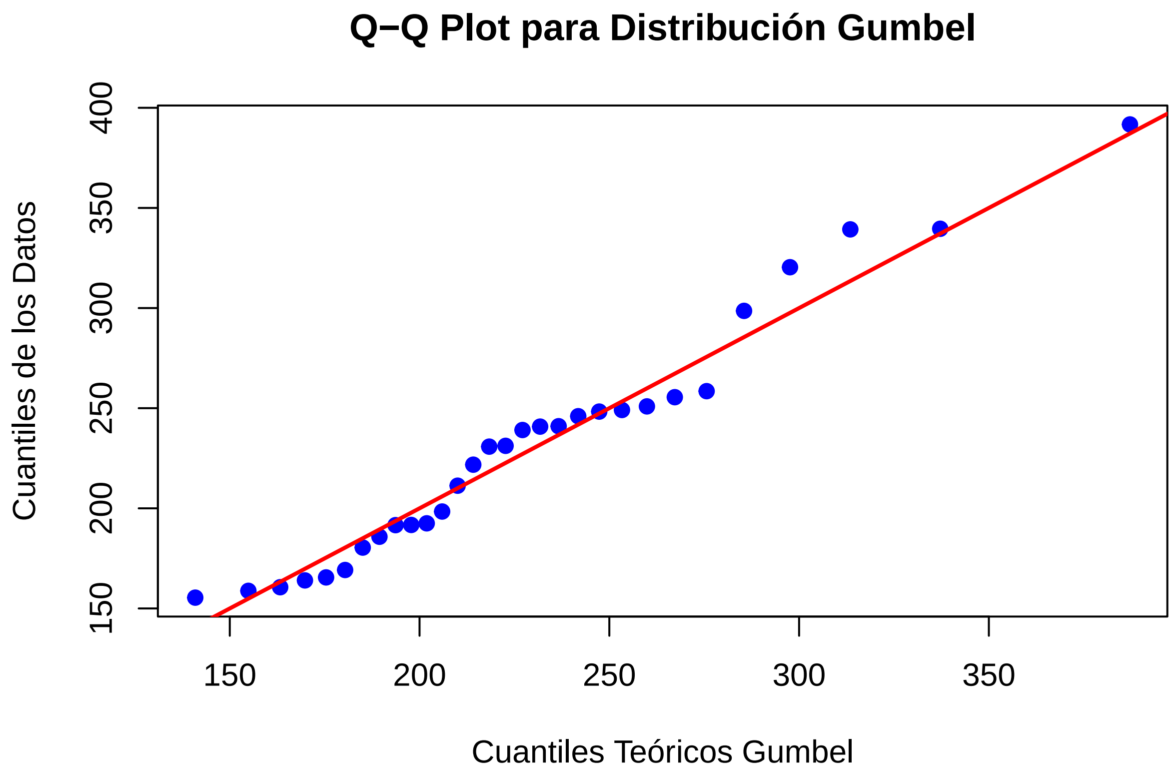
acumulada empírica y teórica y el QQplot.

```r
ecdf_empirica <- ecdf(datos_max_anuales$Lluvia)
plot(ecdf_empirica, main = "Distribución Acumulada Empírica vs. Gumbel Teórica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad Acumulada")
```

```r
# Superponer la curva de la distribución acumulada teórica Gumbel
curve(pgumbel(x, a = gumbel_fit$estimate[1], b = gumbel_fit$estimate[2]), add = TRUE, col = "red", lwd =
legend("bottomright", legend = c("Empírica", "Teórica Gumbel"), col = c("black", "red"), lwd = 2)
```




```r
# Generar el Q-Q Plot para la Distribución Gumbel
# Calcular los cuantiles teóricos usando la función de cuantiles de Gumbel
qqplot(qgumbel(ppoints(length(datos_max_anuales$Lluvia)), a = gumbel_fit$estimate[1], b = gumbel_fit$es
       sort(datos_max_anuales$Lluvia),
       main = "Q-Q Plot para Distribución Gumbel",
       xlab = "Cuantiles Teóricos Gumbel",
       ylab = "Cuantiles de los Datos",
       pch = 19, col = "blue")

# Agregar la línea de identidad (para comparar los cuantiles teóricos y empíricos)
abline(0, 1, col = "red", lwd = 2)
```



Análisis Visual

```
Histograma con Densidad Teórica Superpuesta:
    La gráfica muestra el histograma de la densidad empírica de los datos de precipitación máxima mensua
    Visualmente, la curva teórica (en azul) sigue en general la tendencia de los datos empíricos (en ro

Distribución Acumulada Empírica vs. Teórica:
    La gráfica compara la distribución acumulada empírica con la distribución acumulada teórica de Gumbe
    Las distribuciones empírica y teórica se parecen bastante y muestran una buena alineación, especialm

Q-Q Plot:
    El Q-Q Plot compara los cuantiles teóricos de la distribución Gumbel con los cuantiles empíricos de
    La mayoría de los puntos caen sobre la línea de identidad, lo que indica que los datos se ajustan ra
```

Conclusión

En general, los análisis visuales indican que la distribución Gumbel es un buen modelo para representar los datos de precipitación máxima mensual. Las gráficas de densidad, probabilidad acumulada y el Q-Q Plot sugieren que la Gumbel captura adecuadamente la estructura de los datos, aunque existen pequeñas discrepancias en las colas.

Haz la prueba KS para determinar si los datos se ajustan a una Gumbel Qué información nos da la prueba KS para una Gumbel? ¿Cuál es el valor del estadístico de prueba? ¿Cuál es el p-value de la prueba? ¿Se acepta o se rechaza la hipótesis nula? ¿Podemos concluir que las probabilidades de excedencia de las precipitaciones máximas mensuales siguen una distribución Gumbel? ¿Por qué? ¿Cuántos parámetros tiene la distribución Gumbel? ¿Cuáles son? Estima los parámetros de la Gumbel partiendo de la media y la desviación estándar de los datos y con la fórmula de la media y la desviación estándar de la Gumbel. Compara los valores obtenidos con los estimados con el comando "fitdistrplus" ¿se obtienen los mismos valores por ambos métodos? ¿por qué crees que se dé esta diferencia? ¿por qué crees que se dé esta diferencia?

```
ks_result <- ks.test(datos_max_anuales$Lluvia, "pgumbel", a = gumbel_fit$estimate[1], b = gumbel_fit$es

# Mostrar resultados de la prueba KS
print(ks_result)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datos_max_anuales$Lluvia
## D = 0.10983, p-value = 0.8241
## alternative hypothesis: two-sided
```

Prueba KS para Distribución Gumbel

```
Resultado de la Prueba KS:
    Estadístico de prueba DD: 0.10903
    p-value: 0.8241

Interpretación de la Prueba KS:
    La prueba KS compara la distribución acumulada empírica de los datos de precipitación con la distrib
    Dado que el p-value es mucho mayor que 0.05, no rechazamos la hipótesis nula, lo que indica que no h

Conclusión:
    Con base en el p-value alto, podemos concluir que las probabilidades de excedencia de las precipitac
```

```
# Calcular media y desviación estándar de los datos
media_datos <- mean(datos_max_anuales$Lluvia)
desviacion_datos <- sd(datos_max_anuales$Lluvia)

# Calcular los parámetros de Gumbel (a y b) a partir de la media y desviación estándar
gamma <- 0.5772
b_est <- desviacion_datos * sqrt(6) / pi
a_est <- media_datos - b_est * gamma

# Mostrar los parámetros calculados manualmente
cat("Parámetros calculados manualmente:\n")
```

```
## Parámetros calculados manualmente:
```

```
cat("a (Ubicación) =", a_est, "\n")
```

```
## a (Ubicación) = 203.9791
```

```
cat("b (Escala) =", b_est, "\n")
```

```
## b (Escala) = 46.68678
```

```r
# Comparar con los parámetros obtenidos con fitdistrplus
cat("Parámetros estimados con fitdistrplus:\n")
```

```
## Parámetros estimados con fitdistrplus:
```

```r
print(gumbel_fit$estimate)
```

```
##         a         b
## 204.07290  44.81446
```

Parámetros de la Distribución Gumbel

La distribución Gumbel tiene dos parámetros: - **Parámetro de Ubicación (location)**: denotado como a, que define el punto central de la distribución. - **Parámetro de Escala (scale)**: denotado como b, que determina la dispersión de la distribución.

Estimación de los Parámetros de Gumbel

Para estimar los parámetros a y b de la distribución Gumbel a partir de la media (μ) y desviación estándar (σ) de los datos, utilizamos las siguientes fórmulas:

- **Media de Gumbel**: $\mu = a + b \cdot \gamma$, donde $\gamma \approx 0.5772$ es la constante de Euler-Mascheroni.
- **Desviación Estándar de Gumbel**: $\sigma = \frac{\pi b}{\sqrt{6}}$.

Para obtener los valores de a y b, resolvemos estas ecuaciones en función de μ y σ:

- **Cálculo de b**: $b = \frac{\sigma \cdot \sqrt{6}}{\pi}$
- **Cálculo de a**: $a = \mu - b \cdot \gamma$

Resultados

1. **Parámetros Calculados Manualmente**:
   - a (Ubicación) = 203.9791
   - b (Escala) = 46.68678
2. **Parámetros Estimados con `fitdistrplus`**:
   - a (Ubicación) = 204.07290
   - b (Escala) = 44.81446

Comparación de Métodos

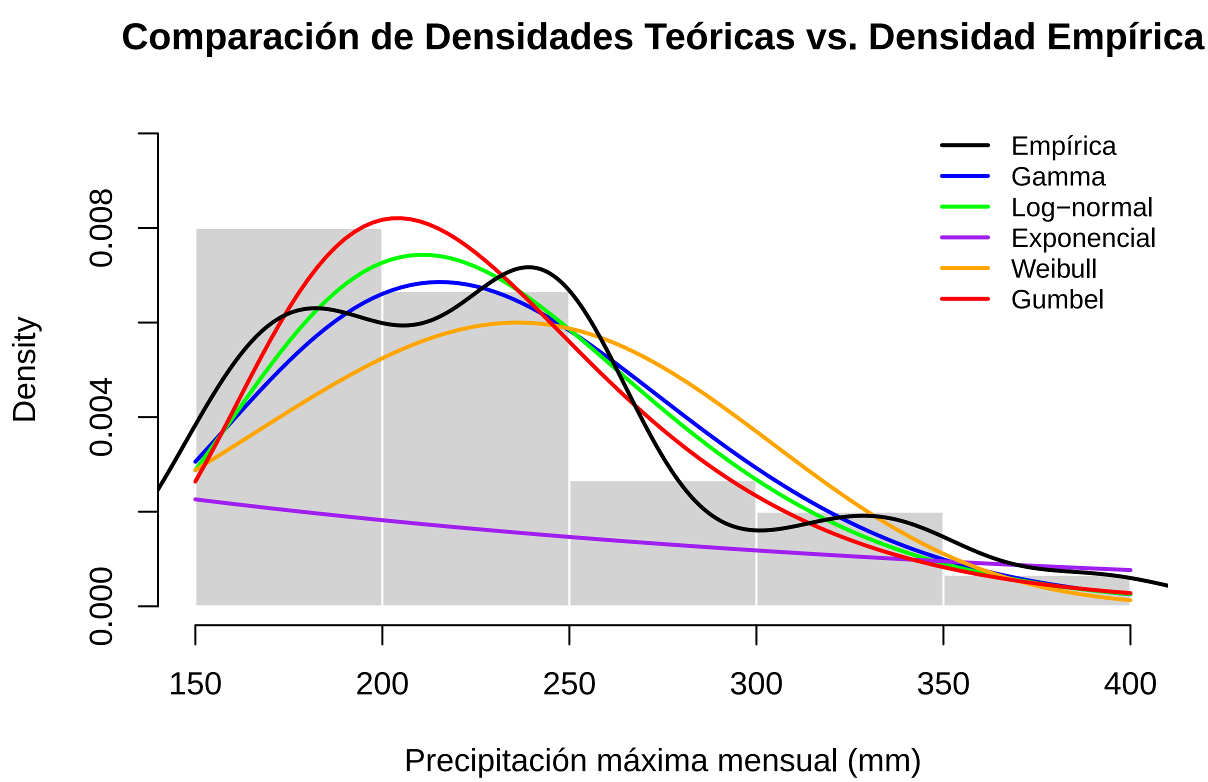
Los valores obtenidos mediante el cálculo manual y el comando `fitdistrplus` son similares pero presentan ligeras diferencias. Esto se debe a los métodos de estimación utilizados:

- **Método de Ajuste de Máxima Verosimilitud**: `fitdistrplus` utiliza máxima verosimilitud para ajustar los parámetros, lo cual es un método estadístico robusto que maximiza la probabilidad de los datos bajo el modelo.
- **Método Basado en Momentos**: El cálculo manual se basa en igualar la media y desviación estándar teóricas con las observadas, un método que puede ser menos preciso en algunos casos.

Compara los ajustes de las distribuciones que analizaste

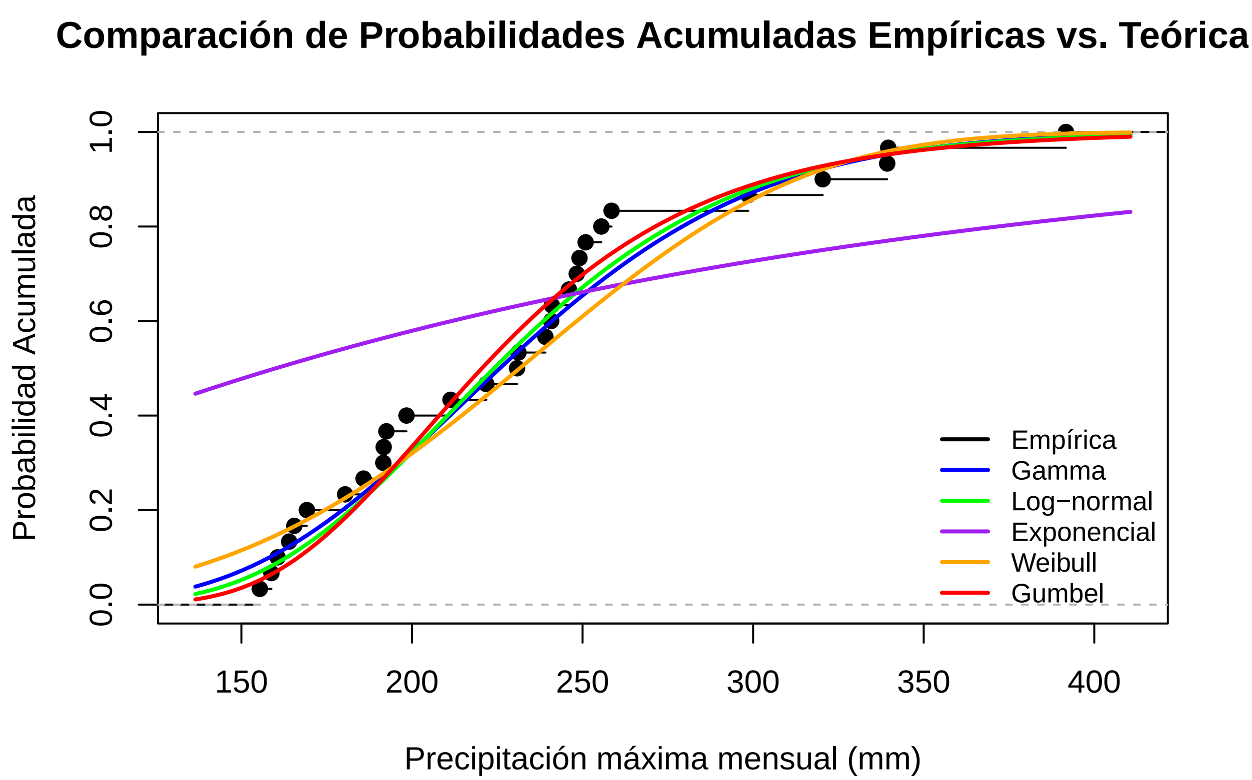
```
# Crear el histograma de densidad empírica
hist(datos_max_anuales$Lluvia, xlab = "Precipitación máxima mensual (mm)", freq = FALSE, ylim = c(0, 0.(
     main = "Comparación de Densidades Teóricas vs. Densidad Empírica",
     col = "lightgray", border = "white")

# Añadir las densidades teóricas para cada distribución
curve(dgamma(x, shape = shape_gamma, scale = scale_gamma), add = TRUE, col = "blue", lwd = 2)
curve(dlnorm(x, meanlog = meanlog_val, sdlog = sdlog_val), add = TRUE, col = "green", lwd = 2)
curve(dexp(x, rate = rate_exp), add = TRUE, col = "purple", lwd = 2)
curve(dweibull(x, shape = shape_weibull, scale = scale_weibull), add = TRUE, col = "orange", lwd = 2)
curve(dgumbel(x, a = gumbel_fit$estimate[1], b = gumbel_fit$estimate[2]), add = TRUE, col = "red", lwd =

# Densidad empírica
lines(density(datos_max_anuales$Lluvia), col = "black", lwd = 2)

# Añadir la leyenda
legend("topright", legend = c("Empírica", "Gamma", "Log-normal", "Exponencial", "Weibull", "Gumbel"),
       col = c("black", "blue", "green", "purple", "orange", "red"), lwd = 2, bty = "n", cex = 0.8)
```

Haz un gráfico comparativo de los histogramas de densidad empírica vs densidad teórica de todas las distribuciones que analizaste (todas las distribuciones en un solo gráfico.



```r
# Distribución acumulada empírica
plot(ecdf(datos_max_anuales$Lluvia), main = "Comparación de Probabilidades Acumuladas Empíricas vs. Teó
     xlab = "Precipitación máxima mensual (mm)", ylab = "Probabilidad Acumulada")

# Añadir las distribuciones acumuladas teóricas
curve(pgamma(x, shape = shape_gamma, scale = scale_gamma), add = TRUE, col = "blue", lwd = 2)
curve(plnorm(x, meanlog = meanlog_val, sdlog = sdlog_val), add = TRUE, col = "green", lwd = 2)
curve(pexp(x, rate = rate_exp), add = TRUE, col = "purple", lwd = 2)
curve(pweibull(x, shape = shape_weibull, scale = scale_weibull), add = TRUE, col = "orange", lwd = 2)
curve(pgumbel(x, a = gumbel_fit$estimate[1], b = gumbel_fit$estimate[2]), add = TRUE, col = "red", lwd =

# Añadir la leyenda
legend("bottomright", legend = c("Empírica", "Gamma", "Log-normal", "Exponencial", "Weibull", "Gumbel")
       col = c("black", "blue", "green", "purple", "orange", "red"), lwd = 2, bty = "n", cex = 0.8)
```

Haz un gráfico comparativo de las probabilidades acumuladas empírica vs teóricas de todas las distribuciones que analizaste (todas las distribuciones en un solo gráfico.




```r
# Prueba KS para la distribución Gamma
ks_gamma <- ks.test(datos_max_anuales$Lluvia, "pgamma", shape = shape_gamma, scale = scale_gamma)
p_valor_gamma <- ks_gamma$p.value
estadistico_gamma <- ks_gamma$statistic

# Prueba KS para la distribución Log-normal
```

```r
ks_lognorm <- ks.test(datos_max_anuales$Lluvia, "plnorm", meanlog = meanlog_val, sdlog = sdlog_val)
p_valor_lognorm <- ks_lognorm$p.value
estadistico_lognorm <- ks_lognorm$statistic

# Prueba KS para la distribución Exponencial
ks_exp <- ks.test(datos_max_anuales$Lluvia, "pexp", rate = rate_exp)
p_valor_exp <- ks_exp$p.value
estadistico_exp <- ks_exp$statistic

# Prueba KS para la distribución Weibull
ks_weibull <- ks.test(datos_max_anuales$Lluvia, "pweibull", shape = shape_weibull, scale = scale_weibul
p_valor_weibull <- ks_weibull$p.value
estadistico_weibull <- ks_weibull$statistic

# Prueba KS para la distribución Gumbel
ks_gumbel <- ks.test(datos_max_anuales$Lluvia, "pgumbel", a = gumbel_fit$estimate[1], b = gumbel_fit$es
p_valor_gumbel <- ks_gumbel$p.value
estadistico_gumbel <- ks_gumbel$statistic

# Crear el dataframe con los p-valores y los estadísticos
resultados <- data.frame(
  Distribucion = c("Gamma", "Log-normal", "Exponencial", "Weibull", "Gumbel"),
  P_valor = c(p_valor_gamma, p_valor_lognorm, p_valor_exp, p_valor_weibull, p_valor_gumbel),
  Estadistico = c(estadistico_gamma, estadistico_lognorm, estadistico_exp, estadistico_weibull, estadis
)

# Mostrar los resultados
print(resultados)
```

```
##   Distribucion      P_valor Estadistico
## 1        Gamma 6.281505e-01   0.1317063
## 2   Log-normal 7.849110e-01   0.1145014
## 3  Exponencial 3.686459e-07   0.4897941
## 4      Weibull 2.877946e-01   0.1741890
## 5       Gumbel 8.240706e-01   0.1098252
```

Define cuál es la mejor distribución que se ajusta a tus datos. Argumenta interpretando la comparación entre los gráficos y analizando las pruebas de ajuste de curva. Análisis Visual

```
Gráfico de Densidades Teóricas vs. Densidad Empírica: En el primer gráfico, la curva de densidad empíri
    La distribución Exponencial (curva púrpura) tiene un ajuste pobre, pues no sigue la forma empírica.
    Las distribuciones Gamma, Log-normal, Weibull y Gumbel parecen ajustarse mejor a la densidad empíri
Gráfico de Probabilidades Acumuladas Empíricas vs. Teóricas: En el segundo gráfico, la curva de la proba
    La Exponencial muestra una diferencia considerable respecto a la empírica.
    Las distribuciones Gamma, Log-normal, Weibull y Gumbel presentan un ajuste visualmente más cercano a
```

2. Análisis de los p-valores de la Prueba KS

   Los p-valores nos indican cuán bien se ajustan los datos a cada distribución, con valores más altos indicando un mejor ajuste. De acuerdo a la tabla de p-valores obtenida: La Exponencial tiene un p-valor extremadamente bajo, lo cual indica que se puede rechazar la hipótesis de ajuste con alta confianza. Las distribuciones Gamma, Log-normal, Weibull, y Gumbel presentan p-valores altos (mayores a 0.05), lo que sugiere un buen ajuste a los datos. Entre estas cuatro distribuciones, Gumbel tiene el p-valor más alto (0.824), seguida de Log-normal (0.785), Gamma (0.628), y Weibull (0.288).

Conclusión

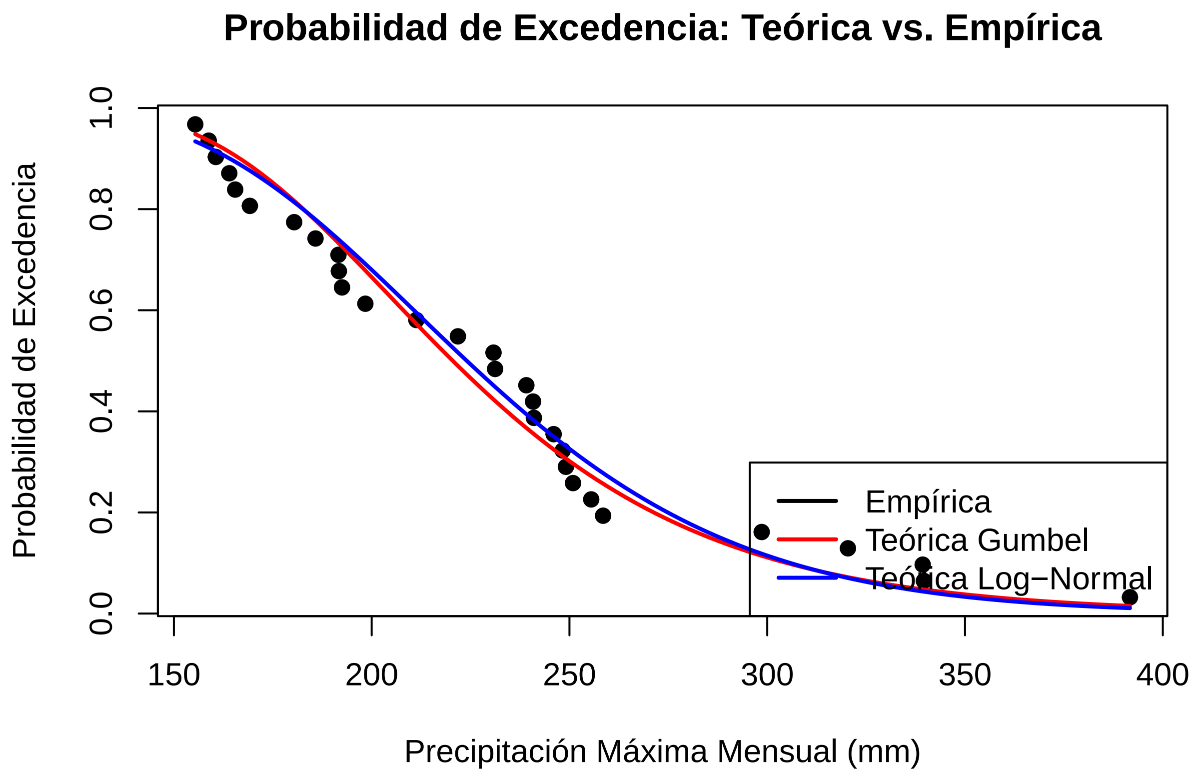
```
Basado en la combinación de análisis visual y los p-valores de las pruebas de bondad de ajuste, la dist
La distribución Gumbel muestra tanto una buena alineación en los gráficos de densidad y acumulada, como
La Log-normal también muestra un buen ajuste y podría ser considerada una alternativa viable, pero lige
```

DISEÑO DE OBRAS HIDRÁULICAS

4. Precipitación de diseño de obras hidráulicas

Se desea diseñar una presa derivadora para una zona de riego mediana. Investiga el periodo de retorno recomendado para esta obra hidráulica, puedes consultarlo en: https://pon.sdsu.edu/periodos_de_retorno_cna.html. ### A. Haz el gráfico comparativo de la probabilidad de excedencia teórica vs empírica. ¿Qué te indica ese gráfico? interpreta y argumenta la certeza de la selección de la distribución elegida.
B. Utilizando el límite inferior del intervalo de periodo de retorno sugerido, encuentra la probabilidad de excedencia o de ocurrencia para ese valor. Recuerda que: $P_{ret} = \frac{1}{P_{exe}}$ ### C. Conociendo la probabilidad de excedencia, calcula su complemento $(1 - P_{exe})$ y utiliza esta probabilidad para encontrar el valor de la precipitación máxima mensual que tendrá ese periodo de retorno. En el código se te da un ejemplo si la distribución de probabilidad a la que mejor se ajustaron los datos fue la Gumbel y deseamos calcular el caudal máximo para un periodo de retorno de 200 años

```r
library(fitdistrplus)

# Definir función de distribución acumulada para Gumbel
pgumbel <- function(q, a, b) exp(-exp((a - q) / b))

# Ajustar la distribución Gumbel
gumbel_fit <- fitdist(datos_max_anuales$Lluvia, "gumbel", start = list(a = 1, b = 1))

# Ajustar la distribución Log-Normal
lognormal_fit <- fitdist(datos_max_anuales$Lluvia, "lnorm")

# Ordenar las precipitaciones de mayor a menor y calcular probabilidades de excedencia y no excedencia
datos_max_anuales <- datos_max_anuales[order(-datos_max_anuales$Lluvia),]
N <- nrow(datos_max_anuales)
datos_max_anuales$Rango <- 1:N
datos_max_anuales$P_exc <- datos_max_anuales$Rango / (N + 1)
datos_max_anuales$P_no_exc <- 1 - datos_max_anuales$P_exc
datos_max_anuales$P_ret <- 1 / datos_max_anuales$P_exc

# Gráfico de probabilidad de excedencia empírica con curvas teóricas
plot(datos_max_anuales$Lluvia, datos_max_anuales$P_exc, type = "p", pch = 19, col = "black",
     main = "Probabilidad de Excedencia: Teórica vs. Empírica",
     xlab = "Precipitación Máxima Mensual (mm)", ylab = "Probabilidad de Excedencia")
curve(1 - pgumbel(x, a = gumbel_fit$estimate[1], b = gumbel_fit$estimate[2]), add = TRUE, col = "red", 
curve(1 - plnorm(x, meanlog = lognormal_fit$estimate[1], sdlog = lognormal_fit$estimate[2]), add = TRUE
legend("bottomright", legend = c("Empírica", "Teórica Gumbel", "Teórica Log-Normal"), col = c("black", 
```




```r
# Visualizar primeras filas del conjunto de datos
head(datos_max_anuales)
```

```
## # A tibble: 6 x 8
## # Groups:   Anio [6]
##    Anio Mes   Estado  Lluvia Rango  P_exc P_no_exc P_ret
##   <int> <chr> <chr>    <dbl> <int>  <dbl>    <dbl> <dbl>
## 1  2021 Ago   Sinaloa   392.     1 0.0323    0.968 31
## 2  2019 Ago   Sinaloa   340.     2 0.0645    0.935 15.5
## 3  2022 Ago   Sinaloa   339.     3 0.0968    0.903 10.3
## 4  2018 Sep   Sinaloa   320.     4 0.129     0.871  7.75
## 5  2013 Sep   Sinaloa   299.     5 0.161     0.839  6.2
## 6  2008 Ago   Sinaloa   258.     6 0.194     0.806  5.17
```

```r
# Definir periodo de retorno y calcular precipitación máxima para la distribución Gumbel
periodo_retorno <- 100 # recomendado inferior
prob_excedencia <- 1 / periodo_retorno
prob_no_excedencia <- 1 - prob_excedencia

precipitacion_gumbel <- gumbel_fit$estimate[1] - gumbel_fit$estimate[2] * log(-log(prob_no_excedencia))
cat("\nPrecipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Gumbel):", p
```

```
##
## Precipitación máxima mensual para un periodo de retorno de 100 años (Gumbel): 410.2261 mm
```

```r
# Calcular precipitación máxima para la distribución Log-Normal
precipitacion_lognormal <- qlnorm(prob_no_excedencia, meanlog = lognormal_fit$estimate[1], sdlog = logno
cat("Precipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Log-Normal):"
```

```
## Precipitación máxima mensual para un periodo de retorno de 100 años (Log-Normal): 394.1259 mm
```

El gráfico de probabilidad de excedencia compara los valores empíricos con las distribuciones teóricas de Gumbel y Log-Normal. La cercanía de ambas curvas teóricas a los puntos empíricos sugiere que ambas distribuciones son adecuadas para modelar la precipitación máxima mensual, aunque la de Gumbel parece ajustarse ligeramente mejor. Este tipo de análisis es crucial en el diseño de obras hidráulicas, ya que permite estimar la probabilidad de que ciertos eventos extremos de precipitación ocurran, facilitando el diseño seguro de infraestructura con base en periodos de retorno sugeridos.

Precipitación máxima mensual para un periodo de retorno de 100 años (Gumbel): 410.2261 mm Precipitación máxima mensual para un periodo de retorno de 100 años (Log-Normal): 391.1259 mm

D. El resultado de este ejemplo será una aproximación del caudal máximo que se tendría en Aguascalientes con un periodo de retorno de 200 años. ¿Qué significa este valor? ¿Qué pasa si incrementamos el periodo de retorno? ¿El caudal máximo para este periodo de retorno será el mismo si utilizamos datos históricos de otro estado? ¿Por qué crees que las obras hidráulicas deben diseñarse en base a periodos de retorno sugeridos? ¿Por qué es importante conocer la distribución de probabilidad a la que mejor se aproximan los datos históricos? Explora otros periodos de retorno diferentes a los que se proporcionan en los periodos sugeridos para contestar esta pregunta. Estos valores sugieren que para un evento que ocurre cada 100 años en promedio, la precipitación máxima mensual esperada sería aproximadamente 410 mm bajo una distribución Gumbel y 391 mm bajo una distribución Log-Normal. Al incrementar el periodo de retorno, estos valores de precipitación máxima aumentarían, ya que representan eventos menos frecuentes y de mayor magnitud. Al disminuir el periodo de retorno, los valores de precipitación disminuirían.

```r
periodo_retorno <- 50
prob_excedencia <- 1 / periodo_retorno
prob_no_excedencia <- 1 - prob_excedencia

precipitacion_gumbel <- gumbel_fit$estimate[1] - gumbel_fit$estimate[2] * log(-log(prob_no_excedencia))
cat("\nPrecipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Gumbel):", p
```

```
##
## Precipitación máxima mensual para un periodo de retorno de 50 años (Gumbel): 378.9362 mm
```

```r
# Calcular precipitación máxima para la distribución Log-Normal
precipitacion_lognormal <- qlnorm(prob_no_excedencia, meanlog = lognormal_fit$estimate[1], sdlog = logno
cat("Precipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Log-Normal):"
```

```
## Precipitación máxima mensual para un periodo de retorno de 50 años (Log-Normal): 368.8836 mm
```

```r
periodo_retorno <- 200
prob_excedencia <- 1 / periodo_retorno
prob_no_excedencia <- 1 - prob_excedencia

precipitacion_gumbel <- gumbel_fit$estimate[1] - gumbel_fit$estimate[2] * log(-log(prob_no_excedencia))
cat("\nPrecipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Gumbel):", p
```

```
##
## Precipitación máxima mensual para un periodo de retorno de 200 años (Gumbel): 441.4019 mm
```

```
# Calcular precipitación máxima para la distribución Log-Normal
precipitacion_lognormal <- qlnorm(prob_no_excedencia, meanlog = lognormal_fit$estimate[1], sdlog = logno
cat("Precipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Log-Normal):"
```

```
## Precipitación máxima mensual para un periodo de retorno de 200 años (Log-Normal): 418.7385 mm
```

```
periodo_retorno <- 1000
prob_excedencia <- 1 / periodo_retorno
prob_no_excedencia <- 1 - prob_excedencia

precipitacion_gumbel <- gumbel_fit$estimate[1] - gumbel_fit$estimate[2] * log(-log(prob_no_excedencia))
cat("\nPrecipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Gumbel):", p
```

```
##
## Precipitación máxima mensual para un periodo de retorno de 1000 años (Gumbel): 513.6178 mm
```

```
# Calcular precipitación máxima para la distribución Log-Normal
precipitacion_lognormal <- qlnorm(prob_no_excedencia, meanlog = lognormal_fit$estimate[1], sdlog = logno
cat("Precipitación máxima mensual para un periodo de retorno de", periodo_retorno, "años (Log-Normal):"
```

```
## Precipitación máxima mensual para un periodo de retorno de 1000 años (Log-Normal): 474.4461 mm
```

Como podemos observar, se obtuvo una precipitación máxima mensual en un periodo de 50 años de aproximadamente 379mm para Gumbel y 369mm para Log-Normal, una precipitación máxima mensual en un periodo de 200 años de aproximadamente 441mm para Gumbel y 419mm para Log-Normal y una precipitación máxima mensual en un periodo de 1000 años de aproximadamente 514mm para Gumbel y 474mm para Log-Normal.

El caudal máximo para un periodo de retorno específico no será el mismo si utilizamos datos históricos de otro estado. Las condiciones climáticas, geográficas y ambientales varían significativamente entre regiones, lo que influye en los patrones de precipitación y, por ende, en los cálculos de caudal. Por lo tanto, es esencial utilizar datos locales para obtener estimaciones precisas y aplicables.

Las obras hidráulicas deben diseñarse en base a periodos de retorno sugeridos porque estos periodos representan la probabilidad de eventos extremos, lo que ayuda a mitigar riesgos de inundaciones y garantizar la seguridad de las estructuras y de las comunidades circundantes. Diseñar con un periodo de retorno adecuado significa que se está considerando la variabilidad climática y las proyecciones a largo plazo, lo cual es crucial para la resiliencia de la infraestructura.

Finalmente, conocer la distribución de probabilidad que mejor se ajusta a los datos históricos permite realizar pronósticos más precisos sobre eventos futuros. La distribución proporciona un marco para entender la frecuencia e intensidad de las precipitaciones, lo que es fundamental para la planificación y el diseño de proyectos hidráulicos.