

Selección de Cartera: Un Enfoque de Sesgos de Comportamiento

Francisco Vargas Serrano - Universidad de Sonora, México

José Arturo Montoya¹  - Universidad de Sonora, México

José del Carmen Jiménez Hernández - Universidad Tecnológica de la Mixteca, México

Resumen

El objetivo es proporcionar carteras financieras óptimas en función de una verosimilitud normal, una distribución *a priori* sobre los parámetros del modelo de valoración de activos y la opinión del inversionista sobre cómo ponderar la verosimilitud y la *a priori* en la construcción de la cartera. Se aplica la metodología de inferencia bayesiana sesgada a la selección de carteras de media-varianza, con diferentes configuraciones de sesgo o ponderación. Los resultados muestran una propuesta eficaz para encontrar carteras óptimas que reflejen ponderaciones hechas sobre la verosimilitud y las creencias *a priori*. Además, incluir sesgos en la selección de carteras puede ser relevante para la optimización de la cartera. La propuesta contribuye al campo de los sesgos de finanzas conductuales y se puede aplicar fácilmente a otros modelos financieros que han sido tratados desde un enfoque bayesiano. En conclusión, la propuesta proporciona ponderaciones óptimas para la cartera que reflejan tanto los datos como las creencias, y la inclusión de sesgos en la optimización de la cartera puede ayudar a construir carteras óptimas que incorporen preferencias de riesgo y objetivos de inversión.

Clasificación JEL: G1, G11, G41, G110, G410, G170

Palabras clave: asignación de activos, optimización de cartera, carteras de media-varianza, carteras de razón de Sharpe óptima, sesgos de comportamiento, inferencia bayesiana sesgada.

Portfolio Selection: A Behavior Biased Approach

Abstract

Here the goal is to provide optimal financial portfolios based on a normal likelihood, an *a priori* distribution on the parameters of the asset valuation model, and the investor's opinion on how to weigh likelihood and *a priori* in portfolio construction. The biased Bayesian inference methodology is applied to the selection of mean-variance portfolios, with different bias or weighting configurations. The results show an effective proposal to find optimal portfolios that reflect weightings made on likelihood and *a priori* beliefs. In addition, including biases in portfolio selection can be relevant to portfolio optimization. The proposal contributes to the field of behavioral finance biases and can be easily applied to other financial models that have been treated from a Bayesian approach. In conclusion, the proposal provides optimal weights for the portfolio that reflect both data and beliefs, and the inclusion of biases in portfolio optimization can help build optimal portfolios that incorporate risk preferences and investment objectives.

JEL Classification: G1, G11, G41, G110, G410, G170

Keywords: asset allocation, portfolio optimization, mean-variance portfolios, Sharpe ratio optimal portfolio, behavioral biases, biased Bayesian inference.

¹ Autor de correspondencia, Universidad de Sonora. E-mail: arturo.montoya@unison.mx

*Sin fuente de financiamiento para el desarrollo de la investigación

1. Introducción

El enfoque bayesiano proporciona un marco de referencia para abordar cuestiones clave en finanzas, como los problemas de selección de cartera (Pástor 2000, Harvey *et al.* 2010, De Franco *et al.* 2019, Kolm *et al.* 2021, Bauder *et al.* 2021, Bodnar *et al.* 2022, Sahamkhadam *et al.* 2022). En este enfoque, quien toma las decisiones puede combinar la verosimilitud de los datos con una distribución *a priori*, o creencia *a priori*, sobre los parámetros del modelo. La verosimilitud es la probabilidad de que un modelo produzca los datos observados, mientras que la distribución *a priori* refleja la información conocida del tomador de decisiones con respecto al grado de incertidumbre sobre los parámetros del modelo. Por lo tanto, la inferencia bayesiana proporciona una distribución posterior de los parámetros del modelo, dada la información sobre la distribución *a priori* de los parámetros y la verosimilitud de los datos observados. En esta distribución posterior, la verosimilitud de los datos tiene el mismo peso que la creencia *a priori* sobre los parámetros del modelo. Sin embargo, los tomadores de decisiones a menudo tienen diferentes niveles de confianza en sus datos, modelos y creencias. Por ejemplo, los inversionistas de cartera con exceso de confianza tienden a dar menos peso a los datos que a sus propias creencias. Además, este tipo de sesgo de comportamiento financiero puede ser la base de una estrategia de selección de cartera rentable. La toma de decisiones conductuales en finanzas se analiza en De Bondt *et al.* (2013). Para una discusión general sobre las consecuencias de los sesgos de comportamiento en las finanzas, también remitimos al lector a Moosa y Ramiah (2017).

Este trabajo trata sobre un enfoque de inferencia bayesiana sesgada para lidiar con el sesgo en la toma de decisiones financieras y se centra en el problema clásico de la selección de carteras. Se muestra cómo seleccionar carteras financieras óptimas en función de la verosimilitud de los datos, una distribución *a priori* sobre los parámetros del modelo de valoración de activos y la visión del inversionista sobre el peso de la verosimilitud y la distribución *a priori* en la construcción de la cartera. Con base en este enfoque, estos pesos pueden verse como un sesgo que cuantifica la importancia relativa de los componentes de la distribución posterior y son especificados por el tomador de decisiones.

El problema de la selección de la cartera es encontrar una asignación óptima de la riqueza en varios activos para lograr una compensación adecuada entre el rendimiento de la inversión y el riesgo. Este problema fue iniciado por Markowitz (1952) con su modelo de selección de cartera de media-varianza. Markowitz (1952) proporciona la base para la teoría moderna de la asignación de activos de cartera. Según Statman (1999), los principios de cartera de Markowitz son uno de los pilares de las finanzas neoclásicas. Por lo tanto, el enfoque propuesto aquí se aplicará al modelo de selección de cartera de media-varianza para facilitar las comparaciones con otras carteras eficientes, como la varianza mínima global y las carteras de razón de Sharpe, y para simplificar la interpretación.

El modelo de Markowitz se basa en los siguientes aspectos: (a) Los rendimientos de los activos de cualquier cartera se consideran como un vector de variables aleatorias, para las cuales el inversionista propone una distribución de probabilidad para el período bajo estudio. El valor esperado estimado del vector de variables aleatorias se utiliza para cuantificar la rentabilidad de la inversión; (b) la matriz de varianza-covarianza se utiliza para medir la dispersión de la cartera, que los gestores de cartera utilizan como medida del riesgo asociado a una cartera en particular; (c) el

comportamiento racional del inversionista le lleva a preferir la composición de una cartera que representa la mayor rentabilidad, para un determinado nivel de riesgo. Por lo tanto, un vector de rendimiento esperado μ , que contiene el rendimiento esperado de cada activo en la cartera, y una matriz de varianza-covarianza Σ , que representa las interrelaciones entre pares de activos, son los parámetros del modelo de fijación de precios de activos de Markowitz.

Las carteras de varianza mínima global y razón de Sharpe son desarrollos adicionales en la teoría del modelo de media-varianza de Markowitz. La cartera de varianza mínima global es la que tiene el riesgo más bajo de todas las carteras factibles. Es decir, a lo largo del conjunto de carteras óptimas que minimizan el riesgo para un nivel dado de rendimiento esperado (frontera eficiente o frontera de cartera), el punto más a la izquierda corresponde a la cartera con la varianza mínima global en relación con todas las demás carteras posibles. En este caso, el parámetro del modelo es la matriz de varianza-covarianza Σ . Por otro lado, la razón de Sharpe es el cociente entre el rendimiento y el riesgo. Por lo tanto, la cartera con la máxima razón de Sharpe ofrece la mayor rentabilidad esperada por unidad de riesgo. En este caso, los parámetros del modelo son el vector de retorno esperado μ y la matriz de varianza-covarianza Σ .

En este trabajo se propone un método basado en la inferencia bayesiana sesgada, utilizando una densidad normal multivariada con parámetros (μ, Σ) para los rendimientos de los activos y una distribución *a priori* conjugada para (μ, Σ) . Esta propuesta considera diferentes configuraciones de sesgo para estimar μ y Σ , el vector de retornos medios y la matriz de varianza-covarianza involucrados en el modelo de cartera de media-varianza. Se obtienen expresiones matemáticas compactas de forma cerrada para los estimadores nuevos de μ y Σ (sesgadas por la creencia del tomador de decisiones) y se utilizan para estudiar sus propiedades asintóticas. Además, las expresiones para los pesos de algunas carteras también se obtienen en forma cerrada. Finalmente, utilizando diferentes configuraciones de sesgo (estimación de μ y Σ con diferentes parámetros de sesgo), se evalúa el desempeño de las carteras.

Este enfoque muestra que es posible ponderar datos y creencias y generar una cartera óptima cuyo rendimiento y riesgo reflejen estos componentes. Por lo tanto, el método propuesto permite encontrar los pesos óptimos de la cartera, incorporando sesgos en la optimización de la cartera. El análisis asintótico y financiero sobre las consecuencias de incluir sesgos en la selección de carteras ilustra características especiales importantes del rendimiento de estas carteras sesgadas. Por ejemplo, los estimadores sesgados heredan propiedades asintóticas de los estimadores de máxima verosimilitud, como asintóticamente insesgados y eficientes. Además, los niveles de media y varianza tienden a variar un poco más en la cartera sesgada de la razón de Sharpe que en la cartera sesgada de varianza mínima global.

Otra contribución importante de este trabajo es la creación de un método formal dentro del campo de los sesgos de finanzas conductuales para la selección de cartera. Además, la metodología propuesta puede aplicarse a otros modelos financieros y puede implementarse fácilmente en aquellos que han sido tratados desde un enfoque bayesiano como el modelo de Black y Litterman (1992), el modelo de Black y Scholes (1973) y modelos de volatilidad estocástica (Andersen *et al.* 2010, Heston 1993).

El documento está estructurado de la siguiente manera. En la Sección 2 se presenta una clase nueva de estimadores para el modelo de selección de cartera de media-varianza utilizando un enfoque de inferencia bayesiana sesgada, que incorpora la opinión del inversionista sobre las

ponderaciones de los componentes de la distribución posterior bayesiana (verosimilitud y distribuciones *a priori*). El desarrollo matemático para obtener dichos estimadores se describe en el Apéndice A de este documento. Además, el fundamento teórico para el enfoque de inferencia propuesto, así como comportamiento asintótico de los estimadores sesgados se muestra en el Apéndice B. En la Sección 3 se calculan las carteras de media-varianza sesgadas y se muestra, a través de expresiones matemáticas cerradas, cómo las carteras clásicas de media-varianza son modificadas por los sesgos. En la Sección 3 también se describe la implementación computacional de la metodología propuesta y se ilustra, con un ejemplo, el impacto de los sesgos en relación con las carteras clásicas de media-varianza. La Sección 4 presenta un análisis de referencia sobre el rendimiento de las carteras sesgadas para mostrar sus posibles beneficios y costos. Por último, las conclusiones se presentan en la Sección 5.

2. Inferencia Bayesiana Sesgada Para Rendimientos Normales de Activos

En esta sección, se describe el enfoque propuesto para generar estimadores sesgados para el modelo de cartera de media-varianza basado en diferentes niveles de importancia relativa (sesgos) que el tomador de decisiones asigna a los datos y sus creencias. Este enfoque se aplica a la distribución normal multivariada para facilitar la comparación con otras carteras de media-varianza (por ejemplo, la razón de Sharpe y las carteras de varianza mínima global) y para simplificar la interpretación, como se verá en la siguiente sección de este documento. Los estimadores sesgados obtenidos aquí, y que serán utilizados en los modelos de optimización de cartera de la Sección 3, se presentan a través de expresiones matemáticas de forma cerrada.

La inferencia bayesiana sesgada fue introducida por Matsumori *et al.* (2018) para explicar los sesgos de decisión psicológica y los defectos en el control cognitivo. La aplicación de esta metodología también aparece en áreas de sistemas cognitivos e interacción entre factores humanos y sistemas computacionales (Rezaei *et al.* 2022, Rastogi *et al.* 2022). Matsumori *et al.* (2018) introdujeron sesgos exponenciales en la inferencia bayesiana de la siguiente manera:

$$P(H|D) \propto [P(D|H)]^\beta [P(H)]^\alpha,$$

donde $P(H|D)$ es la probabilidad de la hipótesis H después de observar los datos D , o posterior de H , $P(D|H)$ es la probabilidad de D como función de H , o verosimilitud de H , $P(H)$ es la probabilidad *a priori* de la hipótesis H , o *a priori* de H , β es el sesgo exponencial de la verosimilitud y α es el sesgo exponencial correspondiente a la probabilidad *a priori*.

Aquí, se considera una formulación paramétrica de la inferencia bayesiana sesgada y dos tipos de sesgos que afectan la distribución *a priori*. Formalmente, sean $\mathbf{x}_1, \dots, \mathbf{x}_n$ las observaciones muestradas a partir de una densidad de probabilidad $f(\mathbf{x}|\boldsymbol{\theta})$, donde $\boldsymbol{\theta} = (\boldsymbol{\phi}, \xi)$ es una parametrización del modelo. Sea $\pi(\boldsymbol{\phi}, \xi) = \pi(\boldsymbol{\phi}|\xi)\pi(\xi)$ la distribución *a priori* sobre $\boldsymbol{\theta} = (\boldsymbol{\phi}, \xi)$, donde $\pi(\boldsymbol{\phi}|\xi)$ es el *a priori* de $\boldsymbol{\phi}$ condicionado sobre ξ y $\pi(\xi)$ es el *a priori* de ξ . Por lo tanto, se propone un enfoque bayesiano sesgado considerando

$$\pi(\boldsymbol{\phi}, \boldsymbol{\xi} | \boldsymbol{x}) \propto [L_x(\boldsymbol{\phi}, \boldsymbol{\xi})]^\beta [\pi(\boldsymbol{\phi} | \boldsymbol{\xi})]^{\alpha_1} [\pi(\boldsymbol{\xi})]^{\alpha_2} \quad (1)$$

como un modelo general sesgado para actualizar las creencias *a priori* a las creencias posteriores dados los datos. Los parámetros β , y α_2 , representan los sesgos en la verosimilitud $L_x(\boldsymbol{\phi}, \boldsymbol{\xi})$, la *a priori* $\pi(\boldsymbol{\phi} | \boldsymbol{\xi})$ y la *a priori* $\pi(\boldsymbol{\xi})$, respectivamente. Más adelante en el Apéndice B (Apoyo teórico) se justifica teóricamente el uso de verosimilitudes ponderadas y *a priori* potenciadas acomodando la inferencia bayesiana sesgada dentro de un marco de inferencia general que se basa en funciones de pérdida en lugar de funciones de distribución (Bissiri *et al.* 2016).

En general, los parámetros de sesgo α_1 y α_2 en (1) controlan la planitud de la densidad de probabilidad a priori original, que se logra especificando $\alpha_1 = \alpha_2 = 1$. Téngase en cuenta que cuando $\alpha_1 = \alpha_2$ la inferencia es como la inferencia bayesiana sesgada habitual. En particular, si $\alpha_1 = \alpha_2 = 1$ entonces obtenemos la distribución posterior de potencia o posterior atemperada. Del mismo modo, el parámetro de sesgo β controla la planitud de la verosimilitud original obtenida con $\beta = 1$. Ver Matsumori *et al.* (2018) para una descripción más detallada del efecto de los parámetros de sesgo en diferentes factores en la inferencia bayesiana.

Aquí, se asume que los parámetros de sesgo son positivos. Además, supóngase que $f(\boldsymbol{x} | \boldsymbol{\theta})$ es una función de densidad normal multivariada de un vector columna aleatorio d -dimensional $\boldsymbol{X} = (X_1, \dots, X_d)^T$, dada por

$$f(\boldsymbol{x} | \boldsymbol{\theta}) \propto [\det(\boldsymbol{\Upsilon})]^{0.5} e^{-0.5(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Upsilon} (\boldsymbol{x} - \boldsymbol{\mu})},$$

con parámetro $\boldsymbol{\theta} = (\boldsymbol{\mu}, \mathcal{T})$, donde $\boldsymbol{\mu}$ es la media del vector \boldsymbol{X} y $\mathcal{T}^{-1} = \Sigma_d$ es la matriz de varianza-covarianza. La *a priori* se elige de modo que tenga la siguiente forma:

$$\pi(\boldsymbol{\mu}, \mathcal{T}) = \pi(\boldsymbol{\mu} | \mathcal{T}) \pi(\mathcal{T}),$$

donde $\pi(\boldsymbol{\mu} | \mathcal{T})$ es una función de densidad normal multivariada dada por:

$$\pi(\boldsymbol{\mu} | \mathcal{T}) \propto [\det(\mathcal{T})]^{0.5} e^{-0.5\lambda_0(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \mathcal{T} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)},$$

y $\pi(\mathcal{T})$ es una densidad de Wishart definida por

$$\pi(\mathcal{T}) \propto [\det(\mathcal{T})]^{0.5(v_0 - d - 1)} e^{-0.5\text{tr}(\boldsymbol{\Psi}_0^{-1} \mathcal{T})}.$$

Los parámetros $(\boldsymbol{\mu}_0, \lambda_0)$ son los hiperparámetros para la *a priori* $\pi(\boldsymbol{\mu} | \mathcal{T})$. El vector medio $\boldsymbol{\mu}_0$ se utiliza para determinar la ubicación de $\pi(\boldsymbol{\mu} | \mathcal{T})$, mientras que el escalar λ_0 afecta la precisión a su alrededor. Por otro lado, los parámetros $(v_0, \boldsymbol{\Psi}_0)$ son los hiperparámetros para la *a priori* $\pi(\mathcal{T})$. El entero positivo v_0 es el número de grados de libertad. La media a priori de $\pi(\mathcal{T})$ es $v_0 \boldsymbol{\Psi}_0$, donde $\boldsymbol{\Psi}_0$ es una matriz simétrica y no singular. Nótese que una opción para $\boldsymbol{\Psi}_0$ sería $(1/v_0)\boldsymbol{\Sigma}_0^{-1}$, donde $\boldsymbol{\Sigma}_0$ es una suposición *a priori* para la matriz de varianza-covarianza. Ver Zhang (2021) para una explicación más completa de la influencia de las *a priori* Wishart.

La distribución posterior resultante sobre $(\boldsymbol{\mu}, \mathcal{T})$ es una densidad Normal-Wishart,

$$\pi(\boldsymbol{\mu}, \mathcal{T}) \propto \{[\det(\mathcal{T})]^{0.5} e^{-0.5\lambda_p(\boldsymbol{\mu}-\boldsymbol{\mu}_p)^T \mathcal{T}(\boldsymbol{\mu}-\boldsymbol{\mu}_p)}\} \{[\det(\mathcal{T})]^{0.5v_p} e^{-0.5\text{tr}(\Psi_p^{-1}\mathcal{T})}\},$$

con parámetros

$$\begin{aligned}\boldsymbol{\mu}_p &= \frac{\lambda_0\alpha_1\boldsymbol{\mu}_0 + n\beta\hat{\boldsymbol{\mu}}}{\lambda_0\alpha_1 + n\beta}, \\ \lambda_p &= \lambda_0\alpha_1 + n\beta, \\ v_p &= n\beta + (\alpha_1 - \alpha_2) + \alpha_2v_0 + d(1 - \alpha_2), \\ \Psi_p &= \left[\alpha_2\Psi_0^{-1} + n\beta\hat{\Sigma}_d + \frac{\lambda_0\alpha_1n\beta}{\lambda_0\alpha_1 + n\beta}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T \right]^{-1},\end{aligned}$$

donde $\hat{\boldsymbol{\mu}}$ y $\hat{\Sigma}_d$ son las estimaciones de máxima verosimilitud (MLE) de $\boldsymbol{\mu}$ y Σ_d , respectivamente, dadas por

$$\hat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad \text{and} \quad \hat{\Sigma}_d = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T.$$

El desarrollo matemático para apoyar estos resultados se describe en el Apéndice A, al final de este documento. Se propone estimar $\boldsymbol{\mu}$ utilizando la estimación de localización posterior dada por

$$\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_p = \frac{\lambda_0\alpha_1\boldsymbol{\mu}_0 + n\beta\hat{\boldsymbol{\mu}}}{\lambda_0\alpha_1 + n\beta}.$$

Por otro lado, si $\mathcal{T} = \Sigma_d^{-1}$ tiene una distribución Wishart con parámetros (v_p, Ψ_p) , entonces $\mathcal{T}^{-1} = \Sigma_d$ tiene una distribución inversa Wishart con parámetros (v_p, Ψ_p^{-1}) . Así, proponemos como estimador para Σ la media de esta distribución:

$$\tilde{\Sigma}_d = \frac{\Psi_p^{-1}}{v_p - d - 1} = \frac{\alpha_2\Psi_0^{-1} + n\beta\hat{\Sigma}_d + \frac{\lambda_0\alpha_1n\beta}{\lambda_0\alpha_1 + n\beta}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T}{n\beta + (\alpha_1 - \alpha_2) + \alpha_2(v_0 - d) - 1}.$$

En la siguiente sección, se utilizarán estos estimadores sesgados para construir carteras sesgadas de media-varianza.

3. Carteras Sesgadas de Media-Varianza

Un hito en la teoría moderna de carteras financieras es el análisis de media-varianza introducido por Markowitz (1952). En su teoría, una cartera eficiente es aquella que ofrece el riesgo mínimo para un valor esperado de rendimiento. Es decir, en la frontera eficiente de Markowitz, cada cartera minimiza

el riesgo para un rendimiento determinado. Por lo tanto, el inversionista puede seleccionar su cartera óptima en función de su nivel de aversión al riesgo en la frontera eficiente.

La cartera de varianza mínima global (VMG) y la cartera óptima de razón de Sharpe (RS) son dos de las carteras eficientes de media-varianza más utilizadas. Ambas carteras desempeñan un papel importante en la asignación de activos; la cartera de VMG corresponde al inversionista que es completamente adverso al riesgo, mientras que la cartera de RS corresponde al inversionista que mide la rentabilidad en relación con el riesgo. En esta sección se utilizan los estimadores sesgados obtenidos en la sección anterior (para la rentabilidad esperada de los activos y la matriz de varianza-covarianza) para calcular expresiones matemáticas de forma cerrada para las ponderaciones y características de carteras eficientes de media-varianza tales como la media y varianza para las carteras VMG y RS. Además, se presentan implementaciones computacionales de estos métodos para la selección sesgada de carteras.

3.1 Modelo Sesgado de Media-Varianza de Markowitz

Supóngase que un inversionista posee una cartera de d activos. Sea $\mathbf{X}_i = (X_{i1}, \dots, X_{id})^T$ con $i = 1, \dots, n$ una muestra aleatoria de los rendimientos de estos activos, y supóngase que $\mathbf{X}_1, \dots, \mathbf{X}_n$ son vectores que siguen una distribución normal multivariada con función de densidad conjunta $f(\mathbf{x}; \boldsymbol{\mu}, \Sigma_d)$, donde $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)$ es el vector de parámetros que representan los rendimientos esperados de los activos de la cartera y Σ_d es la matriz de varianza-covarianza correspondiente. Bajo nuestro enfoque bayesiano sesgado generalizado, dada una muestra observada $\mathbf{x}_1, \dots, \mathbf{x}_n$ de $\mathbf{X}_1, \dots, \mathbf{X}_n$, se calculan las estimaciones sesgadas $\tilde{\boldsymbol{\mu}}$ y $\tilde{\Sigma}_d$ de los parámetros $\boldsymbol{\mu}$ y Σ_d , como fueron descritos en la Sección 2.

En el marco de cartera de media-varianza sesgada, se puede obtener una cartera óptima resolviendo el siguiente problema de optimización:

$$\min_{\mathbf{w}} \mathbf{w}^T \tilde{\Sigma}_d \mathbf{w} \text{ sujeto a } \mathbf{w}^T \tilde{\boldsymbol{\mu}} = \mu^*, \mathbf{1}^T \mathbf{w} = 1, \text{ y } w_i \geq 0 \text{ con } i = 1, \dots, d,$$

donde $\mathbf{w} = (w_1, \dots, w_d)^T$ es un vector de columna de ponderaciones de cartera, μ^* es un nivel particular de rendimiento deseado y $\mathbf{1} = (1, \dots, 1)^T$ es un vector columna cuya dimensión es la misma que la dimensión de \mathbf{w} . La frontera eficiente se obtiene utilizando ponderaciones de cartera para asignaciones de activos que minimizan el riesgo $\mathbf{w}^T \tilde{\Sigma}_d \mathbf{w}$ para cualquier nivel de rendimiento $\mathbf{w}^T \tilde{\boldsymbol{\mu}}$. En este caso, la cartera óptima $\tilde{\mathbf{w}}^*$ correspondiente a μ^* es simplemente un punto de entre el número infinito de puntos en la frontera eficiente, que se obtiene numéricamente. Para un inversionista adverso al riesgo, las ponderaciones sesgadas de la cartera de utilidad esperada se eligen resolviendo el siguiente problema:

$$\text{Max}_{\mathbf{w}} \mathbf{w}^T \tilde{\boldsymbol{\mu}} - \left(\frac{\delta}{2}\right) \mathbf{w}^T \tilde{\Sigma}_d \mathbf{w} \text{ sujeto a } \mathbf{1}^T \mathbf{w} = 1,$$

donde $\delta > 0$ es el parámetro de aversión al riesgo. Así, de forma análoga a Font (2016), el vector de ponderaciones óptimas de la cartera correspondiente a la cartera de utilidad esperada sesgada viene dado por

$$\tilde{\mathbf{w}}_{EU} = \frac{\tilde{\Sigma}_d^{-1} \mathbf{1}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}} + \frac{1}{\delta} \left(\tilde{\Sigma}_d^{-1} - \frac{\tilde{\Sigma}_d^{-1} \mathbf{1} \mathbf{1}^T \tilde{\Sigma}_d^{-1}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}} \right) \tilde{\boldsymbol{\mu}}.$$

El límite de $\tilde{\mathbf{w}}_{EU}$, a medida que δ se acerca al infinito, conduce a ponderaciones sesgadas de la cartera de VMG,

$$\tilde{\mathbf{w}}_{GMV} = \frac{\tilde{\Sigma}_d^{-1} \mathbf{1}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}}.$$

La media muestral y la varianza de esta cartera vienen dadas por

$$\tilde{R}_{GMV} = \frac{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}} \quad \text{y} \quad \tilde{V}_{GMV} = \frac{1}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}},$$

respectivamente. La frontera eficiente sesgada en el espacio de la media (R) y la varianza (V) es la parte superior de la parábola dada por

$$(R - \tilde{R}_{GMV})^2 = \left[\tilde{\boldsymbol{\mu}}^T \left(\tilde{\Sigma}_d^{-1} - \frac{\tilde{\Sigma}_d^{-1} \mathbf{1} \mathbf{1}^T \tilde{\Sigma}_d^{-1}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \mathbf{1}} \right) \tilde{\boldsymbol{\mu}} \right] (V - \tilde{V}_{GMV}).$$

Otra cartera que se encuentra en la frontera eficiente de media-varianza sesgada es la cartera de tangencia, que maximiza la razón de Sharpe sesgada,

$$\max_{\mathbf{w}} \frac{\mathbf{w}^T \tilde{\boldsymbol{\mu}}}{\sqrt{\mathbf{w}^T \tilde{\Sigma}_d \mathbf{w}}} \quad \text{sujeto a} \quad \mathbf{1}^T \mathbf{w} = 1.$$

Las ponderaciones y características de la cartera de RS sesgada vienen dadas por

$$\tilde{\mathbf{w}}_{SR} = \frac{\tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}, \quad \tilde{R}_{SR} = \frac{\tilde{\boldsymbol{\mu}}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}{\mathbf{1}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}, \quad \text{y} \quad \tilde{V}_{SR} = \frac{\tilde{\boldsymbol{\mu}}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}}}{(\mathbf{1}^T \tilde{\Sigma}_d^{-1} \tilde{\boldsymbol{\mu}})^2}.$$

Se concluye esta subsección proporcionando una expresión matemática para $\tilde{\Sigma}_d^{-1}$, que puede ser útil para establecer la relación entre las características de la cartera sesgada y los componentes de la inferencia sesgada. Sea $a_0 = (v_p - d - 1)/n\beta$, $b_0 = \alpha_2/n\beta$, $c_0 =$

$\lambda_0\alpha_1/(\lambda_0\alpha_1 + n\beta)$ y $\mathbf{z} = (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)$. Entonces, usando la identidad de la matriz de Woodbury y la definición de $\tilde{\Sigma}_d$, se puede reescribir $\tilde{\Sigma}_d^{-1}$ de la siguiente manera:

$$\begin{aligned}\tilde{\Sigma}_d^{-1} &= a_0(\hat{\Sigma}_d + b_0\Psi_0^{-1} + c_0\mathbf{z}\mathbf{z}^T)^{-1} \\ &= a_0\left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d(b_0\Psi_0^{-1} + c_0\mathbf{z}\mathbf{z}^T)^{-1}\hat{\Sigma}_d]^{-1}\right\} \\ &= a_0\left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\right\},\end{aligned}$$

Donde

$$\Gamma = b_0^{-1}\Psi_0 - \frac{b_0^{-2}c_0\Psi_0\mathbf{z}\mathbf{z}^T\Psi_0}{1 + b_0^{-1}c_0\mathbf{z}^T\Psi_0\mathbf{z}}.$$

Así, por ejemplo, el vector de ponderaciones óptimas para una cartera de varianza mínima global sesgada se puede reescribir como:

$$\begin{aligned}\tilde{\mathbf{w}}_{GMV} &= \frac{\left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\right\} \mathbf{1}}{\mathbf{1}^T \left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\right\} \mathbf{1}} \\ &= \frac{\hat{\Sigma}_d^{-1}\mathbf{1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\mathbf{1}}{\mathbf{1}^T \left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\right\} \mathbf{1}} \\ &= \frac{\frac{\hat{\Sigma}_d^{-1}\mathbf{1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\mathbf{1}}{\mathbf{1}^T \hat{\Sigma}_d^{-1}\mathbf{1}}}{\mathbf{1}^T \left\{\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\right\} \mathbf{1}} \\ &= \frac{\mathbf{w}_{GMV} - \check{\mathbf{w}}_{GMV}}{1 - \check{s}_{GMV}},\end{aligned}$$

Donde

$$\mathbf{w}_{GMV} = \frac{\hat{\Sigma}_d^{-1}\mathbf{1}}{\mathbf{1}^T \hat{\Sigma}_d^{-1}\mathbf{1}}, \quad \check{\mathbf{w}}_{GMV} = \frac{[\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\mathbf{1}}{\mathbf{1}^T [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\mathbf{1}}, \quad \text{y} \quad \check{s}_{GMV} = \frac{\mathbf{1}^T [\hat{\Sigma}_d + \hat{\Sigma}_d\Gamma\hat{\Sigma}_d]^{-1}\mathbf{1}}{\mathbf{1}^T \hat{\Sigma}_d^{-1}\mathbf{1}}.$$

Observe que \mathbf{w}_{GMV} es el vector de ponderaciones de la cartera óptima VMG calculado utilizando la matriz de varianza-covarianza de la muestra $\hat{\Sigma}_d$, $\check{\mathbf{w}}_{GMV}$ es un vector de

ponderaciones no estandarizadas que depende de los componentes de la inferencia sesgada a través de Γ , y \check{s}_{GMV} es una constante de normalización.

Siguiendo el mismo procedimiento, las ponderaciones para la cartera de RS sesgada óptima se pueden expresar como:

$$\tilde{w}_{SR} = \frac{w_{SR} - \check{w}_{SR} - \bar{w}_{SR}}{1 - \check{s}_{SR} - \bar{s}_{SR}},$$

Donde

$$w_{SR} = \frac{\hat{\Sigma}_d^{-1} \hat{\mu}}{\mathbf{1}^T \hat{\Sigma}_d^{-1} \hat{\mu}}, \quad \check{w}_{SR} = \frac{[\hat{\Sigma}_d + \hat{\Sigma}_d \Gamma \hat{\Sigma}_d]^{-1} \hat{\mu}}{\mathbf{1}^T \hat{\Sigma}_d^{-1} \hat{\mu}},$$

$$\bar{w}_{SR} = \frac{c_0 \left(\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d \Gamma \hat{\Sigma}_d]^{-1} \right) (\hat{\mu} - \mu_0)}{\mathbf{1}^T \hat{\Sigma}_d^{-1} \hat{\mu}},$$

$$\check{s}_{SR} = \frac{\mathbf{1}^T [\hat{\Sigma}_d + \hat{\Sigma}_d \Gamma \hat{\Sigma}_d]^{-1} \hat{\mu}}{\mathbf{1}^T \hat{\Sigma}_d^{-1} \hat{\mu}} \quad \text{y} \quad \bar{s}_{SR} = \frac{c_0 \mathbf{1}^T \left(\hat{\Sigma}_d^{-1} - [\hat{\Sigma}_d + \hat{\Sigma}_d \Gamma \hat{\Sigma}_d]^{-1} \right) (\hat{\mu} - \mu_0)}{\mathbf{1}^T \hat{\Sigma}_d^{-1} \hat{\mu}}.$$

3.2 Implementación Computacional

Las expresiones matemáticas de forma cerrada presentadas en esta sección son implementadas utilizando un software libre para computación estadística y gráficos llamado R. Como ejemplo, a continuación, se proporciona una configuración de parámetros para el modelo de selección de cartera de media-varianza sesgada. Por otro lado, el paquete NMOF de R se puede utilizar para implementar una amplia variedad de carteras de media-varianza sesgadas; véase Gilli *et. al.* (2019) y Schumann (2011–2023). NMOF tiene funciones que calculan rendimientos, volatilidades y composiciones para carteras a lo largo de una frontera eficiente, utilizando sólo la media y la matriz de varianza-covarianza (estimada) como entradas.

Ejemplo: Para el enfoque de modelado de inferencia bayesiana sesgada, los elementos seleccionados para este ejemplo son:

$$n = 104, d = 3, \hat{\mu} = \begin{pmatrix} 0.0012 \\ 0.0054 \\ 0.0004 \end{pmatrix}, \hat{\Sigma}_d = \begin{bmatrix} 0.0003357 & 0.0000926 & -0.0000095 \\ 0.0000926 & 0.0035848 & -0.0000595 \\ -0.0000095 & -0.0000595 & 0.0000611 \end{bmatrix},$$

$$\lambda_0 = n, v_0 = d, \mu_0 = \begin{pmatrix} 0.0012 \\ 0.0200 \\ 0.0004 \end{pmatrix},$$

$$\Psi_0 = \frac{1}{v_0} \begin{bmatrix} 0.0003357 & -0.0000522 & 0.0000160 \\ -0.0000522 & 0.0065156 & -0.0000689 \\ 0.0000160 & -0.0000689 & 0.0000611 \end{bmatrix}^{-1},$$

$$\beta = 1, \alpha_1 = 2, \alpha_2 = 1.$$

Los valores para n , d , $\hat{\mu}$, y $\hat{\Sigma}_d$ fueron calculados utilizando precios en dólares de la página de *yahoo finance* convertidos a rendimientos semanales del Promedio Industrial *Dow Jones* (DJIA), *Bitcoin* (BTC), e *iShares iBoxx* (LQD), por medio del paquete de R, *quantmod*, usando *getSymbols* para el período 2015-01-01 a 2017-01-01 (104 semanas). Con fines ilustrativos, los parámetros para las distribuciones *a priori* se especificaron en función de los rendimientos semanales que se obtuvieron de 104 semanas adicionales. Los parámetros de sesgo se seleccionaron para ilustrar una posible configuración, donde la distribución *a priori* de μ condicionada a $\mathcal{T} = \Sigma_d^{-1}$ tiene un parámetro de sesgo ($\alpha_1 = 2$) mayor que el asignado a la *a priori* de Σ_d ($\alpha_2 = 1$) y a la verosimilitud ($\beta = 1$).

El conjunto de datos arroja los siguientes resultados:

$$\tilde{\mu} = \begin{pmatrix} 0.0012 \\ 0.0151 \\ 0.0004 \end{pmatrix} \text{ y } \tilde{\Sigma}_d = \begin{bmatrix} 0.0003454 & 0.0000911 & -0.0000090 \\ 0.0000911 & 0.0039149 & -0.0000615 \\ -0.0000090 & -0.0000615 & 0.0000629 \end{bmatrix}.$$

Entonces,

$$\tilde{w}_{GMV} = \begin{pmatrix} 0.1572 \\ 0.0217 \\ 0.8211 \end{pmatrix}, \tilde{R}_{GMV} = 0.0008457, \tilde{V}_{GMV} = 0.0000489,$$

$$\tilde{w}_{SR} = \begin{pmatrix} 0.1563 \\ 0.2293 \\ 0.6144 \end{pmatrix}, \tilde{R}_{SR} = 0.0039038, \tilde{V}_{SR} = 0.0002255.$$

La frontera eficiente sesgada y clásica utilizando $(\tilde{\mu}, \tilde{\Sigma}_d)$ y $(\hat{\mu}, \hat{\Sigma}_d)$, respectivamente, se puede observar en la Figura 1. En esta figura, la media y la varianza de las carteras óptimas (carteras VMG y RS) están marcadas en la gráfica de la frontera eficiente correspondiente. Las fronteras eficientes convergen en el punto de menor riesgo, sin embargo, a medida que éste aumenta, la frontera eficiente sesgada tiene un rendimiento esperado cada vez mayor al de la frontera clásica por cada unidad de riesgo. Lo anterior indica que la frontera eficiente sesgada refleja un menor apetito por el riesgo comparada con la frontera clásica.

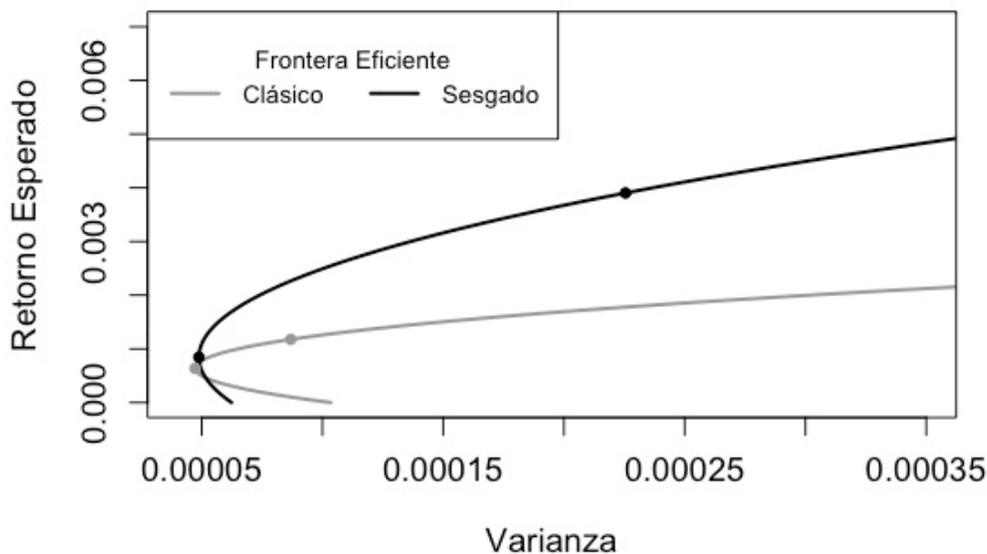


Figura 1. La frontera eficiente sesgada y clásica obtenida con base en los estimadores $(\hat{\mu}, \hat{\Sigma}_d)$ y $(\tilde{\mu}, \tilde{\Sigma}_d)$.

4. Análisis de Referencia de Desempeño Para la Selección de Carteras

En esta sección se analiza el impacto de los sesgos en relación con las carteras clásicas de media-varianza. Las cuestiones consideradas en este análisis se describen a continuación:

- El estado actual y futuro de los precios (una versión estandarizada de los precios ajustados por dividendos y divisiones) de las tres ($d = 3$) variables financieras de interés (DJIA, BTC y LQD) se muestran en la Figura 2(a). Los datos del estado actual (rendimientos semanales durante las primeras 104 semanas) se utilizarán para estimar el rendimiento semanal esperado y la matriz de varianza-covarianza. Los datos de estado futuro (rendimientos semanales para las próximas 104 semanas) se utilizarán para cuantificar el impacto de los sesgos en el rendimiento semanal futuro de las carteras óptimas. En la Figura 2(b) se presentan gráficas y estadísticas descriptivas (histogramas, dispersión y correlación) de estos datos. Se observa que DJIA, BTC y LQD exhiben distribuciones de frecuencias localizadas cerca del valor cero, con medias muestrales 0.001154, 0.005316, 0.000454 y desviaciones estándar 0.018259, 0.059587, 0.007809, respectivamente, donde la distribución correspondiente a BTC es la que presenta el mayor rango y dispersión. La correlación negativa más alta en magnitud se presenta

entre las variables BTC y LQD, con un valor de -0.13, sugiriendo que a medida que aumenta una la otra tiende a disminuir. Por otro lado, la prueba de normalidad multivariada de Henze y Zirkler (1990) proporciona una estadística de prueba $HZ=1.0442$ y un correspondiente p-valor de 0.0267. Con base en esta evidencia se supondrá una normal multivariada para los rendimientos semanales de DJIA, BTC y LQD durante el estado actual

- Usando los rendimientos semanales del estado actual ($n = 104$ semanas), obtenidos a partir de los precios mostrados en la Figura 2(a), las estimaciones del rendimiento semanal esperado $\hat{\mu}$ y la matriz de varianza-covarianza $\hat{\Sigma}_d$ son los presentados en el Ejemplo de la Sección 3. Además, los parámetros $\lambda_0, \nu_0, \mu_0, \Psi_0$ para distribuciones *a priori* son las presentadas en ese ejemplo. Por otro lado, los valores para los parámetros de sesgo para el análisis son $\beta \in [0.5, 2]$, $\alpha \in [0.5, 2]$ y $\alpha_1 = \alpha_2 = \alpha$. Es decir, se presenta una cuadrícula para β y α que va de $[0.5, 2]$ a $[0.5, 2]$.
- Se consideran las carteras óptimas sesgadas basadas en modelos de varianza mínima global y razón de Sharpe. Interesa analizar el rendimiento esperado y la varianza de estas carteras sesgadas en función de los datos del estado actual. Además, utilizando las ponderaciones estimadas de estas carteras, se calcularon y estudiaron los rendimientos semanales acumulados para el estado futuro.



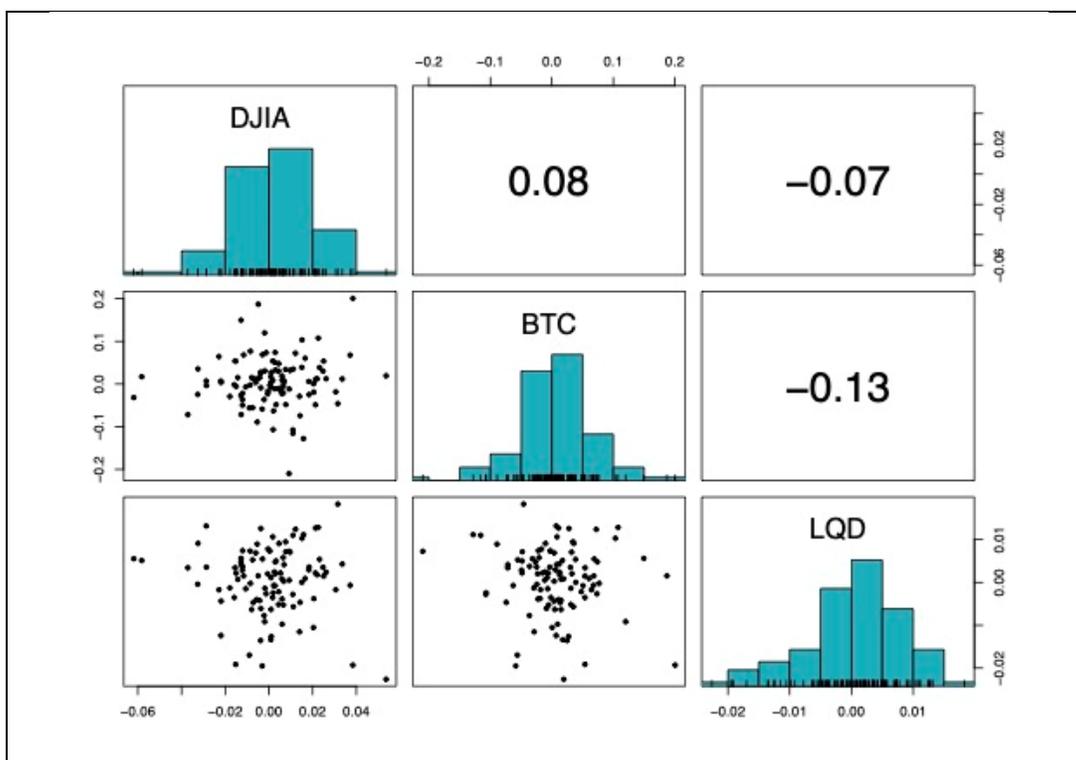


Figura 2. (a) Comportamiento de los precios durante el estado actual y futuro, empleados para calcular rendimientos semanales utilizados en este estudio. (b) Gráficas y estadísticas descriptivas de los rendimientos semanales obtenidos de las variables financieras de interés durante el estado actual.

La Figura 3 muestra gráficos de imagen de superficies correspondientes a la media y varianza obtenida mediante el uso de carteras óptimas sesgadas con $\beta \in [0.5, 2]$, $\alpha \in [0.5, 2]$ y $\alpha_1 = \alpha_2 = \alpha$. Los gráficos de imagen para carteras óptimas sesgadas basadas en el modelo de varianza mínima global se presentan en la Figura 3(a)-(b) mientras que los correspondientes a la razón de Sharpe se muestran en la Figura 3(c)-(d). Se observa una tendencia similar en todas las gráficas, con la media y la varianza de la cartera óptima sesgada aumentando a medida que β decrece y α se incrementa. Sin embargo, los cambios en los respectivos niveles de la media (retorno esperado) y la varianza, calculados utilizando el modelo de varianza mínima global, no parecen variar mucho; ver los valores en la escala de colores de la Figura 3(a)-(b). Además, el sesgo α no parece afectar a la varianza cuando los valores de β son mayores que 1. Los niveles de media y varianza parecen variar un poco más en el modelo de la razón de Sharpe; ver los valores en la escala de colores de la Figura 3(c)-(d). La Figura 4 muestra gráficos de imagen de las superficies correspondientes al rendimiento acumulado semanal para el estado futuro de las carteras óptimas en estudio (varianza mínima global y carteras de la razón de Sharpe). El gráfico de imagen para carteras óptimas sesgadas basadas en el modelo de varianza mínima global se muestra en la Figura 4 (a), mientras que el gráfico correspondiente a la razón de Sharpe se muestra en la Figura 4 (b).

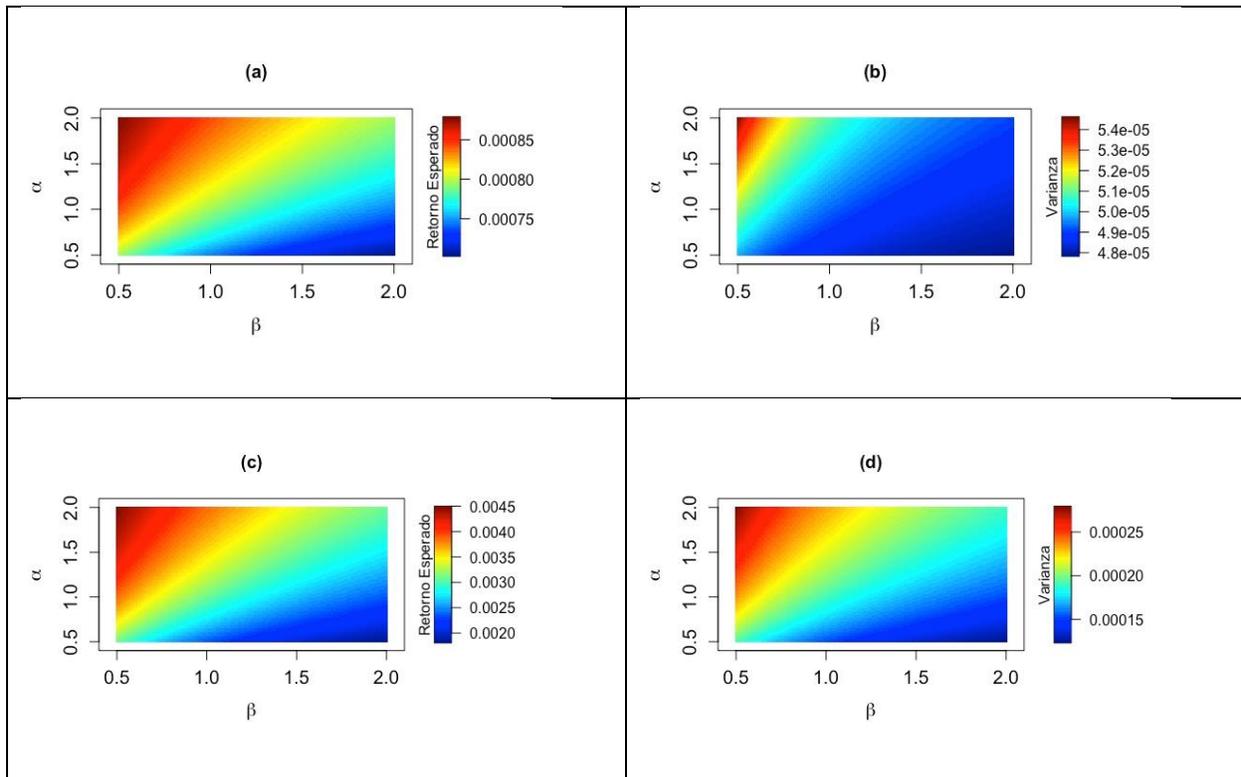


Figura 3. Superficies correspondientes a la media y varianza obtenidas mediante el uso de carteras óptimas sesgadas basadas en modelos de varianza mínima global y razón de Sharpe, con $\beta \in [0.5, 2]$, $\alpha \in [0.5, 2]$, y $\alpha_1 = \alpha_2 = \alpha$. (a) Rendimientos esperados correspondientes a carteras de varianza mínima global óptimas sesgadas; (b) Varianza de carteras de varianza mínima global óptimas sesgadas; (c) Rendimientos esperados correspondientes a carteras sesgadas de razón de Sharpe óptimas; (d) Varianza de las carteras sesgadas de razón de Sharpe óptimas.

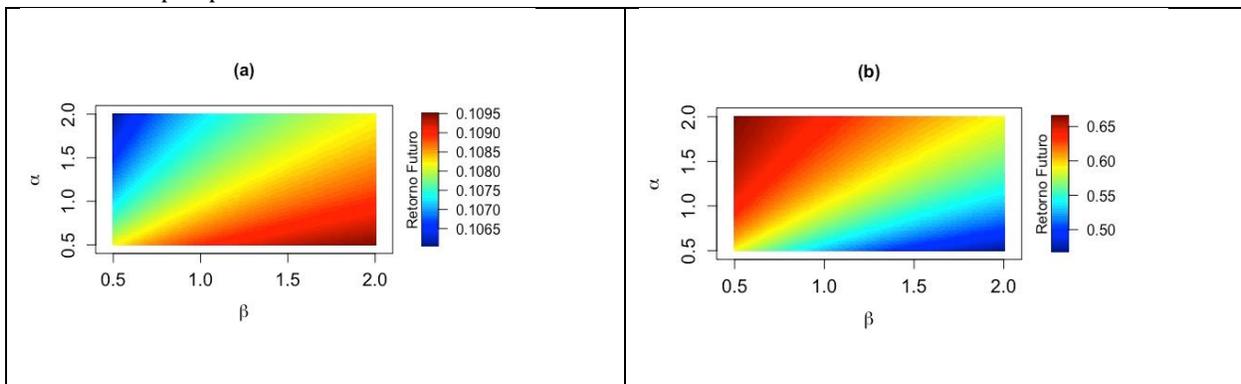


Figura 4. Superficies correspondientes al rendimiento acumulado semanal obtenido mediante el uso de carteras óptimas sesgadas basadas en modelos de varianza mínima global y razón de Sharpe, con $\beta \in [0.5, 2]$, $\alpha \in [0.5, 2]$, y $\alpha_1 = \alpha_2 = \alpha$. (a) Rendimiento acumulado semanal para el estado futuro correspondiente a las carteras de varianza mínima global óptima sesgada; (b) Rendimiento acumulado semanal para el estado futuro correspondiente a las carteras sesgadas de razón de Sharpe.

En la Figura 4(a), el rendimiento acumulado semanal más alto para el estado futuro se obtiene aumentando el valor de β y disminuyendo el valor de α , mientras que en la Figura 4(b) ocurre lo contrario. Además, los cambios en los niveles de rendimiento acumulativo semanal (retorno futuro), calculados utilizando el modelo de varianza mínima global, no parecen variar mucho. Sin embargo, estos niveles parecen variar un poco más en el modelo de razón de Sharpe. Esto ya se había observado para la media y la varianza de ambos modelos.

5. Conclusiones

Este artículo presenta un esfuerzo de investigación centrado en el desarrollo de carteras financieras óptimas mediante la incorporación de inferencia bayesiana y sesgos de los inversionistas en el proceso de construcción de carteras. El estudio utiliza la metodología de inferencia bayesiana sesgada en el proceso de selección de carteras de media-varianza, lo que permite diferentes configuraciones de sesgo o ponderación sobre el modelo estadístico y las creencias a priori. El enfoque propuesto demuestra eficacia en la identificación de carteras óptimas que reflejen la interacción entre las probabilidades basadas en datos y las creencias subjetivas de los inversionistas.

La metodología se aplica específicamente a las carteras de media-varianza utilizando una densidad normal multivariada para la rentabilidad de los activos (parametrizada a través de μ y Σ) y una distribución a priori conjugada para (μ, Σ) . El estudio asegura propiedades asintóticas positivas de los nuevos estimadores de μ y Σ , garantizando consistencia y eficiencia. La inclusión de sesgos en el proceso de estimación se facilita al considerar varias configuraciones de sesgo para las creencias a priori sobre μ y Σ . Además, las expresiones matemáticas de forma cerrada para las ponderaciones de la cartera se derivan para determinados tipos de carteras.

El rendimiento de la cartera se evalúa rigurosamente bajo diferentes configuraciones de sesgos, lo que pone de manifiesto la relevancia de incorporar sesgos en la selección de carteras para su optimización. La investigación contribuye al campo de los sesgos de las finanzas conductuales en la selección de carteras y subraya la adaptabilidad de la metodología propuesta a otros modelos financieros tratados desde una perspectiva bayesiana. Los resultados sugieren que el enfoque puede producir ponderaciones óptimas de la cartera que armonicen tanto los datos objetivos como las creencias subjetivas de los inversionistas. En general, se ha demostrado que la inclusión de sesgos en la optimización de carteras es valiosa para permitir a los inversionistas construir carteras alineadas con sus preferencias de riesgo y objetivos de inversión.

Declaración

Los autores informan que no hay conflictos de intereses que declarar.

Referencias

- [1] Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2010). Parametric and Nonparametric Volatility Measurement. T.G. Andersen. *Handbook of Financial Econometrics: Tools and Techniques* (pp. 67-137). San Diego, United States: North-Holland. <https://doi.org/10.1016/b978-0-444-50897-3.50005-5>
- [2] Bauder, D., Bodnar, T., Parolya, N. and Schmid, W. (2021). Bayesian mean–variance analysis: optimal portfolio selection under parameter uncertainty. *Quantitative Finance*, **21**(2), 221–242. <https://doi.org/10.1080/14697688.2020.1748214>
- [3] Bissiri, P.G., Holmes, C.C. and Walker, S. G. (2016). A General Framework for Updating Belief Distributions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **78**(5), 1103–1130. <https://doi.org/10.1111/rssb.12158>
- [4] Black, F. and Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, **48**(5), 28–43. <https://doi.org/10.2469/faj.v48.n5.28>
- [5] Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**(3), 637–654. <https://doi.org/10.1086/260062>
- [6] Bodnar, T., Lindholm, M., Niklasson, V. and Thorsén, E. (2022). Bayesian portfolio selection using VaR and CVaR. *Applied Mathematics and Computation*, **427**, 127120. <https://doi.org/10.1016/j.amc.2022.127120>
- [7] De Bondt, W., Mayoral, R.M. and Vallelado, E. (2013). Behavioral decision-making in finance: An overview and assessment of selected research. *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, **42**(157), 99–118. <https://doi.org/10.1080/02102412.2013.10779742>
- [8] De Franco, C., Nicolle, J. and Pham, H. (2019). Bayesian learning for the Markowitz portfolio selection problem. *International Journal of Theoretical and Applied Finance*, **22**(7), 1950037. <https://doi.org/10.1142/s0219024919500377>
- [9] Font, B. (2016). Bootstrap estimation of the efficient frontier. *Computational Management Science*, **13**, 541–570. <https://doi.org/10.1007/s10287-016-0257-2>
- [10] Gilli, M., Maringer, D., and Schumann, E. (2019). Numerical Methods and Optimization in Finance, Second edition. Elsevier/Academic Press, Waltham, MA, USA. ISBN 978-0128150658, <http://www.enricoschumann.net/NMOF/>.
- [11] Harvey, C.R., Liechty, J.C., Liechty, M.W. and Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, **10**(5), 469–485. <https://doi.org/10.1080/14697681003756877>
- [12] Henze, N., and Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in statistics-Theory and Methods*, **19**(10), 3595–3617. <https://doi.org/10.1080/03610929008830400>
- [13] Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, **6**(2), 327–343. <https://doi.org/10.1093/rfs/6.2.327>
- [14] Kolm, P.N., Ritter, G. and Simonian, J. (2021). Black–Litterman and Beyond: The Bayesian Paradigm in Investment Management. *The Journal of Portfolio Management*, **47**(5), 91–113. <https://doi.org/10.3905/jpm.2021.1.222>

- [15] Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, **7**, 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- [16] Matsumori, K., Koike, Y. and Matsumoto, K. (2018). A Biased Bayesian Inference for Decision-Making and Cognitive Control. *Frontiers in Neuroscience*, **12**, 734. <https://doi.org/10.3389/fnins.2018.00734>
- [17] Moosa, I.A. and Ramiah, V. (2017). *The Financial Consequences of Behavioural Biases*. London, England: Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-69389-7>
- [18] Pástor, L. (2000). Portfolio Selection and Asset Pricing Models. *The Journal of Finance*, **55**(1), 179–223. <https://doi.org/10.1111/0022-1082.00204>
- [19] Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., & Tomsett, R. (2022). Deciding Fast and Slow: The Role of Cognitive Biases in AI-assisted Decision-making. *Proceedings of the ACM on human-computer interaction*, **6**(CSCW1), 1-22. <https://doi.org/10.1145/3512930>
- [20] Rezaei, Z., Setayeshi, S., & Mahdipour, E. (2022). A Biased Inferential Naivety learning model for a network of agents. *Cognitive Systems Research*, **76**, 1-12. <https://doi.org/10.1016/j.cogsys.2022.07.001>
- [21] Sahamkhadam, M., Stephan, A. and Östermark, R. (2022). Copula-based Black–Litterman portfolio optimization. *European Journal of Operational Research*, **297**(3), 1055–1070. <https://doi.org/10.1016/j.ejor.2021.06.015>
- [22] Schumann, E. (2011–2023). Numerical Methods and Optimization in Finance (NMOF) Manual. Package version 2.8-0). <http://enricoschumann.net/NMOF/>.
- [23] Statman, M. (1999). Behavioral finance: Past Battles and Future Engagements. *Financial Analysts Journal*, **55**(6), 18–27. <https://doi.org/10.2469/faj.v55.n6.2311>
- [24] Zhang, Z. (2021). A Note on Wishart and Inverse Wishart Priors for Covariance Matrix. *Journal of Behavioral Data Science*, **1**(2), 119–126. <https://doi.org/10.35566/jbds/v1n2/p2>

Apéndice A

Si $\mathbf{X} = (X_1, X_2, \dots, X_d)$ es un vector aleatorio distribuido normal multivariado con parámetros $(\boldsymbol{\mu}, \mathcal{T} = \Sigma_d^{-1})$, entonces la verosimilitud de $(\boldsymbol{\mu}, \mathcal{T})$ viene dada por

$$\begin{aligned} L(\boldsymbol{\mu}, \mathcal{T}) &\propto \left\{ \exp \left[-\frac{n}{2} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^T \mathcal{T} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \right] [\det(\mathcal{T})]^{\frac{n}{2}} \exp \left[-\frac{\text{tr}(n\hat{\Sigma}_d \mathcal{T})}{2} \right] \right\}^{\beta} \\ &= \exp \left[-\frac{n\beta}{2} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^T \mathcal{T} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \right] [\det(\mathcal{T})]^{n\beta/2} \exp \left[-\frac{\text{tr}(n\beta\hat{\Sigma}_d \mathcal{T})}{2} \right]. \end{aligned}$$

Asumimos una densidad *a priori* para $(\boldsymbol{\mu}, \mathcal{T})$ dada por

$$\pi(\boldsymbol{\mu}, \mathcal{T}) = \pi(\boldsymbol{\mu}|\mathcal{T})\pi(\mathcal{T}),$$

donde $\pi(\boldsymbol{\mu}|\mathcal{T})$, la distribución de los rendimientos esperados condicionada a \mathcal{T} , es una densidad normal multivariada *a priori* para $\boldsymbol{\mu}$ con parámetros $(\boldsymbol{\mu}_0, \lambda_0 \mathcal{T})$ y $\pi(\mathcal{T})$ es una

densidad Wishart *a priori* para \mathcal{T} con parámetros (ν_0, Ψ_0) . Entonces, la distribución *a priori* para $(\boldsymbol{\mu}, \mathcal{T})$ toma la forma

$$\begin{aligned} \pi(\boldsymbol{\mu}, \mathcal{T}) &\propto \left\{ [\det(\mathcal{T})]^{\frac{1}{2}} \exp \left[-\frac{\lambda_0}{2} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \mathcal{T} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \right] \right. \\ &\quad \left. [\det(\mathcal{T})]^{\frac{(\nu_0-d-1)}{2}} \exp \left[-\frac{1}{2} \text{tr}(\Psi_0^{-1} \mathcal{T}) \right] \right\}^\alpha \\ &= [\det(\mathcal{T})]^{\frac{\alpha}{2}} \exp \left[-\frac{\lambda_0 \alpha}{2} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \mathcal{T} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \right] \\ &\quad [\det(\mathcal{T})]^{\frac{\alpha(\nu_0-d-1)}{2}} \exp \left[-\frac{1}{2} \text{tr}(\alpha \Psi_0^{-1} \mathcal{T}) \right]. \end{aligned}$$

Por lo tanto, la distribución posterior para $(\boldsymbol{\mu}, \mathcal{T})$ dado la muestra observada es:

$$\begin{aligned} \pi(\boldsymbol{\mu}, \mathcal{T}) &\propto [\det(\mathcal{T})]^{\frac{[(\alpha\nu_0+n\beta)+d(1-\alpha)-d-1]}{2}} \exp \left\{ -\frac{1}{2} \text{tr}[(\alpha\Psi_0^{-1} + n\beta\hat{\Sigma}_d)\mathcal{T}] \right\} \\ &\quad [\det(\mathcal{T})]^{1/2} \exp \left[-\frac{n\beta}{2} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^T \mathcal{T} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) - \frac{\lambda_0 \alpha}{2} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \mathcal{T} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \right]. \end{aligned}$$

Obsérvese que

$$\begin{aligned} &n\beta(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^T \mathcal{T} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) + \lambda_0 \alpha (\boldsymbol{\mu} - \boldsymbol{\mu}_0)^T \mathcal{T} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \\ &= (\lambda_0 \alpha + n\beta) \left(\boldsymbol{\mu} - \frac{\lambda_0 \alpha \boldsymbol{\mu}_0 + n\beta \hat{\boldsymbol{\mu}}}{\lambda_0 \alpha + n\beta} \right)^T \mathcal{T} \left(\boldsymbol{\mu} - \frac{\lambda_0 \alpha \boldsymbol{\mu}_0 + n\beta \hat{\boldsymbol{\mu}}}{\lambda_0 \alpha + n\beta} \right) \\ &\quad + \frac{\lambda_0 \alpha n\beta}{\lambda_0 \alpha + n\beta} \text{tr}[(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T \mathcal{T}]. \end{aligned}$$

Entonces, tenemos que

$$\begin{aligned} \pi(\boldsymbol{\mu}, \mathcal{T}) &\propto [\det(\mathcal{T})]^{\frac{[(\alpha\nu_0+n\beta)+d(1-\alpha)-d-1]}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[(\alpha\Psi_0^{-1} + n\beta\hat{\Sigma}_d \right. \right. \\ &\quad \left. \left. + \frac{\lambda_0 \alpha n\beta}{\lambda_0 \alpha + n\beta} (\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)^T \right) \mathcal{T} \right\} [\det(\mathcal{T})]^{1/2} \exp \left[-\frac{1}{2} (\lambda_0 \alpha \right. \\ &\quad \left. + n\beta) \left(\boldsymbol{\mu} - \frac{\lambda_0 \alpha \boldsymbol{\mu}_0 + n\beta \hat{\boldsymbol{\mu}}}{\lambda_0 \alpha + n\beta} \right)^T \mathcal{T} \left(\boldsymbol{\mu} - \frac{\lambda_0 \alpha \boldsymbol{\mu}_0 + n\beta \hat{\boldsymbol{\mu}}}{\lambda_0 \alpha + n\beta} \right) \right]. \end{aligned}$$

Apéndice B

El enfoque inferencial propuesto aquí es estadísticamente riguroso porque cae dentro del marco general de actualización de creencias de Bissiri *et al.* (2016), basado en funciones de pérdida en lugar de funciones de distribución. Bissiri *et al.* (2016) construyen una pérdida esperada (en el espacio de medidas de probabilidad en espacio- θ)

$$\mathcal{L}(v; \pi_0, x) = \int \ell(\theta, x)v(d\theta) + d_{KL}(v, \pi),$$

donde $\ell(\theta, x)$ es alguna función de pérdida y $d_{KL}(v, \pi_0)$ la divergencia de Kullback-Leibler de v a una *a priori* π para el parámetro θ , tal que

$$\pi^* = \underset{v}{\operatorname{argmin}} \mathcal{L}(v; \pi, x)$$

es una selección óptima de la distribución posterior dada una *a priori* $\pi_0(\theta)$ y datos x . Los autores muestran que el minimizador de $\mathcal{L}(v; \pi_0, x)$ viene dado por

$$\pi^*(\theta) = \frac{e^{[-\ell(\theta, x)]}\pi(\theta)}{\int e^{[-\ell(\theta, x)]}\pi(d\theta)},$$

tal que

$$0 < \int e^{[-\ell(\theta, x)]}\pi(d\theta) < \infty.$$

Entonces, la teoría presentada en Bissiri *et al.* (2016) tiene la forma de una actualización bayesiana que utiliza una función de pérdida exponencial negativa en lugar de la función de verosimilitud.

Para nuestros propósitos, $l(\theta, x)$ es la función de pérdida logarítmica dada por

$$l(\theta, x) = -\{\beta \log[L(\theta, x)] + (\alpha - 1)\pi(\theta)\},$$

donde $\log[L(\theta, x)]$ es la función log-verosimilitud para la distribución normal multivariada con parámetro $\theta = (\boldsymbol{\mu}, \Sigma)$, $\pi(\theta) = \pi(\boldsymbol{\mu}|\Sigma^{-1})\pi(\Sigma^{-1})$ es una *a priori* para θ con $\pi(\boldsymbol{\mu}|\Sigma^{-1})$ y $\pi(\Sigma^{-1})$ seleccionadas como densidades normal y Wishart, respectivamente, y β y α son los parámetros de sesgo correspondientes a las ponderaciones de la verosimilitud y la *a priori*, respectivamente. Entonces, en nuestro caso, la posterior óptima π^* es una densidad Normal-Wishart (véase el Apéndice A de este artículo).

Por otro lado, el estudio del desempeño de los estimadores sesgados se puede llevar a cabo teóricamente. Por ejemplo, la esperanza para los estimadores sesgados de $\boldsymbol{\mu}$ y Σ se da como:

$$E(\tilde{\boldsymbol{\mu}}) = E(\hat{\boldsymbol{\mu}}) + c_0 E(\boldsymbol{\mu}_0 - \hat{\boldsymbol{\mu}}),$$

$$E(\tilde{\Sigma}_d) = \left(\frac{1}{a_0}\right) E(\hat{\Sigma}_d) + \left(\frac{b_0}{a_0}\right) \Psi_0^{-1} + \left(\frac{c_0}{a_0}\right) E(\mathbf{z}\mathbf{z}^T).$$

Ahora, en general, puesto que $E(\hat{\boldsymbol{\mu}}) = \boldsymbol{\mu}$, $E(\hat{\Sigma}_d) = (1 - n^{-1})\Sigma$ entonces $\tilde{\boldsymbol{\mu}}$ y $\tilde{\Sigma}_d$ no son estimadores insesgados. Sin embargo, cuando n va al infinito obtenemos:

$$\lim_{n \rightarrow \infty} c_0 = 0, \quad \lim_{n \rightarrow \infty} \frac{(1 - n^{-1})}{a_0} = 1, \quad \lim_{n \rightarrow \infty} \frac{b_0}{a_0} = 0 \quad \text{y} \quad \lim_{n \rightarrow \infty} \frac{c_0}{a_0} = 0.$$

Entonces, tanto $\tilde{\boldsymbol{\mu}}$ como $\tilde{\Sigma}_d$ son estimadores asintóticamente insesgados, aunque no sean insesgados. Además, con base a estos límites, los estimadores sesgados heredan propiedades asintóticas del MLE, como la eficiencia asintótica.