

# Incluyendo saltos en la construcción de portafolios

Autor: Héctor Enrique Ramírez

Asesor: Diego Jara Pinzón

Junio 14, 2021

## Abstract

El insumo principal para la construcción de portafolios tipo Markowitz es la distribución de los retornos de los activos financieros que usualmente se asume normal. La literatura ha demostrado que la incidencia de eventos extremos es más común de lo que predice la distribución gaussiana. En consecuencia, este artículo tiene como objetivo analizar la inclusión de saltos de Poisson que ha demostrado ser de utilidad en el terreno de la valoración de derivados financieros. En primer lugar, se encuentra la ecuación estocástica para el valor del portafolio partiendo de las distribuciones de los retornos de sus componentes. A partir de lo anterior, usando datos del mercado norteamericano durante el periodo 1996-2021, se comprueba que: (i) estadísticamente la nueva distribución resulta más adecuada para modelar los retornos y (ii) se obtienen beneficios en la optimización de portafolios bajo diversos objetivos.

**Palabras clave:** Poisson, optimización de portafolios, leptokurtosis.

**Clasificación JEL:** G10, G11, G12.

# 1 Introducción

Los mercados financieros alrededor del mundo han visto una gran expansión desde inicios del siglo pasado. Particularmente, los mercados accionarios han demostrado ser un vehículo de inversión para objetivos que van desde la generación de capital hasta una alternativa de retiro. Buscando una menor exposición al riesgo idiosincrático de ciertos activos, los administradores de portafolios han recurrido a la diversificación. Típicamente el enfoque escogido para lograr este cometido sigue la idea planteada por Markowitz (1952), asumiendo retornos con distribución normal resuelve un problema de optimización de media - varianza para el portafolio. Sin embargo, un hecho estilizado ampliamente documentado es que estos retornos se alejan de una distribución gaussiana, lo que lleva a problemas de implementación como los altos errores de estimación que llevan a grandes disparidades entre las estimaciones ex-ante y los resultados ex-post.

Un hecho documentado es el de *leptokurtosis* donde los eventos extremos ocurren con mayor frecuencia llevando a colas más pesadas que las documentadas por la distribución gaussiana. Partiendo del hecho descrito anteriormente, surge la pregunta: ¿existe alguna forma de introducir el fenómeno de leptokurtosis en la construcción de portafolios? La literatura ha respondido con enfoques como la introducción de cópulas, el uso de distribuciones elípticas más generales que la normal, mejora en los estimadores, optimización robusta, entre otras.

Siguiendo la misma línea de pensamiento, este trabajo busca aportar introduciendo una distribución usada en la valoración de derivados, que posee un componente continuo gaussiano y otro discreto, fenómeno que resulta bastante intuitivo: cuando el precio no sigue una senda "normal" se observan saltos discretos en el precio, ya sea por efectos como la intervención de un agente con poder de mercado o la aparición de una noticia que afecta el precio en el corto plazo. Al diseñar portafolios bajo este enfoque se espera obtener ventajas respecto al modelo gaussiano como: una mejor identificación de riesgos y mayor precisión al modelar.

El estudio encuentra que utilizar la distribución con saltos añade mayor exactitud de modelamiento de los retornos y en consecuencia, se obtienen mejorías en el diseño de portafolios. El documento está organizado de la siguiente manera: en primer lugar se presenta la revisión de literatura, seguido del marco teórico donde se desarrollan todos los conceptos necesarios para realizar la implementación con datos del mercado norteamericano. Posteriormente se hace una breve descripción de los datos y metodología utilizada. Finalmente, se presentan los resultados y las conclusiones.

## 2 Revisión de literatura

La literatura sobre teoría del portafolio inicia con Markowitz en 1952 en donde se plantea por primera vez la idea de diversificación entre activos financieros. En este trabajo se explota la correlación entre activos para encontrar una frontera eficiente donde un agente puede establecer los pesos a asignar en el portafolio dada su aversión al riesgo. Su introducción generó un marco general sobre el cual trabajar en años posteriores: algunos autores se centraron en los problemas prácticos que implicaba esta teoría mientras que otros desarrollaron otros enfoques.

Entre los problemas prácticos que se encontraron el más notable quizás fue la correcta estimación de los parámetros tales como la matriz varianza-covarianza o el vector de retornos de los activos, como bien se documenta en Ledoit & Wolf (2003). Los autores de este documento lograron una mejora considerable en la estimación de la matriz var-cov, logrando una aproximación bastante apropiada bajo el método estadístico conocido como shrinkage, en donde se añade sesgo al estimador para reducir su varianza, método que resultó bastante exitoso en la práctica. De manera reciente Lopez de Prado (2020) estableció un método más apropiado para eliminar el ruido de la matriz var-cov conocido como denoising, donde basándose en la distribución Marchenko-Pastur se corrigen los valores propios de la matriz. Este método ha demostrado ser efectivo para retornos que se suponen normales.

Por otro lado, la literatura documentó el fenómeno de leptokurtosis que presentaban los retornos financieros, hecho que puso en duda la utilidad de la distribución gaussiana como pilar de la teoría del portafolio y de la fórmula de Black-Scholes. En la valoración de opciones, Merton (1976) añade los saltos de Poisson (con amplitud de salto uniforme), para corregir este fenómeno. Posteriormente la literatura se enfocaría en estudiar la forma más precisa de modelar la amplitud de estos saltos. Synowiec (2008) logra demostrar, mediante una serie de pruebas estadísticas, que el modelo doblemente exponencial resulta ser el más adecuado para modelar los retornos en mercados de divisas. Por otro lado, Hanson & Westman (2003) trabajan aplicaciones en optimización de portafolios del modelo planteado por Merton bajo enfoques dinámicos, con un activo libre de riesgo y otro riesgoso, para maximizar la utilidad esperada (descontada a valor presente) de la riqueza final de un agente en un periodo de tiempo. Este trabajo no logró identificar investigaciones adicionales en la línea de incluir saltos para optimizar portafolios.

Finalmente, cabe resaltar que el enfoque planteado inicialmente por Markowitz, donde la varianza se usa como medida de riesgo fue descartada y de esta crítica nacieron otros enfoques entre los cuales se puede resaltar la rama conocida como Post-Modern Portfolio Theory (PMPT). Varios autores redefinieron el riesgo y se utilizaron medidas más apropiadas como el VaR o el CVaR como es el caso de Krokmal et al. (2001) para construir una frontera eficiente. El enfoque del PMPT redefine el riesgo como la dispersión por debajo de un nivel deseado de retorno, desarrollada principalmente por Frank Sortino a quien se le atribuye el Sortino Ratio, una medida que corrige al Sharpe Ratio por la definición de riesgo como bien indica Mazhar & Srivastava (2018). La inclusión de distribuciones más apropiadas que la normal resulta imprescindible para el éxito de estos enfoques.

### **3 Marco teórico**

La teoría moderna del portafolio y sus refinamientos se refieren a la teoría de inversión que le permite a un inversionista maximizar el retorno esperado para cierto nivel de riesgo. Es una formalización de la diversificación, donde el uso de varios activos resulta menos riesgoso que enfocarse solamente en uno. En el proceso de selección de un portafolio se pueden identificar dos etapas: la primera consiste en la estimación de parámetros de los activos con los que se va a construir el portafolio (conocidos como creencias o *beliefs*) y la segunda se enfoca en el modo en que se

asignan pesos a estos activos. Esta última se encarga de definir un objetivo de optimización como una función de utilidad o una medida de riesgo, y de encontrar la forma de resolverlo.

Este estudio se enfoca en los beliefs, es decir los inputs con los que se calculan los portafolios. Las soluciones que brinda la optimización de portafolios serán eficientes siempre y cuando los beliefs sean correctos. Mediante la introducción del proceso compuesto de Poisson, que modela saltos discontinuos en el precio, se pretende refinar los beliefs sobre los retornos de los activos y en consecuencia obtener mejores resultados que los alcanzados bajo un enfoque gaussiano.

En esta sección se describen todos los aspectos teóricos que se utilizarán en el estudio. En primer lugar, se desarrollan las ecuaciones que describirán el proceso estocástico que sigue el retorno logarítmico para una acción y luego para un portafolio de renta variable, y posteriormente se muestran las distribuciones del retorno logarítmico tanto para el caso de una sola acción como para el portafolio entero. Finalmente, se describirán las pruebas de ajuste estadístico que se usarán en la sección empírica y los problemas de optimización a abordar, con sus respectivas formulaciones y restricciones.

### 3.1 Ecuaciones estocásticas

Los precios de los activos pueden ser identificados mediante procesos estocásticos: sucesiones de variables aleatorias que se definen en función del tiempo, donde cada una presenta una distribución de probabilidad propia. En esta sección se introducirán los procesos a usar, la ecuación para el caso de un solo activo de renta variable y el desarrollo para inferir la ecuación de un portafolio.

#### 3.1.1 Definiciones preliminares

Los procesos estocásticos a usar a lo largo del trabajo son el movimiento browniano y el proceso compuesto de Poisson. Venegas (2008) define el movimiento browniano, también conocido como proceso de Wiener, como una familia de variables aleatorias denotada por  $\{W_t\}_{t \geq 0}$  que cumple las siguientes condiciones:

- (i)  $W_0 = 0$  casi seguramente en todas partes, es decir, el proceso empieza en  $t = 0$  con probabilidad 1.
- (ii) Para cualquier par de tiempos  $0 \leq s \leq t$  los incrementos  $W_t - W_s$  siguen una distribución normal con media 0 y varianza  $t - s$ .
- (iii) Estos incrementos son independientes estocásticamente entre sí para el conjunto de tiempos  $0 \leq t_1 < t_2 < \dots < t_n$ .

Este modelo se ha usado extensamente para modelar los retornos de los activos financieros, particularmente en el contexto de valoración de derivados.

En segundo lugar, el proceso compuesto de Poisson se compone de un proceso simple de Poisson y una variable aleatoria subyacente. El primero es un proceso de conteo: busca contabilizar el número de veces que se repite un proceso aleatorio como la llegada de un cliente a una tienda.

Usualmente este se denota por  $N_t$  con tasa  $\lambda$ . Al igual que el proceso de Wiener cumple una serie de condiciones:

(i)  $N_0 = 0$ .

(ii) Si  $s < t$ , entonces  $N_s \leq N_t$ .

(iii) Las probabilidades de que se presente un salto se describen por:

$$P \{N_{t+dt} - N_t = 1\} = \lambda dt + o(dt)$$

$$P \{N_{t+dt} - N_t = 0\} = 1 - \lambda dt + o(dt),$$

donde  $o(dt)/dt \rightarrow 0$  si  $dt \rightarrow 0$ .

(iv) Los incrementos son estocásticamente independientes para el conjunto de tiempos  $0 \leq t_1 < t_2 < \dots < t_n$ .

Ahora, el proceso compuesto de Poisson con tasa  $\lambda$  y variable aleatoria subyacente  $D_i$  denotado por  $Y_t$  con  $t \geq 0$  es:

$$Y_t = \sum_{i=1}^{N_t} D_i,$$

donde  $N_t$  es un proceso simple de Poisson y  $D_i$  son variables iid, independientes de  $N_t$ .

### 3.1.2 Precio de una acción

Habiendo definido ambos procesos se puede caracterizar el proceso que sigue el precio de una acción por:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + J dN_t,$$

donde  $S_t$  es el precio del activo en el momento  $t$ ,  $\mu$  es la media o drift del retorno del activo,  $\sigma$  la volatilidad,  $dW_t$  es una representación conveniente del movimiento browniano (senda normal),  $J$  es la variable aleatoria subyacente que representa la amplitud de un salto y  $dN_t$  es un proceso de Poisson, donde los dos últimos componen el proceso compuesto de Poisson (Venegas, 2008).

### 3.1.3 Ecuación para un portafolio con saltos no correlacionados

Ahora se construirá la ecuación estocástica que definirá a un portafolio de  $M$  acciones. El valor del portafolio se denotará por  $V_t$ . La dinámica de sus precios está dada por:

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \sigma_i dW_t^i + J_i dN_t^i, \tag{1}$$

para  $i = 1, 2, \dots, M$ , donde  $\mu_i$  representa el drift de la acción  $i$ ,  $\sigma_i$  representa la volatilidad y  $J_i$  amplitud de un salto aleatorio.  $W_t^i$  es un proceso browniano,  $N_t^i$  es un proceso de Poisson con tasa  $\lambda_i$ . Se supondrá que los procesos brownianos entre acciones poseen correlación, es decir  $dW_t^i dW_t^j = \rho_{ij} dt$  donde claramente si  $i = j$   $\rho_{ii} = 1$ . Ahora, los procesos  $J_i$ ,  $dN_t^i$  no poseerán correlación entre ellos ni con los otros procesos; es decir  $dN_t^i dW_t^j = 0$  para todo  $i, j$  y  $dN_t^i dN_t^j = 0$  para  $i \neq j$ . Finalmente, se supone que la correlación entre la amplitud de los saltos es 0.

De este modo se puede establecer que la inversión del portafolio en un activo  $i$  tendrá un peso  $w_i$ , por lo tanto la inversión en este activo será  $w_i V_t$  y luego el cambio total en el valor del portafolio se puede escribir como:

$$dV_t = \sum_{i=1}^M w_i V_t \frac{dS_t^i}{S_t^i},$$

que, reescrito toma la forma:

$$dV_t = V_t \left[ \left( \sum_{i=1}^M \mu_i w_i \right) dt + \left( \sum_{i=1}^M \sigma_i w_i dW_t^i \right) + \left( \sum_{i=1}^M w_i J_i dN_t^i \right) \right].$$

En primer lugar, nótese que la suma de los movimientos brownianos puede ser resumido como un solo movimiento browniano, pues es una suma de variables normales correlacionadas de modo que tienen esperanza cero y varianza:

$$\text{Var} \left[ \sum_{i=1}^M \sigma_i w_i dW_t^i \right] = \sum_{i=1}^M \sum_{j=1}^M w_i w_j \sigma_i \sigma_j dW_t^i dW_t^j = \left( \sum_{i=1}^M \sum_{j=1}^M w_i w_j \sigma_i \sigma_j \rho_{ij} \right) dt,$$

que puede ser escrito en forma cuadrática como  $(\vec{w}^T \Sigma \vec{w}) dt$  donde  $\vec{w}$  es el vector de pesos y  $\Sigma$  es la matriz varianza - covarianza que captura las relaciones entre las volatilidades. Finalmente:

$$\sum_{i=1}^M \sigma_i w_i dW_t^i = \sqrt{\vec{w}^T \Sigma \vec{w}} dW_t.$$

De igual modo para el drift se puede escribir:

$$\sum_{i=1}^M \mu_i w_i = \vec{w}^T \vec{\mu}.$$

Luego el valor del portafolio se puede reescribir como:

$$dV_t = V_t \left[ (\vec{w}^T \vec{\mu}) dt + \sqrt{\vec{w}^T \Sigma \vec{w}} dW_t + \left( \sum_{i=1}^M w_i J_i dN_t^i \right) \right]. \quad (2)$$

Ahora, siguiendo a Privault (2018), se aplica el lema de Itô a (1):

$$df(S_t^i) = \mu_i S_t^i \frac{df}{dS_t^i} dt + \sigma_i S_t^i \frac{df}{dS_t^i} dW_t^i + \frac{\sigma_i^2}{2} S_t^{i2} \frac{d^2 f}{dS_t^{i2}} dt + (f(S_t^i) - f(S_{t-}^i)) dN_t^i, \quad (3)$$

de modo que aplicando  $f = \ln(S^i)$ :

$$d[\ln(S^i)] = \left( \mu_i - \frac{\sigma_i^2}{2} \right) dt + \sigma_i dW_t^i + Q_i dN_t^i,$$

con  $Q_i = \ln(1 + J_i)$ . De manera análoga se puede aplicar el lema de Itô al portafolio para  $f = \ln(V_t)$  teniendo en cuenta que hay múltiples procesos de Poisson involucrados:

$$d[\ln(V_t)] = \left( \vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2} \right) dt + \sqrt{\vec{w}^T \Sigma \vec{w}} dW_t + \sum_{i=1}^M [\ln(V_{t-} + V_{t-} w_i J_i) - \ln(V_{t-})] dN_t^i, \quad (4)$$

donde el término que contiene a los procesos de Poisson se puede tratar:

$$\sum_{i=1}^M [\ln(V_{t-} + V_{t-} w_i J_i) - \ln(V_{t-})] dN_t^i = \sum_{i=1}^M [\ln(1 + w_i J_i)] dN_t^i = \sum_{i=1}^M [\ln(1 + (e^{Q_i} - 1) w_i)] dN_t^i.$$

Ahora, se utilizará una aproximación de Taylor de primer orden usando la función  $g(Q_i) = \ln(1 + (e^{Q_i} - 1) w_i)$  centrada en  $Q_i = 0$ :

$$g(Q_i) \approx g(0) + g'(0) Q_i = \left[ \frac{w_i e^{Q_i}}{1 + (e^{Q_i} - 1) w_i} \Big|_{Q_i=0} \right] Q_i = w_i Q_i,$$

de modo que el proceso estocástico para  $f = \ln(V_t)$  se puede resumir como:

$$d[\ln(V_t)] \approx \left( \vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2} \right) dt + \sqrt{\vec{w}^T \Sigma \vec{w}} dW_t + \sum_{i=1}^M w_i Q_i dN_t^i, \quad (5)$$

que, resumiendo los procesos de Poisson en forma vectorial:

$$d[\ln(V_t)] \approx \left( \vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2} \right) dt + \sqrt{\vec{w}^T \Sigma \vec{w}} dW_t + \vec{w}^T (\vec{Q} \circ d\vec{N}_t), \quad (6)$$

donde el operador  $\circ$  representa el producto Hadamard,  $\vec{Q}$  y  $d\vec{N}_t$  son vectores tales que:  $\vec{Q} = [Q_1, Q_2, \dots, Q_M]$ ,  $d\vec{N}_t = [dN_t^1, dN_t^2, \dots, dN_t^M]$ .

Integrando el proceso (6) entre  $(t, t + \Delta t)$  (con  $\Delta t$  pequeño) se tiene que:

$$\Delta[\ln(V_t)] \approx \left( \vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2} \right) \Delta t + \sqrt{\vec{w}^T \Sigma \vec{w}} \Delta W_t + \vec{w}^T (\vec{Q} \circ \Delta \vec{N}_t). \quad (7)$$

Resulta útil aproximar los procesos  $\vec{N}_t$  por la distribución binomial con un solo experimento (Bernoulli), que posee media  $\vec{\lambda} \Delta t$  donde  $\vec{\lambda}$  representa el vector de tasas de los procesos de Poisson  $\vec{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_M]$ . De este modo se pueden hallar expresiones para el valor esperado del proceso (7):

$$E[\Delta(\ln V_t)] = \left( \vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2} \right) \Delta t + \vec{w}^T (E[\vec{Q}] \circ \vec{\lambda}) \Delta t,$$

y de manera similar la varianza:

$$\begin{aligned} \text{Var} [\Delta (\ln V_t)] &= (\vec{w}^T \Sigma \vec{w}) \Delta t + \text{Var} [\vec{w}^T (\vec{Q} \circ \Delta \vec{N}_t)] \\ &= (\vec{w}^T [\Sigma + \text{diag}(\vec{\lambda} \circ (E[\vec{Q}^2] - E^2[\vec{Q}] \circ \vec{\lambda} \Delta t))] \vec{w}) \Delta t, \end{aligned}$$

donde  $\text{diag}(x)$  se refiere a una matriz diagonal conteniendo el vector  $x$ . Nótese que a la matriz var-cov  $\Sigma$  que está caracterizada por distribuciones normales se le está sumando una diagonal, un proceso parecido al shrinkage que se suele usar para corregir las matrices var-cov (Ledoit & Wolf, 2004). Si se introducen correlaciones entre los saltos se puede incluir una matriz varianza - covarianza  $\Lambda$  similar a  $\Sigma$  que las incluya:

$$\text{Var} [\Delta (\ln V_t)] = (\vec{w}^T [\Sigma + \Lambda] \vec{w}) \Delta t.$$

Esta matriz puede ser calculada como

$$\Lambda = C_\lambda \circ \left\{ \left[ \vec{\lambda} \circ (E[\vec{Q}^2] - E^2[\vec{Q}] \circ \vec{\lambda} \Delta t) \right] \left[ \vec{\lambda} \circ (E[\vec{Q}^2] - E^2[\vec{Q}] \circ \vec{\lambda} \Delta t) \right]^T \right\},$$

donde  $C_\lambda$  es la matriz de correlaciones de eventos extremos. Según Synowiec (2008), el proceso  $Q_i$  se puede identificar por un modelo doblemente exponencial, que ha demostrado ser bastante adecuado para modelar la amplitud de saltos:

$$f_i(x; p, \eta^+, \eta^-) = p\eta^+ e^{-\eta^+ x} \mathbf{1}_{[0, \infty)}(x) + q\eta^- e^{\eta^- x} \mathbf{1}_{(-\infty, 0)}(x) \quad (8)$$

## 3.2 Distribución de los retornos logarítmicos

En esta sección se describirán las fórmulas que caracterizan las distribuciones de los retornos logarítmicos tanto para el caso de un solo activo como para un portafolio, basándose en lo desarrollado en la sección anterior.

### 3.2.1 Caso para un solo activo

Si se tiene un solo activo donde su precio es  $S_t$  para el momento  $t \geq 0$  su correspondiente proceso estocástico es:

$$\Delta [\ln(S_t)] = \left( \mu - \frac{\sigma^2}{2} \right) \Delta t + \sigma \Delta W_t + Q \Delta N_t, \quad (9)$$

donde  $Q$  posee una distribución doblemente exponencial descrita en (8); entonces para tiempos muy pequeños el proceso  $\Delta N_t$  se puede aproximar por una distribución binomial (se da un salto o no). Luego la distribución sería:

$$f_{\Delta \ln(S_t)} \approx (1 - \lambda \Delta t) f_X + \lambda f_{X+Q} \Delta t,$$



donde  $f_X$  es una distribución normal para  $X := \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma \Delta W_t$  y  $f_{X+Q}$  es la distribución cuando se le suma el salto. Synowiec (2008) muestra que esta distribución es equivalente a:

$$\begin{aligned}
f_{\Delta \ln(S_t)}(x) &= \frac{1 - \lambda \Delta t}{\sigma \sqrt{\Delta t}} \varphi \left( \frac{x - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t}{\sigma \sqrt{\Delta t}} \right) \\
&+ \lambda \Delta t \left( p \eta^+ e^{\frac{1}{2} \eta^{+2} \sigma^2 \Delta t} e^{-\left(x - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t\right) \eta^+} \Phi \left( \frac{x - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t - \eta^+ \sigma^2 \Delta t}{\sigma \sqrt{\Delta t}} \right) \right) \\
&+ \lambda \Delta t \left( q \eta^- e^{\frac{1}{2} \eta^{-2} \sigma^2 \Delta t} e^{\left(x - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t\right) \eta^-} \Phi \left( \frac{x - \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \eta^- \sigma^2 \Delta t}{\sigma \sqrt{\Delta t}} \right) \right),
\end{aligned} \tag{10}$$

donde  $\varphi$  es la densidad de la distribución y  $\Phi$  es la función de distribución acumulada de una normal estándar.

### 3.2.2 Caso para un portafolio con saltos no correlacionados

Para un portafolio se puede demostrar que su ecuación es:

$$f_{\Delta \ln(V_t)} \approx \left(1 - \lambda_V \frac{\Delta t}{M}\right) f_X + \lambda_V f_{X+Q} \frac{\Delta t}{M}, \tag{11}$$

con  $\lambda_V = \sum_{i=1}^M \lambda_i \mathbf{1}_{[w_i \neq 0]}$  la tasa del proceso de Poisson del portafolio, con  $X := \left(\vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2}\right)t + \sqrt{\vec{w}^T \Sigma \vec{w}} W_t$  la cual sigue una distribución normal con media  $\left(\vec{w}^T \vec{\mu} - \frac{\vec{w}^T \Sigma \vec{w}}{2}\right)t$  y varianza  $\vec{w}^T \Sigma \vec{w} t$ , y es independiente de  $Q$ , con distribución:

$$f_Q(x) = \sum_{i=1}^M \frac{\lambda_i}{\lambda} f_i \left( x; p_i, \frac{\eta_i^+}{w_i}, \frac{\eta_i^-}{w_i} \right) \mathbf{1}_{[w_i \neq 0]},$$

donde  $f_i$  es la distribución doblemente exponencial de cada activo. De la ecuación (11) se puede ver que la ecuación de un portafolio se puede caracterizar por la suma de un variable normal y un proceso compuesto de Poisson ajustado al tiempo  $\Delta t/M$  pues ahora la frecuencia de saltos es mucho más alta. Para utilizar esta distribución en la práctica se puede utilizar una aproximación mediante la distribución binomial donde se debe cumplir la condición  $\lambda \Delta t/M \leq 1$ . La demostración se encuentra en el apéndice.

### 3.3 Pruebas de ajuste estadístico

Para corroborar que una serie de datos se ajusta a una distribución de probabilidad se pueden utilizar los estadísticos de Anderson-Darling y el estadístico Cramer von-Mises, cada uno con la hipótesis nula de que la distribución empírica es estadísticamente igual a la distribución paramétrica propuesta.

Ambos estadísticos se basan en la distancia entre la función acumulada empírica  $F_n$  y la función a probar  $F$ :

$$n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 w(x) dF(x),$$

donde  $w(x)$  es la función que pesa la observación. En el caso Anderson-Darling se tiene  $w(x) = [F(x)(1 - F(x))]^{-1}$  mientras que para el estadístico Cramer von-Mises se tiene que  $w(x) = 1$ , dándole menos importancia a las observaciones en las colas en comparación con el primero. En el contexto de este trabajo, es más conveniente atender la situación de las colas y por ello resulta más adecuado usar el estadístico Anderson-Darling como prueba.

### 3.4 Optimización de portafolios

La optimización consiste en distribuir un presupuesto entre los activos que compondrán el portafolio. Para ello, se plantea una función objetivo a maximizar o minimizar donde las variables de decisión serán los pesos de los activos.

#### 3.4.1 Media - Varianza

Este enfoque fue desarrollado por Markowitz (1952) y típicamente consiste en:

$$\begin{aligned} \min_{\vec{w}} \quad & \frac{1}{2} \vec{w}^T \Sigma \vec{w} \\ \text{s.t.} \quad & \vec{w}^T \mathbf{1} = 1 \\ & \vec{w}^T \vec{\mu} \geq \mu^* \\ & w_i \geq 0 \quad \forall_i, \end{aligned} \tag{12}$$

donde  $w$  es el vector de pesos del portafolio. Este problema minimiza la varianza que es  $\vec{w}^T \Sigma \vec{w}$ , sujeto a una restricción presupuestaria ( $\vec{w}^T \mathbf{1} = 1$ ), un mínimo retorno medio ( $\vec{w}^T \vec{\mu} \geq \mu^*$ ) y un portafolio solo con posiciones en largo ( $w_i \geq 0$ ). Este problema es de optimización cuadrática, que computacionalmente es bastante eficiente.

#### 3.4.2 Mínimo - CVaR

La medida de volatilidad no es una medida de riesgo coherente a diferencia de la medida conocida como CVaR o *Expected Shortfall*, que adicionalmente resulta más adecuada para el uso de los inversionistas. Esta se define como:

$$ES_{\alpha}(X) = -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_{\alpha}(X)} xf(x)dx,$$

donde  $\alpha$  define el nivel desde el cual al inversionista le interesan las pérdidas y  $\text{VaR}_{\alpha}(X)$  es el percentil  $\alpha$  de la distribución. La minimización del CVaR se expone en Rockafellar & Uryasev (2000) como un problema de programación lineal:

$$\begin{aligned}
& \min_{\vec{w}, z^{\text{VaR}}, \vec{v}} -z^{\text{VaR}} + \frac{1}{(1-\alpha)J} \sum_{j=1}^J v_j \leq c^* \\
& \text{s.t. } v_j \geq -z_i + z^{\text{VaR}} \quad \text{para } j = 1, \dots, J \\
& \quad v_j \geq 0 \quad \text{para } j = 1, \dots, J \\
& \quad z_i = \sum_{n=1}^M w_n r_{jn} \quad \text{para } j = 1, \dots, J \\
& \quad \vec{w}^T \mathbf{1} = 1 \\
& \quad w_i \geq 0 \quad \forall_i,
\end{aligned} \tag{13}$$

donde  $j = 1, \dots, J$  representa simulaciones de Monte Carlo. Este enfoque será de principal interés pues probará el ajuste en las colas de la distribución.

## 4 Metodología y descripción de datos

### 4.1 Descripción de datos

Los datos utilizados para el ejercicio se obtuvieron de la librería de Kenneth R. French de 49 Industry Portfolios para el periodo comprendido desde inicios de 1996 hasta finales de Febrero de 2021. La base de datos está compuesta por los retornos diarios del valor de mercado de las 49 principales industrias de la economía estadounidense. A continuación se presentan estadísticas descriptivas para algunas industrias:

Código	Sector	Retorno	Volatilidad	Skewness	Kurtosis
Agric	Agricultura	5.9	27.8	-0.2	11.8
Food	Comida	7.5	16.5	-0.1	8.7
Soda	Soda y dulces	8.7	23	-0.9	17.1
Beer	Licor	7	18.5	-0.2	12.5
Smoke	Cigarrillo	13.3	24.1	-0.3	10.9
Toys	Juguetes	5.7	26.9	-0.4	7.6
Fun	Recreación	6.4	32.2	-0.4	6.7
Books	Imprenta	-0.7	25	-0.1	10.6
Hshld	Hogar	5.4	18.2	-1.1	25.5

Tabla 1: Principales estadísticas descriptivas calculadas sobre los retornos para 9 de las 49 principales industrias estadounidenses.

Las estadísticas descriptivas para todas las industrias se encuentran en el apéndice, dentro de las cuales se encuentran el retorno anualizado, la volatilidad anualizada, asimetría y kurtosis. De estas se puede notar que los retornos presentan asimetría negativa (mayor probabilidad en la cola izquierda) y su kurtosis indica gran relevancia de los eventos en las colas, por lo cual puede ser indicado utilizar una distribución diferente a la gaussiana.

## 4.2 Metodología

Para comprobar que la distribución planteada presenta una ventaja frente a la distribución normal en diseño de portafolios se realizaron las siguientes pruebas:

- (i) Se realizaron las pruebas estadísticas de Anderson-Darling y Cramer von-Mises para cada acción verificando que las acciones siguen la distribución o por lo menos estas obtienen un estadístico menor comparado a la distribución normal. Este procedimiento se efectúa para toda la muestra.
- (ii) Se realizaron *back-test* históricos bajo enfoques de mínima varianza, media/varianza y mínimo CVaR conformando diversos portafolios: un portafolio creado con el modelo normal, el segundo que se creó con el modelo de Poisson asumiendo que los saltos no se correlacionan, otro donde se tiene en cuenta el proceso normal que predice el modelo de Poisson y finalmente el modelo donde los saltos se correlacionan.

En ambos casos los parámetros se calcularon por máxima verosimilitud. Por ejemplo, para la muestra completa el sector agricultura tuvo como parámetros:

$$(\mu, \lambda, \sigma, p, \eta^+, \eta^-) = (2.72\%, 117.29, 13\%, 54.77\%, 67.77, 61.05),$$

bajo el modelo de poisson. Esto quiere decir que 117.29 de los 260 días de trading ocurren saltos, mientras que el retorno y la volatilidad de la parte continua es de 2.72% y 13% respectivamente. En cambio, bajo el modelo normal se obtiene  $(\mu, \sigma) = (10\%, 27.8\%)$ . Para realizar una mejor comparativa es pertinente calcular el retorno esperado y la volatilidad de cada modelo:

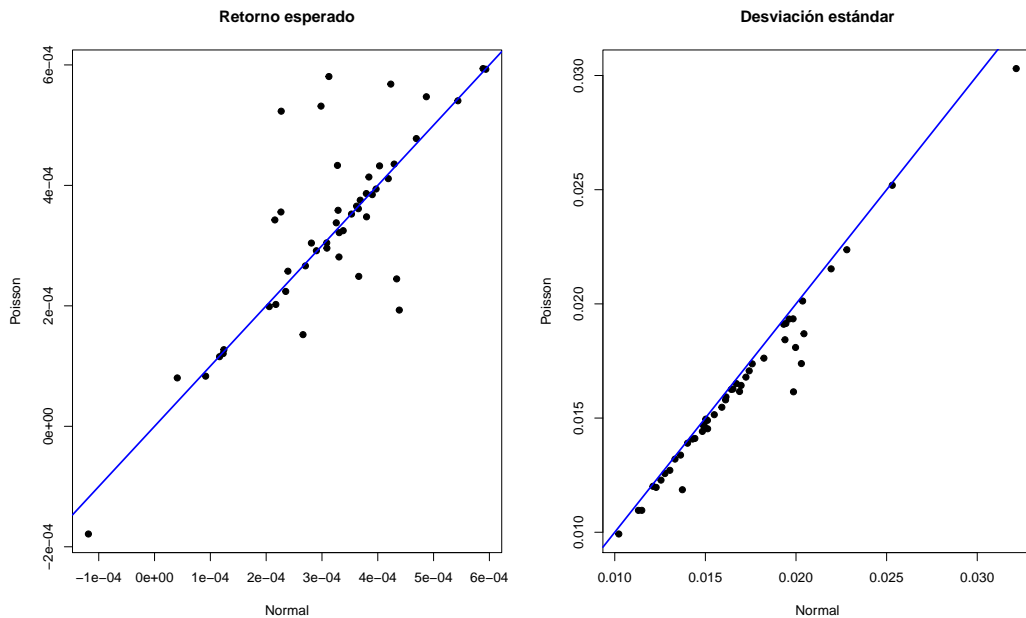


Figura 1: Retorno y volatilidad diaria calculados para las 49 industrias utilizando toda la muestra. Se pone como referencia la línea  $y = x$  en azul.

De aquí se puede notar que los retornos esperados mantienen una relación positiva entre ambos modelos pero difieren en magnitud en algunos casos. Por otro lado, se puede notar que la volatilidad del modelo normal es mayor a la que predice el modelo de saltos.

Para el enfoque de saltos correlacionados se calculó la matriz de correlaciones con el enfoque propuesto en Choi & Shin (2018), el cual calcula las correlaciones de eventos extremos. Esta matriz puede no llegar a ser definida positiva, por lo cual se aproximó a la matriz definida positiva más cercana. Un ejemplo de esta se encuentra en el apéndice.

## 5 Resultados

### 5.1 Ajuste univariado

En primer lugar, se realizó el ajuste del modelo de Poisson para las 49 industrias para toda la muestra. Seguidamente, se realizaron los test estadísticos de Anderson - Darling (AD) y Cramér-von Mises (CV), los resultados se encuentran a continuación:

Distribución	Anderson-Darling	Cramér-von Mises
Poisson	60.4 %	75 %
Normal	0 %	0 %

Tabla 2: Porcentaje de acciones que pasan el test estadístico con un p-valor mayor al 5 %.

Se puede ver que el porcentaje de acciones donde el modelo normal es la distribución generadora de datos es el 0%, mientras que para el 60.4% de las industrias el estadístico AD no rechaza la hipótesis nula de corresponder al modelo de saltos. Esto último indica que los saltos de Poisson ayudan a modelar los eventos extremos, aunque se debe considerar que pueden existir alternativas más apropiadas, investigación que se puede dejar para un trabajo futuro. Finalmente, el estadístico CV presenta un mayor porcentaje respecto al AD.

### 5.2 Backtesting

Posteriormente se presentan los resultados del backtesting para el caso de mínima varianza, mínimo CVaR y media/varianza, calculadas según (12) y (13). Para este ejercicio se rebalanceó cada trimestre, realizando el cálculo de los parámetros para los modelos con una ventana de tiempo de 4 años (cada año contiene 260 días de trading). Las gráficas del backtest se encuentran en el apéndice junto a un ejemplo de las matrices de correlación calculadas.

Las estrategias se describen a continuación:

- (i) **Normal** se refiere al portafolio construido con la distribución normal.
- (ii) **Poisson** está basado en la distribución de Poisson con saltos no correlacionados.

- (iii) **Poisson (C)** posee el mismo enfoque que el anterior con la diferencia que este tiene en cuenta las correlaciones
- (iv) **Normal(P)** es aquel que solo tiene en cuenta los eventos normales inferidos en el modelo de Poisson, es decir, esta estrategia usa los parámetros  $\mu$  y  $\sigma$  encontrados en (1).

Las estadísticas reportadas son:

- **Retorno anualizado** Esta es la tasa compuesta de retorno al final del periodo calculada como:

$$R_a = \left( \prod_{t_1}^T (1 + R_t) \right)^{260/T} - 1$$

Donde  $T$  es el número de días y  $R_t$  el retorno diario del portafolio. Se toman 260 días de trading. Se reporta como porcentaje.

- **Desviación Estándar** Para hallar la volatilidad anualizada de un portafolio se debe ajustar la desviación estándar por el número de periodos:

$$\sigma_a = \sqrt{\text{Varianza}(R_t)} \sqrt{260}$$

Al igual que el retorno anualizado esta se reporta como porcentaje.

- **Sharpe Ratio** Es una medida de riesgo-retorno. Se define como el cociente entre el retorno anualizado y la desviación estándar.
- **Maximum Drawdown** Esta es la caída más grande que tuvo el valor del portafolio durante el periodo de estudio. Comenzando por la primera fecha se van calculando las pérdidas acumuladas hasta que el portafolio encuentre un máximo. Posteriormente se repite el proceso hasta que la muestra acabe. De todas estas perdidas se escoge la mayor. Se reporta como porcentaje.
- **Calmar Ratio** Al igual que el Sharpe Ratio es una medida de riesgo-retorno, sin embargo, en este caso el riesgo se define como el Maximum Drawdown, luego será el coeficiente entre el retorno y la mayor pérdida acumulada.
- **CVaR** Definida anteriormente, esta se interpreta como la pérdida de valor esperada en un día para los  $\alpha$  % peores casos. Para este ejercicio se reporta el CVaR para el 5% de las peores pérdidas. Se reporta como porcentaje.
- **NEC** El número efectivo de componentes es una medida de concentración de los pesos de un portafolio. Esta se calcula como:

$$NEC = \frac{1}{\sum_{i=1}^M w_i^2}$$

Donde si se tiene un portafolio tiene pesos iguales, es decir,  $w_i = 1/M$  entonces  $NEC = M$  que es la máxima diversificación posible. En caso contrario donde solo se invierte en un activo se tiene  $NEC = 1$ . Se reporta el promedio que se obtuvo del NEC para los rebalances y su desviación estándar.

A continuación se presenta el análisis y los resultados del ejercicio. Para cada estadística se subraya aquel portafolio que obtuvo el mejor resultado.

### 5.2.1 Mínima varianza

Se presentan las estadísticas para la minimización de la varianza:

Estadística	Normal	Poisson	Normal (P)	Poisson (C)
<b>Retorno</b>	10.8	10.7	<u>11.3</u>	10.3
<b>Desv. Estándar</b>	18.3	17.4	<u>17.7</u>	16.1
<b>Sharpe Ratio</b>	0.593	0.612	0.639	<u>0.642</u>
<b>Max. Drawdown</b>	53.8	52.7	52.6	<u>48.5</u>
<b>Calmar Ratio</b>	0.201	0.202	<u>0.214</u>	0.212
<b>CVaR</b>	2.7	2.6	2.6	<u>2.4</u>
<b>NEC Promedio</b>	39.7	25.7	24.7	20.7
<b>Desv. NEC</b>	2.4	2.9	2.8	2.3

Tabla 3: Estadísticas calculadas para la metodología de mínima varianza reportado para las 4 estrategias.

De las estadísticas de los portafolios de mínima varianza se puede concluir lo siguiente:

- Al establecer como objetivo minimizar la volatilidad del portafolio se pueden ver mejoras significativas como la reducción de la volatilidad al incluir el modelo con saltos, en su versión correlacionada como no correlacionada.
- El portafolio con saltos correlacionados es el que presenta las mejores estadísticas respecto a volatilidad, retorno-riesgo, máxima pérdida y CVaR. Esto indica que la inclusión de correlación de eventos extremos de manera explícita agrega una ventaja en el diseño de portafolios para este objetivo.
- Finalmente, el portafolio que presenta únicamente el componente normal derivado del modelo con saltos presenta mejores estadísticas que el modelo sin saltos, esto puede ser consecuencia de una correcta identificación por parte del modelo entre eventos extremos (saltos) y comunes (normal).

### 5.2.2 Media-varianza

Para el enfoque de media-varianza se calculó un portafolio equally weighted (de pesos iguales) para establecer su retorno esperado como el mínimo retorno dentro del problema de optimización. Las estadísticas se presentan a continuación:

Estadística	Normal	Poisson	Normal (P)	Poisson (C)	EW
<b>Retorno</b>	9.1	10.9	<u>11.6</u>	11.1	10.4
<b>Desv. Estándar</b>	29	20.9	21.1	<u>19.9</u>	20.3
<b>Sharpe Ratio</b>	0.315	0.52	0.547	<u>0.557</u>	0.511
<b>Max. Drawdown</b>	68.2	50.8	51	<u>46.4</u>	56.5
<b>Calmar Ratio</b>	0.134	0.214	0.227	<u>0.239</u>	0.183
<b>CVaR</b>	4.3	3.1	3.1	<u>2.9</u>	3
<b>NEC Promedio</b>	12.4	16.3	15.6	12.9	49
<b>Desv. NEC</b>	17.1	9.9	9.5	8.1	0

Tabla 4: Estadísticas calculadas para la metodología de media-varianza reportado para un portafolio de pesos iguales (EW) establecido como benchmark y las 4 estrategias.

Para el caso de los portafolios de media/varianza se puede notar lo siguiente:

- Nuevamente todos los portafolios derivados del modelo de Poisson presentan mejores estadísticas que el modelo normal y sobrepasan al benchmark en retorno-riesgo.
- El modelo correlacionado lidera en las estadísticas exceptuando el retorno anualizado. Este resultado reafirma la ventaja que presenta tener en cuenta la correlación entre los saltos. La diferencia en términos de volatilidad y máxima pérdida con el portafolio normal es bastante significativa.

### 5.2.3 Mínimo CVaR

Finalmente, las estadísticas para el caso de la minimización del CVaR:

Estadística	Normal	Poisson	Normal (P)	Poisson (C)
<b>Retorno</b>	11.3	11.2	9.3	<u>12</u>
<b>Desv. Estándar</b>	<u>15.1</u>	16.3	22.9	15.5
<b>Sharpe Ratio</b>	0.75	0.686	0.409	<u>0.778</u>
<b>Max. Drawdown</b>	39.6	44.3	60.2	<u>37.6</u>
<b>Calmar Ratio</b>	0.286	0.253	0.155	<u>0.32</u>
<b>CVaR</b>	<u>2.2</u>	2.4	3.4	2.3
<b>NEC Promedio</b>	5.9	12.1	3.2	5.2
<b>Desv. NEC</b>	1.6	3.4	2	1.9

Tabla 5: Estadísticas calculadas para la metodología de mínimo CVaR reportado para las 4 estrategias.

Por último, para el caso de la minimización del CVaR:

- El portafolio **Normal(P)** presenta el mayor rezago. Como esta vez se tiene un objetivo de minimización de eventos extremos es de esperar que esto suceda si el modelo discrimina bien entre los eventos extremos y comunes.



- El modelo de Poisson sin correlaciones se queda rezagado, lo cual indica que la correlación entre los saltos presenta una mayor relevancia para este tipo de optimizaciones.
- Finalmente, cuando se incluyen las correlaciones se presentan mejoras en indicadores de riesgo-retorno respecto a los demás enfoques. La diferencia entre las variables de riesgo no es mucha entre el modelo normal y el modelo de Poisson correlacionado. El primero presenta un menor CVaR en el agregado, pero el último una menor pérdida máxima, siendo esta última estadística más importante en la práctica.

## 6 Conclusiones

Los administradores de portafolio a menudo se encuentran con varios problemas al diseñar portafolios bajo un enfoque gaussiano como son la poca robustez del retorno esperado, la consistencia de los estimadores de varianza-covarianza, entre otros. En ese orden de ideas el presente estudio realiza varios aportes que buscan dar un soporte teórico acorde con lo observado en el mercado y adicionalmente realiza un ejercicio empírico para mostrar su validez. En primer lugar, se plantearon las ecuaciones para la dinámica de los activos cuando se le añaden saltos de Poisson, seguidamente se encontraron los primeros momentos estadísticos del portafolio y se demostró que el portafolio de saltos no correlacionados sigue un proceso muy similar al de sus activos.

Por otra parte, gracias a los resultados empíricos se puede concluir que el modelo de Poisson presenta un mejor ajuste univariado comparado al modelo normal para los retornos de los activos estudiados. Esto va en línea con los resultados obtenidos bajo los enfoques de media/varianza y mínima varianza, donde el modelo de Poisson sobrepasa en estadísticas al modelo normal. Adicionalmente, se pudo corroborar que las correlaciones de los saltos de Poisson resultan relevantes cuando se desea optimizar en cualquiera de los 3 casos. Esto último puede ser consecuencia del riesgo no diversificable de la economía, que se materializa como una mayor correlación entre los activos en eventos extremos.

Por último, se encontraron varias limitaciones: el modelo de Poisson no es el proceso generador de datos para la totalidad de los activos, lo cual se puede deber a la manera en que se calcularon los parámetros o la distribución que se utilizó para modelar la amplitud de los saltos. Por otra parte, a pesar de que en este estudio el modelo de Poisson logró superar en algunas estadísticas para el caso de la minimización del CVaR, los resultados no son concluyentes y vale la pena hacer más pruebas para determinar si este enfoque posee ventajas.

Como posibles trabajos futuros se puede poner a prueba la robustez de la metodología planteada. En primer lugar, utilizar ventanas de tiempo más cortas para el cálculo de los parámetros, utilizar otros tipos de datos y realizar un rebalanceo más frecuente determinarían si el enfoque en definitiva es mejor. Para ello se pueden mirar otros mercados como el colombiano, que a diferencia del mercado estadounidense es más volátil y posee una menor liquidez, por lo cual los saltos en los precios pueden llegar a ser más frecuentes. Por otro lado, sería correcto enfocarse en los periodos de caída del mercado para visualizar los efectos que tiene sobre la gestión del riesgo de portafolios en estas condiciones.

## Referencias

- Benati, S. (2015). Using medians in portfolio optimization. *Journal of the Operational Research Society*, 66(5), 720–731. <https://doi.org/10.1057/jors.2014.57>
- Choi, J.-E., & Shin, D. W. (2018). Quantile correlation coefficient: A new tail dependence measure. <https://arxiv.org/abs/1803.06200v1>
- Hanson, F., & Westman, J. (2002). Portfolio optimization with jump-diffusions: Estimation of time-dependent parameters and application [ISSN: 0191-2216]. *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, 1, 377–382 vol.1. <https://doi.org/10.1109/CDC.2002.1184523>
- Hunter, D. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*. Retrieved March 1, 2021, from <https://www.risk.net/node/2161203>
- Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8), 1086–1101. Retrieved March 1, 2021, from <https://www.jstor.org/stable/822677>
- Ledoit, O., & Wolf, M. (2004). Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4), 110–119. <https://doi.org/10.3905/jpm.2004.110>
- López de Prado, M. (2020). *Machine learning for asset managers*. Cambridge University Press. <https://doi.org/10.1017/9781108883658>
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- Maronna, R., Martin, D., Salibián-Barrera, M., & Yohai, V. (2006). Measuring robustness. *Robust Statistics* (pp. 51–85). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470010940.ch3>
- Mazhar, S., & Srivastava, P. (2018). Comparative analysis of sharpe and sortino ratio with reference to top ten banking and finance sector mutual funds. *International Journal of Management Studies (IJMS)*, 5(4(2)). Retrieved March 1, 2021, from <http://www.researchersworld.com/index.php/ijms/article/view/1236>
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1), 125–144. [https://doi.org/10.1016/0304-405X\(76\)90022-2](https://doi.org/10.1016/0304-405X(76)90022-2)
- Privault, N. (2018). Capítulo 20: Stochastic calculus for jump processes.
- Roncalli, T. (2013). *Introduction to risk parity and budgeting* (1st edition). Chapman; Hall/CRC.
- Synowiec, D. (2008). Jump-diffusion models with constant parameters for financial log-return processes. *Computers & Mathematics with Applications*, 56(8), 2120–2127. <https://doi.org/10.1016/j.camwa.2008.02.051>
- Venegas, F. (2008). *Riesgos financieros y económicos. productos derivados y decisiones económicas bajo incertidumbre*. (2nd ed.). Cengage Learning Editores.

## A Distribución del retorno logarítmico de un portafolio

En esta sección se llegará a una expresión para hallar la distribución que siguen los retornos logarítmicos. Primero, integrando la ecuación (5) entre  $(0, t)$  se obtiene:

$$\ln(V_t) - \ln(V_0) = \left( w^T \mu - \frac{w^T \Sigma w}{2} \right) t + \sqrt{w^T \Sigma w} W_t + \sum_{i=1}^M \sum_{j=1}^{N_{ii}} [w_i Q_{ij}] , \quad (14)$$

donde la expresión  $Q_{ij}$  se refiere a la amplitud del salto en el momento  $j$ , de la acción  $i$ . Recordemos que por cada acción se tiene un proceso compuesto de Poisson que llamaremos  $Y_i = \sum_{j=1}^{N_{ii}} [w_i Q_{ij}]$  con  $N_t$  que sigue un proceso de Poisson con parámetro constante  $\lambda_i$  y  $Q_{ij}$  es una colección de variables iid para todo  $j$  (es decir por acción sigue la misma distribución con los mismos parámetros).

Para llegar a la distribución de la ecuación (14), se usarán proposiciones que serán demostradas a continuación.:

**Proposición 1** Sea  $Y = \sum_{i=1}^{N_t} [X_i]$  un proceso compuesto de Poisson con parámetro  $\lambda$  y  $X_i$  iid con distribución  $f_x$ . Entonces  $f_y$  se puede escribir como:

$$f_y(u) = \sum_{n=0}^{\infty} f_x^{*n}(u) \frac{e^{-\lambda t} (\lambda t)^n}{n!} , \quad (15)$$

donde  $f_x^{*n}$  hace referencia a la  $n$ -ésima iteración de la convolución de  $f_x$ , es decir la suma de  $X_1 + X_2 + \dots + X_N$ , donde  $f_x^{*0} = \delta_0$  que es la distribución de la delta de Dirac.

**Demostración** Sabemos que la variable  $Y$ :

$$Y = \begin{cases} 0 & \text{con } P(N_t = 0) \\ X_1 & \text{con } P(N_t = 1) \\ X_1 + X_2 & \text{con } P(N_t = 2) \\ \vdots & \end{cases}$$

Luego la probabilidad de que  $y$  tome cierto valor  $u$  será:

$$P(Y = u) = P(u = 0) P(N_t = 0) + P(u = X_1) P(N_t = 1) + P(u = X_1 + X_2) P(N_t = 2) + \dots ,$$

de modo que  $P(u = X_1 + X_2 + \dots + X_k)$  se puede representar como una  $k$ -convolución debido a que las  $X$  son iid y  $N_t$  tiene distribución de Poisson con parámetro  $\lambda$ . De este modo se llega a la ecuación (15).

De ahora en adelante, la distribución  $f_x$  asociada al proceso compuesto de Poisson  $Y$  llevará por nombre distribución subyacente.

**Proposición 2** Si se tiene una colección de  $M$  procesos compuestos de Poisson independientes entre sí, la suma de estos sigue una distribución compuesta de Poisson con parámetro  $\lambda = \sum_{i=1}^M \lambda_i$  y distribución subyacente  $f = \sum_{i=1}^M \frac{\lambda_i}{\lambda} f_i$

**Demostración** Un proceso compuesto de Poisson posee una distribución característica:

$$\varphi_{Y_i}(m) = E \left[ e^{imY_i} \right] = \exp(\lambda_i t (\varphi_{X_i}(m) - 1)) ,$$

con  $\varphi_{X_i}(m)$  la función característica de la distribución subyacente asociada al proceso  $Y_i$ . Ahora, como la función característica de la suma es la productoria de las funciones características individuales, si  $Y$  es la colección de los  $M$  procesos denotados por  $Y_i$  entonces:

$$\varphi_Y(m) = \prod_{i=1}^M \varphi_{X_i}(m) = \exp \left( \sum_{i=1}^M \lambda_i t [\varphi_{X_i}(m) - 1] \right) = \exp \left( \lambda t \left( \sum_{i=1}^M \left[ \frac{\lambda_i}{\lambda} \varphi_{X_i}(m) \right] - 1 \right) \right) ,$$

con  $\lambda = \sum_{i=1}^M \lambda_i$ . Ahora usando la transformada inversa de Fourier se puede recuperar la distribución subyacente:

$$f = \frac{1}{2\pi} \int_{\mathbb{R}} e^{imx} \overline{\varphi_X(m)} dm .$$

Usando  $\varphi_X(m) = \sum_{i=1}^M \left[ \frac{\lambda_i}{\lambda} \varphi_{X_i}(m) \right]$ :

$$f = \frac{1}{2\pi} \int_{\mathbb{R}} e^{imx} \overline{\sum_{i=1}^M \left[ \frac{\lambda_i}{\lambda} \varphi_{X_i}(m) \right]} dm = \frac{1}{2\pi} \int_{\mathbb{R}} e^{imx} \sum_{i=1}^M \left[ \frac{\lambda_i}{\lambda} \overline{\varphi_{X_i}(m)} \right] dm = \sum_{i=1}^M \frac{\lambda_i}{\lambda} \left[ \frac{1}{2\pi} \int_{\mathbb{R}} e^{imx} \overline{\varphi_{X_i}(m)} dm \right] ,$$

de donde se concluye:

$$f = \sum_{i=1}^M \frac{\lambda_i}{\lambda} f_i .$$

**Proposición 3** Si  $X$  sigue una distribución doblemente exponencial con parámetros  $p$ ,  $\eta^+$  y  $\eta^-$  y  $a$  es una constante real diferente de 0 entonces  $aX$  tiene distribución doblemente exponencial con parámetros  $p$ ,  $\frac{\eta^+}{a}$  y  $\frac{\eta^-}{a}$

**Demostración** La variable  $X$  sigue una función de distribución:

$$f_X(x; p, \eta^+, \eta^-) = p\eta^+ e^{-\eta^+ x} \mathbf{1}_{[0, \infty)}(x) + (1-p)\eta^- e^{\eta^- x} \mathbf{1}_{(-\infty, 0)}(x) ,$$

con función característica:

$$\varphi_X(t) = p \frac{\eta^+}{\eta^+ - it} + (1-p) \frac{\eta^-}{\eta^- + it} .$$

Luego  $aX$ :

$$\varphi_{aX}(t) = \varphi_X(at) = p \frac{\eta^+}{\eta^+ - iat} + (1 - p) \frac{\eta^-}{\eta^- + iat} = p \frac{(\eta^+/a)}{(\eta^+/a) - it} + (1 - p) \frac{(\eta^-/a)}{(\eta^-/a) + it} .$$

Y de este modo queda demostrado.

Usando estas proposiciones se puede ver que la ecuación (14) sigue una distribución de la suma de una variable normal y una compuesta de Poisson, si los saltos siguen distribuciones doblemente exponenciales con parámetros  $p_i, \eta_i^+, \eta_i^-$ . Cada acción posee un proceso compuesto de Poisson  $Y_j$ , luego por la proposición 2 y 3, la distribución subyacente de la suma de los procesos compuestos de Poisson (la notaremos como  $Y$ ) se puede escribir como:

$$f_X(x) = \sum_{i=1}^M \frac{\lambda_i}{\lambda} f_i \left( x; p_i, \frac{\eta_i^+}{w_i}, \frac{\eta_i^-}{w_i} \right) \mathbf{1}_{[w_i \neq 0]} ,$$

con  $\lambda = \sum_{i=1}^M \lambda_i \mathbf{1}_{[w_i \neq 0]}$ . Luego como  $X = \left( w^T \mu - \frac{w^T \Sigma w}{2} \right) t + \sqrt{w^T \Sigma w} W_t$  sigue una distribución normal con media  $\left( w^T \mu - \frac{w^T \Sigma w}{2} \right) t$  y varianza  $w^T \Sigma w t$ , y es independiente de  $Y$  entonces  $\ln \left( \frac{V_t}{V_0} \right) = \Delta \ln(V_t) = X + Y$  tiene una distribución:

$$f_{\Delta \ln(V_t)} = \sum_{n=0}^{\infty} (f_X * f_Y^{*n}) \frac{e^{-\lambda t} (\lambda t)^n}{n!}$$

Esta se puede aproximar a la ecuación (11).

## B Estadísticas descriptivas de los activos

<b>Código</b>	<b>Sector</b>	<b>Retorno</b>	<b>Volatilidad</b>	<b>Skewness</b>	<b>Kurtosis</b>
Agric	Agricultura	5.9	27.8	-0.2	11.8
Food	Comida	7.5	16.5	-0.1	8.7
Soda	Soda y dulces	8.7	23	-0.9	17.1
Beer	Licor	7	18.5	-0.2	12.5
Smoke	Cigarrillo	13.3	24.1	-0.3	10.9
Toys	Juguetes	5.7	26.9	-0.4	7.6
Fun	Recreación	6.4	32.2	-0.4	6.7
Books	Imprenta	-0.7	25	-0.1	10.6
Hshld	Hogar	5.4	18.2	-1.1	25.5
Clths	Ropa	9.7	26	-0.3	6.9
Hlth	Salud	7.9	23.3	-1.1	12.3
MedEq	Equipo Médico	9.7	20.2	-0.5	8
Drugs	Farmaceútica	6.3	19.5	-0.1	5.5
Boxes	Contenedores	7.3	24.2	-0.3	5.2
Trans	Transporte	7.5	24	-0.5	7.7
Whlsl	Al por mayor	6.6	20.6	-0.5	8.8
Rtail	Al por menor	7.5	21.5	-0.1	5.8
Meals	Restaurantes y Hoteles	10.5	21	-0.3	11.7
Banks	Bancos	0.7	31.3	-0.1	14.7
Insur	Asegurador	5.7	24.4	-0.4	14.6
RIEst	Finca raíz	3.1	31.6	-0.3	11.6
Fin	Financiero	1.9	32	-0.1	9.7
Other	Otros	0.3	24.1	-0.1	8.7
Chems	Químicos	5.6	25.7	-0.5	7.6
Rubbr	Caucho y Plástico	7.6	22.6	-0.4	5.7
Txtls	Textil	1.9	32.7	-0.6	15.6
BldMt	Construcción (Materiales)	6.2	26	-0.4	8.3
Cnstr	Construcción	5.6	32.8	-0.3	6.8
Steel	Metales	-5.6	36.8	-0.4	6.2
FabPr	Fabricados	0.4	33	-0.7	10.6
Mach	Maquinaria	7.5	28.1	-0.3	6.5
ElcEq	Equipo Eléctrico	3.9	27.2	-0.2	7.4
Autos	Automóvil	3.2	31.2	-0.4	7.1
Aero	Aeronáutica	5.9	27.4	-0.7	15.4
Ships	Naval	10.6	28.4	-0.3	4
Guns	Armas	13.1	24.4	-0.1	8.5

<b>Código</b>	<b>Sector</b>	<b>Retorno</b>	<b>Volatilidad</b>	<b>Skewness</b>	<b>Kurtosis</b>
Gold	Oro	-2.9	40.8	0.2	4.5
Mines	Minería	4.3	35.4	-0.3	6.7
Coal	Carbón	-15.4	51.8	-0.4	5.3
Oil	Petroleo y Gas	1.3	29.4	-0.5	13.8
Util	Servicios	6.9	19.8	-0.2	14.7
Telcm	Comunicación	0.6	22	-0.1	8.9
PerSv	Servicios Personales	3.4	23.9	-0.7	8.1
BusSv	Negocios	5.4	22.1	-0.5	9.3
Hardw	Hardware	-1.7	31.4	0.2	7.5
Softw	Software	2.6	26.5	-0.1	6.6
Chips	Equipo Electrónico	0.8	32	0	5.4
LabEq	Equipo de laboratorios	6.5	26.6	-0.2	4.2
Paper	Material de oficina	4.5	21	-0.3	6.1

Tabla 5: Principales estadísticas calculadas sobre los retornos de las 49 principales industrias del mercado norteamericano.

## C Gráficas

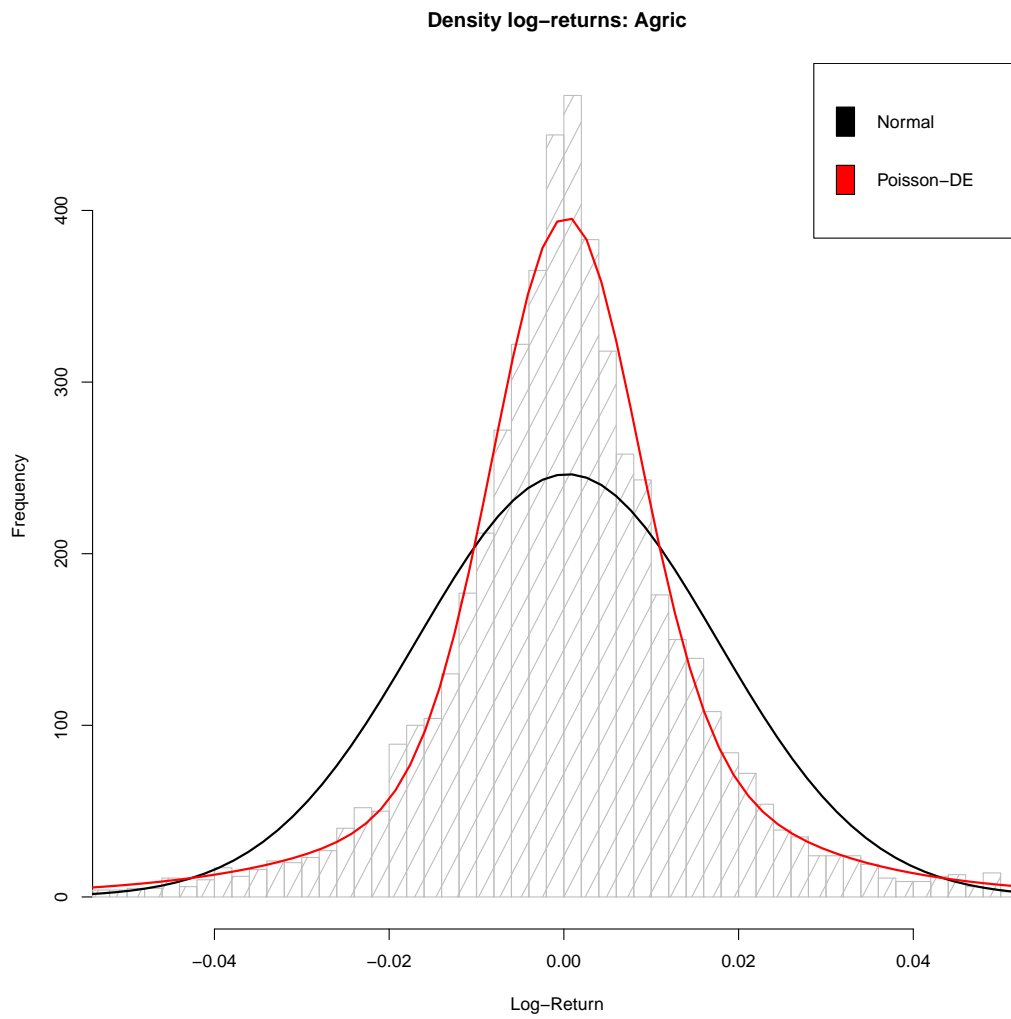


Figura 2: Distribución y ajuste de los retornos logarítmicos del sector Agricultura.



### Minimum Variance

2000-04-03 / 2021-02-26

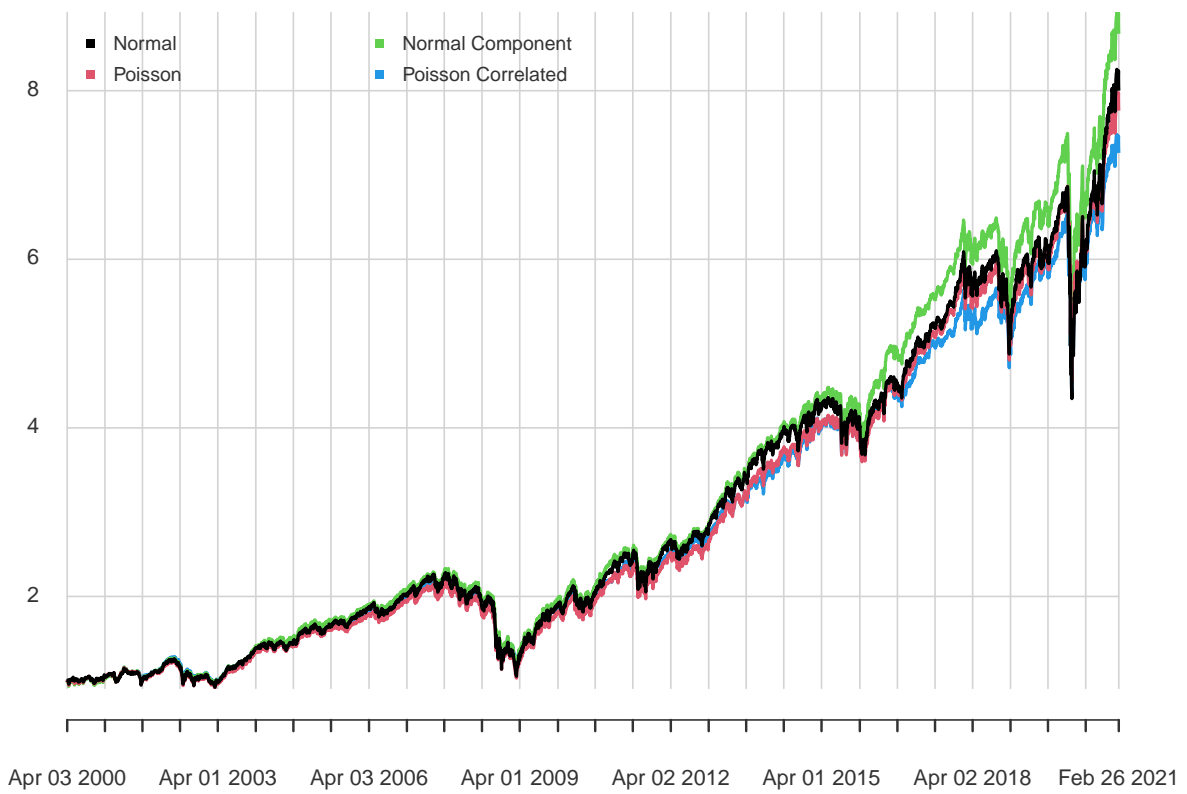


Figura 3: Evolución de la inversión de un dólar para los diversos portafolios en el caso mínima varianza.

### Mean-Variance

2000-04-03 / 2021-02-26

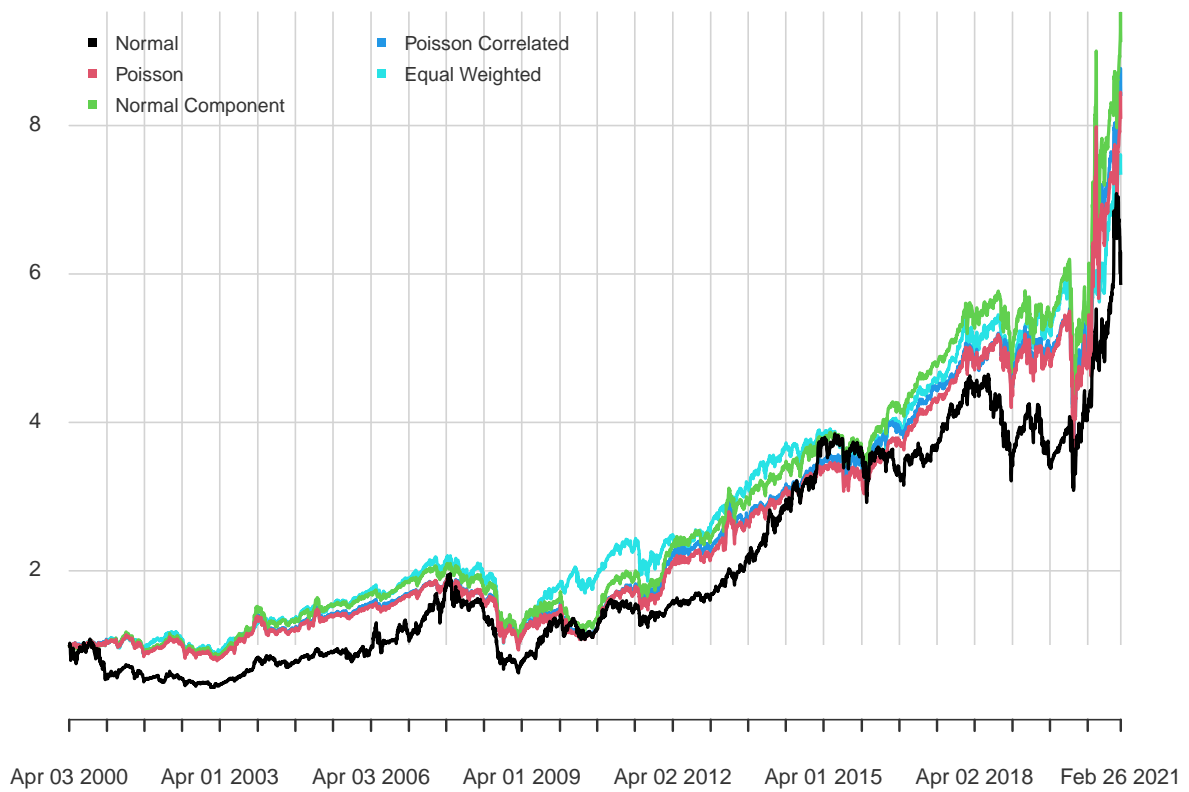


Figura 4: Evolución de la inversión de un dólar para los diversos portafolios en el caso media-varianza.

### Minimum CVaR

2000-04-03 / 2021-02-26

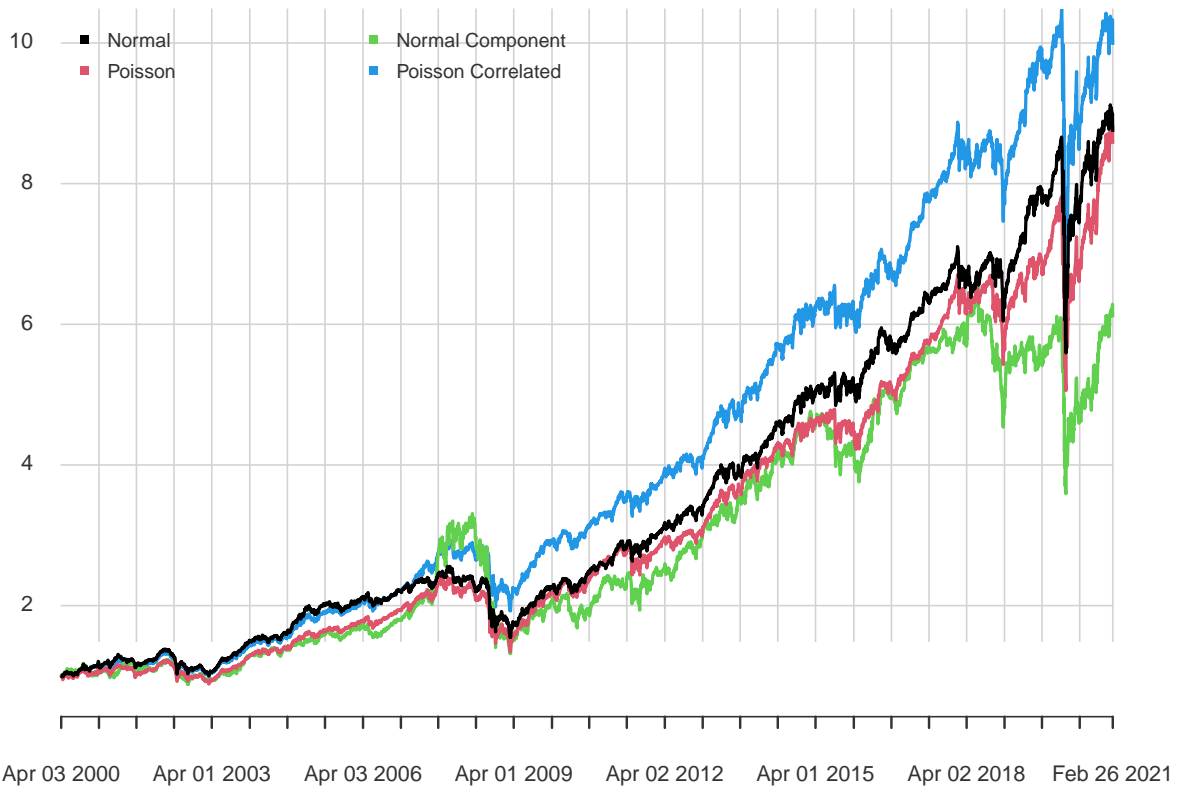


Figura 5: Evolución de la inversión de un dólar para los diversos portafolios en el caso mínimo CVaR.

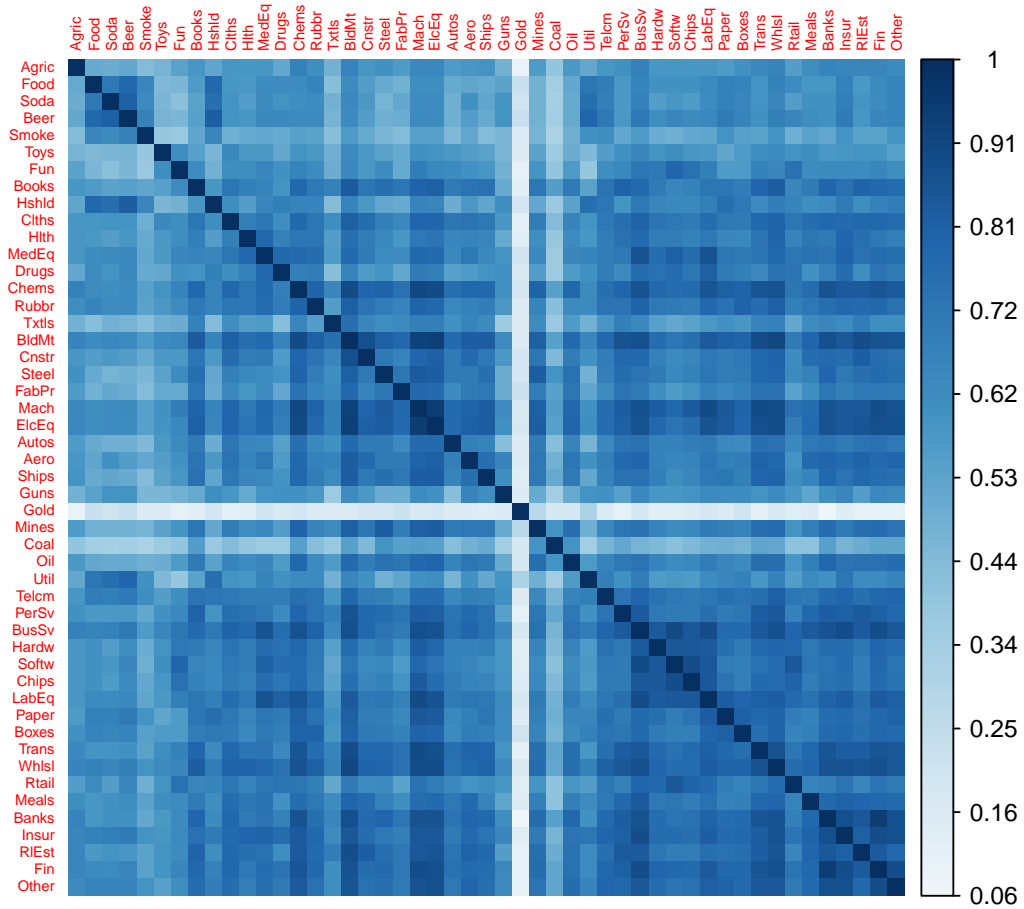


Figura 6: Matriz de correlación de pearson para la fecha 30-06-2020.

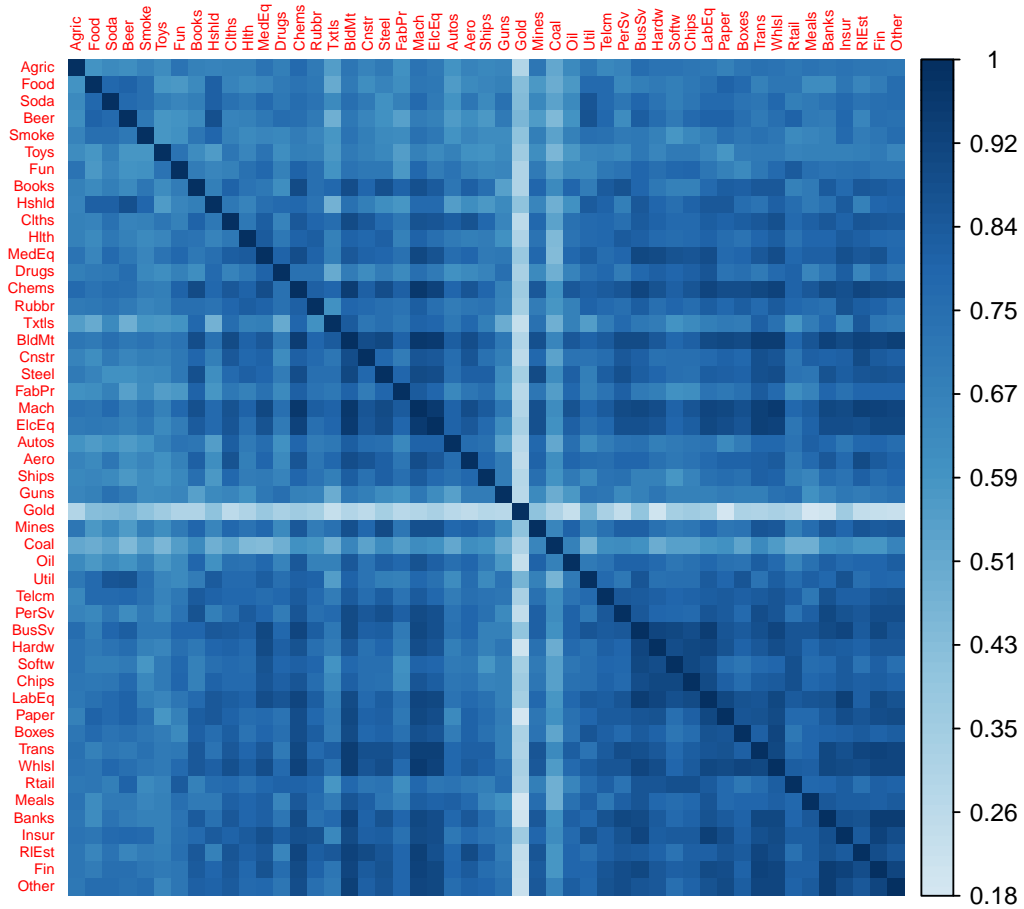


Figura 7: Matriz de correlaciones extremas para la fecha 30-06-2020.