Proyecto eGARCH

Isaac Lopez

2025-04-13

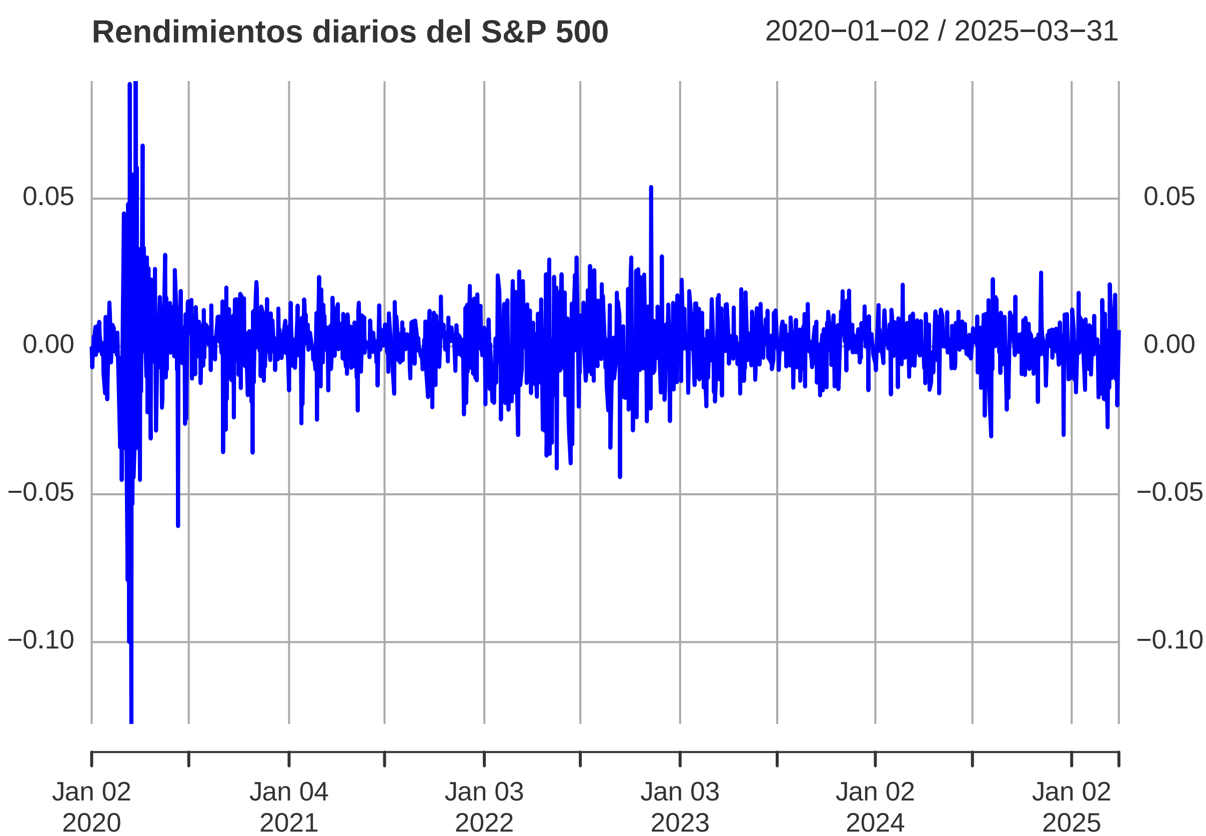
Introducción

En este trabajo se analiza el comportamiento de los rendimientos diarios del índice S&P 500 entre 2020 y 2025 mediante un modelo EGARCH(1,1). El objetivo principal es estimar la volatilidad condicional y evaluar su utilidad en la predicción del Value at Risk (VaR), así como verificar si el modelo es útil para gestión de riesgo financiero a través de un proceso de backtesting.

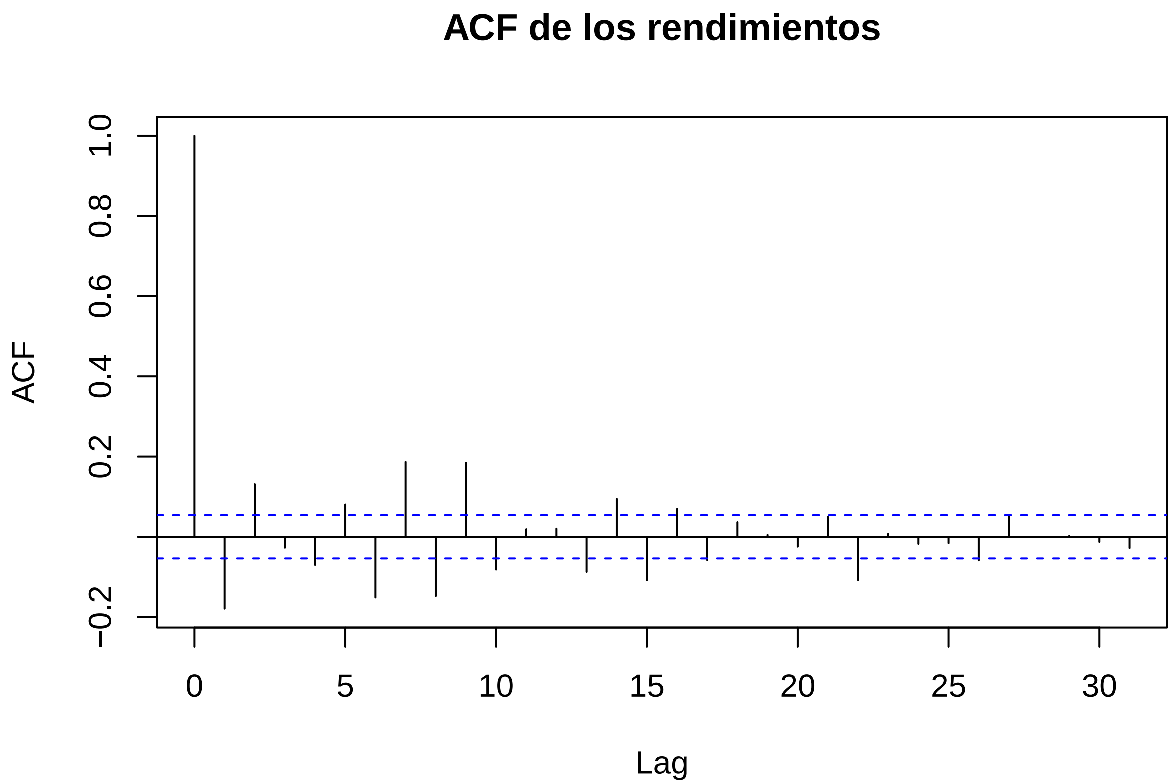
Visualización y estructura temporal de los rendimientos

Se graficaron los rendimientos logarítmicos diarios para observar su comportamiento. También se usaron funciones ACF y PACF para explorar posibles patrones de autocorrelación. Aunque los rendimientos en sí no presentan autocorrelación fuerte, sí se observa agrupamiento de volatilidad, lo cual sugiere la presencia de heterocedasticidad condicional.

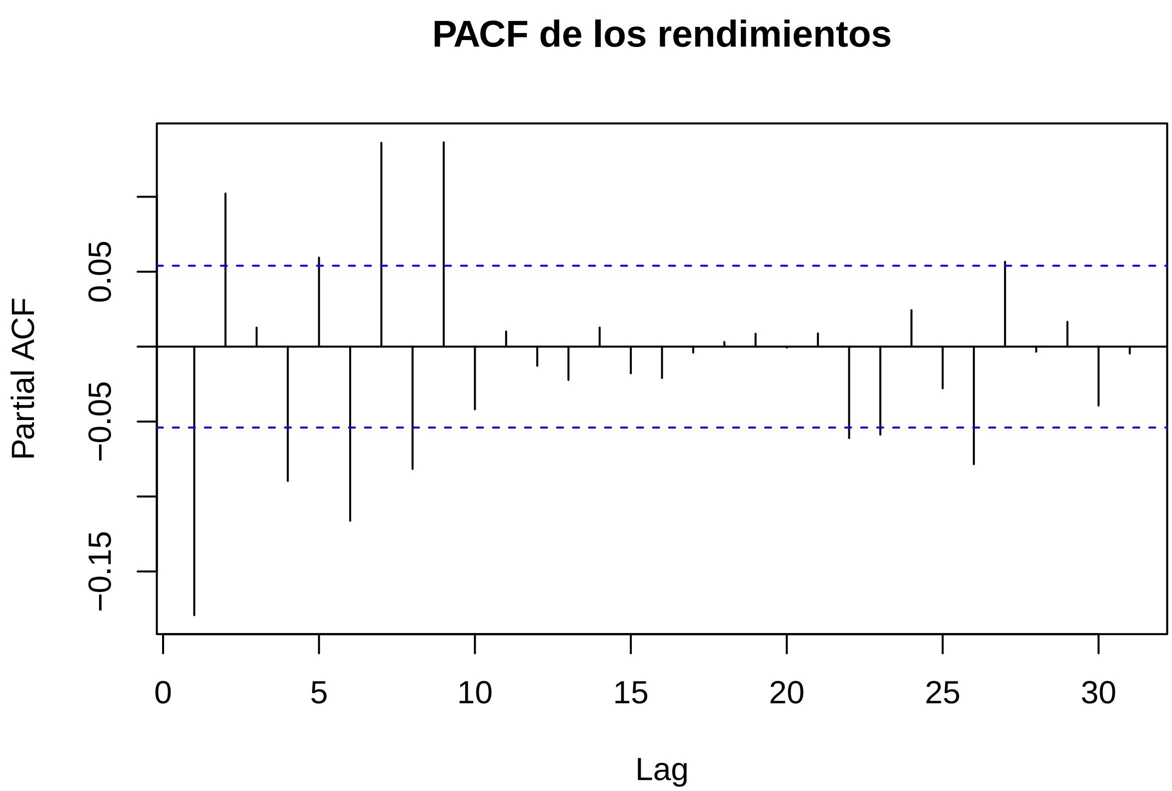
```
plot(rendimientos, main = "Rendimientos diarios del S&P 500", col = "blue")
```



```r
acf(rendimientos, main = "ACF de los rendimientos")
```




```r
pacf(rendimientos, main = "PACF de los rendimientos")
```



Al observar los gráficos de la función de autocorrelación (ACF) y autocorrelación parcial (PACF) de los rendimientos diarios del S&P 500, no se detectan autocorrelaciones significativas más allá del primer rezago. Esto es consistente con la hipótesis de que los retornos financieros siguen un proceso cercano a ruido blanco.

Sin embargo, este comportamiento no implica que no exista estructura temporal en la varianza. De hecho, en los datos financieros es común que los rendimientos no estén correlacionados, pero sí presenten agrupamiento de volatilidad, lo cual justifica analizar modelos GARCH o EGARCH.

Prueba de heterocedasticidad (ARCH Test)

Se aplicó el test de Engle para detectar efectos ARCH en los datos. El resultado fue altamente significativo (p-valor < 2.2e-16), lo cual indica que existe heterocedasticidad. Por lo tanto, es válido usar modelos GARCH o EGARCH para modelar la varianza

```r
ArchTest(rendimientos, lags = 12)
```

```
##
##  ARCH LM-test; Null hypothesis: no ARCH effects
##
## data:  rendimientos
## Chi-squared = 532.89, df = 12, p-value < 2.2e-16
```

Estimación del modelo EGARCH(1,1)

Se ajustó un modelo **EGARCH(1,1)** con distribución **t-student** para capturar tanto la heterocedasticidad condicional como las colas pesadas observadas en los rendimientos diarios del índice S&P 500. El modelo fue estimado sobre los datos de entrenamiento (in-sample) dejando 500 observaciones fuera de muestra para el backtesting del VaR.

Todos los coeficientes resultaron **estadísticamente significativos**, con p-valores cercanos a cero, lo cual respalda la validez del modelo.

Criterios de información:

Para evaluar el ajuste del modelo, se consultaron los siguientes indicadores:

- **AIC**: -6.355

- **BIC**: -6.331

Valores negativos y bajos de AIC y BIC indican un modelo ajustado en relación con su complejidad. Estos resultados respaldan el uso del modelo EGARCH frente a otras opciones más simples como GARCH(1,1), que usualmente tienen peores valores en estos criterios al no capturar asimetrías ni colas pesadas.

```r
modelo_especificacion <- ugarchspec(
  variance.model = list(model = "eGARCH", garchOrder = c(1,1)),
  mean.model     = list(armaOrder = c(0,0), include.mean = TRUE),
  distribution.model = "std"
)

modelo_ajustado <- ugarchfit(spec = modelo_especificacion, data = rendimientos)

show(modelo_ajustado)
```

```
##
## *---------------------------------*
## *          GARCH Model Fit        *
## *---------------------------------*
```

```
##
## Conditional Variance Dynamics
## -----------------------------------
## GARCH Model  : eGARCH(1,1)
## Mean Model   : ARFIMA(0,0,0)
## Distribution : std
##
## Optimal Parameters
## ------------------------------------
##         Estimate  Std. Error    t value Pr(>|t|)
## mu      0.000754    0.000231     3.2664 0.001089
## omega  -0.250294    0.005633   -44.4306 0.000000
## alpha1 -0.168142    0.015712   -10.7012 0.000000
## beta1   0.973504    0.000028 34299.0186 0.000000
## gamma1  0.136604    0.012726    10.7345 0.000000
## shape   6.653102    1.083093     6.1427 0.000000
##
## Robust Standard Errors:
##         Estimate  Std. Error    t value Pr(>|t|)
## mu      0.000754    0.000223     3.3779 0.000731
## omega  -0.250294    0.006207   -40.3226 0.000000
## alpha1 -0.168142    0.021608    -7.7813 0.000000
## beta1   0.973504    0.000034 28899.9625 0.000000
## gamma1  0.136604    0.017462     7.8230 0.000000
## shape   6.653102    1.277022     5.2099 0.000000
##
## LogLikelihood : 4194.022
##
## Information Criteria
## ------------------------------------
##
## Akaike       -6.3551
## Bayes        -6.3315
## Shibata      -6.3552
## Hannan-Quinn -6.3463
##
## Weighted Ljung-Box Test on Standardized Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      1.672  0.1960
## Lag[2*(p+q)+(p+q)-1][2]     1.682  0.3215
## Lag[4*(p+q)+(p+q)-1][5]     3.120  0.3855
## d.o.f=0
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## ------------------------------------
##                         statistic p-value
## Lag[1]                      2.012  0.1560
## Lag[2*(p+q)+(p+q)-1][5]     3.391  0.3403
## Lag[4*(p+q)+(p+q)-1][9]     4.035  0.5830
## d.o.f=2
##
## Weighted ARCH LM Tests
```

```
## ------------------------------------
##             Statistic Shape Scale P-Value
## ARCH Lag[3]  0.003192 0.500 2.000  0.9549
## ARCH Lag[5]  1.029521 1.440 1.667  0.7242
## ARCH Lag[7]  1.313956 2.315 1.543  0.8578
##
## Nyblom stability test
## ------------------------------------
## Joint Statistic:  1.6375
## Individual Statistics:
## mu     0.1710
## omega  0.2141
## alpha1 0.1087
## beta1  0.1819
## gamma1 0.7730
## shape  0.1745
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:          1.49 1.68 2.12
## Individual Statistic:     0.35 0.47 0.75
##
## Sign Bias Test
## ------------------------------------
##                    t-value   prob sig
## Sign Bias           1.2339 0.2175
## Negative Sign Bias  0.3912 0.6957
## Positive Sign Bias  1.1243 0.2611
## Joint Effect        1.8817 0.5973
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## ------------------------------------
##   group statistic p-value(g-1)
## 1    20     58.54    6.570e-06
## 2    30     63.32    2.345e-04
## 3    40     71.74    1.083e-03
## 4    50     97.55    4.579e-05
##
##
## Elapsed time : 0.1254821
```

Evaluación del ajuste: residuos estandarizados

Se aplicó la prueba Ljung-Box a los residuos estandarizados para verificar que no quedara autocorrelación. El p-valor fue 0.1227, por lo que no se rechaza la hipótesis de independencia. Esto sugiere que el modelo EGARCH ajusta bien la estructura temporal de la serie.

```r
residuos_estandar <- residuals(modelo_ajustado, standardize = TRUE)

# Prueba Ljung-Box
Box.test(residuos_estandar, lag = 20, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
```

```
## data:  residuos_estandar
## X-squared = 27.465, df = 20, p-value = 0.1227
```

Cálculo del VaR condicional diario (1% y 5%)

A partir del modelo estimado, se generó un pronóstico rolling de la desviación estándar condicional para 500 observaciones fuera de muestra. Con esta desviación y los grados de libertad estimados se calculó el VaR al 1% y 5% usando la distribución t. Estos valores representan pérdidas máximas esperadas bajo cierto nivel de confianza.

```r
#backtesting
fuera_muestra <- 500

#quitando para backtesting
modelo_ajustado <- ugarchfit(
  spec = modelo_especificacion,
  data = rendimientos,
  out.sample = fuera_muestra
)

#rolling
pronostico <- ugarchforecast(
  modelo_ajustado,
  n.ahead = 1,
  n.roll = fuera_muestra - 1
)

desv_condicional <- sigma(pronostico)

grados_libertad <- coef(modelo_ajustado)["shape"]

#significancia
nivel_1 <- 0.01
nivel_5 <- 0.05

VaR_1 <- -qt(nivel_1, df = grados_libertad) * desv_condicional * sqrt((grados_libertad - 2) / grados_li
VaR_5 <- -qt(nivel_5, df = grados_libertad) * desv_condicional * sqrt((grados_libertad - 2) / grados_li

rend_fuera_muestra <- tail(rendimientos, fuera_muestra)

tabla_var <- data.frame(
  fecha = index(rend_fuera_muestra),
  rendimiento = as.numeric(rend_fuera_muestra),
  VaR_1 = as.numeric(VaR_1),
  VaR_5 = as.numeric(VaR_5)
)
```

Backtesting del VaR

Se compararon los VaRs estimados con los rendimientos observados para identificar violaciones (cuando las pérdidas superan el VaR). Se calcularon las tasas de violación observadas y se aplicó el test de Kupiec para validar si la frecuencia de violaciones es coherente con los niveles esperados. El modelo pasó el test, lo que respalda su validez estadística.

```r
tabla_var$violacion_1 <- tabla_var$rendimiento < -tabla_var$VaR_1
tabla_var$violacion_5 <- tabla_var$rendimiento < -tabla_var$VaR_5

tasa_violacion_1 <- mean(tabla_var$violacion_1, na.rm = TRUE)
tasa_violacion_5 <- mean(tabla_var$violacion_5, na.rm = TRUE)

cat("Tasa de violaciones al VaR 1%: ", round(tasa_violacion_1, 4), "\n")
```

```
## Tasa de violaciones al VaR 1%:  0.01
```

```r
cat("Tasa de violaciones al VaR 5%: ", round(tasa_violacion_5, 4), "\n")
```

```
## Tasa de violaciones al VaR 5%:  0.04
```

Prueba de Kupiec para el VaR 1%

La prueba de Kupiec evalúa si el número de violaciones observadas al VaR coincide con el número esperado estadísticamente. En este caso, el modelo esperaba 5 violaciones y se observaron exactamente 5 violaciones reales, lo que ya visualmente sugiere un buen ajuste.

En la parte Uncondicional (UC) del test, el estadístico fue 0 y el p-valor fue 1, por lo que claramente se acepta la hipótesis nula de que las violaciones son correctas en proporción.

Además, en la parte Condicional (CC) —que evalúa no solo cuántas violaciones hay, sino si están bien distribuidas en el tiempo (sin dependencia)— el p-valor fue 0.95, también muy alto. Se concluye que las violaciones no solo son proporcionales, sino también independientes, lo que refuerza la validez del modelo.

En resumen, el modelo EGARCH(1,1) genera estimaciones del VaR que superan las pruebas de backtesting, tanto en cantidad como en distribución lo que valida su uso para gestión de riesgo financiero.

```r
#kuiper
VaRTest(alpha = 0.01, actual = tabla_var$rendimiento, VaR = -tabla_var$VaR_1)
```

```
## $expected.exceed
## [1] 5
##
## $actual.exceed
## [1] 5
##
## $uc.H0
## [1] "Correct Exceedances"
##
## $uc.LRstat
## [1] 0
##
## $uc.critical
## [1] 3.841459
##
## $uc.LRp
## [1] 1
##
## $uc.Decision
## [1] "Fail to Reject H0"
##
## $cc.H0
## [1] "Correct Exceedances & Independent"
##
```

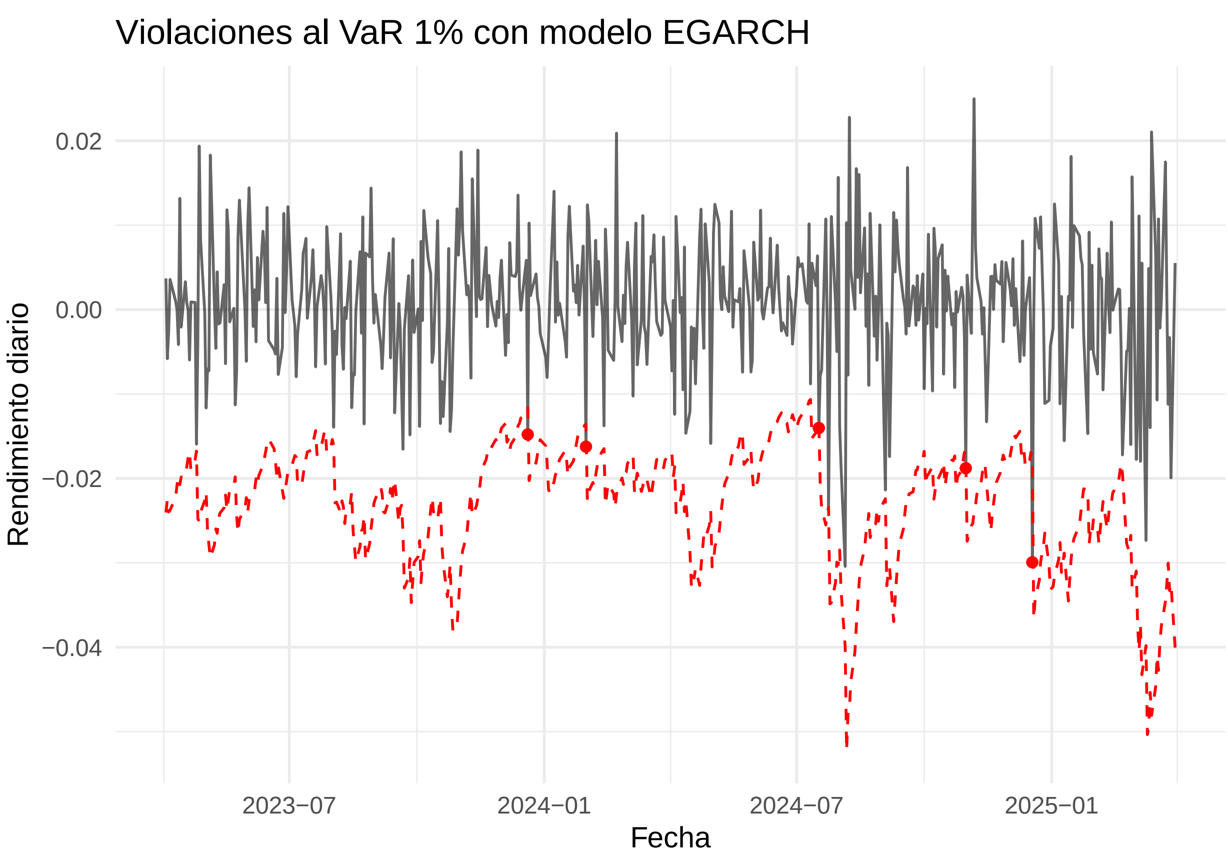
```
## $cc.LRstat
## [1] 0.1012163
##
## $cc.critical
## [1] 5.991465
##
## $cc.LRp
## [1] 0.9506511
##
## $cc.Decision
## [1] "Fail to Reject H0"
```

Visualización de las violaciones al VaR

Se graficó la serie de rendimientos junto con el VaR al 1%. Los días en que el rendimiento fue menor al VaR se marcaron como violaciones. Este gráfico ayuda a visualizar cuándo el modelo subestimó el riesgo extremo.

```r
library(ggplot2)

ggplot(tabla_var, aes(x = fecha)) +
  geom_line(aes(y = rendimiento), color = "black", alpha = 0.6) +
  geom_line(aes(y = -VaR_1), color = "red", linetype = "dashed") +
  geom_point(data = filter(tabla_var, violacion_1 == TRUE),
             aes(y = rendimiento), color = "red", size = 1.5) +
  labs(title = "Violaciones al VaR 1% con modelo EGARCH",
       y = "Rendimiento diario", x = "Fecha") +
  theme_minimal()
```



El modelo EGARCH(1,1) fue capaz de capturar correctamente la volatilidad condicional del S&P 500 y produjo estimaciones confiables del VaR en un periodo fuera de muestra. Las pruebas estadísticas respaldan la especificación del modelo, y el backtesting muestra tasas de violación razonables. Por lo tanto, se concluye que este enfoque puede ser útil en contextos de pronóstico de riesgo financiero sobre índices financieros