UNIVERSIDAD NACIONAL DE COLOMBIA

PREGRADO EN ESTADISTICA

Departamento de Estadística
Facultad de Ciencias

Actuaria de Contingencias de Vida

Profesor:
Norman Diego Giraldo Gomez
Segundo Trabajo de Actuaria

Integrantes:

Ricardo William Salazar Espinal C.C. 1017219472

Medellín, Colombia

Medellin, septiembre 2 de 2024

Índice

Índice de figuras

Índice de cuadros

1 Objetivo

Un departamento de estructuracion de un Banco diseña una anualidad a n = 10 años, financiada mediante tasas de rendimiento aleatorio de un fondo de inversiones. La anualidad es de tipo Lineal, con pagos trimestrales (m = 4) vencido, con valor inicial de C = 2.5 unidades monetarias. Se pacta una tasa de incremento de costo de vida anual, de iq efectiva anual. Desarrolle los siguientes puntos.

Para el analisis de los dos primeros puntos consideramos datos entre $2014-08-26$ y $2024-08-23$ datos del fondo "Vanguard LifeStrategy Growth Fund".

2 Punto a

Con los datos de rendimientos diarios del archivo asignado, agreguelos a trimestre. Ajuste el modelo NIG las tasas geometricas $\delta_m(k) = log(1 + im(k))$. Reporte: los parametros estimados, la grafica de ajuste de log-densidades observadas versus estimadas y la varianza $\sigma^2 = Var(\delta_m(k))$(para VG use vgVar() en VarianceGamma y nigVar para la NIG en GeneralizedHyperbolic).

2.1 Solucion

Definimos que la variable aleaotria $X = \mu + \delta X_e \sim NIG(\mu, \alpha, \beta, \delta)$ esta tiene una distribucion normal inversa Gaussiana (NIG). esta tiene una funcion de distribucion de la forma:

$$f(x) = \frac{\alpha\delta(\alpha(\sqrt{\delta^2 + (x-\mu)^2})}{\pi\sqrt{\delta^2 + (x-\mu)^2}}$$

En esta primera parte del trabajo estimaremos estos parametros mediante algunas librerias y calculos para ello usaremos los rendimientos diarios del fondo a estudiar y posteriormente observaremos el comportamiento como una distribucion NIG,de la cual obtendremos unos parametros estimados para posteriormente compararlos con la distribucion de $log(1 + im_k(k)) \sim iddNIG(\mu, \alpha, \beta, \delta)$, donde estos logaritmos tambien deberian seguir una distribuccion **NIG**.

A continuacion podemos observar en la tabla los valores para los parametros estimado de una **NIG**

```
        mu       delta       alpha        beta
 0.03881244  0.04042249 15.40577649 -9.17142223
```

```
[1] 0.005058298
```

Tabla 1: Datos Estimados NIG Trimestre

	Variable	Valor
mu	mu	0.0388124
delta	delta	0.0404225
alpha	alpha	15.4057765
beta	beta	-9.1714222

Tenemos adicionalmente la varianza ($\delta_m(k)$) como:

Tabla 2: Varianza de $delta_m(k)$

Variable	Valor
var(delta_m(k))	0.0050583

En el analisis de los rendimientos trimestrales, hemos calculado la varianza de $\delta_m(k)$, que es la transformacion logaritmica de la suma de los rendimientos trimestrales, y obtenemos un valor de 0.005058298 .Este valor de varianza relativamente bajo sugiere que los rendimientos trimestrales tienden a mostrar una variabilidad moderada cuando se transforman logaritmicamente. La transformacion logaritmica ayuda a reducir la influencia de los valores extremos, proporcionando una medida mas estable de la variabilidad de los rendimientos.

Ahora presentamos un grafico de la tendencia de las log densidades a lo largo de un periodo de 10 años.

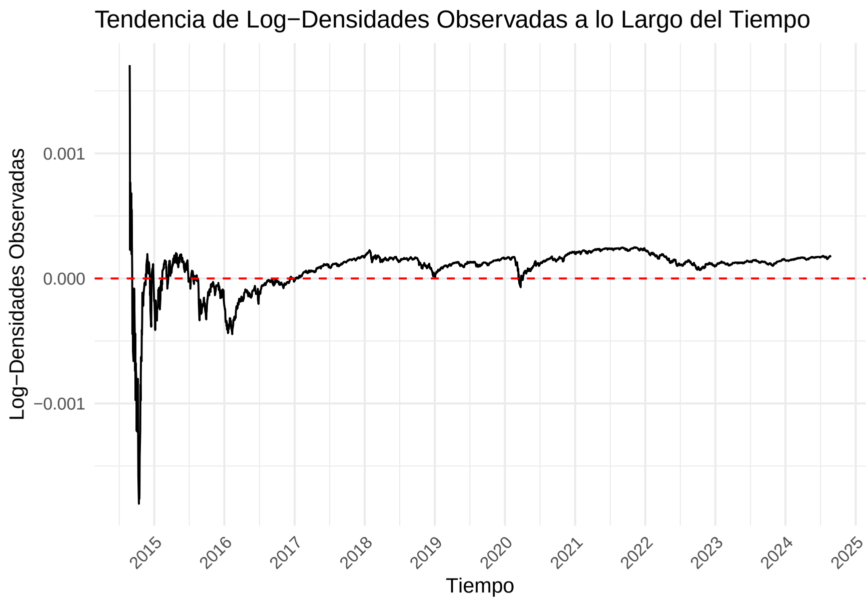

Figura 1: Log-Densidades Observadas

podemos observar que en los años del 2014 al 2016 la tendencia es a tener una alta obsilacion, esto se puede deber a que al principio tenemos una precencia de pocos datos, luego de los años posteriores al 2017 podemos observar una tendencia a tener poca obsilacion teniendo una gran estabilidad.

A continuacion realizaremos un grafico de ajuste de log-densidades observadas versus estimadas, donde podemos observar que:

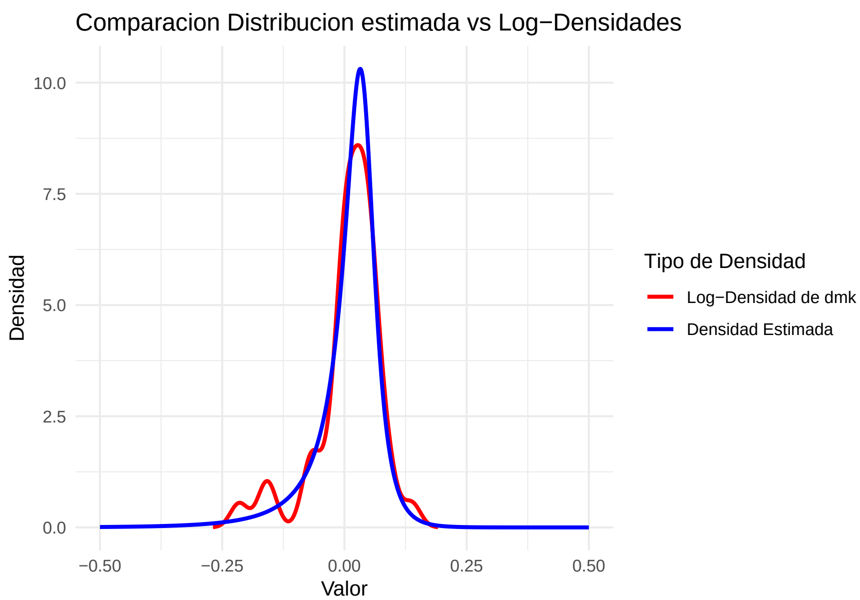

Figura 2: Comparacion Distribucion estimada vs Log-Densidades

Ambas curvas parecen tener una forma muy similar, lo que sugiere que la Log-Densidad de dmk y la Densidad Estimada representan de manera similar la distribucion subyacente de los datos. Ambas densidades tienen un pico muy pronunciado alrededor de 0, lo cual es tipico de distribuciones centradas.

En los extremos (tanto negativos como positivos), se puede observar que la curva verde (Log-Densidad de dmk) tiene mas oscilaciones en comparacion con la curva morada (Densidad Estimada), que es mas suave. Esto podria indicar que la Log-Densidad de dmk tiene mas variabilidad o ruido en los valores de baja probabilidad.

Al convertir rendimientos diarios en trimestrales, se estan sumando o promediando muchos rendimientos individuales. Este proceso puede generar variaciones adicionales en la densidad estimada, especialmente si los rendimientos diarios presentan fluctuaciones significativas.

3 Punto b

Calcule valor de la anualidad Geometrica o Lineal, cierta, asumiendo: una tasa efectiva anual de ia, equivalente a la media delas tasas diarias positiva del fondo (ver procedimiento en los ejemplos). Escoja una tasa de inflacion efectiva anual $0 < iq < ia$, de tal forma que los pagos se incrementan anualmente, q = 1, a esta tasa.

3.1 Solucion

Lo primero es calcular una tasa ia adecuada, para ello consideraremos los periodos de tiempo donde los $i_m(k) > 0$,llamaremos a esta $mi_m(k)$ luego promediamos estos valores para encontrar una tasa ia adecuada.

$$i_a = (1 + E(mi_m(k))^{360/1} - 1$$

donde como tenemos nuestros rendimientos diarios por lo tanto consideramos los 360 dias y tenemos un aumento cada años por lo cual tomamos $q = 1$.

Al realizar este calculo encontramos que $i_a = 13.59141\%$, por lo tanto proponemos una tasa de inflacion menor a esta para realizar el computo de los pagos,proponemos $iq = 1\%$ la cual es una tasa de inflacion adecuada esto considerando que este es una valor aproximado en la cual oscila en estados unidos, teniendo encuenta que una sobre estimacion en esta puede indicar un pago mayor al previsto, por lo tanto es importante proponer una cuota acorde con el mercado.

definimos $C_j = C * (1 + i_q)^{\frac{\lfloor \frac{t \cdot q}{m} \rfloor}{q}}$

donde:

- C C=2.5 (valor base de cada cuota)
- i_q 0.01 (1%de incremento anual)
- m es el numero de periodos por año.
- n es el numero total de años.
- q es el incremento seria 1 ya que se incrementa anualmente.

Podemos calcular como serian los pagos de las cuotas C_j para cada uno de los siguientes trimestres durante 10 años. Podemos ver como serian estos años en el siguiente grafico.

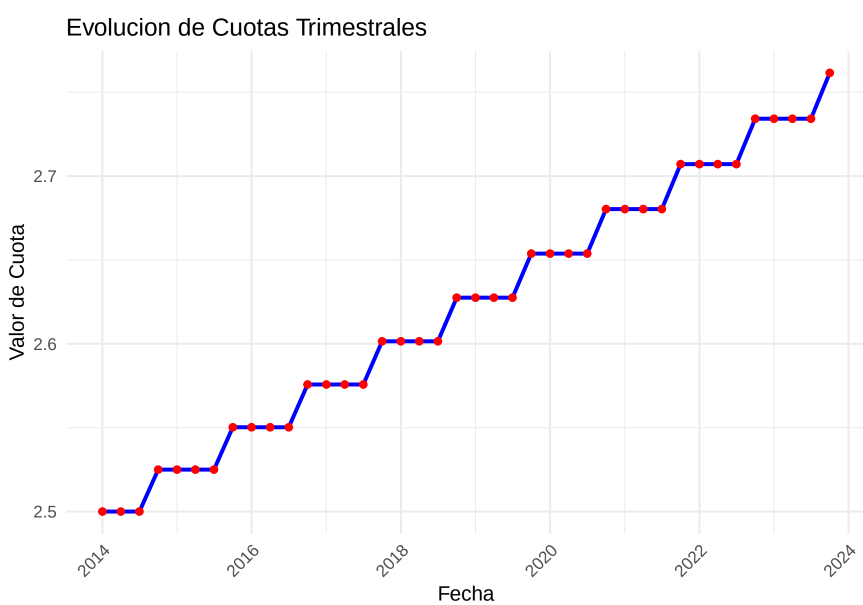

Figura 3: Evolucion de Cuotas Trimestrales

Para calcular el valor de la anualidad con incrementos lineales consideramos el valor presente $F(0)$, el cual tiene como formula:

$$\left(L^{(q)}a\right)\frac{(m)}{n|} = \frac{1}{m}\sum_{t=1}^{n\cdot m}(1+i)^{-\frac{t}{m}}\left(1+\frac{\rho}{q}\left\lfloor\frac{(t-1)\cdot q}{m}\right\rfloor\right)$$

donde:

- PV es el valor presente ajustado.
- i es la tasa de interes.
- m es el numero de periodos por año.
- n es el numero total de años.
- ρ es el incremento anual.
- q es el numero de periodos en un año (en nuestro caso 4 para trimestres).
- $\lfloor\cdot\rfloor$ denota la funcion piso, que devuelve el mayor entero menor o igual al argumento.

donde el valor de anualidad seria:

$$C_p = m * n * \left(L^{(q)}a\right)\frac{(m)}{n|}$$

el cual tiene un valor de 57.567469 el cual es el valor de la anualidad con 40 pagos vencidos.

4 Punto c

Calcule la matriz de varianzas-covarianzas de las $\wedge(j)$ y simule N = 500 valores de estas variables con el procedimiento de simulacion de una copula Gaussiana con marginales NIG.Reporte algunos histogramas y diagramas de dispersion con dos de estas variables. Genere la exponencial de estas variables y luego genere la suma S de estas exponenciales, que es una LogNIG. Muestre el histograma.

4.1 Solucion

para dar solucion a este problema usando simulacion debemos seguir los siguientes pasos:

Asumiremos $\delta_m(k) \sim$ i.i.d $\mathrm{NIG}(\mu,\alpha,\beta,\gamma)$, luego dado la propiedad reproductiva $\wedge(j) \sim \mathrm{NIG}(j\mu,\alpha,\beta,j\gamma)$. esto nos permitira realizar la simulacion de manera correcta, ya que tenemos distribuciones **NIG** es suficiente con utilizar los parametros obtenidos en la primera parte del trabajo para proceder con la simulacion.

Para la comparacion entre simulaciones consideraremos la simulacion numero 1 y la simulacion numero 40 en estas podemos observar que:

4.1.1 Analisis de comparacion simulaciones 1 y 40

Simulacion 1:

La grafica muestra una distribucion con un claro sesgo a la derecha. Esta asimetria es consistente con lo que se espera de la distribucion **NIG**, donde uno de los parametros β controla la asimetria.

La presencia de colas largas en el lado izquierdo sugiere que hay una mayor probabilidad de valores extremos negativos, una caracteristica tipica de las colas pesadas en la distribucion NIG.

La menor dispersion y la concentracion de valores cerca de cero podrian indicar un parametro δ mas bajo, lo que reduce la varianza de la distribucion.

Simulacion 40:

La grafica de la Simulacion 40 es mas simetrica, lo que indica que la asimetria controlada por β en esta simulacion es menor o diferente en comparacion con la Simulacion 1.

Aun asi, se observa cierta dispersion y las colas de la distribucion son mas pronunciadas que en una distribucion normal estandar, reflejando de nuevo las colas pesadas caracteristicas de la NIG.

La mayor dispersion sugiere un parametro δ mas alto, lo que lleva a una mayor varianza en los datos simulados. Esta diferencia puede ser el resultado de la variabilidad natural en las simulaciones o de variaciones en los parametros de la funcion *NIG*.

Comparacion entre Simulaciones:

La diferencia en la forma y dispersion entre las simulaciones 1 y 40 refleja como diferentes instancias de la distribucion **NIG** pueden generar resultados significativamente diferentes, incluso si estan basadas en la misma estructura de correlacion **Rho**.

Esto puede ser crucial en aplicaciones practicas, como la evaluacion de riesgos, donde diferentes realizaciones de un modelo NIG pueden llevar a conclusiones diferentes sobre el riesgo o el retorno esperado de un portafolio de activos.

Interpretacion en Contexto Financiero:

Si estas simulaciones representan retornos de activos, por ejemplo, la Simulacion 1 podria estar mostrando un activo con retornos relativamente estables pero con un riesgo significativo de grandes perdidas (gran acomulacion a la derecha).

La Simulacion 40 podria estar representando un activo mas volatil, con mayor variabilidad en sus retornos y una distribucion de riesgos mas balanceada, aunque todavia existe la posibilidad de eventos extremos en ambos lados de la distribucion.

En resumen, las graficas proporcionan una vision visual de como los parametros de la distribucion *NIG* y la estructura de correlacion afectan las distribuciones simuladas. Este tipo de analisis es crucial para entender la gama de posibles resultados en un contexto de gestion de riesgos o simulaciones financieras.

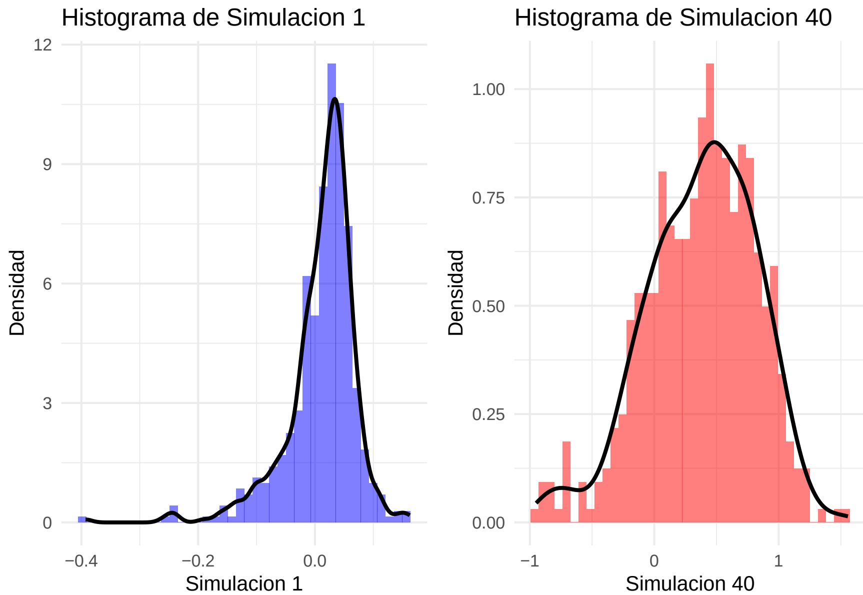

Figura 4: Comparacion Simulaciones numero 1 y 40

Ahora analizaremos un diagrama de dispercion entre las simulaciones numero 1 y numero 40.

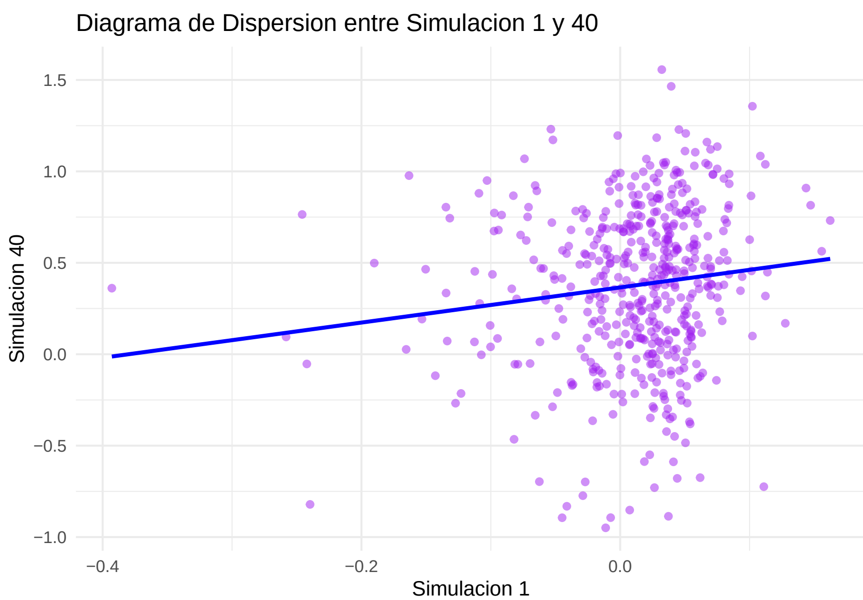

Figura 5: diagrama de dispercion simulacion 1 y 40

4.1.2 Analisis diagrama de dispersion simulacion 1 y 40

Relacion Lineal:

La presencia de una recta de regresion sugiere que hay una relacion lineal entre las simulaciones 1 y 40. Dado que la pendiente de la recta es positiva (*aunque bastante cercana* *a cero*), esto indica una relacion positiva debil, a medida que los valores en la Simulacion 1 aumentan, los valores en la Simulacion 40 tambien tienden a aumentar ligeramente.

Dispersion de los Datos:

Observamos que los puntos estan bastante dispersos alrededor de la linea de regresion, lo que indica que la relacion lineal entre las dos simulaciones es debil. La alta dispersion sugiere que hay otros factores o variaciones aleatorias que estan influyendo en las diferencias entre estas dos simulaciones.

La dispersion sugiere que, aunque existe una tendencia general, hay una considerable variabilidad que no es capturada por la simple relacion lineal. Esto podria ser debido a las diferencias en los parametros **NIG** utilizados en cada simulacion o simplemente a la naturaleza aleatoria de los procesos.

Colas Pesadas y Outliers:

Se observan algunos puntos mas alejados de la nube principal de datos, lo que podria ser indicativo de colas pesadas, una caracteristica propia de la distribucion **NIG**. Estos puntos representan eventos extremos, que son menos comunes pero mas probables en distribuciones con colas pesadas.

Interpretacion Financiera:

En un contexto financiero, si las simulaciones representan diferentes escenarios de retorno de activos, la baja correlacion entre las simulaciones 1 y 40 podria implicar que, bajo diferentes escenarios de mercado, los rendimientos no estan fuertemente correlacionados, lo cual puede ser relevante para la diversificacion de portafolios o la evaluacion de riesgos.

La existencia de outliers o valores extremos tambien es importante en la evaluacion de riesgos, ya que estos puntos podrian representar situaciones de mercado extremas que podrian impactar significativamente en la toma de decisiones.

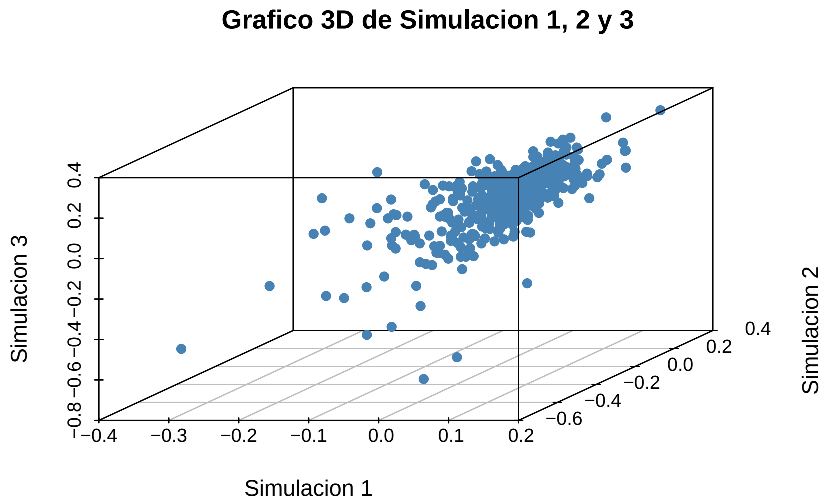

Figura 6: Grafica en 3d primeras 3 simulaciones

4.1.3 Analisis grafico 3D primeras 3 simulaciones

Correlaciones entre Simulaciones:

Tendencias Generales: Se puede observar una tendencia positiva general entre las simulaciones, donde los puntos tienden a alinearse diagonalmente en el espacio 3D. Esto sugiere que hay una correlacion positiva entre las simulaciones 1, 2 y 3.

Dispersion: Sin embargo, los puntos no estan estrictamente alineados, lo que indica que aunque hay una correlacion positiva, no es perfecta. Esto es similar a lo observado en el diagrama de dispersion anterior, pero ahora en tres dimensiones.

Relacion entre las Simulaciones:

Simulacion 1 vs Simulacion 2 vs Simulacion 3: La forma de la nube de puntos sugiere que todas las simulaciones estan relacionadas de manera positiva entre si, pero hay variabilidad. Esto podria ser un reflejo de la estructura de dependencia impuesta por la matriz de correlacion **Rho** que se esta utilizando en la simulacion.

Visualizacion de la Dependencia: El grafico 3D es util para visualizar la relacion multivariada entre las tres primeras simulaciones. La correlacion positiva sugiere que, aunque las simulaciones son independientes en cada ejecucion, hay una estructura subyacente de dependencia dictada por la matriz de correlacion **Rho**.

Impacto de la Distribucion NIG: Los puntos mas alejados del centro reflejan la naturaleza de las colas pesadas de la distribucion NIG. Estos eventos extremos deben ser considerados especialmente en aplicaciones financieras donde el riesgo de perdidas significativas es una preocupacion.

Utilidad en Modelado: se puede obtener escenarios de riesgo, ya que este tipo de visualizaciones ayuda a identificar no solo la tendencia central, sino tambien la variabilidad y los posibles eventos extremos que un modelo podria prever.

Por ultimo consideraremos la suma S de las exponenciales.

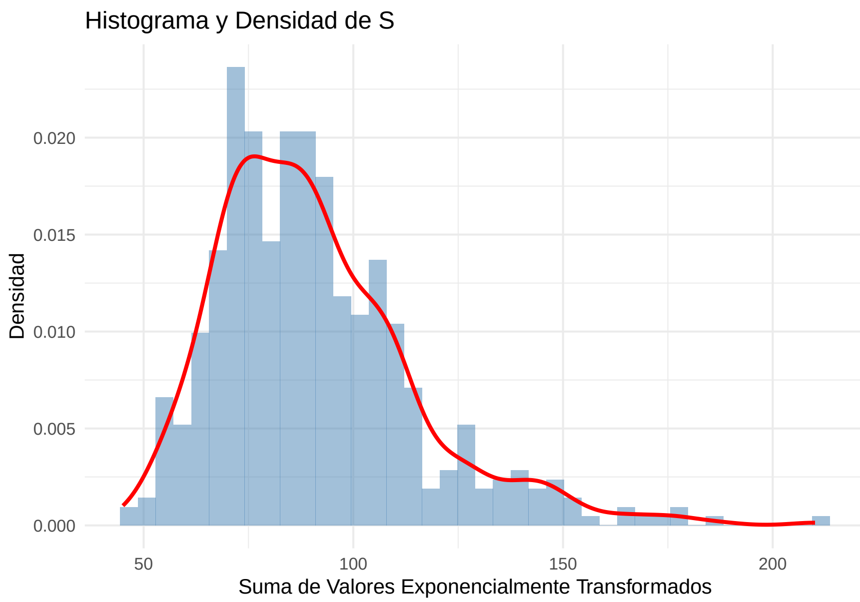

Figura 7: suma de Valores Exponencialmente Transformados

En este grafico podemos notar que la grafica tienda a tener una cola muy pesada a la derecha lo cual es algo comun en las distribucion **NIG*.

5 Punto d

Encuentre una distribucion que ajuste los datos de la variable S, en el caso NIG con la libreria gamlss (posiblemente sea una distribucion Gamma Inversa). Con estas aproximaciones calcule el cuantil de $qS(0.9)$ como el costo con recargo. Muestre el ajuste superponiendo al histograma de S la densidad estimada.Encuentre el porcentaje de recargo con respecto a la media $qS(0.9) = (1 + \theta)E(X)$.

5.1 Solucion

Lo primero que realizaremos sera una verificacion donde veremos cual es le mejor modelo mediante la tecnica **AIC** esta es una medida que combina la bondad de ajuste del modelo con su complejidad. El objetivo es encontrar un modelo que no solo ajuste bien los datos, sino que también sea lo más sencillo posible, para evitar el sobreajuste, por lo tanto al aplicar el modelo obtenemos que los 5 mejores modelos son:

Tabla 3: Mejores 5 Ajustes

Modelo	AIC
IGAMMA	4498.480
BCCG	4499.452
BCCGo	4499.452
GG	4499.464
exGAUS	4500.058

Como podemos observar el mejor modelo mediante **AIC** es una modelo **gamma inversa(IGAMMA)**, por lo tanto proponemos calcular el cuantil 90, de una gamma inversa con los parametros μ, σ, estimados mediante maxima verosimilitud.

Tabla 4: Parámetros Estimados Mediante S para la Gamma Inversa

Parámetro	Valor
Mu	80.4881822
Sigma	0.2440417

Luego calculamos la media para S que es la suma de las exponenciales de nuestras simulaciones, la cual tiene un valor de 90.73752, donde podemos notar que la probabilidad de que una variable distribuida IGAMMA con los parametros μ, σ sea mayor que la media es de 0.4326525, es decir un 43.26%

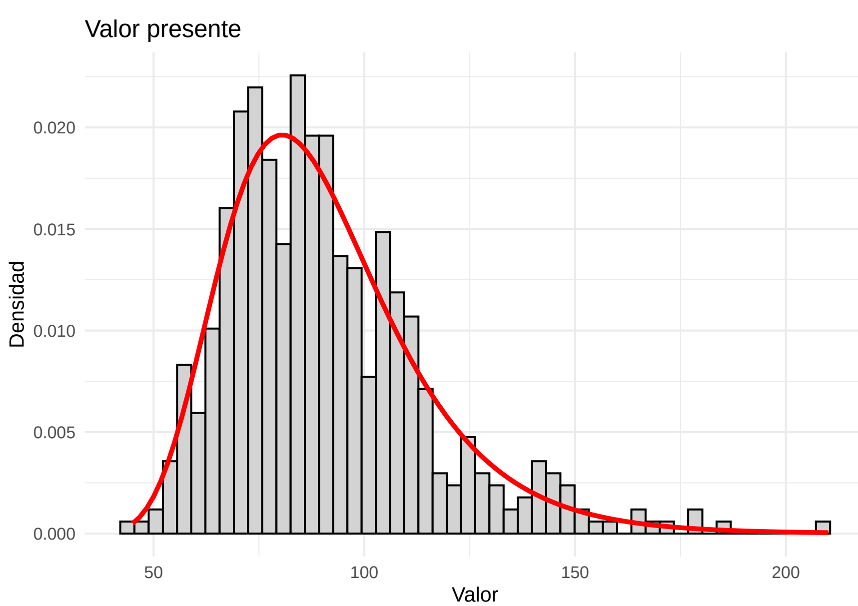

Figura 8: Valores presentes mediante cupulas

Ahora calculamos el cuantil $q_s(90)$ el cual seria el valor con recargo el cual tiene un valor de 121.3534, como ultimo calculamos el porcentaje de recargo usando la formula $E(X)(1+\theta) = q_s(90)$, de donde obtenemos que $\theta = 0.3374114$ lo cual indica que tenemos un porcentaje de recargo del 33.74114%. esto nos sugiere que el precio indicado seria de 121.3534 y se tiene un 33 de recargo adicional en base a la media esperada.

En conclusion a un que ambos metodos dan valores sugeridos diferentes, comparar ambos puede no ser lo adecuado, si no que se debe tener presente en que tipo de escenario esta ocurriendo, ya que uno se hizo con un metodo mas clasico mientras que el segundo es un metodo realizado mediante simulaciones que podrian indicar diversos escenarios.

6 Apendice de Codigos

6.1 Codigo punto a

```r
np = nrow(D)-1


#----------rendimientos diarios m = 360
imk = diff(log(D$Close),1,1)
dmk = log(1+imk)
fechas = as.Date(D$Date[-1],format="%Y-%m-%d")
ejex.mes = seq(fechas[1],fechas[np], "months")
ejex.ano = seq(fechas[1],fechas[np],"years")

#-----------agregar desde dia a mes
library(highfrequency)
library(xts)
imk = xts(x=imk, order.by = fechas)

ts.treemonth <- apply.quarterly(imk,FUN=sum)
dmk = log(1+ts.treemonth)
```

```r
# Crear el data frame con tus datos
datos <- data.frame(
  mu = 0.03881244,
  delta = 0.04042249,
  alpha = 15.40577649,
  beta = -9.17142223
)

# Transponer los datos para que cada fila sea una columna
```

```r
datos_transpuestos <- t(datos)
colnames(datos_transpuestos) <- "Valor"

# Convertir a data frame
datos_transpuestos_df <- as.data.frame(datos_transpuestos)
datos_transpuestos_df$Variable <- rownames(datos_transpuestos_df)

# Reordenar las columnas
datos_transpuestos_df <- datos_transpuestos_df[c("Variable", "Valor")]

# Mostrar la tabla con kable y agregar el titulo
kable(datos_transpuestos_df,
      caption = "Datos Estimados NIG Trimestre",
      align = "l",
      format.args = list(big.mark = ","))  # Formatea los numeros si es necesario
```

```r
# Crear el data frame con el valor de la varianza
sigma2 <- 0.005058298
varianza <- data.frame(
  Variable = "var(delta_m(k))",
  Valor = sigma2
)

# Mostrar la tabla con kable
kable(varianza,
      caption = "Varianza de $delta_m(k)$",  # Usar sintaxis LaTeX para delta
      align = "l",
      format.args = list(big.mark = ","))
```

```r
# Supongamos que fechas e imk ya estan definidos
data <- data.frame(fechas = fechas, cummean_imk = GMCM:::cummean(imk))

ggplot(data, aes(x = fechas, y = cummean_imk)) +
  geom_line() +  # Linea del grafico
  scale_x_date(
    date_breaks = "1 year",          # Intervalos de 1 año para las etiquetas
    date_labels = "%Y",              # Mostrar solo el año
    breaks = seq(from = min(data$fechas), to = max(data$fechas), by = "1 year")  # Mostr
  ) +
  labs(
    y = "Log-Densidades Observadas",  # Etiquetas de los ejes
    x = "Tiempo",
    title = "Tendencia de Log-Densidades Observadas a lo Largo del Tiempo"  # Titulo del
  ) +
```

```r
  theme_minimal() +  # Tema limpio para el grafico
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)  # Rotar etiquetas del eje x para
  ) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "red")
```

```r
library(ggplot2)



# Crear el grafico
ggplot() +
  geom_line(data = df_densitiobv, aes(x = x, y = y, color = "Densidad Estimada"), size =
  geom_line(data = df_graphic, aes(x = x, y = y, color = "Log-Densidad de dmk"), size =
  labs(title = "Comparacion Distribucion estimada vs Log-Densidades",
       x = "Valor",
       y = "Densidad") +
  scale_color_manual(name = "Tipo de Densidad",
                     values = c("Log-Densidad de dmk" = "purple", "Densidad Estimada" =
                     labels = c("Log-Densidad de dmk", "Densidad Estimada")) +
  theme_minimal() +
  theme(legend.position = "right")  # Puedes ajustar la posicion de la leyenda si lo de
```

6.2 Codigo punto b

```r
library(ggplot2)

# Valores de cuotas cj
cj <- c(2.500000, 2.500000, 2.500000, 2.525000, 2.525000, 2.525000, 2.525000, 2.550250,
        2.550250, 2.550250, 2.550250, 2.575753, 2.575753, 2.575753, 2.575753, 2.601510,
        2.601510, 2.601510, 2.601510, 2.627525, 2.627525, 2.627525, 2.627525, 2.653800,
        2.653800, 2.653800, 2.653800, 2.680338, 2.680338, 2.680338, 2.680338, 2.707142,
        2.707142, 2.707142, 2.707142, 2.734213, 2.734213, 2.734213, 2.734213, 2.761555)

# Crear un vector de fechas para 10 años con 4 trimestres por año
fechas <- seq.Date(from = as.Date("2014-01-01"), by = "3 months", length.out = length(cj

# Crear el data frame
df_cuotas <- data.frame(
  Fecha = fechas,
  Cuota = cj
)
```

```
# Crear el grafico
ggplot(df_cuotas, aes(x = Fecha, y = Cuota)) +
  geom_line(color = "blue", size = 1) +
  geom_point(color = "red") +
  labs(
    title = "Evolucion de Cuotas Trimestrales",
    x = "Fecha",
    y = "Valor de Cuota"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1)  # Rotar etiquetas del eje x para
  )
```

```
Lavqmn = function(i,m,q,n,rho){
try(if(m%%q != 0) stop("m no es divisible por q"))
v = 1/(1+i)
t = seq(1,n*m,1)
res = (1/m)*sum((1+i)^(-t/m)*(1+(rho/q)*floor((t-1)*q/m)))
return(res)}


Cp = C*m*Lavqmn(ia,m,q,n,rho)
```

6.3 Codigo punto C

```
n<-10
m<-4
ndim = n*m
R = matrix(0,ndim,ndim)
for(i in 1:ndim){
for(j in 1:ndim){
R[i,j] = ifelse(i <= j, i,j)}}
R = sigma2*R
Rho = cov2cor(R)

library(mvtnorm)
# opciones para la matriz Sigma = Rho,
# rmvnorm(n, mean = mu, sigma = Sigma, method = "eigen"),
# rmvnorm(n, mean = mu, sigma = Sigma, method = "svd"),
# rmvnorm(n, mean = mu, sigma = Sigma, method = "chol")
set.seed(1256)
```

```r
sim.GC <- function(n, Rho, qnig){
dat <- rmvnorm(n=n, mean = rep(0,nrow(Rho)), sigma = Rho)
for(j in 1:nrow(Rho)){
dat[,j] <- qnig(pnorm(dat[,j]),param =  c(j,j,1,1)*nig.est)
}
return(dat)
}

Lambda.j = sim.GC(500,Rho,qnig)
```

```r
library(ggplot2)
library(gridExtra)
# Convertir el conjunto de datos a un formato adecuado para ggplot
Lambda.j.df <- data.frame(
  Sim1 = Lambda.j[,1],
  Sim40 = Lambda.j[,40]
)

# Crear los histogramas con curvas de densidad usando ggplot
hist_sim1 <- ggplot(Lambda.j.df, aes(x = Sim1)) +
  geom_histogram(aes(y = ..density..), bins = 40, fill = "blue", alpha = 0.5) +
  geom_density(color = "black", size = 1) +
  labs(title = "Histograma de Simulacion 1", x = "Simulacion 1", y = "Densidad") +
  theme_minimal()

hist_sim40 <- ggplot(Lambda.j.df, aes(x = Sim40)) +
  geom_histogram(aes(y = ..density..), bins = 40, fill = "red", alpha = 0.5) +
  geom_density(color = "black", size = 1) +
  labs(title = "Histograma de Simulacion 40", x = "Simulacion 40", y = "Densidad") +
  theme_minimal()



# Mostrar los histogramas uno al lado del otro y el diagrama de dispersion por separad
grid.arrange(hist_sim1, hist_sim40, ncol = 2)  # Combina
```

```r
# Convertir el conjunto de datos a un formato adecuado para ggplot
# Crear el diagrama de dispersion con colores diferenciados para los puntos
scatter_plot <- ggplot(Lambda.j.df, aes(x = Sim1, y = Sim40)) +
  geom_point(color = "purple", alpha = 0.5) +
  geom_smooth(method = "lm", color = "blue", se = FALSE) +  # Linea de ajuste
  labs(title = "Diagrama de Dispersion entre Simulacion 1 y 40", x = "Simulacion 1", y =
  theme_minimal()
```

```r
print(scatter_plot)
```

```r
library(scatterplot3d) # load
scatterplot3d(Lambda.j[,1:3],
              pch = 16,                    # Tipo de punto
              color = "steelblue",         # Color de los puntos
              xlab = "Simulacion 1",       # Etiqueta del eje x
              ylab = "Simulacion 2",       # Etiqueta del eje y
              zlab = "Simulacion 3",       # Etiqueta del eje z
              main = "Grafico 3D de Simulacion 1, 2 y 3")  # Titulo del grafico
```

```r
library(ggplot2)


Y = exp(-Lambda.j)%*%cj
S = apply(Y,1,sum)
S_df <- data.frame(S = S)

# Crear el histograma y la curva de densidad
ggplot(S_df, aes(x = S)) +
  geom_histogram(aes(y = ..density..), bins = 40, fill = "steelblue", alpha = 0.5) +
  geom_density(color = "red", size = 1) +
  labs(title = "Histograma y Densidad de S", x = "Suma de Valores Exponencialmente Trans
  theme_minimal()
```

6.4 Codigo punto D

```r
library(gamlss)
Sfit1 <- fitDist(y = S, type = "realplus")

------las  mejores 5 con respecto al AIC
```

```r
# Crear los datos de la tabla
data <- data.frame(
  Modelo = c("IGAMMA", "BCCG", "BCCGo", "GG", "exGAUS"),
  AIC = c(4498.480, 4499.452, 4499.452, 4499.464, 4500.058)
)

kable(data, caption = "Mejores 5 Ajustes", align = c("l", "r"), format = "latex") %>%
  kable_styling(latex_options = c("striped", "hold_position"))
```

```r
 Estimación del modelo IGAMMA (deberías tener esta parte ya hecha)
H <- gamlssML(S, family = IGAMMA)
coef.IGAMMA <- c(mu = H$mu, sigma = H$sigma)

# Crear un data frame para ggplot
S_df <- data.frame(S = S)
```

```r
ggplot(S_df, aes(x = S)) +
  geom_histogram(aes(y = ..density..), bins = 50, fill = 'lightgray', color = 'black') +
  stat_function(
    fun = dIGAMMA,
    args = list(mu = coef.IGAMMA[1], sigma = coef.IGAMMA[2]),
    color = 'red',
    size = 1.2
  ) +
  labs(
    title = 'Valor presente',
    x = 'Valor',
    y = 'Densidad'
  ) +
  theme_minimal() +
  theme(legend.position = 'top') +
  scale_fill_manual(values = c('darkgray', 'red'), name = 'Leyenda', labels = c('datos',
  guides(fill = guide_legend(override.aes = list(linetype = c(0, 1), size = c(0.5, 1.2))
```