

Estocástica aplicada en diferentes contextos

Jhonier Rangel

Wilson Gordillo Thiriat

Wilson Jairo Pinzón Casallas



© **Jhonier Rangel**

Universidad ECCI

jrangelg@ecci.edu.co

Wilson Gordillo Thiriat

Universidad Distrital Francisco José de Caldas

wgordillot@udistrital.edu.co

Wilson Jairo Pinzón Casallas

Universidad Distrital Francisco José de Caldas

wjpinzonc@udistrital.edu.co

© Editorial Grupo Compás, 2025

Guayaqui, Ecuador

www.grupocompas.com

<http://repositorio.grupocompas.com>

Primera edición, 2025

ISBN: 978-9942-33-910-2

Distribución online

 Acceso abierto

Cita

Rangel J., Gordillo, W., Pinzón, W. (2025) Estocástica aplicada en diferentes contextos. Editorial Grupo Compás

Este libro ha sido debidamente examinado y valorado en la modalidad doble par ciego con fin de garantizar la calidad de la publicación. El copyright estimula la creatividad, defiende la diversidad en el ámbito de las ideas y el conocimiento, promueve la libre expresión y favorece una cultura viva. Quedan rigurosamente prohibidas, bajo las sanciones en las leyes, la producción o almacenamiento total o parcial de la presente publicación, incluyendo el diseño de la portada, así como la transmisión de la misma por cualquiera de sus medios, tanto si es electrónico, como químico, mecánico, óptico, de grabación o bien de fotocopia, sin la autorización de los titulares del copyright.

Índice

Capítulo 1. Dependencias entre variables.....	3
Capítulo 2. Inferencia sobre las cadenas de Markov en \mathbb{R}	18
Capítulo 3. Ecuaciones con ruidos aleatorios.....	31
Referencias.....	42

Capítulo 1. Dependencias entre variables

Tanto en las ciencias exactas como en las ciencias humanas el análisis de la correlación entre variables ayuda a evaluar dependencias. Sin embargo, en algunos casos estas dependencias pueden ser no lineales y no cumplir los supuestos básicos para hacer un análisis, por ese motivo es necesario explorar otro tipo de coeficientes de correlación no lineal que permite flexibilizar el supuesto de normalidad bi-variante.

En (Martínez r., et al., 2009) se presenta una caracterización por rangos del coeficiente de correlación de Spearman, el cual permite evaluar la relación entre variables de manera no lineal, pero con presencia de monotonía, tanto creciente como decreciente. Por otro lado, (Bastías et al., 2013) destacan la importancia de las Buenas Prácticas de Manufactura, señalando que en áreas como la medicina, muchas variables, aunque se representan numéricamente, no son necesariamente variables de razón. (Santabárbara, 2019) Explica que en muchos estudios solo se presenta el valor estimado del coeficiente de variación. Sin embargo, el autor argumenta que es una buena práctica registrar un intervalo de confianza que acompañe la estimación puntual, ya que en muchos casos el coeficiente aparente puede estar muy alejado de cero. No obstante, al calcular el intervalo, este puede incluir el cero o tener un límite inferior cercano a cero, lo que indicaría que no existe una correlación significativamente alta. El autor también destaca la ventaja de utilizar SPSS para calcular el coeficiente de correlación de manera precisa, y en este artículo se explora el uso del programa R (<https://www.R-project.org/>).

En Schober et al. (2018) se indica que el coeficiente de correlación lineal de Pearson refleja la asociación entre variables, donde un cambio en una variable está vinculado a un cambio en otra, ya sea en la misma (correlación positiva) o en la dirección opuesta (correlación negativa). Para datos continuos y normalmente distribuidos, se utiliza comúnmente este coeficiente. Sin embargo, para datos no distribuidos normalmente, datos ordinales o aquellos con valores atípicos relevantes, se emplea el coeficiente de correlación de rangos de Spearman, que mide la asociación monótona. Ambos coeficientes tienen un rango de -1 a +1, donde 0 indica ausencia de asociación. Además, se pueden realizar pruebas de hipótesis y calcular intervalos de confianza para evaluar la significancia y la fuerza de la relación en la población de origen de los datos. No obstante, en esta investigación también se examinan escenarios en los que no se cumple el supuesto de normalidad, así como situaciones en las que los datos siguen distribuidos con comportamientos distribucionales diferentes.

En este artículo también se exponen las técnicas de remuestreo Leave One Out (LOO), que consiste en dejar un individuo fuera en cada remuestreo, y el Bootstrap, que implica extraer muestras mediante muestreo aleatorio simple. Con estas muestras, se calcula un estimador; en este caso, se obtiene el coeficiente de correlación para los individuos que pertenecen a la submuestra seleccionada. (Bishara & Hittner, 2012) ya abordaron el tema, considerando distribuciones distintas a la normal. Sin embargo, como diferencia clave de esta investigación, en este trabajo se considerará la distribución T multivariante, contaminada con diferentes proporciones de datos que siguen otro parámetro de centralidad y una matriz de varianzas y covarianzas distinta.

En el trabajo de (Serna-Morales, J. K., et al., 2024) se expone que los datos de falla bivariados son comunes en estudios de confiabilidad y supervivencia, donde la estimación de la fuerza de dependencia es a menudo un paso importante en el análisis de los datos. En la literatura, se ha establecido que los coeficientes de correlación miden la relación lineal entre dos variables, pero también pueden existir relaciones no lineales fuertes entre ellas. El coeficiente de concordancia τ de Kendall se ha convertido en una herramienta útil para el análisis de datos bivariados, la cual es usada en pruebas no paramétricas de independencia y como una medida complementaria de asociación. En el análisis de datos de confiabilidad, hay un fenómeno que ocurre cuando el valor de las observaciones se conoce parcialmente, lo cual se conoce como censura. En este trabajo, se comparan vía simulación dos métodos de estimación del τ de Kendall, una de ellas suponiendo normalidad en las distribuciones marginales y ajustándose individualmente, y la otra basada en cópulas (Gaussiana y Clayton), donde los datos bivariados están censurados a intervalo.

Otra alternativa que se expone en este artículo es el *Bootstrap*, cuyo fundamento teórico se detalla en la literatura. En este caso, se aborda su aplicación en el contexto multivariante. Es importante señalar que las muestras bootstrap univariantes para cada variable no son equivalentes a las muestras multivariantes bootstrap, y de hecho, sería erróneo tratar de relacionarlas directamente. Este matiz es crucial para comprender cómo utilizar correctamente el método en escenarios multivariantes, evitando confusiones y errores en la interpretación de los resultados (Zientek, L. R., & Thompson, B., 2007).

Por otro lado, se llevará a cabo una simulación para examinar el impacto de los datos atípicos en la inferencia del coeficiente clásico. El objetivo es investigar cómo la presencia de valores extremos puede modificar los resultados obtenidos mediante métodos estadísticos tradicionales, como el cálculo del coeficiente de correlación. A través de esta simulación, se analizará cómo distintas proporciones de datos atípicos afectan la precisión y validez de las estimaciones, lo que permitirá obtener una comprensión más profunda de las limitaciones de los enfoques clásicos cuando se enfrentan a datos que no siguen distribuciones convencionales.

Se analizan los coeficientes de correlación más conocidos, tales como el coeficiente de correlación lineal de Pearson, el coeficiente de monotonía de Spearman, el coeficiente de concordancia de Kendall (Shi, X. et al., 2024) y el coeficiente de fiabilidad kappa de Cohen (Fleiss, J. L., et al., 1969)).

$$r = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{\sqrt{[n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2][n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}}$$

Un teorema importante que ayuda a interpretar el coeficiente de correlación lineal de Pearson es el siguiente:

Teorema: El coeficiente de correlación lineal de Pearson de dos variables cuantitativas es el coseno euclidiano de los dos vectores de observaciones centrados por la media.

Demostración:

El coeficiente de correlación de Pearson se puede interpretar como el coseno del ángulo entre los vectores que representan las variables centradas. A continuación, se muestra cómo demostrar esta relación

$$r = \frac{\widehat{cov}(X,Y)}{\widehat{\sigma}_X \widehat{\sigma}_Y}$$

$$\cos(\theta) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Muchos métodos estadísticos requieren el uso de supuestos de distribución (Gutiérrez, J. S. R. , 2024) en este caso este coeficiente tiene el supuesto de la distribución normal bivariada, cuya expresión es la siguiente:

$$X \sim N(\mu, \Sigma)$$

Es importante aclarar que el vector de medias no tiene restricciones en el espacio vectorial de dimensión n . La matriz de varianzas y covarianzas debe ser definida positiva (Smania, G., & Jonsson, E. N., 2021).

$$f(X) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp \exp \left(-\frac{1}{2} (X - \mu)^T \Sigma^{-1} (X - \mu) \right)$$

Donde el vector de medias es un objeto que organiza las medias de cada variable, mientras que la matriz de varianzas y covarianzas contiene las varianzas en su diagonal y las covarianzas fuera de ella.

Otra distribución multivariante importante es la distribución t de Student multivariante, que, al igual que la anterior, requiere un vector de medias y una matriz de varianzas y covarianzas.

Un vector de medias, una matriz definida positiva y \mathbf{v} caracterizan la distribución t de Student multivariante. La distribución t multivariante es:

$$f_X(x) = \left(\frac{\Gamma\left(\frac{p+v}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \right) \left(\frac{1}{\pi} \right)^{\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \left(1 + \frac{1}{v} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)^{-\frac{v+p}{2}}$$

Esta distribución multivariante tiene presencia de datos atípicos (Azzalini, A., & Capitanio, A. ,2003)).

Para visualizar esto se crea una grilla con cuatro paneles donde se simulan los valores de distribuciones normales bivariadas y t de Student bivariantes con 5, 10 y 15 grados de libertad, es posible usar Python junto con bibliotecas como Matplotlib para la visualización y NumPy para la generación de los datos en la extensión de Google Drive (<https://colab.research.google.com/>).

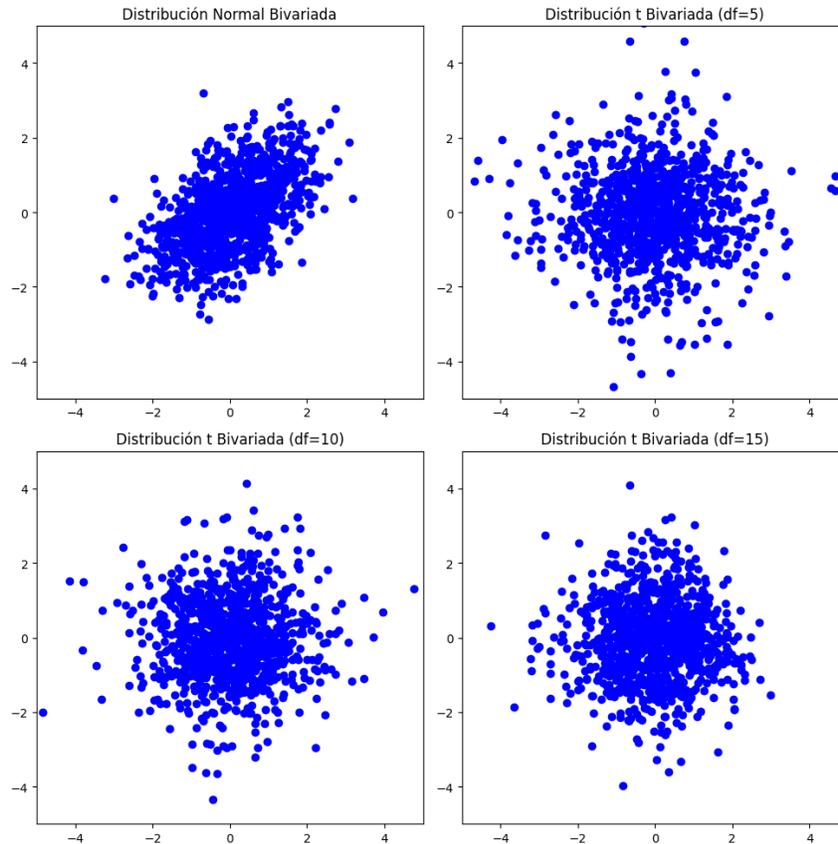


Figura 1. Distribuciones bivalentes generadas con 100 puntos para cada caso. El primer panel muestra una distribución normal bivariada, mientras que los tres paneles siguientes muestran distribuciones *t* de Student bivariantes con 5, 10 y 15 grados de libertad, respectivamente. Todos los gráficos tienen la misma escala en los ejes X e Y para facilitar la comparación visual.

En la figura 1. Se ve que la distribución *T* multivariante se diferencia de la distribución normal bivariada principalmente en su manejo de los datos atípicos y en su forma analítica. La distribución normal bivariada se concentra alrededor de la media y, debido a su menor cola pesada, es menos propensa a generar valores extremos. En cambio, la distribución *T* multivariante, especialmente con grados de libertad bajos, tiene colas más pesadas, lo que aumenta la probabilidad de generar valores atípicos o extremos. Analíticamente, la distribución *T* depende de un parámetro adicional, los grados de libertad, que controla la forma de las colas. A medida que los grados de libertad disminuyen, las colas se hacen más gruesas y, por lo tanto, la probabilidad de observar datos atípicos aumenta significativamente.

Este comportamiento contrasta con la distribución normal bivariada, en la cual los datos atípicos son menos frecuentes y la variabilidad se distribuye de manera más homogénea.

En este apartado se presentan las simulaciones y los resultados obtenidos a partir de datos normales bivariantes y de datos con distribución T bivariada, contaminados por diferentes proporciones de valores atípicos. Se analiza cómo estas contaminaciones afectan la inferencia basada en el supuesto de normalidad. Además, se realiza una aplicación utilizando los índices de cobertura primaria y secundaria en los municipios de Colombia durante el año 2023, con el fin de evaluar el impacto de estas distribuciones en el análisis y los resultados obtenidos.

Se generó un análisis detallado de la relación entre dos variables bajo condiciones normales y contaminadas utilizando datos bivariantes con una distribución normal. Para ello, se generaron datos multivariantes con una media de $[0,0]$ y una matriz de covarianzas que reflejaba un coeficiente de correlación de -0.8 . Posteriormente, se crearon muestras contaminadas a diferentes niveles (5%, 10%, 15%, y 20%) mediante la modificación de una porción de las observaciones originales, lo que simuló la presencia de anomalías o perturbaciones en los datos. Además, se realizó una prueba de Mardia para evaluar la normalidad multivariante de cada conjunto de datos contaminado, obteniendo los valores p correspondientes para el sesgo y la curtosis, con el fin de examinar la influencia de la contaminación en la distribución de los datos.

En términos de correlación, se calcularon tres tipos de coeficientes: Kendall, Spearman y Pearson, para cada nivel de contaminación. El coeficiente de Pearson se utilizó para evaluar las relaciones lineales y normales entre las variables, mientras que los coeficientes de Kendall y Spearman fueron empleados para medir las relaciones no lineales o monotónicas. Los coeficientes de correlación, junto con los valores p obtenidos de la prueba de Mardia, fueron presentados en una tabla. Además, se generaron gráficos que ilustran la distribución de las observaciones originales y contaminadas, facilitando la interpretación visual del impacto de la contaminación en la relación entre las variables.

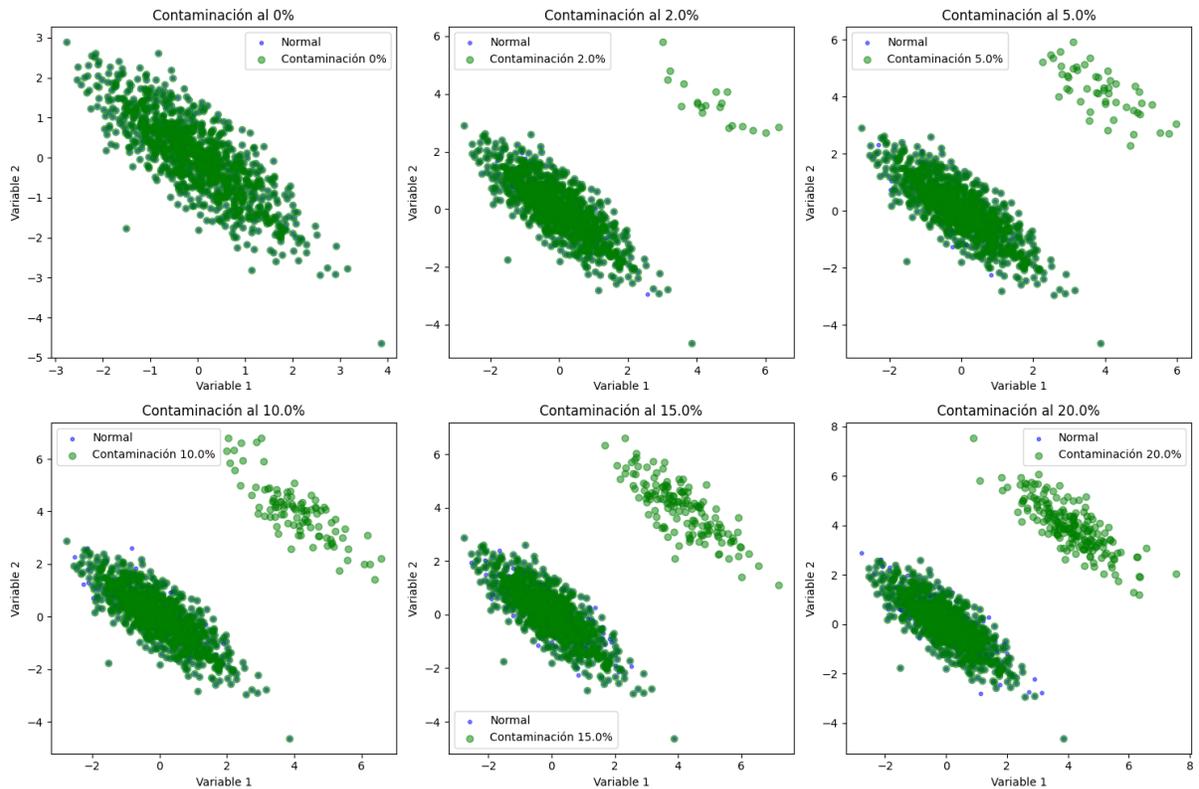


Figura 2. Distribución de los datos originales y contaminados a diferentes niveles de contaminación (5%, 10%, 15%, 20%) en función de la variable 1 y variable 2. Los puntos azules representan los datos originales, mientras que los puntos verdes muestran los datos contaminados. Los gráficos ilustran cómo la contaminación afecta la relación entre las dos variables.

Se evidencia que, aunque localmente hay presencia de correlación negativa, lo más probable es que estemos ante la presencia del efecto Simpson (Aiyou C., et al., 2009), ya que globalmente puede haber una correlación positiva. Este fenómeno ocurre cuando una tendencia observada en varios grupos desaparece o se invierte cuando los datos se combinan, lo que puede llevar a conclusiones erróneas si no se considera la estructura de los datos y la interacción entre las variables a nivel global y local.

Tabla 1. Resultados de los valores p para los tests de sesgo y curtosis, junto con los coeficientes de correlación de Kendall, Spearman y Pearson, para distintos niveles de contaminación.

Contaminación (%)	Coef. Kendall	Coef. Spearman	Coef. Pearson
0.00	-0.585578	-0.780344	-0.790705
0.02	-0.522967	-0.675797	-0.356312
0.05	-0.432012	-0.526342	-0.008416
0.10	-0.300268	-0.300594	0.269511
0.15	-0.177898	-0.100592	0.413265
0.20	-0.079512	0.067052	0.492009

En la tabla 1. se evidencia que el coeficiente de correlación de Pearson es especialmente sensible a las relaciones lineales entre dos variables, lo que lo hace más susceptible a los cambios en la magnitud de los datos. En la tabla proporcionada, se puede observar que a medida que aumenta el nivel de contaminación en los datos (de 0% a 20%), el valor de Pearson pasa de ser negativo (-0.79 en 0% de contaminación) a positivo (0.49 en 20% de contaminación), lo que indica una transición de una relación inversa a una directa a medida que las observaciones se desvían de la normalidad. Esta sensibilidad refleja cómo Pearson reacciona a las alteraciones en la relación lineal entre las variables, mientras que otros coeficientes como Kendall y Spearman, aunque también afectados, son menos sensibles a tales cambios y pueden captar relaciones no lineales o con más ruido.

En los datos contaminados, se puede observar cómo al aumentar el porcentaje de contaminación, los coeficientes de correlación cambian drásticamente. Lo interesante es que al contaminar los datos con un vector alejado del origen, los datos parecen mostrar una relación creciente entre las variables, que en realidad no existe en los datos originales. Este fenómeno es una manifestación clara del Efecto Simpson, que ocurre cuando, al combinar datos de diferentes grupos contaminados, se genera una ilusión de crecimiento en la relación entre las variables. Sin embargo, esta relación no es real, ya que al observar los datos por separado (sin contaminación o con poca contaminación), se evidencia que no hay una correlación positiva o incluso puede haber una relación negativa o inversa entre las variables. La contaminación de los datos al alejar el vector del origen produce un cambio significativo en la forma en que las variables parecen estar relacionadas,

lo que puede generar conclusiones incorrectas si no se tiene cuidado al interpretar los resultados.

A continuación, se presentan las muestras generadas utilizando una distribución t multivariante con un vector de medias en el origen y una matriz de varianzas y covarianzas. Estas muestras se derivan de una mezcla de distribuciones normal y gamma, lo que da como resultado una distribución t multivariante. Para cada nivel de contaminación (0%, 2%, 5%, 10%, 15%, 20%), se generan datos contaminados a partir de esta mezcla, permitiendo observar los efectos de la contaminación en la estructura de los datos. El objetivo es estudiar cómo los datos contaminados afectan la distribución original y cómo los coeficientes de correlación, como los de Kendall, Spearman y Pearson, responden a la introducción de diferentes niveles de contaminación.

La metodología de mezcla entre distribuciones normal y gamma se utiliza para simular un escenario en el que las observaciones fuera de lo común son añadidas a los datos originales. Esta simulación ayuda a comprender el impacto de los valores atípicos en la distribución y permite evaluar cómo las correlaciones entre las variables cambian a medida que se aumenta la contaminación. En particular, la introducción de estos datos contaminados permite explorar los efectos del Efecto Simpson, el cual puede inducir a una interpretación incorrecta de las relaciones entre las variables, especialmente cuando los datos atípicos distorsionan la tendencia general.

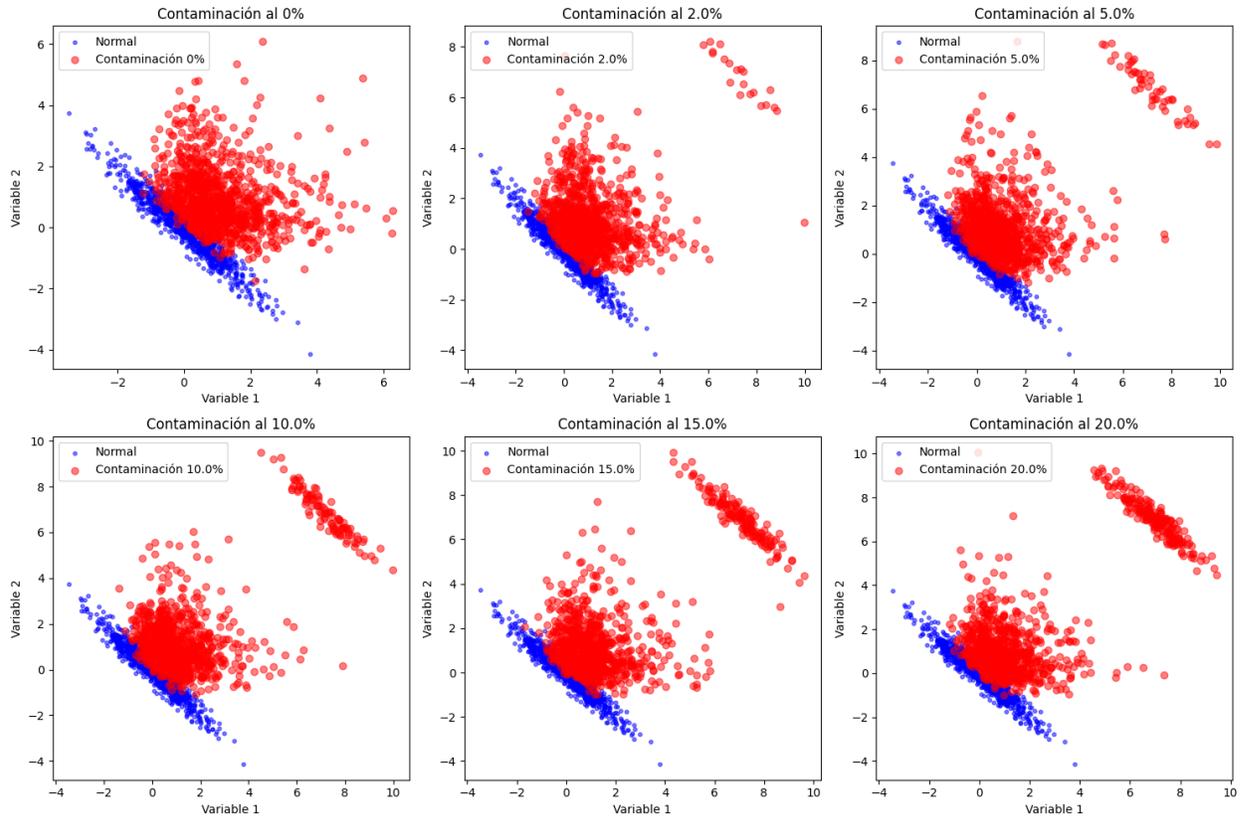


Figura 3. Visualización de los datos originales (en azul) y los datos contaminados (en rojo) a diferentes niveles de contaminación. Cada subgráfica corresponde a un nivel específico de contaminación, variando del 0% al 20%. Los puntos azules representan los datos generados sin contaminación, mientras que los puntos rojos muestran cómo los datos se ven alterados por la mezcla de distribuciones normal y gamma, evidenciando el efecto de la contaminación en la estructura de los datos.

Tabla 2. Coeficientes de correlación de Kendall, Spearman y Pearson, junto con los valores de sesgo (Skewness) y curtosis (Kurtosis) para diferentes niveles de contaminación. Los valores de sesgo y curtosis están presentados con los límites inferior (Lim Inf) y superior (Lim Sup). Los coeficientes de correlación indican la relación entre las variables en función del porcentaje de contaminación en los datos.

%	Kendall	Spearman	Pearson	Sesg			
				(2.5%)	Sesgo (97.5%)	Curtosis (2.5%)	Curtosis (97.5%)
0.	-0.011	-0.015	-0.024	1.05	1.235	1.724	2.271
0.	0.011	0.017	0.034	1.96	2.260	5.499	7.692
0.	-0.014	-0.022	-0.026	2.08	2.249	4.786	5.692
0.	-0.048	-0.073	-0.064	1.83	1.867	2.741	2.754
0.	0.031	0.047	0.049	1.42	1.451	0.916	0.919
0.	0.005	0.009	0.014	1.18	1.209	0.065	-0.007

En la tabla 2. se observan variaciones en los coeficientes de correlación, sesgo y curtosis a medida que aumenta el nivel de contaminación de los datos. En cuanto a los coeficientes de correlación, el valor de Kendall, Spearman y Pearson muestra cambios moderados con el aumento de la contaminación. En particular, los coeficientes tienden a volverse más negativos o menos pronunciados cuando la contaminación es baja (como en el caso de 0% o 2%), sugiriendo que, a niveles bajos de contaminación, las relaciones entre las variables aún se mantienen relativamente fuertes, aunque con una ligera tendencia al decrecimiento. A medida que se incrementa la contaminación (especialmente a 10% y 15%), los coeficientes

comienzan a estabilizarse o incluso a cambiar su signo, indicando una pérdida de la correlación entre las variables.

En cuanto a los valores de sesgo y curtosis, los datos muestran un comportamiento claro con la contaminación. Al principio, en niveles de contaminación bajos (0% a 5%), los valores de sesgo y curtosis se mantienen dentro de rangos razonables, pero a medida que la contaminación aumenta (especialmente a niveles más altos, como 10% y 15%), se observa un aumento significativo en la curtosis, lo que sugiere una mayor concentración de los datos alrededor de la media o la presencia de valores extremos más marcados. Este cambio en la curtosis es consistente con los efectos de la contaminación, que distorsionan la distribución original. En general, la contaminación tiene un impacto considerable sobre la normalidad de los datos, especialmente a medida que aumenta el nivel de contaminación, lo que se refleja en el cambio de los coeficientes y las medidas de dispersión como el sesgo y la curtosis.

Ahora, aplicaremos técnicas de remuestreo a los datos de deserción en primaria y secundaria en Colombia correspondientes al año 2023. Los datos utilizados están disponibles en el sitio web del Ministerio de Educación Nacional de Colombia y abarcan las tasas de deserción por cada departamento tanto en primaria como en bachillerato. En la gráfica se presenta la estimación de la regresión kernel, destacando la posible tendencia creciente observada en las tasas de deserción.

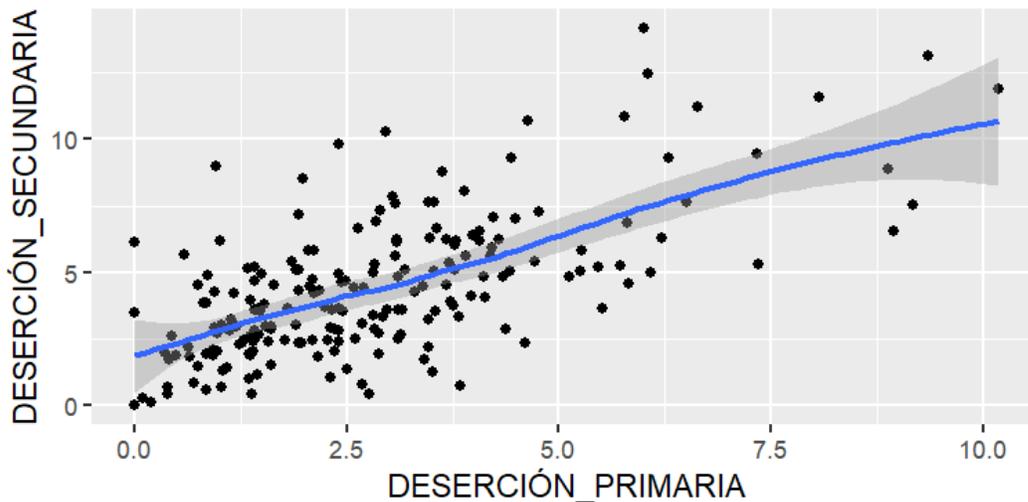


Figura 4. Deserción en secundaria en relación con la deserción en primaria, representada mediante un gráfico de densidad kernel con intervalos de confianza del 95%.

En la figura 4 se observa una relación positiva directa entre las variables. Sin embargo, para evaluar adecuadamente la correlación, es necesario recurrir a un índice de correlación y, en este caso, al coeficiente de correlación lineal de Pearson. No obstante, antes de utilizar este coeficiente, es fundamental someter los datos a una prueba de normalidad multivariante. En este estudio, se empleó el test de Mardia (Wulandari, D., et al., 2021), cuyo resultado arrojó un valor p de 0.012. Dado que este valor es menor al nivel de significancia del 5%, se rechaza la hipótesis nula, lo que indica que los datos no siguen una distribución normal multivariante. Por lo tanto, el uso del coeficiente de correlación lineal de Pearson no es adecuado en este caso.

En este artículo, se ha utilizado un enfoque basado en técnicas de remuestreo para analizar la relación entre la deserción en la educación en primaria y en secundaria en Colombia, a partir de datos proporcionados por el Ministerio de Educación Nacional en 2023. Las técnicas de remuestreo aplicadas incluyen "Leave One Out" (LOO), que permite evaluar la estabilidad de las estimaciones obtenidas al excluir sucesivamente una observación en cada iteración. Esta metodología es útil para obtener intervalos de confianza y evaluar el comportamiento de los estimadores sin depender de supuestos de normalidad, los cuales no se cumplen en este caso según lo indicado por los resultados de las pruebas de normalidad multivariante.

En lugar de utilizar un único conjunto de datos, la técnica Leave One Out permite crear muestras repetidas al eliminar una observación de cada vez y calcular las estadísticas de interés, como la media y la correlación, para cada muestra. Este enfoque nos proporciona una visión más robusta de los resultados, ya que reduce el sesgo que podría generarse si se utilizara el conjunto de datos completo. De esta manera, se obtiene un intervalo de confianza del 95% para las estadísticas calculadas, como el coeficiente de correlación de Spearman entre la deserción primaria y secundaria. Además, se calcula el valor de la correlación en cada iteración para observar la variabilidad de los estimadores. Estas técnicas, combinadas con gráficos de distribución, permiten evaluar la confiabilidad de las conclusiones obtenidas sobre la relación entre las variables de deserción escolar.

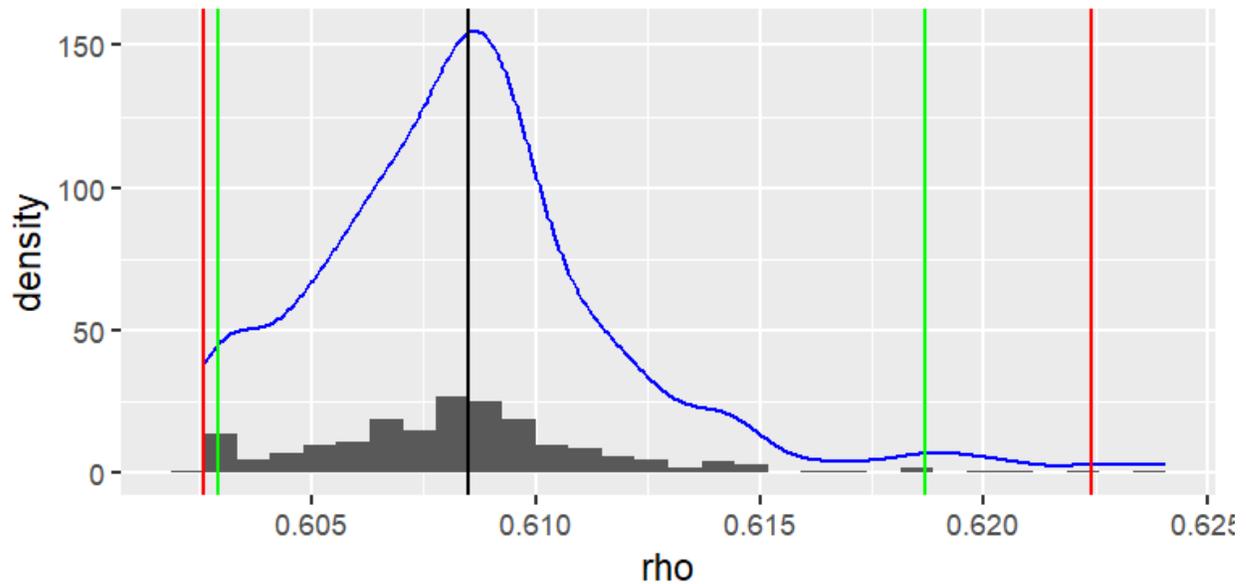


Figura 5. Distribución de los coeficientes de correlación de Spearman para las diferentes muestras generadas mediante la técnica "Leave One Out". En la gráfica se muestra el histograma junto con la densidad estimada, así como los intervalos de confianza al 95% (en verde) y los percentiles del 0.5% y 99.5% (en rojo). La línea negra indica la media de los coeficientes de correlación.

En la Figura 5 se observa que los intervalos de confianza generados mediante las metodologías de remuestreo no son simétricos. Esto contrasta con los intervalos derivados bajo la suposición de un modelo distribucional previamente definido, los cuales sí presentan simetría. Esta asimetría sugiere que asumir un modelo de distribución específico podría llevar a una aproximación más precisa. Además, se puede concluir que existe una correlación directa significativa entre la deserción primaria y secundaria, ya que el intervalo de confianza no incluye el valor cero. De hecho, el intervalo obtenido es un subconjunto que supera el umbral de 0.5, lo que refuerza la presencia de una relación positiva entre ambas variables.

Esta investigación subraya la importancia de un uso adecuado del coeficiente de correlación, especialmente cuando se consideran los supuestos subyacentes de los métodos estadísticos. El análisis reveló que, en situaciones donde los datos no cumplen con los supuestos de normalidad multivariante o presentan atípicos, el uso del coeficiente de correlación lineal de Pearson puede no ser apropiado, ya que podría generar interpretaciones erróneas. Ante esta situación, las alternativas de remuestreo, como el método Leave-One-Out y el bootstrap, se presentan como

herramientas robustas, ya que permiten obtener estimaciones de intervalos de confianza sin depender de distribuciones paramétricas. Estas técnicas no solo ofrecen una mayor flexibilidad, sino que proporcionan estimaciones más confiables al evitar las restricciones asociadas con los métodos tradicionales. En el contexto colombiano, los resultados indican que existe una correlación significativa y directa entre la deserción escolar en primaria y secundaria. Este hallazgo es crucial para los responsables de las políticas educativas, ya que sugiere que las intervenciones en el nivel primario podrían tener un impacto directo en la reducción de la deserción en niveles educativos posteriores. En este sentido, es fundamental que las autoridades educativas utilicen estos análisis para diseñar políticas más eficaces que aborden de manera integral los factores que afectan la permanencia de los estudiantes en el sistema educativo.

Capítulo 2. Inferencia sobre las cadenas de Markov en R

Las cadenas de Markov en tiempo discreto constituyen un marco matemático esencial en la teoría de probabilidad, siendo utilizadas históricamente para modelar sistemas que evolucionan en pasos discretos a través de estados definidos. Dichos modelos tienen diversas aplicaciones; por ejemplo, en marketing, (Styan, 1964) expone cómo aplicar cadenas de Markov en estrategias de marketing. En medicina, (Pegels, 1970) presenta una aplicación sobre el estado de la sangre en pacientes. También se presentan aplicaciones actuales, como la de (Lin, 2021), que evalúa el estado de la velocidad de un motor. Este enfoque encuentra aplicaciones significativas en una variedad de disciplinas, desde ingeniería y finanzas hasta biología y aprendizaje automático. Sin embargo, en dichas aplicaciones no se ha explotado completamente el potencial inferencial que dichos modelos poseen y que aporta significativamente a dichas investigaciones.

En ingeniería de comunicaciones, las cadenas de Márkov son empleadas para analizar la transmisión de datos en canales ruidosos y protocolos de red. En el ámbito financiero, se utilizan para modelar la dinámica de precios en mercados y gestionar riesgos. En biología, estas cadenas son fundamentales para entender la evolución de poblaciones y la propagación de enfermedades. Además, en aprendizaje de máquina, los modelos ocultos de Márkov son esenciales en reconocimiento de patrones y procesamiento del lenguaje natural. Por ejemplo (Mo, 2022) expone una aplicación interesante de las cadenas de Márkov en sistemas de comunicación.

En la investigación de operaciones, las cadenas de Márkov se aplican en la optimización de sistemas y la toma de decisiones, por ejemplo “Describimos nuestro algoritmo para identificar sitios de unión de factores de transcripción (TFBS) a través de Optimización de la cadena de Markov. La clave es un algoritmo para ordenar a la cadena de Markov que capture la dependencia más significativa entre posiciones en el enlace del objetivo secuencia.” (Yang, 2012, pág. 4), también se usa para abordar problemas como la planificación de inventarios y la gestión de recursos. En resumen, estas cadenas proporcionan un marco robusto para modelar y entender procesos estocásticos en distintos contextos, ofreciendo una herramienta valiosa para analizar, simular y optimizar sistemas complejos que evolucionan en pasos discretos en el tiempo.

Como se ha indicado anteriormente las cadenas de Markov en tiempo discreto son un marco teórico esencial en la teoría de probabilidad y la modelización estocástica. Estas cadenas son utilizadas para describir sistemas que evolucionan a lo largo de una secuencia de instantes discretos en el tiempo, donde la transición de un estado a otro está determinada por probabilidades condicionales. Un aspecto clave de las cadenas de Markov es su propiedad de "sin memoria", lo que significa que la probabilidad de moverse a un estado futuro depende únicamente del estado actual y no de la secuencia de eventos previos.

La representación matemática de las cadenas de Markov según (Saghafian, 2018, pág. 357) se realiza a través de la matriz de cambio de estados, comúnmente denotada como P , que describe las probabilidades de transición entre los distintos estados en un solo paso temporal. Esta matriz es fundamental para entender la dinámica a corto plazo de la cadena. Además, las cadenas de Markov exhiben un comportamiento a largo plazo que se caracteriza por la distribución límite, π , que representa la distribución estacionaria a medida que la cadena evoluciona indefinidamente.

En conjunto, las cadenas de Markov proporcionan un marco versátil y poderoso para modelar y analizar sistemas dinámicos y estocásticos en una amplia gama de disciplinas, desde ingeniería y finanzas hasta biología y aprendizaje automático. (Zucchini, 2009) explica que su capacidad para capturar la evolución temporal de un sistema mediante probabilidades de transición y describir su comportamiento estacionario las convierte en una herramienta valiosa para abordar problemas complejos en la modelización y simulación de sistemas dinámicos.

Una característica importante intrínseca a un modelo basado en una cadena de Markov es la distribución estacionaria, "Suponga que tenemos N estados en nuestro sistema. La matriz de cambio de estados, comúnmente denotada como P , es una matriz cuadrada $N \times N$ donde cada elemento P_{ij} representa la probabilidad de transición del estado i al estado j . Es importante destacar que cada fila de la matriz P suma a 1, ya que las probabilidades de transición desde un estado dado deben sumar 1" (Meyer, 1994, pág 1). El autor explica que es necesario tener la matriz de transición antes de proceder al cálculo de la distribución estacionaria haciendo uso del contexto. Sin embargo, en la vida real en muchas ocasiones nos encontramos con una observación parcial del proceso estocástico y con estos datos se quiere realizar inferencia sobre los parámetros.

Por otro lado, teniendo la matriz de transición se hace necesario el cálculo de la distribución límite, (Meyn ,1993) resalta que este es un concepto clave que describe el comportamiento a largo plazo de la cadena de Markov. Para hacer cálculo de dicha distribución se puede seguir el siguiente proceso, "a largo plazo la distribución de probabilidad sobre los estados deja de cambiar en su evolución temporal. Denotemos esta distribución límite como π , y es única para una cadena de Markov irreducible y aperiódica. La distribución límite π se puede encontrar resolviendo la ecuación $\pi P = \pi$, donde P es la matriz de cambio de estados. Esta ecuación establece que la distribución límite es invariante cuando la cadena se encuentra en equilibrio." (Castañeda, 2012, pág. 369).

El propósito de este estudio consiste en llevar a cabo un análisis comparativo entre los métodos inferenciales clásicos de máximo verosimilitud y el método de los momentos aplicados a cadenas de Markov con espacio de estados finito y tiempo discreto. Durante este análisis, se destacarán algunas limitaciones inherentes a estos enfoques tradicionales, lo cual servirá como base para la exploración y justificación de la aplicación de inferencias bayesianas o no paramétricas en estos modelos. Este abordaje busca ofrecer una perspectiva más amplia y robusta, superando las posibles desventajas identificadas en los métodos clásicos, y promoviendo así un entendimiento más profundo de la dinámica subyacente en las cadenas de Markov consideradas.

La aplicación de la inferencia clásica a las cadenas de Markov requiere una sólida revisión de conceptos fundamentales de la estadística matemática. Entre estos conceptos, los momentos poblacionales desempeñan un papel esencial al proporcionar medidas descriptivas de la forma y tendencia central de una distribución, resultandos fundamentales para comprender la dinámica de las transiciones entre estados en las cadenas de Markov. Además, la familiaridad con las distribuciones multinomial y binomial es crucial, ya que estas modelan los eventos en el contexto de múltiples categorías o resultados binarios, respectivamente, aspectos que pueden ser inherentes a las cadenas de Markov. La aplicación de métodos de estimación, tales como el método de momentos y el método de máximo verosimilitud, se revela como un paso indispensable para determinar los parámetros que caracterizan la evolución temporal de la cadena. Complementariamente, el método de Montecarlo, mediante la generación de muestras aleatorias, proporciona un enfoque valioso para simular el comportamiento de la cadena y obtener resultados estadísticos robustos. En conjunto, este conjunto de herramientas estadísticas provee la base necesaria para

explorar y comprender a fondo las cadenas de Markov, facilitando así una inferencia rigurosa sobre su comportamiento en diversas situaciones.

En (Zhang, 2020) se explica que el primer momento es la esperanza matemática definida como sigue:

$$E(X) = \left\{ \int_{-\infty}^{\infty} xf(x)dx \sum_x xP(X = x) \right. \quad (1)$$

El primer momento describe la centralidad de la distribución, también en (Zhang, 2020) se explica la necesidad del segundo momento como fuente de variabilidad de la variable aleatoria, dicho segundo momento centrado en la esperanza se define como:

$$Var(X) = \left\{ \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx \sum_x (x - E(X))^2 P(X = x) \right. \quad (2)$$

Estos dos momentos juegan un papel crucial al realizar inferencia, y son fundamentales para llevar a cabo la estimación. En el proceso de inferencia, se inicia recordando el método de momentos como una herramienta valiosa para realizar estimaciones. (Andrews, 2010) destaca la robustez de este método bajo la condición de que se cumplan los supuestos distribucionales necesarios. Es importante señalar que el método de momentos opera mediante la igualación de momentos teóricos y empíricos. En este contexto particular, nos limitamos a considerar la estimación de la esperanza y la varianza. Este enfoque, al centrarse en estos momentos específicos, simplifica el proceso de inferencia al abordar directamente las características centrales de la distribución en cuestión. Así, al aplicar el método de momentos a la estimación de la esperanza y la varianza, se establece un marco robusto para la inferencia en el contexto de cadenas de Markov con espacio de estados finito y tiempo discreto.

$$\hat{\mu} = \frac{\sum_i x_i}{n} \quad (3)$$

$$\widehat{\sigma^2} = \frac{\sum_i (x_i - \bar{x})^2}{n - 1} \quad (4)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n - 1}} \quad (5)$$

En las cadenas de Markov la inferencia propuesta bajo el método de los momentos consiste en: hallar las frecuencias de transición entre los estados de la cadena de Markov. Estas frecuencias representan el número de veces que la cadena pasa de un estado a otro durante un conjunto de observaciones, luego se definimos los momentos empíricos en términos de estas frecuencias de pasos. Por ejemplo, el primer momento empírico puede representar la media de las transiciones entre estados, y el segundo momento empírico puede corresponder a la varianza de estas transiciones.

Otro método importante a la hora de realizar estimación es el método de máximo verosimilitud, según (Li, 2020, pág. 4) el método de máximo verosimilitud se define como:

$$\hat{\theta} := \operatorname{argmáx}_{\theta \in \Theta} \{L(\theta | X_1, \dots, X_n)\} \quad (6)$$

Suponga que la muestra aleatoria es una familia de variables aleatorias independientes e idénticamente distribuidas

$$\hat{\theta} := \operatorname{argmáx}_{\theta \in \Theta} \left\{ \prod_i f(x_i | \theta) \right\} \quad (7)$$

Se invita al lector demostrar que bajo la distribución binomial el estimador por el método de los momentos y por máximo verosimilitud coincide y es

$$\hat{\pi}_j = \frac{n_j}{n} \quad (8)$$

A su vez bajo un estado en el tiempo t , la estimación de pasar a otro estado por el método de los momentos es

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} \quad (9)$$

Una pregunta muy interesante que nos planteamos es: ¿ p_{ij} es estadísticamente significativo?, luego planteamos el siguiente sistema de hipótesis:

$$\{H_0: p_{ij} = 0 \quad H_a: p_{ij} > 0\} \quad (10)$$

Este contraste de hipótesis puede ser sometido bajo diversos métodos, el que será considerado en este caso es la prueba bajo teorema de límite central expuesto en (Wasserman, 2020)

$$z = \frac{\widehat{p}_{ij}}{\sqrt{\frac{\widehat{p}_{ij}(1 - \widehat{p}_{ij})}{n_i}}} \quad (11)$$

Si z es mayor al cuantil de la distribución normal de probabilidad correspondiente a $1 - \alpha$, se rechaza la hipótesis nula. Esta prueba presenta algunos problemas de consistencia, los cuales se explican en (Bonkhoff, 2020) o en (Linde, 2023). Aunque este sea un problema, requiere un estudio minucioso y se deja como un tema abierto a discusión.

La necesidad de realizar inferencia sobre los parámetros de una cadena de Markov radica en obtener un modelo más parsimonioso. Recordemos que el criterio de parsimonia sugiere que un modelo es mejor cuando maximiza la información retenida al minimizar el número de parámetros. Al considerar la cantidad de parámetros de una cadena de Markov, observamos que se deben estimar $k \times k$ parámetros, lo cual es una cantidad bastante significativa. Si utilizamos métodos inferenciales para evaluar qué parámetros son significativos, reduciremos el número de parámetros a estimar. Para profundizar en el criterio filosófico de parsimonia, se puede consultar (Sober, 1981). Para la simulación se debe tener cuidado ya que globalmente la simulación que haremos no es sobre solo una distribución dada. En la siguiente tabla se expone el algoritmo en pseudocódigo para realizar inferencia sobre una cadena de Markov. Si desea profundizar las reglas generales del pseudocódigo pue ver (Oda, 2015). Para simular la cadena se sigue el siguiente código

Tabla 1. Pseudocódigo para realizar la simulación de una Cadena de Markov

Algoritmo: Simulación de Cadena de Markov
Entradas:
- Matriz de Transición (matriz_transicion)
- Estado Inicial (estado_actual)
- Número de Pasos (num_pasos)
Salida:
- Secuencia de Estados Simulados

-
1. Definir la Matriz de Transición:
 - `matriz_transicion`: matriz NxN con probabilidades de transición.

 2. Definir el Estado Inicial:
 - `estado_actual`: elegir un estado inicial.

 3. Número de Pasos de Simulación:
 - `num_pasos`: determinar la cantidad de pasos a simular.

 4. Simulación:
 - Para cada paso desde 1 hasta `num_pasos`:
 - a. Imprimir o almacenar el estado actual.
 - b. Elegir el próximo estado basado en la matriz de transición y actualizar estado actual.
-

En esta tabla se expone el pseudocódigo para realizar la simulación de una cadena de Markov. Este pseudocódigo puede adaptarse a cualquier lenguaje de programación, como R o Python. Para ver la implementación en R puede ver (Rangel, 2023) y en Python ver (Rangel, 2023).

```
# Definir la matriz de transición

trans_matrix <- matrix(c(0.7, 0.2, 0.1, # Probabilidades de transición de A
                        0.3, 0.5, 0.2, # Probabilidades de transición de B
                        0.1, 0.4, 0.5), # Probabilidades de transición de C
                      nrow = 3, byrow = TRUE)

# Estados posibles

states <- c("A", "B", "C")

# Función para generar las observaciones

generate_observations <- function(trans_matrix, states, n_obs) {
  current_state <- sample(states, 1) # Estado inicial aleatorio
```

```

observations <- character(n_obs) # Vector para almacenar las observaciones

for (i in 1:n_obs) {
  observations[i] <- current_state

  probs <- trans_matrix[states == current_state, ] # Obtener las probabilidades
de transición del estado actual

  current_state <- sample(states, 1, prob = probs) # Transición al siguiente
estado
}

return(observations)
}

# Generar las observaciones
observations <- generate_observations(trans_matrix, states, 100)

observations

```

```

## [1] "B" "B" "A" "A" "A" "C" "B" "A" "A" "A" "A" "A" "A" "A" "C" "B" "A"
"A" "B"

## [19] "A" "B" "B" "B" "C" "C" "C" "C" "A" "A" "A" "B" "B" "C" "B" "B"
"A" "B"

```

```

## [37] "B" "A" "A" "B" "A" "A" "A" "A" "A" "A" "B" "B" "A" "B" "B" "B"
"A" "B"

## [55] "B" "B" "A" "A" "A" "B" "B" "B" "C" "B" "C" "C" "C" "A" "B" "B"
"B" "B"

## [73] "B" "C" "C" "B" "C" "C" "B" "B" "A" "A" "A" "A" "B" "C" "C" "A"
"A" "A"

## [91] "B" "B" "A" "A" "A" "A" "A" "A" "A" "A" "C"

# Función para estimar la matriz de transición
estimate_transition_matrix <- function(observations, states) {
  n_states <- length(states)
  trans_matrix_est <- matrix(0, nrow = n_states, ncol = n_states)

  for (i in 1:(length(observations) - 1)) {
    current_state <- observations[i]
    next_state <- observations[i + 1]

    current_index <- which(states == current_state)
    next_index <- which(states == next_state)

    trans_matrix_est[current_index, next_index] <- trans_matrix_est[current_index,
next_index] + 1
  }
}

```

```

trans_matrix_est <- trans_matrix_est / rowSums(trans_matrix_est)

return(trans_matrix_est)
}

# Estimar la matriz de transición
estimated_trans_matrix <- estimate_transition_matrix(observations, states)
estimated_trans_matrix

##      [,1] [,2] [,3]
## [1,] 0.6590909 0.2727273 0.06818182
## [2,] 0.3157895 0.5000000 0.18421053
## [3,] 0.1764706 0.3529412 0.47058824

```

Observemos que la generación de aleatorios pide un estado inicial, esto se debe a que por definición, los procesos markovianos se definen bajo la distribución condicional, note que dicha distribución condicional si se comporta como un vector aleatorio aleatoria multinomial del tamaño del conjunto de estados de la cadena, esta relación la puede ver en (Zhou, 2009)

Otra perspectiva más ajustada a la realidad y de interés es la exploración del método bayesiano, este trabajo solo considera inferencia clásica sin embargo se puede adaptar la metodología propuesta por (Kelter, 2020). Se destaca la naturaleza secuencial de la estadística bayesiana en el aprendizaje estadístico, según se discute en (Waldmann, 2022). Este enfoque se plantea como un trabajo futuro de gran interés, ya que integra tres aspectos fundamentales de la estadística: procesos estocásticos, estadística matemática para la estimación de parámetros y

pruebas de hipótesis, así como la estadística bayesiana, que permite incorporar experiencia pasada para realizar predicciones con los datos. En ámbitos financieros, como el modelo de Markowitz, ya se han implementado mecanismos bayesianos, como se evidencia en (Chen, 2020). Como perspectiva para futuros estudios, se propone realizar un análisis comparativo entre el estadístico de significancia propuesto en este trabajo y el intervalo de credibilidad propuesto por la estadística bayesiana.

SIMULACIÓN

```
# Modelo M/M/1/N
```

```
####PRIMER PUNTO A)####
```

```
l<-.5
```

```
m<-.3
```

```
N<-5
```

```
r<-l/m
```

```
r
```

```
## [1] 1.666667
```

```
#La oferta de servicio supera a la oferta de demanda.
```

```
p0<-(1-r)/(1-(r^(6)))
```

```
p0
```

```
## [1] 0.03262621
```

```
####PRIMER PUNTO B)####
```

```
l<-.3
```

```
m<-.5
```

```
N<-5
```

```

r<-l/m
r
## [1] 0.6
#La oferta de servicio supera a la oferta de demanda.
p0<-(-1-r)/(1-(r^(6)))
p0
## [1] 0.4195757
#####
####SEGUNDO PUNTO####
set.seed(666)
TF<-5000
t<-seq(0,TF, length=1001)

L1 <- 5-(5*exp(-t/5))+(t*t)/10
l1 <- diff(L1)
length(l1)
## [1] 1000
dx <- NULL
for(i in 1:length(l1)) {
  dx[i] <- rpois(1,l1[i])
}
dx

```

Se propone el uso de métodos de estadística clásica para la estimación y simulación de una cadena de Markov en tiempo discreto con espacio de estados finito. Además, se abordan algunas dificultades asociadas a los métodos clásicos, lo cual subraya la necesidad de recurrir a la estadística bayesiana para la estimación. Dos motivos principales respaldan esta elección: en primer lugar, la naturaleza secuencial de los procesos estocásticos y el enfoque de aprendizaje que ofrece el método bayesiano.

Adicionalmente, se propone la realización de pruebas de significancia clásica con el objetivo de reducir el número de parámetros, cumpliendo así el criterio de parsimonia. Este principio sugiere que un modelo es considerado mejor si logra explicar de manera óptima los datos con la menor cantidad de parámetros posible. La aplicación de pruebas de significancia clásica permitirá evaluar la relevancia estadística de cada parámetro, contribuyendo a la simplificación del modelo y facilitando su interpretación. Esta estrategia no solo busca mejorar la eficiencia del modelo, sino también proporcionar una base más sólida para la toma de decisiones y la generalización de los resultados en diferentes contextos.

Capítulo 3. Ecuaciones con ruidos aleatorios

Las ecuaciones diferenciales estocásticas (EDE) desempeñan un papel fundamental en diversos campos científicos y aplicaciones prácticas. Su importancia radica en su capacidad para modelar y comprender sistemas complejos sujetos a la incertidumbre y variabilidad inherentes a fenómenos naturales o comportamientos aleatorios. A diferencia de las ecuaciones diferenciales clásicas (EDC) que asumen condiciones perfectamente deterministas, las EDE consideran la aleatoriedad, permitiendo así capturar y analizar la variabilidad observada en muchos procesos del mundo real. Este enfoque es crucial en campos como la física, la biología, la economía y la ingeniería financiera, donde los fenómenos involucrados están influenciados por factores aleatorios o fluctuaciones que deben ser comprendidos y modelados con precisión. Las EDE proporcionan herramientas poderosas para simular, predecir y tomar decisiones en sistemas complejos y dinámicos, ofreciendo una visión más realista y completa de los fenómenos que las rodean.

Según Exarchos, I., et al. (2018), las ecuaciones diferenciales estocásticas y la simulación son de gran importancia para resolver problemas de control óptimo estocástico. Estas ecuaciones permiten modelar y analizar sistemas dinámicos sujetos a incertidumbre y ruido. Dichas ecuaciones diferenciales han sido objeto de gran estudio en el área de la teoría de procesos estocásticos. Los últimos avances teóricos han intentado llevar las propiedades que tienen las ecuaciones diferenciales clásicas a las EDE.

La simulación de estas ecuaciones es otra parte interesante que se discutirá en este artículo. En particular, se expondrán e implementarán los métodos de Euler-Maruyama y de Runge-Kutta. Estos métodos generarán trayectorias de estado y evaluarán expectativas numéricamente. Estas herramientas son fundamentales para abordar problemas de control óptimo estocástico de manera eficiente y escalable, y han demostrado ser útiles en diversos campos, como finanzas, robótica y control de procesos. Según Suescun D, et al. (2021), el método de Runge-Kutta clásico se puede generalizar para las EDE de manera muy similar a como se generaliza el método de Euler-Maruyama con el método de Euler para las EDC.

Hay problemas en los cuales las EDC se adaptan muy bien. Sin embargo, estas ecuaciones no consideran el ruido de las observaciones al modelo. Por ejemplo, un experimento que se realiza en cursos básicos de física experimental es dejar caer

una pelota y ajustarle el modelo del movimiento uniformemente acelerado. En la teoría, se ajusta un modelo completamente caracterizado por la posición inicial, la velocidad inicial y la aceleración constante, siguiendo la ley que indica que una parábola solo pasa por tres puntos. Los puntos observados de la posición de la pelota en diferentes tiempos fluctúan alrededor del modelo teórico; este error se puede modelar mediante un modelo estocástico. Además, en situaciones que involucren modelos basados en EDC, puede ocurrir que por efectos ambientales o azarosos la realidad fluctúe alrededor del modelo. En estos casos, es recomendable considerar un modelo de EDE que sí considera dicho ruido.

Para un entendimiento profundo del trabajo con EDE se debe realizar un resumen acerca de los resultados más importantes del análisis estocástico, este resumen se barará en Castañeda, L. B., et al. 2012 y Rincón, L., 2006

El modelo matemático fundamental en la teoría de la probabilidad es el espacio de probabilidad, que se compone de un conjunto Ω , una colección A de subconjuntos de Ω , y una función $P: A \rightarrow [0, 1]$. El conjunto Ω representa los posibles resultados de un experimento aleatorio, mientras que A , llamada σ -álgebra, está constituida por subconjuntos de Ω que cumplen ciertas propiedades de cierre bajo operaciones específicas. Los elementos de A se denominan eventos o conjuntos medibles. La función P , conocida como medida de probabilidad, cumple con axiomas establecidos por A. Kolmogorov en 1933, los cuales indican que la probabilidad de Ω es igual a 1, la probabilidad de cualquier evento en F es mayor o igual a cero, y la función es aditiva sobre sucesiones infinitas de eventos disjuntos dos a dos.

La función de densidad de probabilidad de una distribución normal, representada por $f_X(x) := \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{\sigma^2}} dt$, describe la probabilidad de que una variable aleatoria continua tome un valor menor o igual a x , dados ciertos parámetros como la media y la desviación estándar. Esta fórmula es fundamental en estadística y teoría de probabilidades, ya que modela con precisión una variedad de fenómenos naturales y en aplicaciones prácticas, ofreciendo una representación gráfica de la distribución de probabilidad que sigue una curva simétrica en forma de campana, conocida como la curva de campana gaussiana.

En la teoría de la probabilidad, se emplea una definición de integral más amplia que la integral convencional de Riemann. Esta integral, conocida como integral de Lebesgue, desempeña un papel fundamental en la teoría de procesos estocásticos.

Para profundizar en este campo, se recomienda referirse a la obra de Resnick, S. (2019).

Un proceso estocástico consiste en una colección de variables aleatorias $\{X(t), t \in T\}$ definidas en un espacio de probabilidad (Ω, A, P) . Es importante notar que la letra t representa el índice y T denota el conjunto de índices. Esto se debe a que los procesos estocásticos, en su esencia, surgen al estudiar la evolución temporal de fenómenos aleatorios. Esta designación no presupone nada acerca de la numerabilidad de T .

Un proceso estocástico se puede entender como una función que depende de dos variables: el tiempo (representado por t) y los resultados posibles de un experimento (representados por ω). Al fijar un valor específico para ω , por ejemplo, para un ω , obtenemos una trayectoria del proceso, denotada como $X(\cdot, \omega)$. Esta trayectoria es la representación gráfica de cómo evoluciona el proceso en el tiempo para un resultado particular del experimento. Dicha función de dos variables puede verse como una colección de variables aleatorias indexadas en el tiempo, es decir, $\{X(t, \omega): t \in T, \omega \in \Omega\} = \{X(t): t \in T\}$ y cada $X(t)$ es (Ω, A, P) variable aleatoria.

Los procesos estocásticos tienen dos características importantes, la función de media y la función de autocorrelación si estos son finitos se dice que dichos procesos pertenecen al espacio $L^2(\Omega, A, P)$ Claro, aquí están las definiciones de la función media y la autocovarianza de un proceso estocástico. La función media de un proceso estocástico, denotada como $\mu(t)$, es la esperanza matemática o el valor esperado de la variable aleatoria en un tiempo específico t . Se define como $\mu(t) = E[X(t)]$, donde $E[\cdot]$ representa el operador de esperanza o valor esperado. Esta función proporciona información sobre la tendencia central o el valor promedio del proceso en el tiempo t . Por otro lado, para realizar regiones o intervalos de confianza se usa la función de auto covarianza de un proceso estocástico que mide la covariación entre dos valores del proceso en diferentes tiempos. Se define como $\sigma(s, t) = \text{Cov}[X_s, X(t)]$, donde $\text{Cov}[\cdot]$ representa la covarianza y X_s y $X(t)$ son dos variables aleatorias del proceso en los tiempos s y t , respectivamente. Esta medida cuantifica cómo varían conjuntamente dos instantes diferentes del proceso, proporcionando información sobre la relación de dependencia entre el proceso en diferentes puntos temporales.

Hay un proceso en particular que es de mucho interés para las ecuaciones diferenciales estocásticas dicho modelo es la caminata aleatoria. Una caminata

aleatoria centrada en c es un proceso estocástico discreto que describe el cambio secuencial en una variable a través de pasos sucesivos aleatorios. Sean $Y(1), \dots, Y(n)$ variables aleatorias independientes e idénticamente distribuida (IID) con esperanza cero y varianza finita, se define recursivamente $\{X(t)\}$ como:

- $X(0) = c$
- $X(n) = X(n-1) + Y(n)$

Si se toma la partición de un intervalo $[0, t]$ y se modela una caminata aleatoria, cuando la norma de dicha partición tiende a cero el proceso estocástico recibe el nombre de movimiento browniano (MB), y se puede demostrar que es un proceso con incrementos independientes y estacionarios, por otro lado se puede verificar que sin importar la distribución de las $Y(t)$ si el segundo momento es finito, el proceso tiene distribuciones finito dimensionales gaussianas. El movimiento Browniano fue introducido por el botánico Brown en 1828 ver Brown, R., 1828 donde explica el movimiento de partículas sumergidas en un líquido, ahora en 1905 Einstein lo usó en la teoría de la cinética molecular. posteriormente en 1923 Wiener demuestra la existencia matemática de dicho proceso realizando la partición diádica que usó Lebesgue para construir la integral sobre los reales. La definición matemática de este proceso es:

Un movimiento Browniano estándar unidimensional es un proceso estocástico $\{B(t) : t \geq 0\}$ que cumple con las siguientes propiedades:

- a) $B(0) = 0$ casi con certeza.
- b) Las trayectorias $t \rightarrow B(t)(\omega)$ son continuas para todo $\omega \in \Omega$.
- c) El proceso tiene incrementos independientes.
- d) La variable $B(t) - B(s)$ tiene distribución normal con media 0 y varianza $t - s$ para $0 \leq s < t$.

Este proceso es fundamental ya que constituye la variable integradora en una EDE.

Una ecuación diferencial estocástica (EDE) es una ecuación que involucra tanto términos deterministas como aleatorios. Se define $dX(t) = f(t, X(t)) dt + g(t, X(t)) dW(t)$

Donde:

- $X(t)$ es la función desconocida que se desea encontrar.
- $f(t, X(t))$ representa la derivada parcial de $X(t)$ respecto al tiempo t y la función $X(t)$.
- $g(t, X(t))$ es la derivada parcial de $X(t)$ respecto a la función estocástica $W(t)$, que es un proceso de Wiener o movimiento Browniano.
- dt es el diferencial determinista.
- $dW(t)$ es el diferencial estocástico, que representa la variación aleatoria en el tiempo t .

El movimiento Browniano geométrico es un tipo específico de proceso estocástico que se deriva del movimiento Browniano estándar. Se define mediante la ecuación diferencial estocástica: $dX(t) = \mu X(t) dt + \sigma X(t) dW(t)$

Donde:

- $X(t)$ representa el valor del activo o proceso en el tiempo t .
- μ es la tasa de crecimiento esperada.
- σ es la volatilidad.
- $dW(t)$ es un proceso de Wiener o movimiento Browniano, que describe la variación aleatoria en el tiempo t .

La característica principal del movimiento Browniano geométrico es que el crecimiento no es lineal, sino exponencial, lo que significa que el cambio en $X(t)$ no es proporcional a $X(t)$, sino que depende también del valor actual $X(t)$. Este modelo se utiliza ampliamente en finanzas para modelar el crecimiento de activos financieros, como acciones o inversiones, donde el rendimiento se considera como una tasa de interés compuesta continua, también es usado para modelar crecimientos poblacionales y en general para modelar crecimientos exponenciales.

Otra ecuación que será de importancia en este artículo es el puente browniano cuya EDE en el intervalo $[0,1]$ se define como:

$$dX(t) = (X(t)/(1-t))dt + dB(t)$$

Donde:

- $X(t)$ es el proceso estocástico en el tiempo t .
- $dB(t)$ es el diferencial estocástico (cambio aleatorio) en el tiempo t .
- La ecuación implica una dependencia en $X(t)$ y está definida en el intervalo de tiempo $[0,1]$.

En las Ecuaciones Diferenciales Estocásticas (EDE), al igual que en las Ecuaciones Diferenciales Ordinarias (EDO), existe un teorema de existencia y unicidad que establece condiciones bajo las cuales una solución única existe y es única en un sentido específico.

Para EDE, la noción de unicidad y existencia se relaciona con la continuidad de las funciones involucradas, así como las propiedades de acotación y suavidad. La existencia de soluciones puede garantizarse bajo ciertas condiciones en las funciones de deriva y difusión.

Para el Puente Browniano y el Movimiento Browniano, es posible verificar que cumplen con propiedades de regularidad y suavidad. Además, en muchos casos, se puede demostrar la existencia y unicidad de soluciones para las EDE asociadas a estos procesos estocásticos. Esto se logra mediante técnicas matemáticas como la teoría de martingalas, cálculo estocástico y teoremas de existencia y unicidad adaptados para EDE específicas.

En muchos casos, encontrar soluciones analíticas para EDE puede ser extremadamente desafiante o incluso imposible. Debido a esta dificultad, se recurre a métodos numéricos para aproximarse a las soluciones y encontrar posibles trayectorias de estos procesos estocásticos. Al igual que en las EDC, donde los métodos numéricos son esenciales para obtener soluciones aproximadas en casos complejos o no lineales, en las EDE estos métodos son fundamentales. Algunos de los métodos numéricos comunes utilizados para resolver EDE incluyen: Método de Euler-Maruyama y Métodos de Runge-Kutta Estocásticos. Adaptaciones de los métodos de Runge-Kutta utilizados para resolver EDO, extendidos para manejar términos estocásticos. Estos métodos numéricos ayudan a calcular trayectorias aproximadas y proporcionan soluciones aproximadas para EDE cuando no es factible encontrar soluciones analíticas.

Demos inicio con el método de Euler Maruyama: El método de Euler-Maruyama aproxima la solución de esta EDE utilizando pasos discretos de tiempo Δt . La fórmula del método de Euler-Maruyama se expresa de la siguiente manera:

$$X(n+1) = X(n) + f(t_n, X_n) \Delta t + g(t_n, X_n) \Delta W_t$$

donde:

- X_n es la aproximación de la solución en el tiempo t_n .
- Δt es el tamaño del paso de tiempo.
- ΔW_t es la variación del proceso de Wiener entre los tiempos t_n y $t(n+1)$, es decir, $\Delta W_t = W_{t(n+1)} - W(t_n)$
- $\{W(t)\}$ es un movimiento browniano.

El método de Euler-Maruyama es un enfoque de primer orden para aproximar soluciones de EDE y se basa en una aproximación lineal de la EDE a través de pequeños intervalos de tiempo. Si bien es sencillo de implementar, en ocasiones puede generar errores significativos en situaciones donde la EDE exhibe comportamientos no lineales o donde se requiere una alta precisión en la aproximación.

El método de Runge-Kutta Estocástico de orden dos utiliza dos etapas para calcular la aproximación de la solución en un paso de tiempo Δt . La fórmula del método de Runge-Kutta Estocástico de orden dos se puede expresar como:

$$X(n+1) = X(n) + a_1 * k_1 + a_2 * k_2$$

donde:

- X_n es la aproximación de la solución en el tiempo t_n .
- k_1 y k_2 son aproximaciones intermedias que se calculan de la siguiente manera:
 - $k_1 = f(t_n, X_n) \Delta t + g(t_n, X_n) \Delta W_n$
 - $k_2 = f(t_n + c \Delta t, X_n + c k_1) \Delta t + g(t_n + c \Delta t, X_n + c k_1) \Delta W_n$
- a_1 y a_2 son coeficientes asociados a las aproximaciones k_1 y k_2 .

- c es un parámetro en el intervalo $[0,1]$ que determina la propagación de la solución.

Este método mejora la precisión del método de Euler al considerar una corrección en el paso intermedio utilizando la derivada en el punto medio. Aunque más complejo que el método de Euler, el método de Runge-Kutta Estocástico de orden dos proporciona una mejor aproximación para ciertas EDE que exhiben comportamientos no lineales.

En primer lugar, se llevaron a cabo 100 simulaciones del Puente Browniano con volatilidades de uno, cuatro y cero punto uno, desde el punto $(1,1)$ al punto $(10,9)$. Estas simulaciones se muestran en la gráfica resaltadas en color azul. Además, se estimará un intervalo de confianza del 95%. Esta región se obtiene considerando, para cada punto de la simulación, el cuantil del 2.5% y del 97.5% para las tres volatilidades mencionadas.

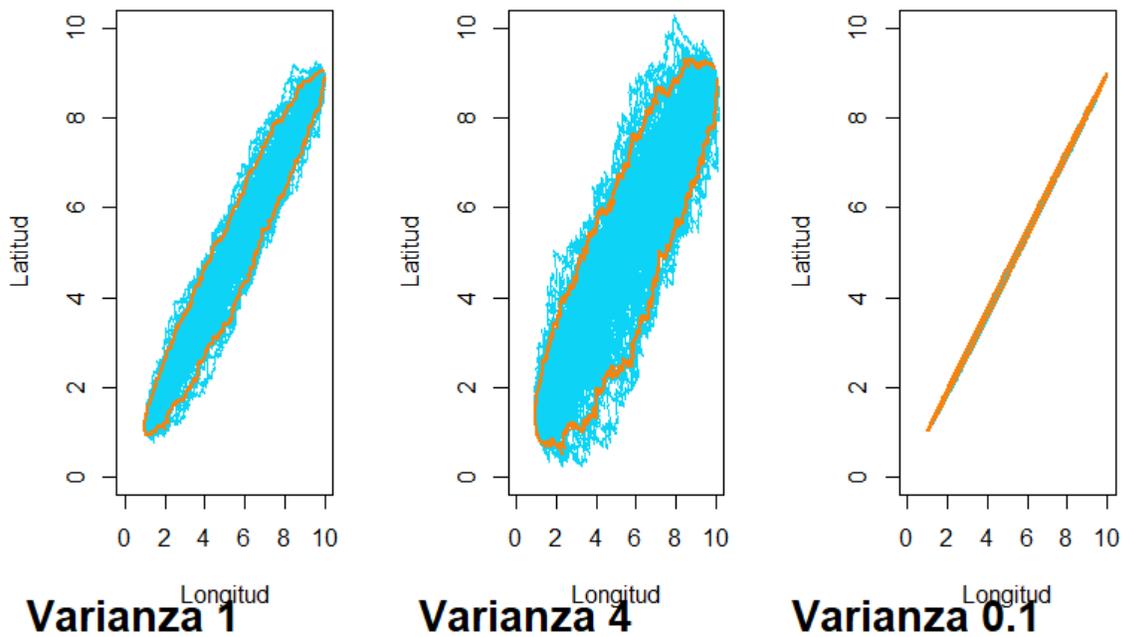


Figura 1. Simulación de Puentes Estocásticos desde $(1,1)$ hasta $(10,9)$ con Diferentes Parámetros de Volatilidad Utilizando Euler-Maruyama

La simulación representa un movimiento migratorio desde un punto inicial hacia otro. Se observa que cuando los individuos parten desde un punto fijo, la distancia entre ellos no es considerable. Asimismo, al acercarse al nuevo punto, la variabilidad tampoco es significativa. Por otro lado, se nota que el modelo clásico de segmento de recta muestra una menor variabilidad, lo que destaca una de las ventajas de utilizar Ecuaciones Diferenciales Estocásticas (EDC), ya que consideran la volatilidad. En este caso, se muestran tres movimientos migratorios con la misma tendencia. Sin embargo, el Puente Browniano permite considerar la volatilidad alrededor de la trayectoria media, a diferencia del modelo clásico.

Parece que tu solicitud está incompleta al mencionar "Las simulaciones por el método de Runge Kutta de orden 2 son:

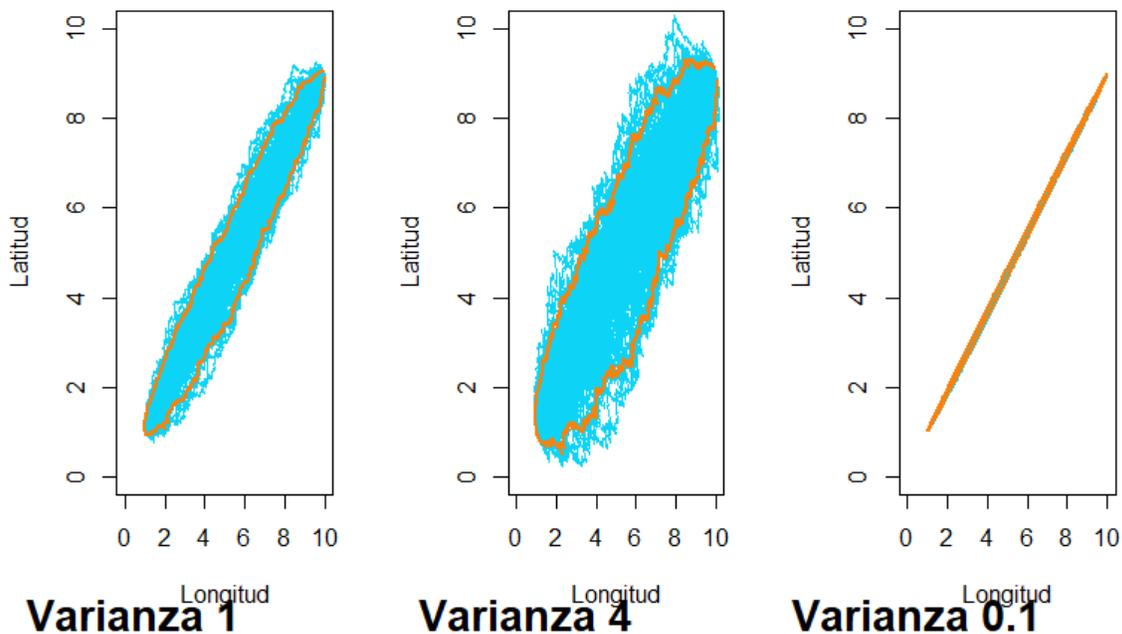


Figura 2. Simulación de Puentes Estocásticos desde (1,1) hasta (10,9) con Distintos Parámetros de Volatilidad Utilizando Runge-Kutta de Orden 2

Se ha observado en la Figura 2. una ausencia significativa de diferencias entre ambos métodos de simulación. Esto se atribuye a la ejecución de 1000 iteraciones en cada una de las 100 simulaciones, así como a la relativa simplicidad de la

Ecuación Diferencial Estocástica (EDE). A pesar de esto, se ha constatado que el método de Runge-Kutta de orden 2 presenta una convergencia más rápida en comparación con el método de Euler-Maruyama.

En vista de la falta de diferencias significativas entre los dos métodos de simulación, especialmente cuando se trata de EDE 'sencillas', se ha decidido emplear el método de Euler-Maruyama para simular el Movimiento Browniano Geométrico y para calcular los límites de confianza del 95%.

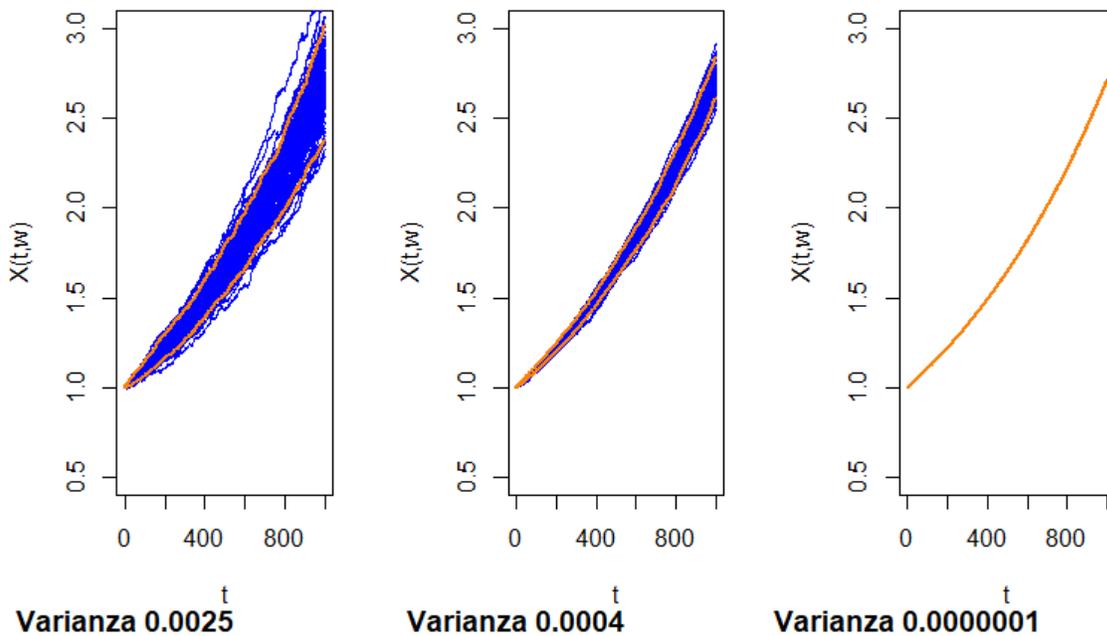


Figura 3. Simulación de Movimientos Brownianos Geométricos con Diferentes Parámetros de Volatilidad mediante el Método de Euler-Maruyama

Se observa en la Figura 3. que al igual que ocurre con el Puente Browniano, cuando la varianza es pequeña, el modelo tiende al clásico. Sin embargo, el Movimiento Browniano Geométrico permite modelar la volatilidad y generar regiones de confianza para la trayectoria media. Estos modelos se aplican comúnmente en el estudio de crecimientos poblacionales o en el modelado de activos financieros. No

obstante, es crucial tener en cuenta que existen momentos en los que la tasa de crecimiento puede ser tanto positiva como negativa. Para una comprensión más profunda y adecuada del Movimiento Browniano Geométrico, se recomienda revisar el trabajo de Soriano, A., et al. 2022

En primer lugar, se demostró que las ecuaciones diferenciales estocásticas (EDE) son útiles para modelar la volatilidad alrededor del modelo clásico. Esta ventaja es notable ya que permite establecer regiones de confianza, lo que mejora la apreciación del modelo. También se resaltó que, al igual que en las Ecuaciones Diferenciales Clásicas (EDC), existen métodos estocásticos para simular estas EDE.

Además, se presentó la utilización de los puentes brownianos como un posible modelo para el movimiento de animales. Se destacó que cuando el parámetro de volatilidad tiende a ser pequeño, el modelo de EDE converge al modelo de EDC. Es importante señalar que este artículo solo ofrece una perspectiva sobre dos tipos de ecuaciones diferenciales estocásticas. Como trabajo futuro, resulta interesante explorar otros modelos de EDE.

Para futuras investigaciones, se propone examinar el comportamiento de otras EDE y encontrar métodos adecuados que permitan estimar sus parámetros, lo que facilitaría la generación de regiones de confianza mediante simulaciones. También se sugiere incluir un análisis sobre ecuaciones diferenciales estocásticas multivariadas, describiendo las características de sus regiones de confianza y sus ventajas en comparación con sistemas de ecuaciones clásicos. Por último, se recomienda aplicar estos modelos a conjuntos de datos del mundo real, teniendo en cuenta cada supuesto y contexto específico.

Referencias

- Andrews, D. W., & Soares, G. (2010). Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica*, 78(1), 119-157. <https://doi.org/10.3982/ECTA7502>
- Bonkhoff, A. K., Hope, T., Bzdok, D., Guggisberg, A. G., Hawe, R. L., Dukelow, S. P., ... & Bowman, H. (2020). Bringing proportional recovery into proportion: Bayesian modelling of post-stroke motor impairment. *Brain*, 143(7), 2189-2206. <https://doi.org/10.1093/brain/awaa146>
- Castañeda, L. B., Arunachalam, V., & Dharmaraja, S. (2012). Introduction to probability and stochastic processes with applications. John Wiley & Sons. <https://doi.org/10.1002/9781118344972>
- Chen, S. D., & Lim, A. E. (2020). A Generalized black–litterman model. *Operations research*, 68(2), 381-410. <https://doi.org/10.1287/opre.2019.1893>
- Kelter, R. (2020). Bayesian alternatives to null hypothesis significance testing in biomedical research: a non-technical introduction to Bayesian inference with JASP. *BMC Medical Research Methodology*, 20(1), 1-12. <https://doi.org/10.1186/s12874-020-00980-6>
- Li, M., & Liu, X. (2020). Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique. *International Journal of Control, Automation and Systems*, 18(6), 1581-1592. <https://doi.org/10.1007/s12555-019-0191-5>
- Lin, X., Zhang, G., & Wei, S. (2021). Velocity prediction using Markov Chain combined with driving pattern recognition and applied to Dual-Motor Electric Vehicle energy consumption evaluation. *Applied Soft Computing*, 101, 106998. <https://doi.org/10.1109/eSmarTA52612.2021.9515733>
- Linde, M., Tendeiro, J. N., Selker, R., Wagenmakers, E. J., & van Ravenzwaaij, D. (2023). Decisions about equivalence: A comparison of TOST, HDI-ROPE, and

- the Bayes factor. *Psychological Methods*, 28(3), 740.
<https://psycnet.apa.org/doi/10.1037/met0000402>
- Meyer, C. D. (1994). Sensitivity of the stationary distribution of a Markov chain. *SIAM Journal on Matrix Analysis and Applications*, 15(3), 715-728.
<https://doi.org/10.1137/S0895479892228900>
- Mo, J. (2022). Performance modeling of communication networks with Markov chains. Springer Nature. <https://doi.org/10.1109/TCOMM.2003.809787>
- Meyn, S., & Tweedie, R. L. (1993). Markov Chains and stochastic stability. En Springer eBooks. <https://doi.org/10.1007/978-1-4471-3267-7>
- Oda, Y., Fudaba, H., Neubig, G., Hata, H., Sakti, S., Toda, T., & Nakamura, S. (2015, November). Learning to generate pseudo-code from source code using statistical machine translation. In 2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE) (pp. 574-584). IEEE. <https://doi.org/10.1109/ASE.2015.36>
- Pegels, C. C., & Jelmert, A. E. (1970). An evaluation of blood-inventory policies: A Markov chain application. *Operations Research*, 18(6), 1087-1098.
<https://doi.org/10.1287/opre.18.6.1087>
- Rangel, J. (2023). Estimación de los parámetros de una cadena de Markov. Recuperado de <https://rpubs.com/JhonierRangel/1127722>
- Rangel, J. (2023). SIMULACIÓN CADENA DE MARKOV EN TIEMPO DISCRETO. Recuperado de <https://github.com/jrangelg/Artificial-intelligence/blob/main/CadenaPy.ipynb>
- Saghafian, S. (2018). Ambiguous partially observable Markov decision processes: Structural results and applications. *Journal of Economic Theory*, 178, 1-35.
<https://doi.org/10.1016/j.jet.2018.08.006>
- Sober, E. (1981). The principle of parsimony. *The British Journal for the Philosophy of Science*, 32(2), 145-156. <https://doi.org/10.2307/2413239>
- Styan, G. P., & Smith Jr, H. (1964). Markov chains applied to marketing. *Journal of Marketing Research*, 1(1), 50-55.
<https://doi.org/10.1177/002224376400100109>

- Waldmann, M. R., & Martignon, L. (2022, May). A Bayesian network model of causal learning. In Proceedings of the twentieth annual conference of the Cognitive Science Society (pp. 1102-1107). Routledge. <https://doi.org/10.4324/9781315782416-198>
- Wasserman, L., Ramdas, A., & Balakrishnan, S. (2020). Universal inference. *Proceedings of the National Academy of Sciences*, 117(29), 16880-16890. <https://doi.org/10.1073/pnas.1922664117>
- Yang, X., Zheng, X. Q., & Lv, L. N. (2012). A spatiotemporal model of land use change based on ant colony optimization, Markov chain and cellular automata. *Ecological Modelling*, 233, 11-19. <https://doi.org/10.1016/j.ecolmodel.2012.03.011>
- Zhang, X., Low, Y. M., & Koh, C. G. (2020). Maximum entropy distribution with fractional moments for reliability analysis. *Structural Safety*, 83, 101904. <https://doi.org/10.1016/j.strusafe.2019.101904>
- Zhou, H., & Lange, K. (2009). Composition Markov chains of multinomial type. *Advances in Applied Probability*, 41(1), 270-291. <https://doi.org/10.1239/aap/1240319585>
- Zucchini, W., & MacDonald, I. L. (2009). Hidden Markov models for time series. En Chapman and Hall/CRC eBooks. <https://doi.org/10.1201/9781420010893>
- Martínez Ortega, R. M., Tuya Pendás, L. C., Martínez Ortega, M., Pérez Abreu, A., & Cánovas, A. M. (2009). El coeficiente de correlación de los rangos de Spearman caracterización. *Revista Habanera de Ciencias Médicas*, 8(2), 0-0.
- Bastías, J. M., Cuadra, M., Muñoz, O., & Quevedo, R. (2013). Correlación entre las buenas prácticas de manufactura y el cumplimiento de los criterios microbiológicos en la fabricación de helados en Chile. *Revista chilena de nutrición*, 40(2), 161-168.
- Santabàrbara, J. (2019). Cálculo del intervalo de confianza para los coeficientes de correlación mediante sintaxis en SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), 1-14.
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, 126(5), 1763-1768.

<https://doi.org/10.1213/ane.0000000000002864>

- Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological methods*, 17(3), 399.
- Zientek, L. R., & Thompson, B. (2007). Applying the bootstrap to the multivariate case: Bootstrap component/factor analysis. *Behavior research methods*, 39(2), 318-325. <https://doi.org/10.3758/BF03193163>
- Fleiss, J. L., Cohen, J., & Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72(5), 323–327. <https://doi.org/10.1037/h0028106>
- Serna-Morales, J. K., Elorza, M. C. J., & Lopera-Gómez, C. M. (2024). Comparación de algunas estimaciones del t de Kendall para datos bivariados con censura a intervalo. *Ciencia En Desarrollo*, 15(1), 130–140. <https://doi.org/10.19053/01217488.v15.n1.2024.15586>
- Smania, G., & Jonsson, E. N. (2021). Conditional distribution modeling as an alternative method for covariates simulation: comparison with joint multivariate normal and bootstrap techniques. *CPT: pharmacometrics & systems pharmacology*, 10(4), 330-339. <https://doi.org/10.1002/psp4.12613>
- Gutiérrez, J. S. R. (2024). Ventajas del uso de las ecuaciones diferenciales estocásticas: puente browniano y movimiento browniano geométrico. *Multidisciplinary & Health Education Journal*, 6(1), 830-838.
- Gutierrez, J. S. R. (2024). Métodos inferenciales sobre cadenas de Markov en tiempo discreto con espacio de estados finito. In *Actas del II Congreso Internacional de Innovación, Ciencia y Tecnología INUDI-UH, 2024* (pp. 277-288). Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú. <https://doi.org/10.35622/inudi.c.02.15>
- Azzalini, A., & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(2), 367-389. <https://doi.org/10.1111/1467-9868.00391>
- Chen, A., Bengtsson, T., & Ho, T. K. (2009). A regression paradox for linear models: Sufficient conditions and relation to Simpson's paradox. *The American*

Statistician, 63(3), 218-225. <https://doi.org/10.1198/tast.2009.08220>

- Wulandari, D., Sutrisno, S., & Nirwana, M. B. (2021). Mardia's skewness and kurtosis for assessing normality assumption in multivariate regression. *Enthusiastic: International Journal of Applied Statistics and Data Science*, 1-6. <https://doi.org/10.20885/enthusiastic.vol1.iss1.art1>
- Shi, X., Jiang, Y., Du, J., & Miao, Z. (2024). An adaptive test based on Kendall's tau for independence in high dimensions. *Journal of Nonparametric Statistics*, 36(4), 1064-1087. <https://doi.org/10.1080/10485252.2023.2296521>
- Kranstauber, B., Kays, R., LaPoint, S. D., Wikelski, M., & Safi, K. (2012). A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology*, 81(4), 738-746.
- Exarchos, I., & Theodorou, E. A. (2018). Stochastic optimal control via forward and backward stochastic differential equations and importance sampling. *Automatica*, 87, 159-165.
- Suescun D, D., Ule D, G., & Rojas A, O. (2021). Runge-Kutta implicit stochastic of order 1.5 applied to the equations of point kinetics.
- Castañeda, L. B., Arunachalam, V., & Dharmaraja, S. (2012). *Introduction to probability and stochastic processes with applications*. John Wiley & Sons.
- Rincón, L. (2006). *Introducción a las ecuaciones diferenciales estocásticas*. UNAM. México.
- Resnick, S. (2019). *A probability path*. Springer.
- Brown, R. (1828). XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The philosophical magazine*, 4(21), 161-173.
- Soriano, A., & Loro, H. (2022). Un problema en Econofísica: Predicción de activos financieros mediante el Movimiento Browniano Geométrico, dentro del mercado bursátil.
- Gutiérrez, J. S. R. (2024). Ventajas del uso de las ecuaciones diferenciales estocásticas: puente browniano y movimiento browniano geométrico. *Multidisciplinary & Health Education Journal*, 6(1), 830-838.

- Gutierrez, J. S. R. (2024). Métodos inferenciales sobre cadenas de Markov en tiempo discreto con espacio de estados finito. In Actas del II Congreso Internacional de Innovación, Ciencia y Tecnología INUDI-UH, 2024 (pp. 277-288). Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú.
- Rangel, J., Suárez, L. A. A., Lara-Cristancho, U., Alazate, J., & Cabezas-Chacón, P. J. (2025). Alternativas para el análisis de correlación cuando no se cumple el supuesto de normalidad bivariante: Simulación y ejemplos en R. *Ciencia Latina Revista Científica Multidisciplinar*, 9(1), 11311-11327.

ISBN: 978-9942-33-910-2



Compás
capacitación e investigación