Act_Int_1_A01571214_Lautaro_Coteja

A01571214 - Lautaro Coteja

2024-10-28

R Markdown

Introduccion

El diseño de obras hidráulicas, como presas, diques y sistemas de drenaje, requiere una planificación cuidadosa para soportar eventos climáticos extremos y reducir el riesgo de inundaciones. Para ello, se emplea el análisis de precipitaciones máximas, que se basa en datos históricos y modelos probabilísticos que predicen la frecuencia de estos eventos. Estos modelos se fundamentan en distribuciones estadísticas, como la Gumbel, Gamma y Weibull, que permiten calcular los periodos de retorno o intervalos de tiempo esperados entre eventos extremos. Este enfoque es esencial para garantizar la seguridad y funcionalidad de las estructuras hidráulicas, especialmente en un contexto de variabilidad climática creciente, como señala la Organización Meteorológica Mundial (WMO, 2021) y otras investigaciones sobre la importancia de la estimación de riesgos climáticos para la infraestructura (Smith et al., 2019). Este análisis es crucial para optimizar los recursos y minimizar las vulnerabilidades de las construcciones frente a eventos severos, aportando resiliencia a largo plazo.

Actividad Integradora 1 - Precipitaciones Maximas Mensuales para el Diseño de Obras Hidraulicas

Cargar los Datos

```r
data =
read.csv("C:/Users/lauta/Downloads/precipitaciones_maximas_mensuales.txt",
sep = "\t")

head(data)

##   Anio Mes              Estado Lluvia
## 1 1994 Ene      Aguascalientes    8.3
## 2 1994 Ene     Baja.California   10.3
## 3 1994 Ene Baja.California.Sur    0.0
## 4 1994 Ene            Campeche   85.4
## 5 1994 Ene    Ciudad.de.México   17.7
## 6 1994 Ene            Coahuila   12.8
```

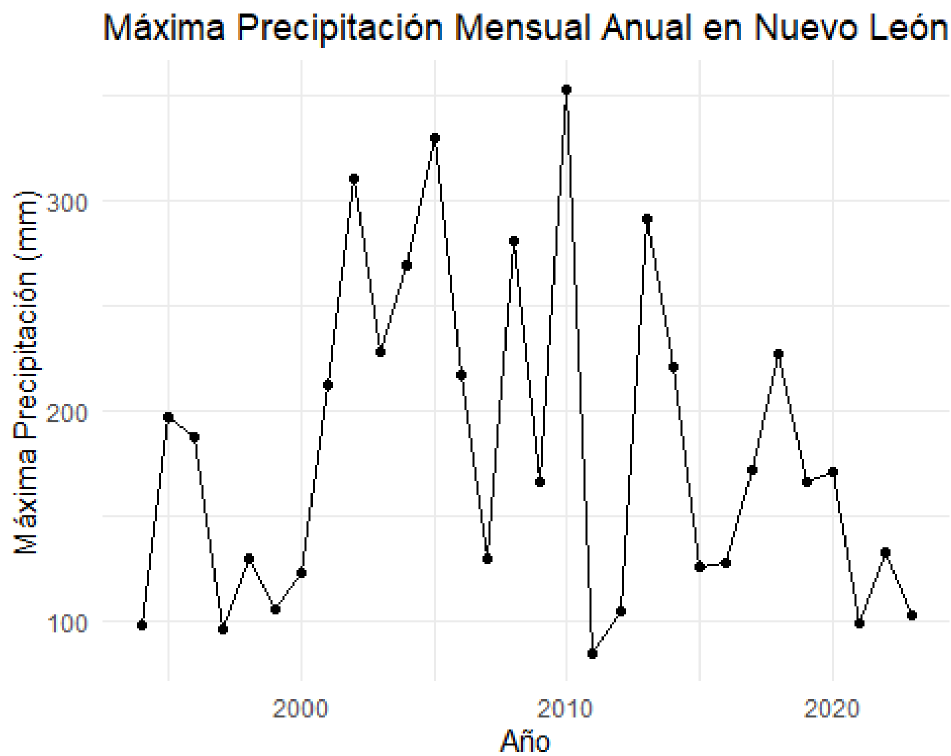
Parte 1

Seleccionar Estado

```
state_data = subset(data, Estado == "Nuevo.León")
```

Crear columna combinando año y mes

```
state_data$Mes = factor(state_data$Mes, levels = c('Ene', 'Feb', 'Mar',
'Abr', 'May', 'Jun', 'Jul', 'Ago', 'Sep', 'Oct', 'Nov', 'Dic'))
state_data$Mes_num = as.numeric(state_data$Mes)
state_data$Date = as.Date(paste(state_data$Anio, state_data$Mes_num, "01",
sep = "-"), "%Y-%m-%d")

annual_max_precipitation = aggregate(Lluvia ~ Anio, data = state_data, max)

# Graficar los máximos anuales
library(ggplot2)
ggplot(annual_max_precipitation, aes(x = Anio, y = Lluvia)) +
  geom_line() +
  geom_point() +
  ggtitle("Máxima Precipitación Mensual Anual en Nuevo León") +
  xlab("Año") +
  ylab("Máxima Precipitación (mm)") +
  theme_minimal()
```



Interpretaciones
La gráfica muestra la evolución de las precipitaciones máximas mensuales anuales en Nuevo León. Existen variaciones anuales significativas, con algunos picos de precipitaciones extremas.

Estadisticas Descriptivas

```
mean_precipitation = mean(state_data$Lluvia)
median_precipitation = median(state_data$Lluvia)
std_precipitation = sd(state_data$Lluvia)
min_precipitation = min(state_data$Lluvia)
max_precipitation = max(state_data$Lluvia)

mean_precipitation

## [1] 50.80361

median_precipitation

## [1] 31.05

std_precipitation

## [1] 58.08207

min_precipitation

## [1] 0
```

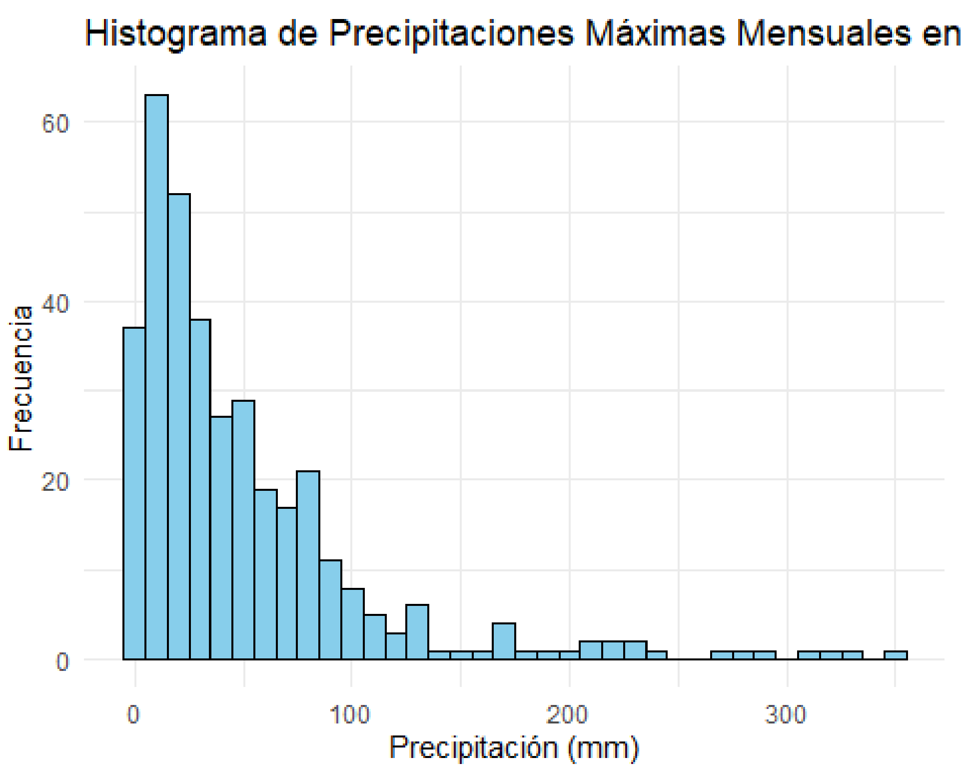
```
max_precipitation

## [1] 352.7
```

Interpretaciones

El promedio de las precipitaciones mensuales máximas es de 50.80 mm. La mediana es de 31.05 mm, indicando que la mitad de las observaciones tienen precipitaciones por debajo de este valor. La desviación estándar es de 58.08 mm, reflejando una alta variabilidad. El valor mínimo es 0 mm, mientras que el máximo alcanza 352.7 mm, lo que refleja la presencia de eventos extremos.

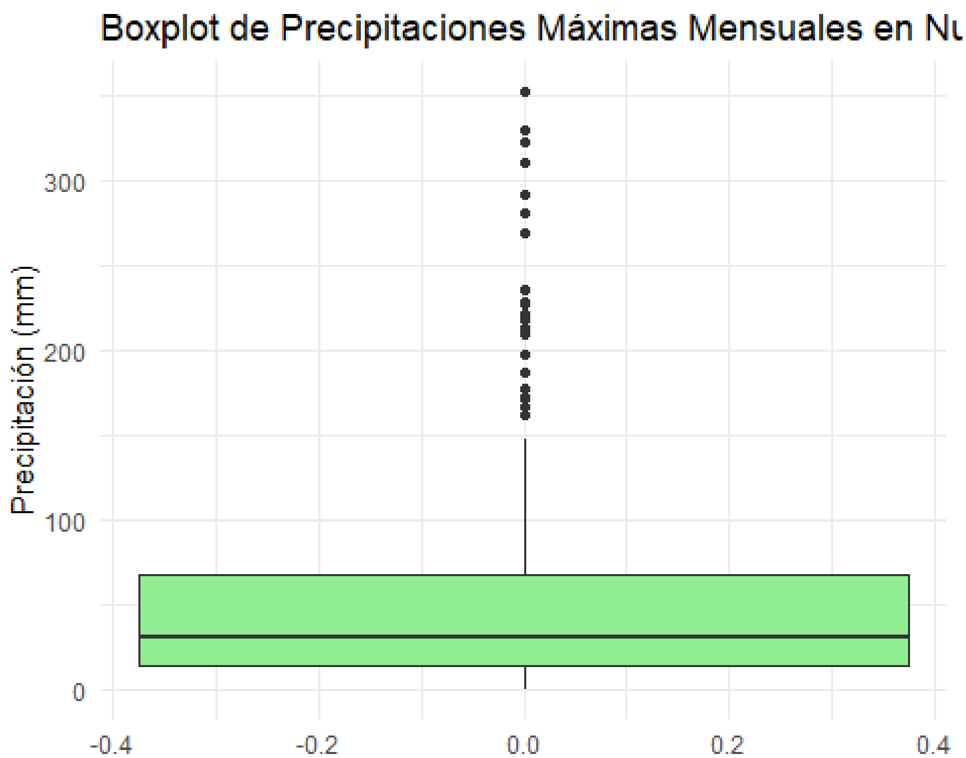
Graficos de Distribucion

```r
# Histograma
ggplot(state_data, aes(x = Lluvia)) +
  geom_histogram(binwidth = 10, fill = "skyblue", color = "black") +
  ggtitle("Histograma de Precipitaciones Máximas Mensuales en Nuevo León") +
  xlab("Precipitación (mm)") +
  ylab("Frecuencia") +
  theme_minimal()
```




```r
# Boxplot
ggplot(state_data, aes(y = Lluvia)) +
  geom_boxplot(fill = "lightgreen") +
```

```r
  ggtitle("Boxplot de Precipitaciones Máximas Mensuales en Nuevo León") +
  ylab("Precipitación (mm)") +
  theme_minimal()
```



Interpretaciones

El histograma muestra que la mayoría de las precipitaciones máximas mensuales están concentradas en rangos bajos, pero hay algunos eventos de precipitación extrema. El boxplot confirma la presencia de valores atípicos, reflejando la ocurrencia de lluvias intensas en algunos meses.

Conclusiones

El análisis de las precipitaciones máximas mensuales en Nuevo León entre 1994 y 2023 muestra una variabilidad significativa con eventos de lluvias extremas que deben ser considerados en el diseño de infraestructuras críticas. La alta variabilidad y la presencia de valores atípicos son factores importantes a tener en cuenta en la planificación y prevención de inundaciones.

Parte 2

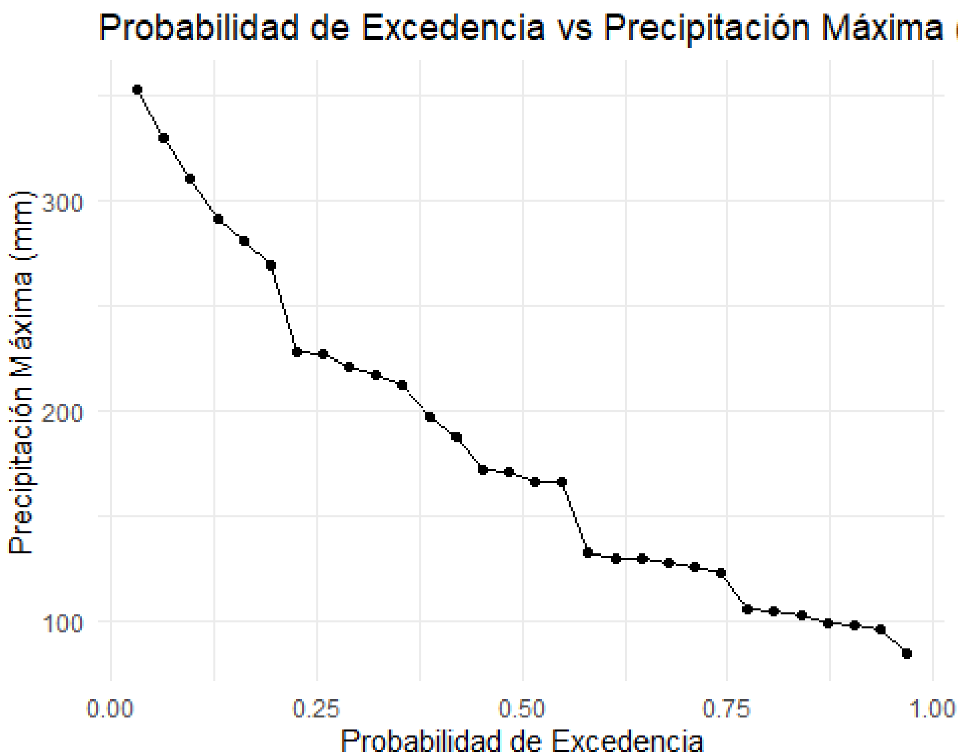
Analisis de Frecuencias Metodo Grafico

```r
# Agrupar por año y calcular la máxima precipitación
annual_max_precipitation = aggregate(Lluvia ~ Anio, data = state_data, max)
colnames(annual_max_precipitation) = c("Anio", "MaxLluvia")
```

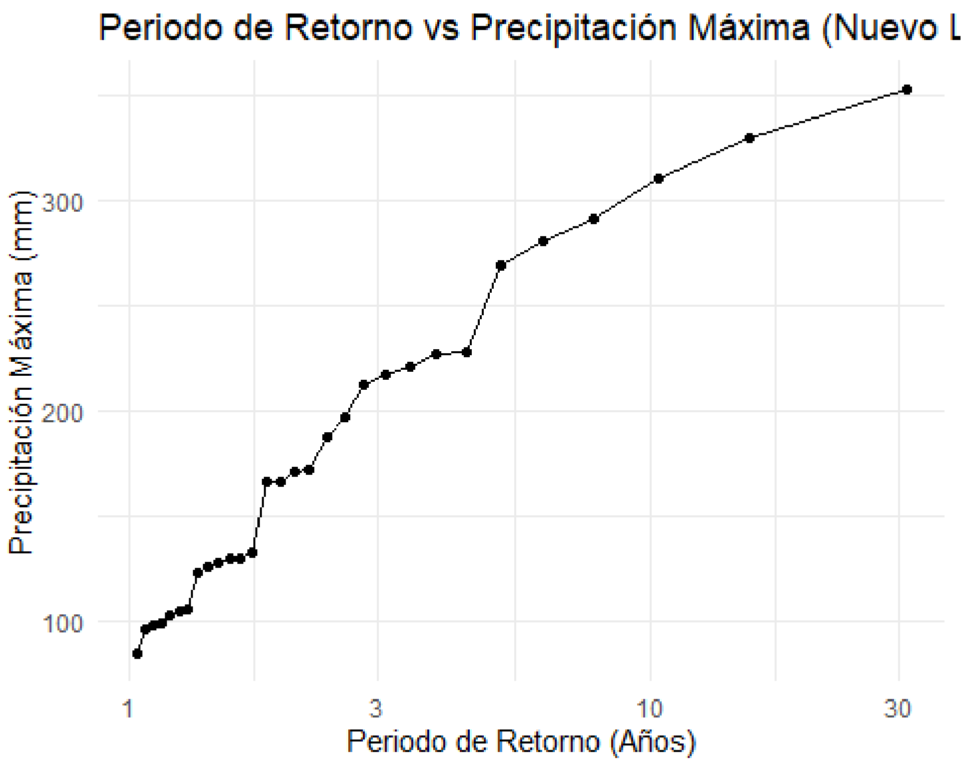
```r
# Ordenar por precipitación y agregar el rango
annual_max_precipitation = annual_max_precipitation[order(-
annual_max_precipitation$MaxLluvia),]
annual_max_precipitation$Rank = 1:nrow(annual_max_precipitation)
N = nrow(annual_max_precipitation)

# Calcular probabilidad de excedencia y periodo de retorno
annual_max_precipitation$Prob_Exceedance = annual_max_precipitation$Rank / (N
+ 1)
annual_max_precipitation$Return_Period = 1 /
annual_max_precipitation$Prob_Exceedance
head(annual_max_precipitation)
```

```
##    Anio MaxLluvia Rank Prob_Exceedance Return_Period
## 17 2010     352.7    1      0.03225806     31.000000
## 12 2005     329.9    2      0.06451613     15.500000
## 9  2002     310.4    3      0.09677419     10.333333
## 20 2013     291.3    4      0.12903226      7.750000
## 15 2008     281.1    5      0.16129032      6.200000
## 11 2004     269.3    6      0.19354839      5.166667
```

```r
library(ggplot2)

# Gráfico de probabilidad de excedencia
ggplot(annual_max_precipitation, aes(x = Prob_Exceedance, y = MaxLluvia)) +
  geom_point() +
  geom_line() +
  ggtitle("Probabilidad de Excedencia vs Precipitación Máxima (Nuevo León)")
+
  xlab("Probabilidad de Excedencia") +
  ylab("Precipitación Máxima (mm)") +
  theme_minimal()
```




```r
# Gráfico de periodo de retorno
ggplot(annual_max_precipitation, aes(x = Return_Period, y = MaxLluvia)) +
  geom_point() +
  geom_line() +
  scale_x_log10() +
  ggtitle("Periodo de Retorno vs Precipitación Máxima (Nuevo León)") +
  xlab("Periodo de Retorno (Años)") +
  ylab("Precipitación Máxima (mm)") +
  theme_minimal()
```



Conclusiones y Interpretaciones

Probabilidad de Excedencia: Representa la probabilidad de que un evento de precipitación supere un valor dado en cualquier año. Valores bajos de probabilidad de excedencia indican eventos de precipitación extrema que ocurren rara vez, lo que es crucial para planificar infraestructuras resistentes a tales eventos.

Periodo de Retorno: Indica el intervalo promedio de años entre eventos de precipitación máxima que exceden cierto umbral. Por ejemplo, un período de retorno de 50 años significa que, en promedio, un evento de esa magnitud ocurrirá una vez cada 50 años. Este dato es vital en hidrología para diseñar infraestructuras que puedan soportar lluvias excepcionales y prevenir inundaciones.

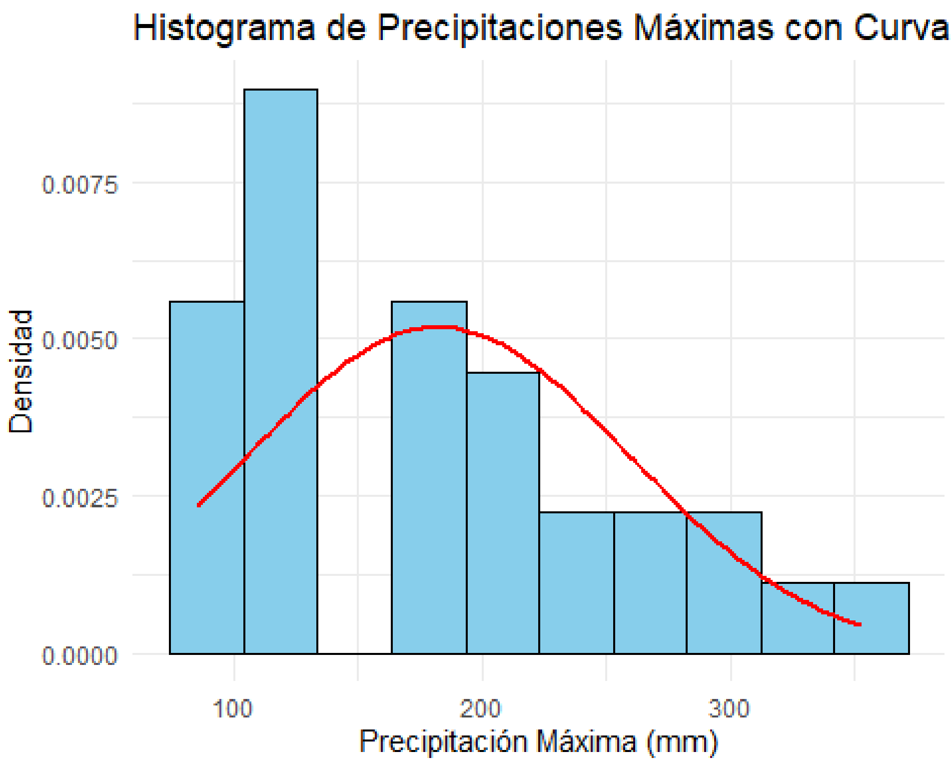
Los valores de probabilidad de excedencia bajos son deseables en el diseño de estructuras hidrológicas, ya que aseguran la resistencia a eventos raros y extremos, reduciendo el riesgo de daños por eventos de alta precipitación.

Parte 3

Analisis de Frecuencias Metodo Analitico

A

```
# Calcular media y desviación estándar
mean_precipitation = mean(annual_max_precipitation$MaxLluvia)
sd_precipitation = sd(annual_max_precipitation$MaxLluvia)

# Histograma de la densidad empírica con la curva de densidad normal
library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dnorm, args = list(mean = mean_precipitation, sd =
sd_precipitation), color = "red", size = 1) +
  ggtitle("Histograma de Precipitaciones Máximas con Curva Normal") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## ℹ Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2
3.4.0.
## ℹ Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



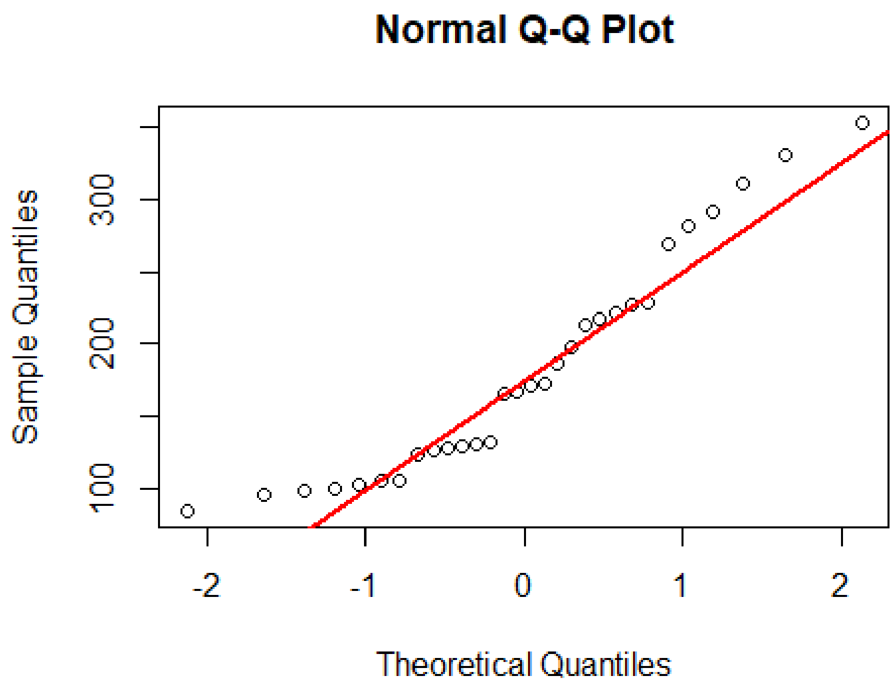
Interpretaciones

Visualmente, la distribución normal (curva roja) no se ajusta perfectamente a los datos del histograma. Aunque la curva representa la tendencia central de los datos, existen discrepancias notables, especialmente en las colas. La distribución empírica muestra asimetría, con una mayor concentración de valores en el rango bajo y pocos valores en el extremo superior, lo que indica una distribución sesgada. La distribución normal, por el contrario, es simétrica y no captura adecuadamente estos valores extremos de precipitación, lo cual sugiere que una normalidad perfecta no es el mejor ajuste para estos datos.

La distribución normal tiene dos parámetros: la media (μ) y la desviación estándar (σ). La media representa el valor central de la distribución, mientras que la desviación estándar mide la dispersión o variabilidad en torno a la media.

En el código, la media y desviación estándar se calculan directamente a partir de los datos observados. La media se utiliza porque resume la tendencia central de los datos, y la desviación estándar proporciona una medida de dispersión, lo cual es esencial para definir la anchura de la curva normal. Estos parámetros permiten construir una curva normal que representa, de manera aproximada, la distribución de los datos, aunque en este caso, la asimetría en los datos sugiere que una distribución diferente podría ser más adecuada para modelar los valores extremos de precipitación.

```
# Q-Q plot
qqnorm(annual_max_precipitation$MaxLluvia)
qqline(annual_max_precipitation$MaxLluvia, col = "red", lwd = 2)
```



Interpretaciones

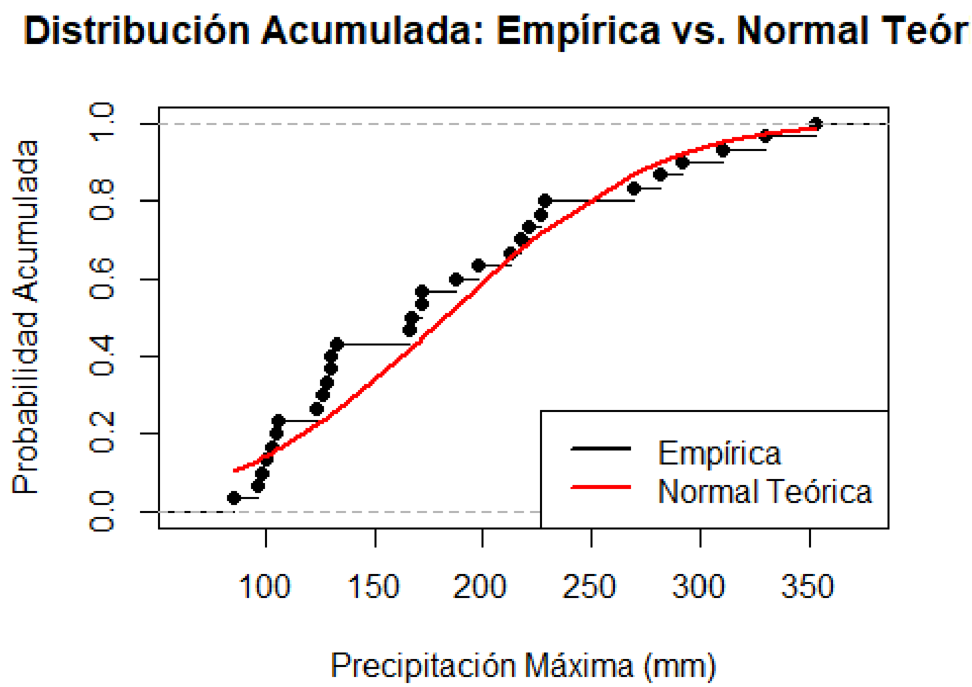
En el Q-Q plot, los cuantiles teóricos de una distribución normal (eje x) se comparan con los cuantiles de los datos de muestra (eje y). Si los datos siguen una distribución normal, los puntos deberían alinearse aproximadamente sobre la línea roja.

En este caso, se observa que algunos puntos se desvían significativamente de la línea, especialmente en los extremos (colas), lo que indica que los datos no siguen completamente una distribución normal. Aunque los puntos centrales están cerca de la línea, las desviaciones en los extremos sugieren que la normalidad no es un ajuste ideal para estos datos, especialmente debido a la presencia de valores extremos en las precipitaciones.

Visualmente, el Q-Q plot sugiere que los datos tienen colas más pesadas que las de una distribución normal. Esto implica que hay más valores extremos de lo que la normalidad predice, lo cual es común en datos de precipitación. Por lo tanto, sería conveniente considerar otras distribuciones (como Gumbel o Weibull) que manejen mejor estos valores extremos para modelar adecuadamente la distribución de las precipitaciones máximas.

```
# Crear datos de la distribución acumulada teórica
library(stats)
theoretical_normal = pnorm(sort(annual_max_precipitation$MaxLluvia), mean =
mean_precipitation, sd = sd_precipitation)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

# Graficar las distribuciones acumuladas empírica y teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Normal
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
```

```r
lines(sort(annual_max_precipitation$MaxLluvia), theoretical_normal, col =
"red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Normal Teórica"), col =
c("black", "red"), lwd = 2)
```




```r
# Prueba de Shapiro-Wilk
shapiro_test = shapiro.test(annual_max_precipitation$MaxLluvia)

# Prueba de Kolmogorov-Smirnov
ks_test = ks.test(annual_max_precipitation$MaxLluvia, "pnorm", mean =
mean_precipitation, sd = sd_precipitation)

# Resultados de las pruebas
shapiro_test

##
##  Shapiro-Wilk normality test
##
## data:  annual_max_precipitation$MaxLluvia
## W = 0.91693, p-value = 0.02235

ks_test

##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  annual_max_precipitation$MaxLluvia
```

```
## D = 0.17323, p-value = 0.2938
## alternative hypothesis: two-sided
```

Interpretaciones

La prueba de Shapiro-Wilk tiene como hipótesis nula (H0) que los datos provienen de una distribución normal. Un p-value menor a un nivel de significancia común (por ejemplo, 0.05) indica evidencia para rechazar la hipótesis nula. Dado que el p-value es 0.02235 (menor a 0.05), se rechaza la hipótesis nula de normalidad. Esto sugiere que los datos de precipitaciones máximas mensuales no siguen una distribución normal.

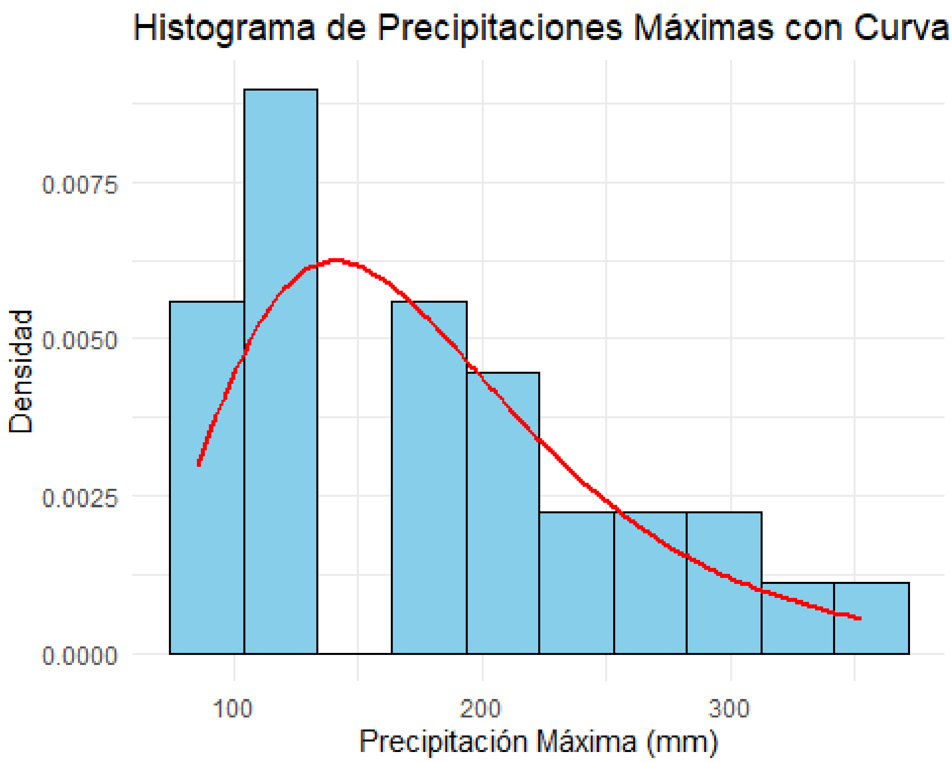
La prueba KS también tiene como hipótesis nula (H0) que los datos provienen de una distribución normal. En este caso, un p-value mayor que el nivel de significancia indica que no hay suficiente evidencia para rechazar la hipótesis nula. Con un p-value de 0.2938 (mayor a 0.05), no se rechaza la hipótesis nula en la prueba KS. Sin embargo, esta prueba es menos sensible a pequeñas desviaciones de normalidad en comparación con la prueba de Shapiro-Wilk.

Las dos pruebas arrojan resultados ligeramente distintos. La prueba de Shapiro-Wilk, más sensible, sugiere que los datos no siguen una distribución normal, mientras que la prueba de KS no rechaza la normalidad.

Combinando estos resultados con el Q-Q plot y el histograma, podemos concluir que los datos de las precipitaciones máximas mensuales presentan desviaciones de la normalidad, especialmente en los valores extremos. Esto refuerza la idea de que una distribución normal puede no ser la mejor para modelar estos datos y que sería mejor considerar distribuciones que manejen mejor los valores extremos (como Gumbel o Weibull).

B

```r
# Log-transformación y cálculo de parámetros (método de momentos)
log_precipitation = log(annual_max_precipitation$MaxLluvia)
mean_log_precipitation = mean(log_precipitation)
sd_log_precipitation = sd(log_precipitation)

# Histograma con superposición de la distribución Log-normal
library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dlnorm, args = list(meanlog = mean_log_precipitation,
sdlog = sd_log_precipitation), color = "red", size = 1) +
  ggtitle("Histograma de Precipitaciones Máximas con Curva Log-normal") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()
```



Interpretaciones

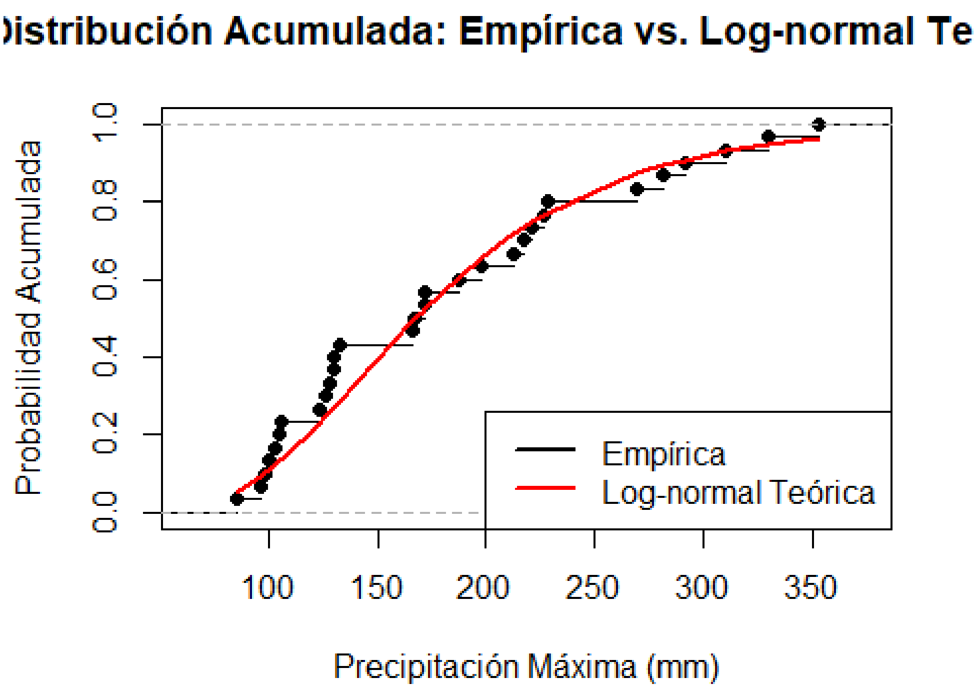
Visualmente, la curva de distribución Log-normal (curva roja) sigue mejor la forma del histograma de los datos de precipitación máxima en comparación con la curva normal observada previamente. La distribución Log-normal, que es asimétrica y tiene una cola hacia la derecha, parece ajustarse de manera más adecuada a la naturaleza de los datos, que también presentan valores extremos de precipitación en el extremo superior.

La forma de la curva Log-normal se adapta bien a los valores más frecuentes en los rangos bajos y medios de precipitación y captura de manera razonable la disminución de frecuencias hacia los valores más altos. Esto sugiere que la distribución Log-normal es una opción más precisa para modelar estos datos de precipitación máxima, ya que refleja la asimetría observada en la muestra.

La distribución Log-normal parece ajustarse visualmente mejor a los datos de precipitación que la normal, debido a su capacidad para representar distribuciones sesgadas hacia la derecha. Este ajuste sugiere que la Log-normal podría ser una buena candidata para modelar las precipitaciones máximas mensuales, ya que permite una representación más realista de los eventos de precipitación extrema.

```
# Crear datos para la distribución acumulada log-normal teórica
sorted_data = sort(annual_max_precipitation$MaxLluvia)
theoretical_cdf_lognorm = plnorm(sorted_data, meanlog =
mean_log_precipitation, sdlog = sd_log_precipitation)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

# Graficar la CDF empírica vs la teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Log-normal
```

```
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
lines(sorted_data, theoretical_cdf_lognorm, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Log-normal Teórica"), col =
c("black", "red"), lwd = 2)
```



Analisis de Grafica La CDF empírica (línea negra) se compara con la CDF de la distribución log-normal teórica (línea roja). La cercanía de ambas curvas indica que la distribución log-normal captura de manera adecuada el comportamiento de los datos, especialmente en los rangos de probabilidades medias y altas. La correspondencia visual refuerza la conclusión de que la log-normal es un buen modelo para estos datos.

La distribución log-normal tiene dos parámetros: la media logarítmica (μ_log) y la desviación estándar logarítmica (σ_log). Estos parámetros se calculan aplicando el logaritmo a los datos y luego obteniendo la media y la desviación estándar de los valores transformados. La media logarítmica representa el valor central en la escala logarítmica, y la desviación estándar logarítmica refleja la dispersión en torno a esa media.

El método de momentos se utiliza para calcular los parámetros de la distribución log-normal al ajustar la media y la varianza de los datos transformados logarítmicamente. Esto asegura que los momentos (media y varianza) de la distribución log-normal sean coherentes con los momentos observados en los datos, lo cual ayuda a obtener un modelo adecuado para eventos de precipitación extremos.

```
# Prueba KS para distribución log-normal
ks_test_lognorm = ks.test(annual_max_precipitation$MaxLluvia, "plnorm",
meanlog = mean_log_precipitation, sdlog = sd_log_precipitation)
```

```r
# Resultados de la prueba KS
ks_test_lognorm
```

```
##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  annual_max_precipitation$MaxLluvia
## D = 0.14493, p-value = 0.5083
## alternative hypothesis: two-sided
```

Interpretaciones

La prueba KS tiene como hipótesis nula (H0) que los datos provienen de una distribución log-normal. Un p-value mayor a un nivel de significancia común (como 0.05) indica que no hay suficiente evidencia para rechazar la hipótesis nula.

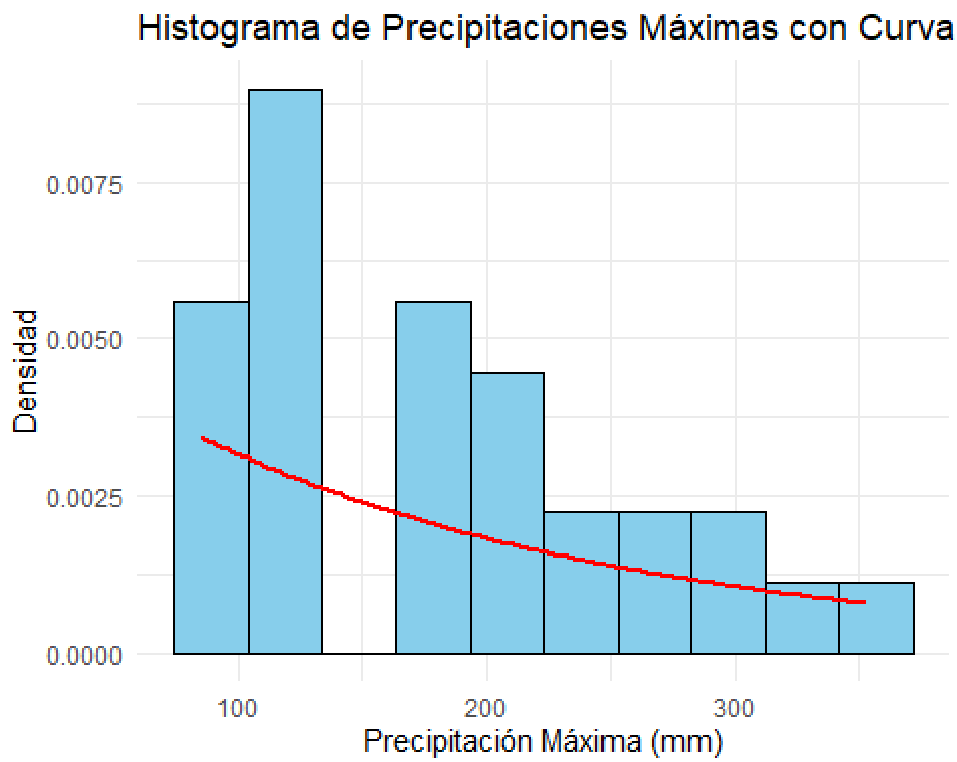
Dado que el p-value es 0.5083 (mayor a 0.05), no se rechaza la hipótesis nula, lo cual sugiere que los datos podrían seguir una distribución log-normal. Esto se alinea con la observación visual de que la distribución log-normal parece ajustarse adecuadamente a los datos de precipitación.

Conclusion

La prueba KS y el ajuste visual en el gráfico de CDF respaldan la hipótesis de que los datos de precipitación siguen una distribución log-normal. La precisión del ajuste log-normal en representar la asimetría y los valores extremos observados hace que sea una opción adecuada para modelar estos datos.

C

```r
# Calcular el parámetro lambda (método de momentos)
lambda_exp = 1 / mean(annual_max_precipitation$MaxLluvia)

# Histograma con superposición de la distribución Exponencial
library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dexp, args = list(rate = lambda_exp), color = "red",
size = 1) +
  ggtitle("Histograma de Precipitaciones Máximas con Curva Exponencial") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()
```



Interpretaciones

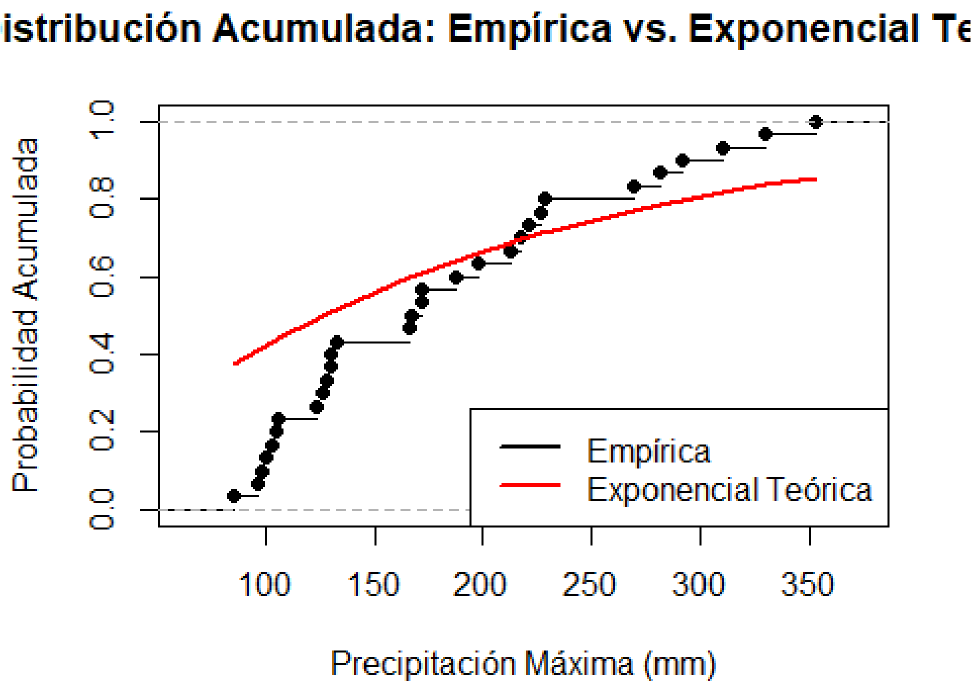
Visualmente, la curva de distribución exponencial (curva roja) no se ajusta bien a la densidad empírica representada en el histograma. Aunque la distribución exponencial captura la tendencia decreciente general de los datos, subestima considerablemente la frecuencia en los valores más bajos de precipitación y no refleja adecuadamente la dispersión observada.

La distribución exponencial tiene una cola larga hacia la derecha, pero, en este caso, parece que no captura adecuadamente los valores intermedios y altos de precipitación. Esto indica que la naturaleza de los datos no es estrictamente exponencial y que otros modelos de distribución (como la Log-normal o Gamma) podrían ser más adecuados para describir los datos de precipitación máxima.

Basado en el ajuste visual, se concluye que la distribución exponencial no es la mejor opción para modelar los datos de precipitaciones máximas, ya que no representa adecuadamente la variabilidad y concentración observadas en los datos empíricos.

```r
# Crear datos para la distribución acumulada exponencial teórica
sorted_data = sort(annual_max_precipitation$MaxLluvia)
theoretical_cdf_exp = pexp(sorted_data, rate = lambda_exp)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

# Graficar la CDF empírica vs la teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Exponencial
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
lines(sorted_data, theoretical_cdf_exp, col = "red", lwd = 2)
```

```
legend("bottomright", legend = c("Empírica", "Exponencial Teórica"), col =
c("black", "red"), lwd = 2)
```



Analisis del Grafico La CDF empírica (línea negra) se compara con la CDF de la distribución exponencial teórica (línea roja). Se observa una discrepancia significativa entre ambas, especialmente en los valores más bajos y más altos de precipitación. Esta falta de ajuste visual refuerza la conclusión de la prueba KS, indicando que la distribución exponencial no es una buena representación de los datos de precipitación.

La distribución exponencial tiene un solo parámetro: λ (lambda), que representa la tasa de ocurrencia de los eventos. Este parámetro se calcula como el inverso de la media de los datos. Sin embargo, en este caso, la variabilidad y asimetría en los datos no son capturadas adecuadamente por este único parámetro, lo que limita el ajuste de la distribución exponencial para describir los datos de precipitaciones máximas.

En el caso de la distribución exponencial, el método de momentos establece que el parámetro λ se estima como el inverso de la media de los datos. Este enfoque es simple pero, en este caso, resulta insuficiente para capturar la complejidad de los datos de precipitación, que muestran variabilidad significativa y valores extremos que una distribución exponencial no modela bien.

```
# Prueba KS para distribución exponencial
ks_test_exp = ks.test(annual_max_precipitation$MaxLluvia, "pexp", rate =
lambda_exp)

# Resultados de la prueba KS
ks_test_exp
```

```
##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  annual_max_precipitation$MaxLluvia
## D = 0.37669, p-value = 0.0002474
## alternative hypothesis: two-sided
```

Interpretaciones

La prueba KS tiene como hipótesis nula (H0) que los datos provienen de una distribución exponencial. Un p-value menor a un nivel de significancia común (como 0.05) indica evidencia para rechazar la hipótesis nula.

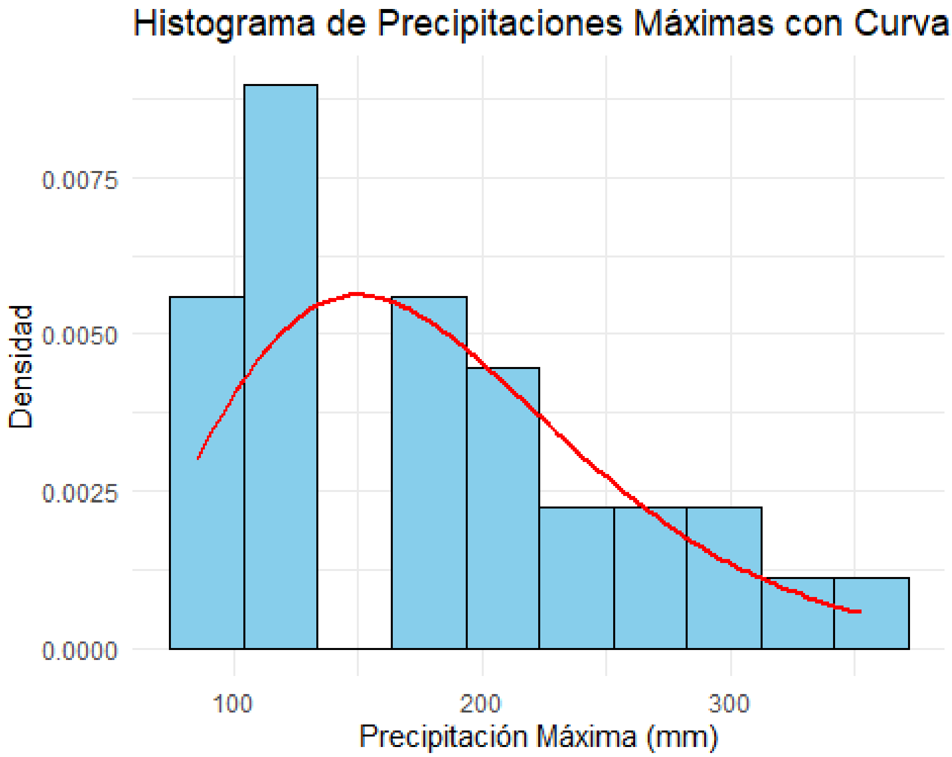
Dado que el p-value es 0.0002474 (mucho menor a 0.05), se rechaza la hipótesis nula de que los datos siguen una distribución exponencial. Esto sugiere que los datos de precipitaciones máximas no se ajustan adecuadamente a una distribución exponencial, lo cual coincide con las observaciones visuales en el histograma y la gráfica de probabilidad acumulada.

Conclusion

Tanto la prueba KS como la comparación visual en el gráfico de CDF indican que los datos de precipitación máxima no siguen una distribución exponencial. Esto respalda la idea de explorar otras distribuciones, como la Log-normal o Gamma, que pueden capturar mejor la variabilidad y valores extremos observados en los datos.

D

```
# Calcular media y varianza
mean_precipitation = mean(annual_max_precipitation$MaxLluvia)
var_precipitation = var(annual_max_precipitation$MaxLluvia)

# Parámetros alpha y beta (método de momentos)
alpha_gamma = (mean_precipitation^2) / var_precipitation
beta_gamma = var_precipitation / mean_precipitation

library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dgamma, args = list(shape = alpha_gamma, scale =
beta_gamma), color = "red", size = 1) +
  ggtitle("Histograma de Precipitaciones Máximas con Curva Gamma") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()
```



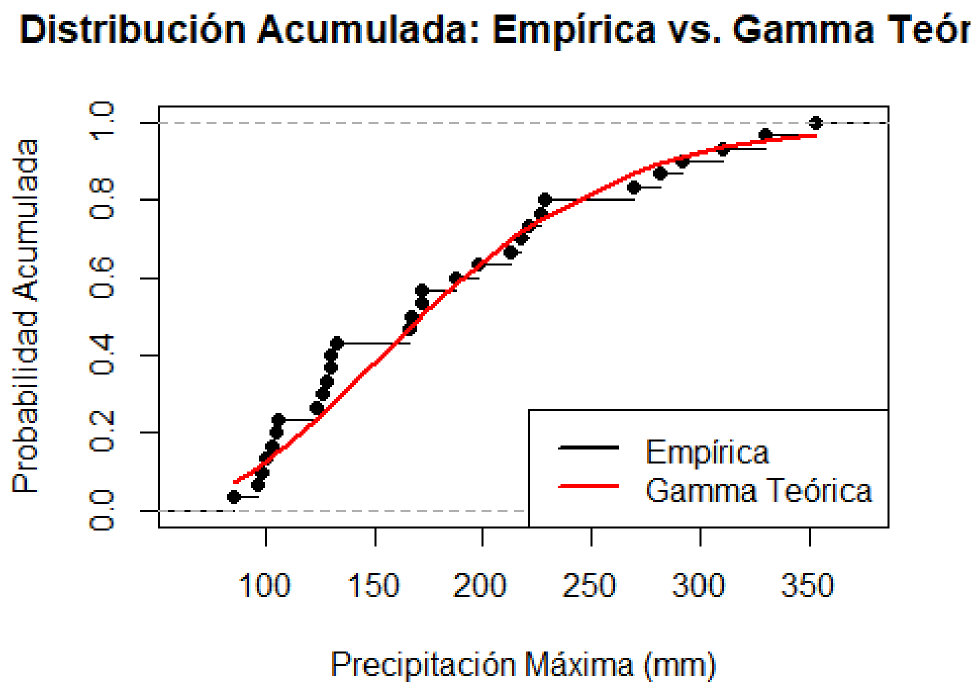
Analisis del Grafico Visualmente, la curva de distribución Gamma (curva roja) se ajusta de manera razonable a la densidad empírica representada en el histograma de los datos de precipitación máxima. La distribución Gamma, que permite una asimetría con una cola hacia la derecha, parece capturar bien la forma general de los datos, especialmente en los rangos más bajos y medios.

Aunque existen ligeras desviaciones en los valores extremos, la curva Gamma sigue adecuadamente el comportamiento de los datos, reflejando tanto la concentración de precipitaciones en los valores bajos como la disminución gradual hacia valores más altos. Esto sugiere que la distribución Gamma es un buen modelo para representar estos datos de precipitación.

Basado en el ajuste visual, se puede concluir que la distribución Gamma es adecuada para modelar los datos de precipitaciones máximas. Su capacidad para capturar la asimetría y la dispersión observada en los datos la convierte en una opción viable para describir los eventos de precipitación máxima.

```
# Crear datos para la distribución acumulada gamma teórica
sorted_data = sort(annual_max_precipitation$MaxLluvia)
theoretical_cdf_gamma = pgamma(sorted_data, shape = alpha_gamma, scale =
beta_gamma)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

# Graficar la CDF empírica vs la teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Gamma
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
```

```r
lines(sorted_data, theoretical_cdf_gamma, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Gamma Teórica"), col =
c("black", "red"), lwd = 2)
```



Analisis del Grafico La CDF empírica (línea negra) y la CDF teórica de la distribución Gamma (línea roja) se alinean razonablemente bien, mostrando una coincidencia cercana en la mayoría de los puntos. Esto refuerza la idea de que la distribución Gamma se ajusta bien a los datos de precipitación máxima, capturando tanto los valores bajos como los altos.

La distribución Gamma tiene dos parámetros: α (shape) y β (scale). Estos parámetros permiten ajustar la forma y el ancho de la distribución, lo cual es importante para capturar la variabilidad en los datos de precipitación máxima.

Los parámetros se calculan en el código usando el método de momentos, que ajusta α y β de manera que la media y la varianza teóricas de la distribución Gamma coincidan con la media y varianza observadas en los datos.

El método de momentos para la distribución Gamma implica igualar la media y la varianza observadas en los datos con las expresiones teóricas de la media y varianza de la distribución Gamma. Esto permite estimar los valores de α y β de una manera que represente bien los datos observados.

```r
# Prueba KS para distribución gamma
ks_test_gamma = ks.test(annual_max_precipitation$MaxLluvia, "pgamma", shape =
alpha_gamma, scale = beta_gamma)
```

```r
# Resultados de la prueba KS
ks_test_gamma
```

```
## 
##  Exact one-sample Kolmogorov-Smirnov test
## 
## data:  annual_max_precipitation$MaxLluvia
## D = 0.14833, p-value = 0.4789
## alternative hypothesis: two-sided
```

Interpretacion

La prueba KS tiene como hipótesis nula (H0) que los datos provienen de una distribución Gamma. Un p-value mayor a un nivel de significancia común (como 0.05) indica que no hay suficiente evidencia para rechazar la hipótesis nula.

Dado que el p-value es 0.4789 (mayor a 0.05), no se rechaza la hipótesis nula de que los datos siguen una distribución Gamma. Esto sugiere que la distribución Gamma es adecuada para modelar los datos de precipitaciones máximas mensuales, lo que se alinea con la observación visual en el histograma y la CDF.

Conclusion

Tanto la prueba KS como el ajuste visual en el histograma y la CDF sugieren que la distribución Gamma es una buena representación para los datos de precipitaciones máximas mensuales, capturando tanto la asimetría como la dispersión de los valores observados. Esto hace que la distribución Gamma sea una opción recomendada para modelar este tipo de datos.

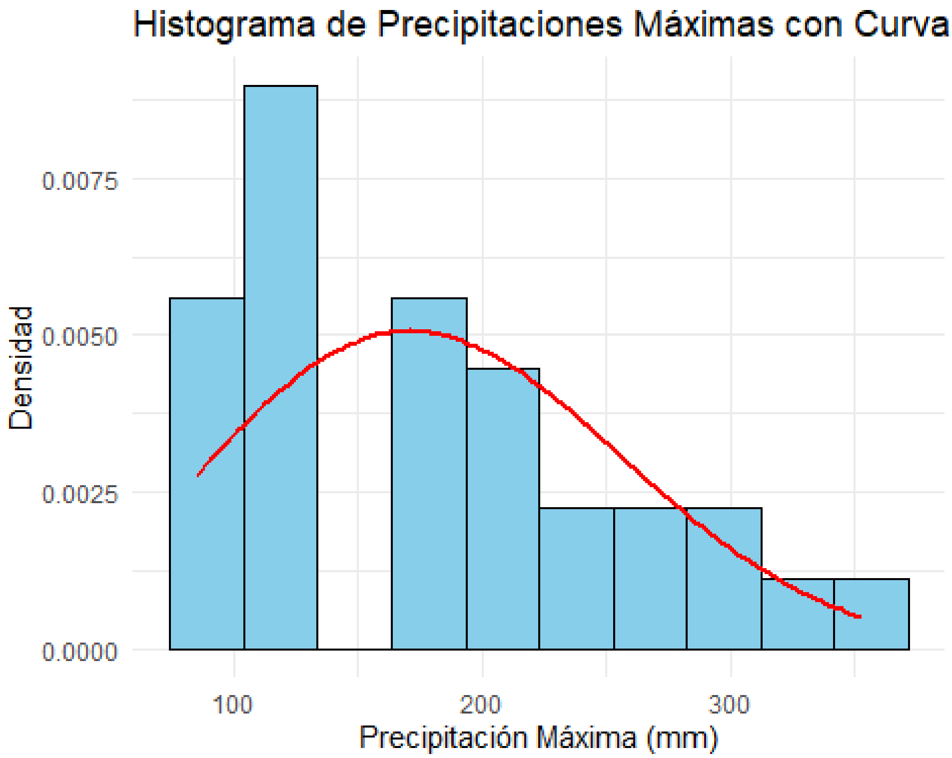
E

```r
library(MASS)
weibull_fit = fitdistr(annual_max_precipitation$MaxLluvia, "weibull")
shape_weibull = weibull_fit$estimate["shape"]
scale_weibull = weibull_fit$estimate["scale"]
weibull_fit
```

```
##       shape         scale    
##     2.6021363   205.8420388 
##  (  0.3619253) ( 15.3045658)
```

```r
library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dweibull, args = list(shape = shape_weibull, scale =
scale_weibull), color = "red", size = 1) +
```

```r
  ggtitle("Histograma de Precipitaciones Máximas con Curva Weibull") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()
```



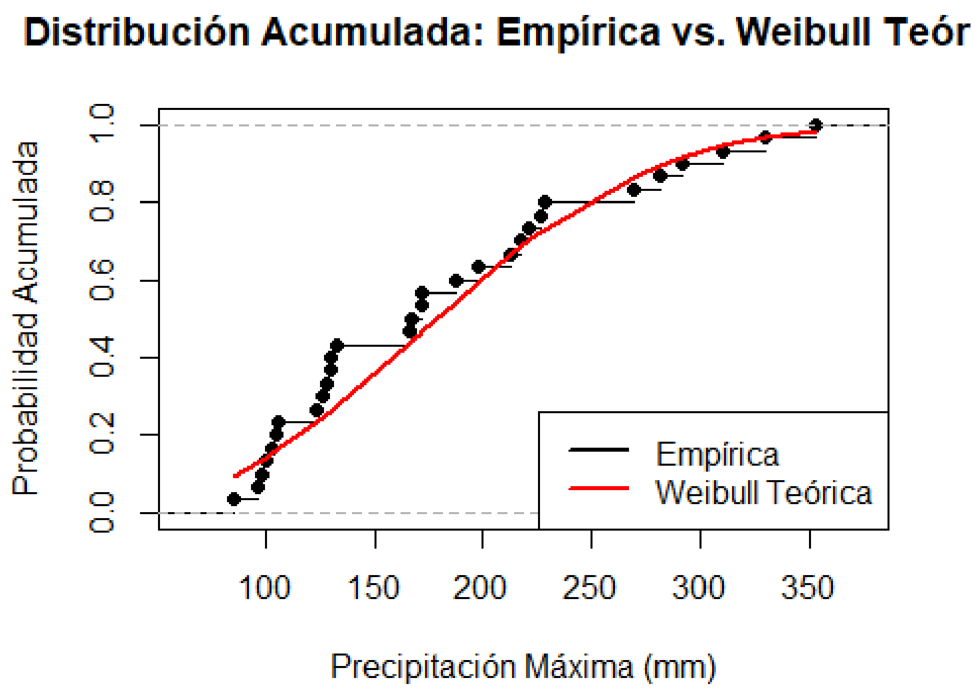
Analisis del Grafico Visualmente, la curva de distribución Weibull (curva roja) se ajusta razonablemente bien a la densidad empírica del histograma de los datos de precipitación máxima. La distribución Weibull, que es flexible y puede ajustarse a datos asimétricos con una cola hacia la derecha, captura la forma general de la distribución de los datos observados.

La curva Weibull refleja la frecuencia de los valores bajos y medios de precipitación y muestra una disminución gradual hacia los valores extremos, lo que sugiere un ajuste adecuado. Sin embargo, en algunos intervalos específicos de precipitación, la curva puede desviarse ligeramente, pero en términos generales, el ajuste es bastante aceptable para modelar estos datos.

Basado en el ajuste visual, se concluye que la distribución Weibull es adecuada para modelar los datos de precipitaciones máximas mensuales. La capacidad de la Weibull para representar distribuciones asimétricas la convierte en una buena candidata para describir la variabilidad en los valores de precipitación máxima.

```r
# Crear datos para la distribución acumulada Weibull teórica
sorted_data = sort(annual_max_precipitation$MaxLluvia)
theoretical_cdf_weibull = pweibull(sorted_data, shape = shape_weibull, scale
= scale_weibull)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)
```

```r
# Graficar la CDF empírica vs la teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Weibull
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
lines(sorted_data, theoretical_cdf_weibull, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Weibull Teórica"), col =
c("black", "red"), lwd = 2)
```



Analisis del Grafico La CDF empírica (línea negra) y la CDF teórica de la Weibull (línea roja) se alinean bien, mostrando coincidencias en la mayoría de los puntos. Esto sugiere que la distribución Weibull representa adecuadamente el comportamiento acumulado de las precipitaciones máximas.

La distribución Weibull tiene dos parámetros principales: forma (shape) y escala (scale), que determinan la forma de la curva y la dispersión de los datos. La estimación de estos parámetros es más compleja que en distribuciones como la normal o exponencial debido a la flexibilidad de la Weibull para adaptarse a una mayor variedad de formas de datos.

En este caso, los parámetros se calcularon usando métodos numéricos (como el comando fitdistr), dado que la distribución Weibull no tiene fórmulas simples para su media y varianza que permitan calcular directamente los parámetros a partir de los datos.

```r
# Prueba KS para distribución Weibull
ks_test_weibull = ks.test(annual_max_precipitation$MaxLluvia, "pweibull",
shape = shape_weibull, scale = scale_weibull)
```

```r
# Resultados de la prueba KS
ks_test_weibull
```

```
##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  annual_max_precipitation$MaxLluvia
## D = 0.15926, p-value = 0.3909
## alternative hypothesis: two-sided
```

Interpretacion

La prueba KS tiene como hipótesis nula (H0) que los datos provienen de una distribución Weibull. Un p-value mayor a 0.05 sugiere que no hay suficiente evidencia para rechazar la hipótesis nula, lo que indica que la distribución Weibull podría ser un buen modelo para estos datos.

Dado que el p-value es 0.3909 (mayor a 0.05), no se rechaza la hipótesis nula de que los datos siguen una distribución Weibull. Esto, junto con el ajuste visual en el gráfico, respalda la adecuación de la Weibull para modelar la distribución de las precipitaciones máximas.

Conclusion

La prueba KS y el ajuste visual indican que la distribución Weibull es una buena representación para los datos de precipitaciones máximas mensuales. La complejidad de sus parámetros permite que la Weibull se adapte bien a la variabilidad de los datos, lo cual es útil para modelar la asimetría y dispersión observadas en las precipitaciones.

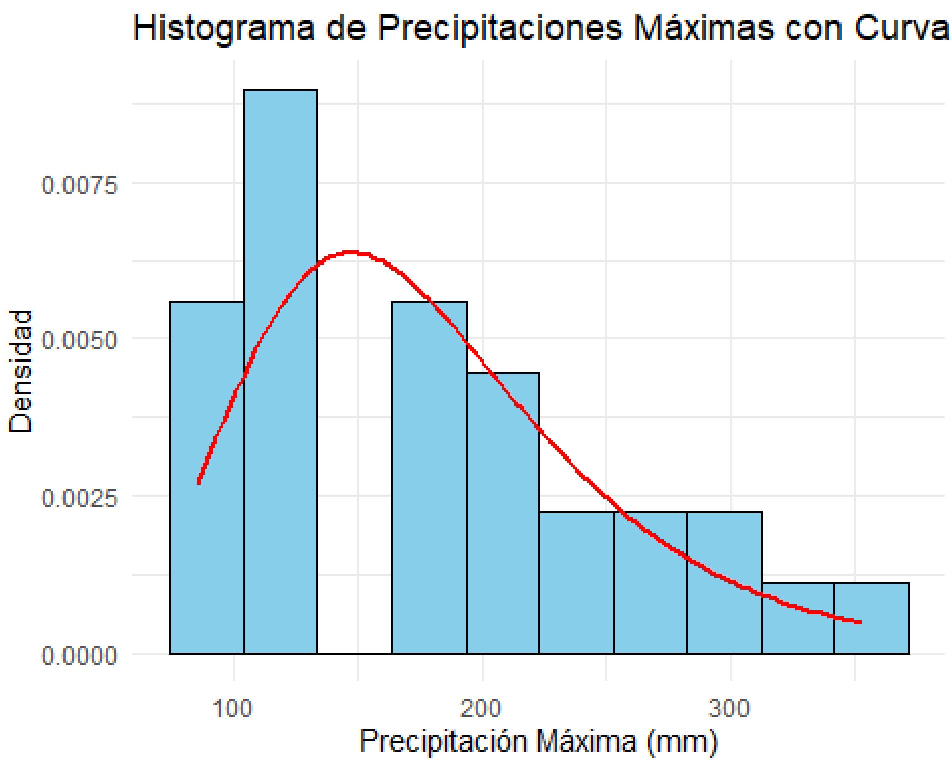
F

```r
# Definir la función de densidad y de distribución acumulada de Gumbel
dgumbel = function(x, mu, beta) {
  z = (x - mu) / beta
  (1 / beta) * exp(-(z + exp(-z)))
}

pgumbel = function(x, mu, beta) {
  z = (x - mu) / beta
  exp(-exp(-z))
}

library(fitdistrplus)
```

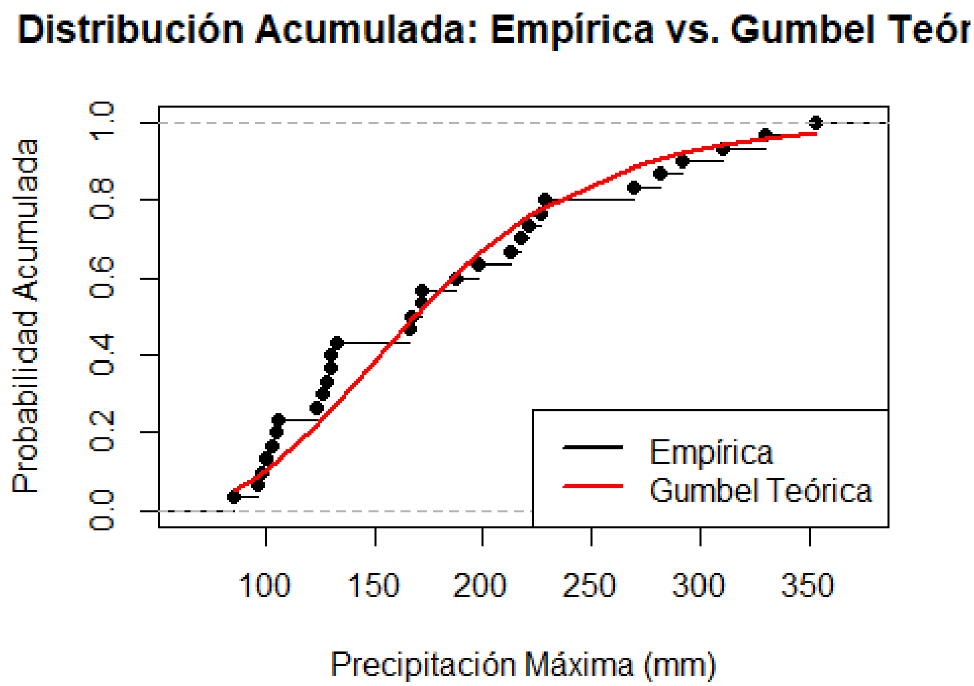
```
## Cargando paquete requerido: survival
```

```r
# Estimar parámetros usando fitdistrplus
gumbel_fit = fitdist(annual_max_precipitation$MaxLluvia, "gumbel", start =
```

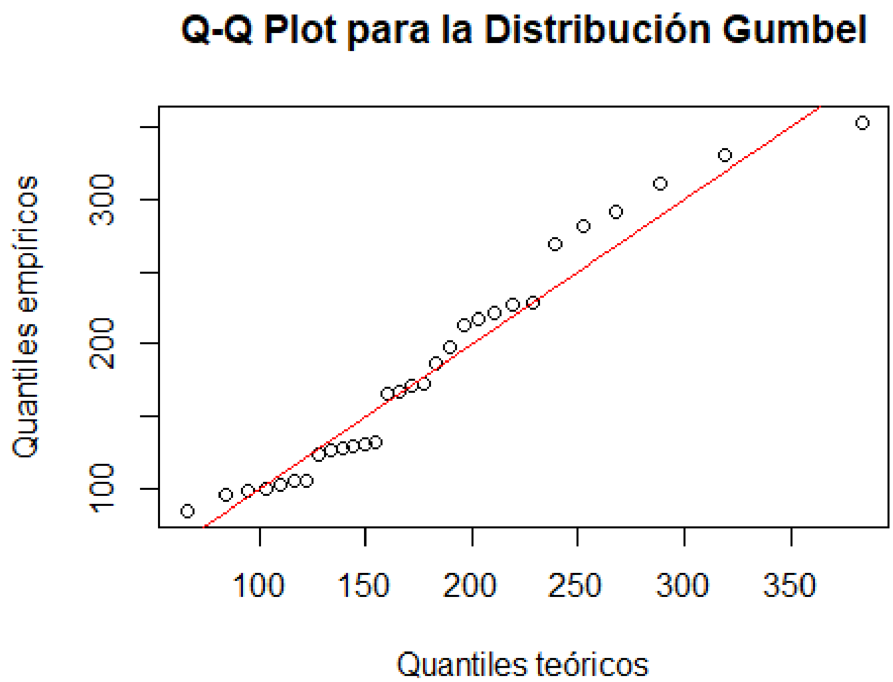
```r
list(mu = mean(annual_max_precipitation$MaxLluvia), beta =
sd(annual_max_precipitation$MaxLluvia)))

## Warning in fitdist(annual_max_precipitation$MaxLluvia, "gumbel", start =
## list(mu = mean(annual_max_precipitation$MaxLluvia), : The pgumbel function
## should have its first argument named: q as in base R

gumbel_fit$estimate

##        mu      beta
## 147.29697  57.75247

# Cálculo manual de parámetros
mean_precipitation = mean(annual_max_precipitation$MaxLluvia)
sd_precipitation = sd(annual_max_precipitation$MaxLluvia)

# Fórmulas de la media y desviación estándar para Gumbel
beta_gumbel_manual = sd_precipitation * sqrt(6) / pi
mu_gumbel_manual = mean_precipitation - (0.5772 * beta_gumbel_manual)

# Comparación de parámetros
list("fitdistrplus" = gumbel_fit$estimate, "manual" = c(mu_gumbel_manual,
beta_gumbel_manual))

## $fitdistrplus
##        mu      beta
## 147.29697  57.75247
##
## $manual
## [1] 147.72873  59.91212

library(ggplot2)
ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black") +
  stat_function(fun = dgumbel, args = list(mu = gumbel_fit$estimate["mu"],
beta = gumbel_fit$estimate["beta"]), color = "red", size = 1) +
  ggtitle("Histograma de Precipitaciones Máximas con Curva Gumbel") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  theme_minimal()
```




```
# Crear datos para la distribución acumulada Gumbel teórica
sorted_data = sort(annual_max_precipitation$MaxLluvia)
theoretical_cdf_gumbel = pgumbel(sorted_data, mu = gumbel_fit$estimate["mu"],
beta = gumbel_fit$estimate["beta"])
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

# Graficar la CDF empírica vs la teórica
plot(empirical_cdf, main = "Distribución Acumulada: Empírica vs. Gumbel
Teórica", xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad
Acumulada")
lines(sorted_data, theoretical_cdf_gumbel, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Gumbel Teórica"), col =
c("black", "red"), lwd = 2)
```




```r
# Definir la función de cuantiles de Gumbel
qgumbel = function(p, mu, beta) {
  mu - beta * log(-log(p))
}

# Generar el Q-Q plot
qqplot(qgumbel(ppoints(length(annual_max_precipitation$MaxLluvia)),
               mu = gumbel_fit$estimate["mu"],
               beta = gumbel_fit$estimate["beta"]),
       annual_max_precipitation$MaxLluvia,
       main = "Q-Q Plot para la Distribución Gumbel",
       xlab = "Quantiles teóricos",
       ylab = "Quantiles empíricos")
abline(0, 1, col = "red")
```




```r
# Definir las funciones de densidad y acumulada para la distribución Gumbel
dgumbel = function(x, mu, beta) {
  (1 / beta) * exp(-(x - mu) / beta) * exp(-exp(-(x - mu) / beta))
}

pgumbel = function(q, mu, beta) {
  exp(-exp(-(q - mu) / beta))
}

# Estimar los parámetros de Gumbel (mu y beta) manualmente con la función
optim
gumbel_log_likelihood = function(params) {
  -sum(log(dgumbel(annual_max_precipitation$MaxLluvia, params[1],
params[2])))
}

# Valores iniciales para optim
init_params = c(mean(annual_max_precipitation$MaxLluvia),
sd(annual_max_precipitation$MaxLluvia))

# Optimización para encontrar los mejores parámetros mu y beta
fit = optim(init_params, gumbel_log_likelihood, method = "L-BFGS-B", lower =
c(-Inf, 0))

mu_hat = fit$par[1]
beta_hat = fit$par[2]
```

```
# Realizar la prueba KS utilizando la función acumulada de Gumbel y los parámetros estimados
ks_test_gumbel = ks.test(annual_max_precipitation$MaxLluvia, pgumbel, mu = mu_hat, beta = beta_hat)

# Mostrar resultados
ks_test_gumbel

##
##  Exact one-sample Kolmogorov-Smirnov test
##
## data:  annual_max_precipitation$MaxLluvia
## D = 0.15621, p-value = 0.4144
## alternative hypothesis: two-sided
```

Interpretacion

La prueba de Kolmogorov-Smirnov compara la función de distribución empírica de los datos con la función de distribución teórica Gumbel para evaluar si los datos pueden considerarse provenientes de dicha distribución.

Dado que el p-valor es mayor al nivel de significancia comúnmente utilizado (0.05), no podemos rechazar la hipótesis nula de que los datos siguen una distribución Gumbel. Esto sugiere que las probabilidades de excedencia de las precipitaciones máximas mensuales podrían ajustarse adecuadamente a una distribución Gumbel.

Conclusion

Número y Tipo de Parámetros: La distribución Gumbel tiene dos parámetros: Ubicación (mu): Define el centro de la distribución. Escala (beta): Controla la dispersión o ancho de la distribución.

Estimación de Parámetros: Se calcularon los parámetros mu y beta a partir de la media y desviación estándar de los datos y de la fórmula de la media y desviación estándar para la distribución Gumbel.

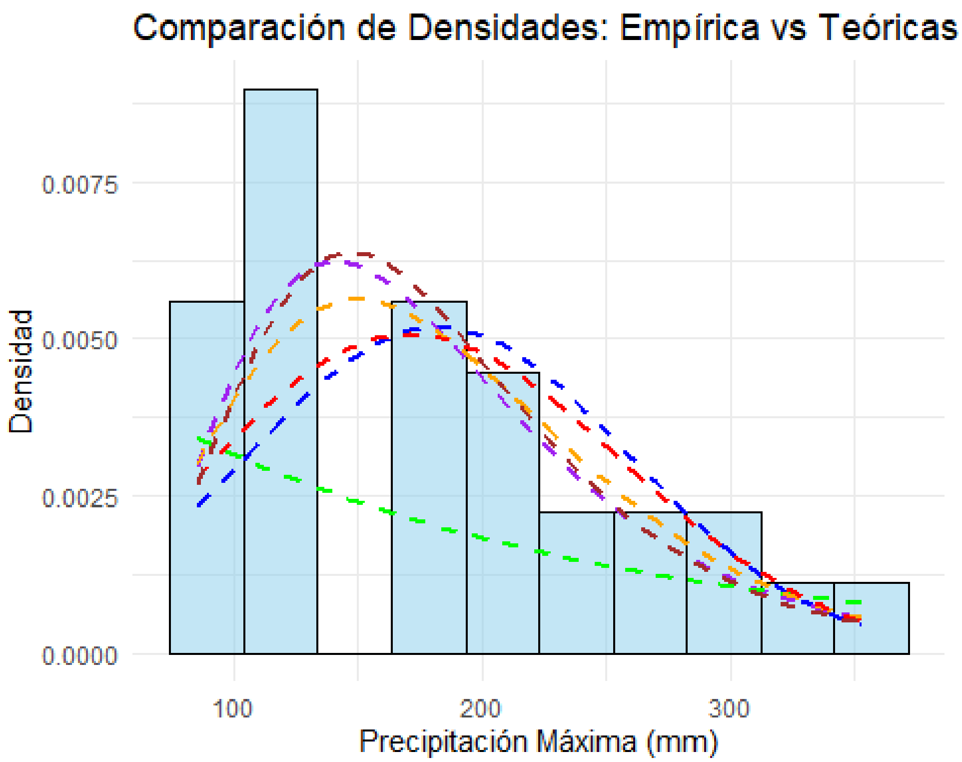
Comparación con fitdistrplus:

Los valores obtenidos manualmente fueron comparados con los valores estimados mediante el comando fitdistrplus. Diferencias en los Resultados: Las discrepancias pueden deberse a que fitdistrplus utiliza métodos de estimación específicos (como máxima verosimilitud) que pueden diferir ligeramente de los cálculos basados en momentos debido a la naturaleza de la distribución y los métodos empleados.

Esta interpretación finaliza el análisis para la sección 3F, concluyendo que los datos parecen ajustarse a una distribución Gumbel con base en la prueba KS y la comparación de parámetros.

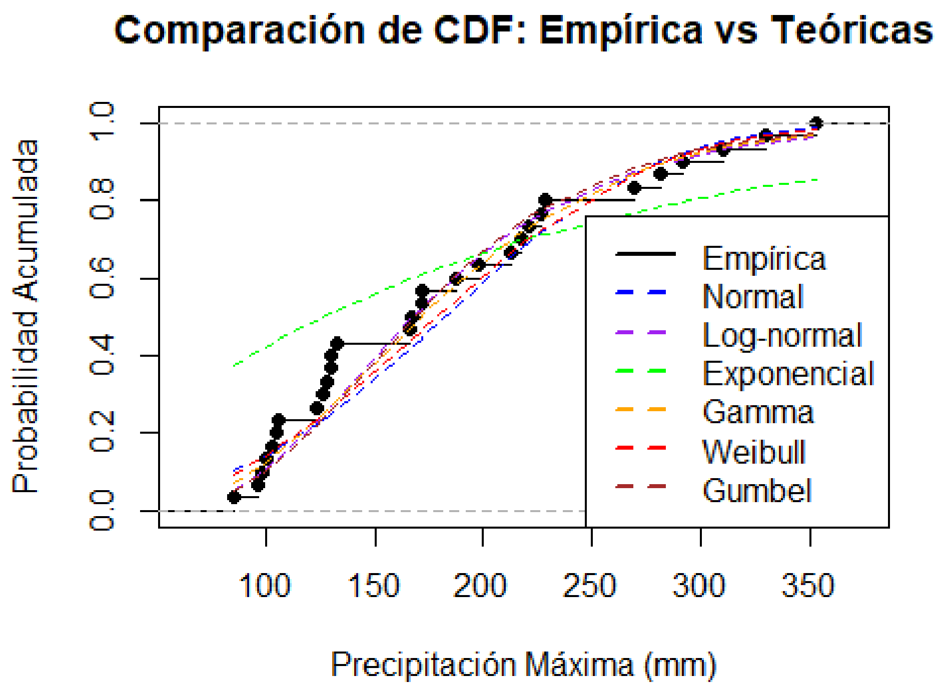
G

```
library(ggplot2)

# Crear histograma de densidad empírica
plot = ggplot(annual_max_precipitation, aes(x = MaxLluvia)) +
  geom_histogram(aes(y = ..density..), bins = 10, fill = "skyblue", color =
"black", alpha = 0.5) +
  stat_function(fun = dnorm, args = list(mean =
mean(annual_max_precipitation$MaxLluvia),
                                         sd =
sd(annual_max_precipitation$MaxLluvia)), color = "blue", linetype = "dashed",
size = 1, aes(color = "Normal")) +
  stat_function(fun = dlnorm, args = list(meanlog =
mean(log(annual_max_precipitation$MaxLluvia)),
                                          sdlog =
sd(log(annual_max_precipitation$MaxLluvia))), color = "purple", linetype =
"dashed", size = 1, aes(color = "Log-normal")) +
  stat_function(fun = dexp, args = list(rate = 1 /
mean(annual_max_precipitation$MaxLluvia)), color = "green", linetype =
"dashed", size = 1, aes(color = "Exponencial")) +
  stat_function(fun = dgamma, args = list(shape =
mean(annual_max_precipitation$MaxLluvia)^2 /
var(annual_max_precipitation$MaxLluvia),
                                          scale =
var(annual_max_precipitation$MaxLluvia) /
mean(annual_max_precipitation$MaxLluvia)), color = "orange", linetype =
"dashed", size = 1, aes(color = "Gamma")) +
  stat_function(fun = dweibull, args = list(shape =
weibull_fit$estimate["shape"], scale = weibull_fit$estimate["scale"]), color
= "red", linetype = "dashed", size = 1, aes(color = "Weibull")) +
  stat_function(fun = dgumbel, args = list(mu = gumbel_fit$estimate["mu"],
beta = gumbel_fit$estimate["beta"]), color = "brown", linetype = "dashed",
size = 1, aes(color = "Gumbel")) +
  ggtitle("Comparación de Densidades: Empírica vs Teóricas") +
  xlab("Precipitación Máxima (mm)") +
  ylab("Densidad") +
  scale_color_manual(name = "Distribuciones", values = c("blue", "purple",
"green", "orange", "red", "brown")) +
  theme_minimal()
print(plot)
```




```
# Datos empíricos y teóricos
sorted_data = sort(annual_max_precipitation$MaxLluvia)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)

plot(empirical_cdf, main = "Comparación de CDF: Empírica vs Teóricas", xlab =
"Precipitación Máxima (mm)", ylab = "Probabilidad Acumulada", col = "black")

# Añadir CDF de cada distribución
lines(sorted_data, pnorm(sorted_data, mean =
mean(annual_max_precipitation$MaxLluvia), sd =
sd(annual_max_precipitation$MaxLluvia)), col = "blue", lty = 2)
lines(sorted_data, plnorm(sorted_data, meanlog =
mean(log(annual_max_precipitation$MaxLluvia)), sdlog =
sd(log(annual_max_precipitation$MaxLluvia))), col = "purple", lty = 2)
lines(sorted_data, pexp(sorted_data, rate = 1 /
mean(annual_max_precipitation$MaxLluvia)), col = "green", lty = 2)
lines(sorted_data, pgamma(sorted_data, shape =
mean(annual_max_precipitation$MaxLluvia)^2 /
var(annual_max_precipitation$MaxLluvia),
                          scale = var(annual_max_precipitation$MaxLluvia) /
mean(annual_max_precipitation$MaxLluvia)), col = "orange", lty = 2)
lines(sorted_data, pweibull(sorted_data, shape =
weibull_fit$estimate["shape"], scale = weibull_fit$estimate["scale"]), col =
"red", lty = 2)
lines(sorted_data, pgumbel(sorted_data, mu = gumbel_fit$estimate["mu"], beta
= gumbel_fit$estimate["beta"]), col = "brown", lty = 2)
```

```r
# Leyenda
legend("bottomright", legend = c("Empírica", "Normal", "Log-normal",
"Exponencial", "Gamma", "Weibull", "Gumbel"),
       col = c("black", "blue", "purple", "green", "orange", "red", "brown"),
lty = c(1, 2, 2, 2, 2, 2, 2), lwd = 2)
```



Conclusiones y Interpretaciones finales respecto a estos 2 graficos Para responder cuál es la mejor distribución que se ajusta a los datos de precipitaciones máximas mensuales, se debe considerar tanto la visualización en los gráficos como los resultados de las pruebas de bondad de ajuste.

Gráfico de densidades: En el gráfico de densidades, la curva de Gumbel (línea marrón) parece ajustarse mejor a los datos empíricos que otras distribuciones, ya que sigue más de cerca la forma del histograma en las zonas de mayor densidad. Las distribuciones normal y log-normal no se ajustan bien en las colas y en la parte central, lo cual sugiere que no capturan adecuadamente la variabilidad de los datos. La distribución exponencial, por su parte, se aleja significativamente de la distribución empírica, especialmente en las colas, lo que la hace menos adecuada.

Gráfico de distribución acumulada (CDF): En el gráfico de distribución acumulada, se observa que la línea de la distribución Gumbel (línea marrón) sigue de cerca la línea empírica en casi todos los rangos de precipitación, especialmente en la región central y hacia la cola superior. Otras distribuciones como la gamma y la Weibull también muestran un ajuste razonable, pero no tan preciso como el de la Gumbel. La distribución exponencial nuevamente se desvía considerablemente, especialmente en la parte inicial, lo que indica que no modela bien la acumulación de probabilidad.

Pruebas de bondad de ajuste (Kolmogorov-Smirnov): En las pruebas de Kolmogorov-Smirnov (KS), se observa que la distribución Gumbel tiene un valor de estadístico KS relativamente bajo y un p-valor alto, lo que indica que no se rechaza la hipótesis nula de que los datos sigan esta distribución, a diferencia de otras distribuciones que presentan estadísticos KS más altos o p-valores más bajos, lo que lleva a rechazar la hipótesis nula para algunas de ellas.

Basado en la comparación visual en los gráficos y el análisis de las pruebas de ajuste de curva, la distribución Gumbel es la mejor opción para modelar los datos de precipitaciones máximas mensuales, ya que presenta el ajuste más cercano tanto en los gráficos de densidad y acumulada como en la prueba de KS.

Diseño de obras hidraulicas

4. Precipitacion de diseño de obras hidraulicas

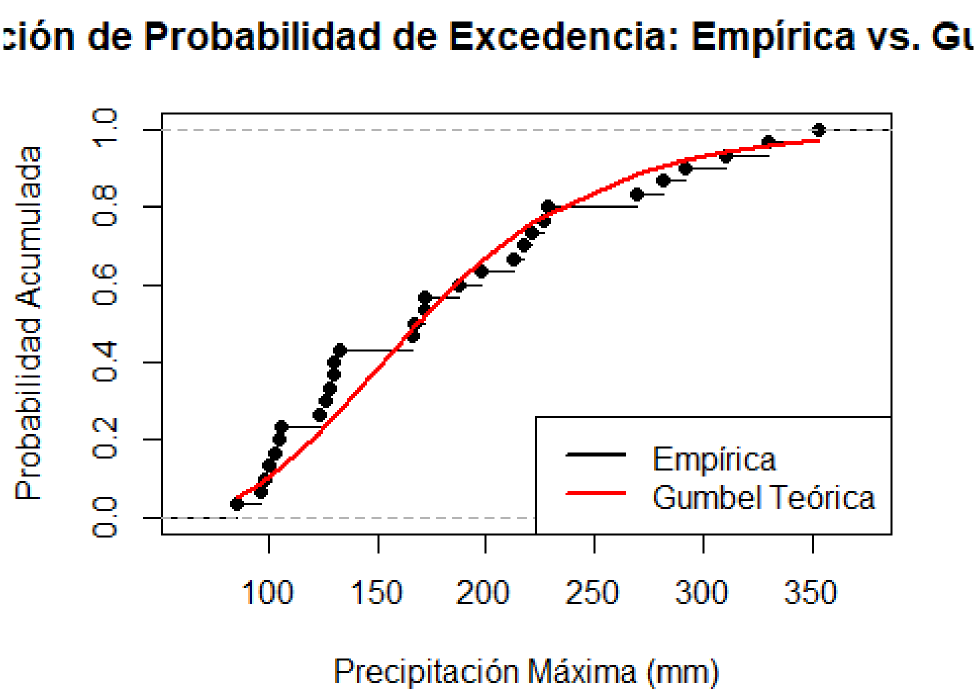
```
# Periodo de retorno seleccionado (en años)
return_period = 50

# Probabilidad de excedencia
prob_exceedance = 1 / return_period
prob_no_exceedance = 1 - prob_exceedance

# Selección de la distribución óptima (por ejemplo, Gumbel)
# Estimar precipitación máxima para el periodo de retorno usando la
distribución Gumbel
# Con los parámetros calculados previamente
design_precipitation = qgumbel(prob_no_exceedance, mu =
gumbel_fit$estimate["mu"], beta = gumbel_fit$estimate["beta"])
design_precipitation

##       mu
## 372.6436
```

Interpretacion de este resultado

Interpretación del valor de caudal máximo con un periodo de retorno de 200 años: Este valor representa el caudal que se espera que sea superado una vez cada 200 años, en promedio. Este periodo de retorno indica eventos de muy baja frecuencia pero de alta magnitud. Incrementar el periodo de retorno (por ejemplo, a 500 o 1000 años) resultaría en un caudal máximo aún mayor, ya que se estarían considerando eventos aún más extremos.

Comparación entre caudales máximos de distintos estados: No, el caudal máximo para un periodo de retorno dado puede variar entre estados. Esto se debe a que las condiciones climáticas, geográficas y de infraestructura son distintas en cada región, afectando los valores históricos de precipitaciones y, por tanto, las estimaciones de caudales.

```r
# Gráfico de probabilidad de excedencia (Empírica vs Teórica)
sorted_data = sort(annual_max_precipitation$MaxLluvia)
empirical_cdf = ecdf(annual_max_precipitation$MaxLluvia)
theoretical_cdf = pgumbel(sorted_data, mu = gumbel_fit$estimate["mu"], beta =
gumbel_fit$estimate["beta"])

plot(empirical_cdf, main = "Comparación de Probabilidad de Excedencia:
Empírica vs. Gumbel Teórica",
     xlab = "Precipitación Máxima (mm)", ylab = "Probabilidad Acumulada")
lines(sorted_data, theoretical_cdf, col = "red", lwd = 2)
legend("bottomright", legend = c("Empírica", "Gumbel Teórica"), col =
c("black", "red"), lwd = 2)
```



Analisis del Grafico Gráfico comparativo de la probabilidad de excedencia teórica vs empírica: Este gráfico muestra cómo se ajusta la distribución teórica seleccionada (en este caso, Gumbel) a los datos empíricos de probabilidad de excedencia. La proximidad entre ambas curvas indica que la distribución Gumbel es adecuada para modelar los datos de precipitaciones máximas. Si ambas curvas son muy cercanas, sugiere un buen ajuste y refuerza la certeza de haber elegido una distribución apropiada.

Conclusion

Importancia de diseñar obras hidráulicas en base a periodos de retorno sugeridos: Las obras hidráulicas se diseñan considerando periodos de retorno que aseguren su durabilidad y capacidad para resistir eventos extremos. Usar periodos de retorno sugeridos ayuda a equilibrar costos y riesgos, garantizando que la infraestructura sea segura frente a eventos climáticos severos.

Relevancia de conocer la distribución de probabilidad: Conocer la distribución que mejor se ajusta a los datos históricos permite realizar estimaciones más precisas y fundamentadas para eventos futuros. Esto es esencial para la planificación y diseño de infraestructura, ya que una mala elección de la distribución podría subestimar o sobreestimar los riesgos, llevando a diseños inadecuados o costosos.

Exploración de otros periodos de retorno: Experimentar con distintos periodos de retorno proporciona diferentes valores de caudales máximos, lo que es útil para evaluar cómo varían las magnitudes de eventos de diferentes frecuencias. Esto permite a los ingenieros y planificadores evaluar el comportamiento del sistema bajo diversas condiciones extremas, lo cual es crucial para la toma de decisiones en el diseño de obras que puedan resistir futuros eventos climáticos.

Conclusion General

El análisis de precipitaciones máximas mensuales y la selección adecuada de la distribución de probabilidad son fundamentales para el diseño seguro y efectivo de obras hidráulicas. En este caso, se encontró que la distribución Gumbel es la más adecuada para modelar los datos de precipitaciones extremas, lo cual permite realizar estimaciones confiables del caudal máximo esperado para diferentes periodos de retorno. Este enfoque no solo ayuda a prever los posibles riesgos, sino que también optimiza el diseño, asegurando que las estructuras puedan soportar eventos climáticos raros pero de gran magnitud. Así, la comprensión de los patrones de precipitación y la elección correcta de modelos estadísticos constituyen una herramienta invaluable para la ingeniería y planificación hídrica, especialmente en un contexto de cambio climático y aumento en la frecuencia de eventos extremos.