

Dependencia espacial y severidad extrema de pérdidas por desastres naturales en Costa Rica

CA0307 Estadística Actuarial II, Escuela de Matemática, Universidad de Costa Rica Sede Rodrigo Facio, San José, Costa Rica

Jose Andrey Prado Rojas C36174^{*1}, Joseph Romero Chinchilla C37006^{*2}
Holmar Adrian Rivera Castellón B86564^{*3}, Dixon Montero Hernández B99109^{*4}

**Estudiante de Ciencias Actuariales, Escuela de Matemática, Universidad de Costa Rica. San José, Costa Rica. Noviembre, 2025*

¹joseandrey.prado@ucr.ac.cr, ²joseph.romero@ucr.ac.cr, ³holmar.rivera@ucr.ac.cr,
⁴dixon.montero@ucr.ac.cr

Resumen

Costa Rica enfrenta de manera recurrente desastres naturales como inundaciones, deslizamientos y sismos, cuyos impactos económicos se concentran en un número reducido de eventos de gran magnitud. Este trabajo propone un marco probabilístico multivariado para cuantificar el riesgo financiero asociado a estos fenómenos, utilizando los registros de pérdidas económicas reportadas a la Comisión Nacional de Prevención de Riesgos y Atención de Emergencias (CNE) para el período 2005–2023 (de Emergencias (CNE), 2025). Tras un proceso de depuración y reclasificación de la base, se realiza un análisis exploratorio para caracterizar la distribución de las pérdidas por provincia, sector y categoría de desastre, identificando patrones espaciales y sectoriales consistentes con estudios previos sobre el impacto de los desastres en el país (MIDEPLAN, 2019; Ministerio de Hacienda, 2020).

El modelado de la severidad se aborda mediante un esquema semiparamétrico cuerpo–cola, combinando distribuciones clásicas de severidad y estimadores de densidad kernel en la parte central con modelos de teoría de valores extremos en la cola, tales como las distribuciones Burr, Pareto y generalizada de Pareto (GPD) (Embrechts et al., 1997; McNeil et al., 2015). Los parámetros se estiman bajo un enfoque bayesiano mediante el algoritmo de Metropolis–Hastings, lo que permite obtener distribuciones posteriores por combinación de provincia–sector y provincia–categoría (Robert & Casella, 2004). La dependencia entre dimensiones clave del riesgo se modela mediante cópulas bivariadas (Clayton, Gumbel, Frank, Gauss y t–Student), seleccionadas con base en criterios de información y medidas de concordancia no lineal (Joe, 2014; Nelsen, 2006; Patton, 2012).

Los resultados muestran que, aunque el comportamiento central de las pérdidas varía entre provincias, sectores y categorías, la región extrema presenta de forma sistemática colas pesadas, con umbrálares de inicio de comportamiento extremo y proporciones de cola comparables entre grupos. Asimismo, se observa una dependencia positiva de magnitud baja a moderada en varias combinaciones asociadas a eventos hidrometeorológicos y a los sectores infraestructura y social, indicando que las pérdidas elevadas tienden a ocurrir de manera conjunta en estas dimensiones. A partir de simulaciones conjuntas se estiman el Valor en Riesgo (VaR) y el Valor en Riesgo Condicional (CVaR) al 95 % y 99 %, evidenciando que una fracción importante del riesgo agregado se concentra en escenarios extremos conjuntos, con implicaciones directas para el diseño de estrategias de aseguramiento, inversión en infraestructura resiliente y mecanismos de financiamiento del riesgo de desastres en Costa Rica.

Palabras Clave: desastres naturales, teoría de valores extremos, cópulas, simulación de Monte Carlo, Valor en Riesgo, Costa Rica.

Introducción

En el mundo, los desastres naturales constituyen una de las principales fuentes de riesgo físico y económico, tanto por los daños directos a la infraestructura como por sus efectos sobre el desarrollo y las finanzas públicas. En el caso de Costa Rica, su localización en una zona de alta actividad sísmica y volcánica, aunada a la exposición recurrente a fenómenos hidrometeorológicos como inundaciones y deslizamientos, implica que los eventos extremos representen una carga relevante para el presupuesto gubernamental, al punto de requerir reasignaciones de recursos y la creación de mecanismos específicos de financiamiento para la prevención y la atención de emergencias (Ministerio de Hacienda, 2020).

Diversos estudios han documentado que estos eventos han provocado daños recurrentes en infraestructura, viviendas y actividades productivas, generando impactos fiscales y sociales significativos a lo largo de las últimas décadas (Ministerio de Hacienda, 2020). En particular, MIDEPLAN (2019) muestra que, para el período 1988–2018 y considerando los daños asociados a sismos y fenómenos hidrometeorológicos, las provincias costeras de Limón y Puntarenas concentran una proporción importante del costo económico total, con montos acumulados del orden de miles de millones de colones.¹ Cuando se analizan únicamente los eventos hidrometeorológicos, como lluvias intensas y sequías, Puntarenas, Limón, Guanacaste y San José figuran entre las provincias más afectadas (MIDEPLAN, 2019).

A nivel sectorial, la evidencia disponible indica que la infraestructura vial y el sector agropecuario se encuentran entre los más vulnerables, registrando pérdidas que superan los miles de millones de colones en el período de estudio, lo que refleja la dependencia del país de actividades fuertemente expuestas a las variaciones climáticas y a los eventos extremos (MIDEPLAN, 2019). Estos resultados subrayan la importancia de contar con herramientas cuantitativas que permitan evaluar de forma integrada la frecuencia y severidad de los desastres, así como su distribución espacial y sectorial, con el fin de apoyar la toma de decisiones en materia de inversión en infraestructura resiliente, seguros y políticas de reducción del riesgo (Ministerio de Hacienda, 2020).

Frente a este panorama, Costa Rica ha impulsado marcos institucionales y de política pública orientados a la gestión prospectiva y correctiva del riesgo, entre ellos el Plan Nacional de Gestión del Riesgo 2021–2025 liderado por la Comisión Nacional de Prevención de Riesgos y Atención de Emergencias (CNE). Paralelamente, diversas entidades han desarrollado módulos y estudios específicos para cuantificar las pérdidas históricas ocasionadas por fenómenos naturales, con el fin de dimensionar su efecto sobre el desarrollo y orientar decisiones de inversión y aseguramiento. Estos esfuerzos resaltan la importancia de pasar de una lógica reactiva, centrada en la atención de emergencias, a un enfoque preventivo apoyado en evidencia estadística y modelos de riesgo robustos.

En este contexto, el objetivo de este trabajo es desarrollar y aplicar un marco probabilístico multivariado que permita modelar de manera conjunta la severidad y la dependencia espacial y sectorial de las pérdidas económicas por desastres naturales en Costa Rica. En particular, se busca: (i) caracterizar la distribución de las pérdidas por provincia, sector y categoría de desastre; (ii) ajustar modelos semiparamétricos cuerpo-cola que capturen adecuadamente el comportamiento extremo de las pérdidas; (iii) modelar la dependencia entre dimensiones clave del riesgo mediante cópulas; y (iv) estimar medidas de riesgo conjunto como el VaR y el CVaR para apoyar la planificación financiera y la gestión del riesgo de desastres.

Metodología

Datos y depuración

Se utilizó la base de datos correspondiente a los registros oficiales de daños y pérdidas económicas reportados ante la Comisión Nacional de Prevención de Riesgos y Atención de Emergencias (CNE), disponible

¹Las cifras exactas pueden variar según la metodología de valoración empleada en cada estudio.

en su portal de datos abiertos (de Emergencias (CNE), 2025). Estos registros abarcan el período 2005–2023 e incluyen información sobre el tipo de evento, la localización geográfica, el sector afectado y los montos de pérdida reportados, lo que permite un análisis conjunto a nivel de provincia, sector económico y categoría de desastre. A partir de esta base se llevó a cabo un proceso de depuración y estandarización de la información, con el fin de asegurar la consistencia de las series y la comparabilidad entre eventos.²

El proceso de depuración consistió, en primer lugar, en la eliminación de observaciones con valores nulos o montos de pérdida iguales a cero, dado que el análisis se realizó en escala logarítmica y se trabaja con pérdidas económicas que, en algunos casos, alcanzan órdenes de magnitud de miles de millones de colones. Posteriormente, se identificó la presencia de montos muy pequeños que generaban inestabilidad en los modelos propuestos; por esta razón se excluyeron los registros con pérdidas inferiores a $e^{14} \approx 1,2$ millones de colones, al considerarse económicamente marginales frente a los eventos de mayor severidad. Adicionalmente, la base presenta descripciones muy detalladas del sector afectado y del tipo de desastre natural (por ejemplo, “Tormenta tropical en la zona húmeda del Caribe Sur”), por lo que se procedió a reclasificar y agrupar estas descripciones en un conjunto reducido y coherente de categorías de desastre y sectores, con el objetivo de facilitar el análisis estadístico y evitar problemas de clases excesivamente fragmentadas (MIDEPLAN, 2019).

Análisis exploratorio de datos

Para estudiar la severidad de las pérdidas según el sector afectado, se construyeron diagramas de caja del monto total de daño en escala logarítmica para los cinco sectores definidos en la base de datos: infraestructura, social, otros, hídrico y productivo (véase figura 1). Estos gráficos muestran que, aunque la mediana de las pérdidas resulta relativamente similar entre sectores, todos presentan distribuciones marcadamente asimétricas con colas derechas largas, lo que evidencia la coexistencia de numerosos eventos de baja severidad con pocos eventos de pérdida extrema. Destaca, no obstante, que los sectores de infraestructura y productivo exhiben medianas ligeramente superiores y una mayor concentración de valores atípicos altos, lo que sugiere una exposición relativamente mayor a pérdidas severas en comparación con los sectores social, hídrico y otros (MIDEPLAN, 2019).

Un comportamiento análogo se observa al analizar las pérdidas por provincia. Los diagramas de caja para Cartago, Heredia, Limón, Alajuela, Guanacaste, Puntarenas y San José indican medianas de pérdida en rangos comparables; sin embargo, se aprecian diferencias en la dispersión y en la longitud de las colas superiores (véase la figura 2). En particular, las provincias de Limón, Puntarenas y San José presentan colas más pesadas y una mayor cantidad de valores extremos, en consonancia con estudios previos que señalan una concentración de eventos de gran impacto en estas regiones costeras y del Valle Central (MIDEPLAN, 2019; Ministerio de Hacienda, 2020). De forma global, las distribuciones entre provincias pueden considerarse relativamente homogéneas en su parte central, pero con diferencias relevantes en la cola, aspecto clave para el análisis de riesgo extremo.

Al examinar las pérdidas por categoría de desastre (hidrometeorológica, geológica, hídrica y otras), se observa nuevamente un patrón de fuerte asimetría y colas derechas largas (figura 3). Las categorías hidrometeorológica y geológica concentran la mayoría de los eventos de mayor severidad, lo que es consistente con la exposición del país a fenómenos climatológicos extremos y a actividad sísmica recurrente. Estos hallazgos exploratorios motivan el uso de modelos de severidad con capacidad para capturar colas pesadas y comportamientos extremos.

²La base original del CNE se encuentra descrita en MIDEPLAN (MIDEPLAN, 2019), donde se detalla la metodología de registro y consolidación de daños.

Ajuste de las distribuciones marginales

El modelado de las pérdidas económicas se abordó mediante un esquema semiparamétrico de cuerpo-cola, en el que se distingue entre el comportamiento central de la distribución y su región extrema. Este enfoque es consistente con la teoría de valores extremos, que sugiere tratar explícitamente la cola de las distribuciones de severidad cuando los eventos extremos tienen un peso relevante en el riesgo agregado (McNeil et al., 2015). En primer lugar, se ajustaron modelos para el cuerpo de la distribución utilizando, por un lado, estimadores de densidad kernel y, por otro, distribuciones paramétricas clásicas de severidad. El uso combinado de ambos enfoques permite capturar formas flexibles en la parte central de la distribución sin perder la interpretabilidad y la estructura teórica que ofrecen los modelos paramétricos (Silverman, 1986).

En cuanto a los estimadores no paramétricos, se consideraron cuatro núcleos: Epanechnikov, exponencial, triangular y gaussiano. El kernel de Epanechnikov es óptimo en el sentido de mínima varianza integrada dentro de la clase de núcleos de soporte compacto y produce estimaciones suaves que reducen la influencia de observaciones muy alejadas (Silverman, 1986). El kernel gaussiano, de soporte infinito, genera densidades particularmente suaves y es robusto frente a pequeñas irregularidades en los datos, mientras que los núcleos triangular y exponencial permiten representar distribuciones con cambios más marcados de pendiente y cierta asimetría en la zona central. Estos estimadores resultan útiles cuando la forma empírica de la distribución no se ajusta bien a una familia paramétrica específica, pero se desea conservar una aproximación flexible a la densidad subyacente.

De forma complementaria, para el cuerpo de la distribución se ajustaron modelos paramétricos de tipo Fisk, lognormal, gamma y Weibull, todos ellos habituales en el modelado de montos de pérdida y tiempos de vida (Klugman et al., 2012). La distribución lognormal es apropiada cuando el logaritmo de las pérdidas presenta un comportamiento aproximadamente normal, lo que implica colas derechas moderadas y una clara asimetría positiva. Las distribuciones gamma y Weibull permiten una mayor flexibilidad en la forma de la cola y en la curvatura de la función de densidad, siendo especialmente útiles para datos positivos con diferentes grados de asimetría y concentración alrededor de un modo. Por su parte, la distribución de Fisk (también conocida como log-logística) se caracteriza por colas más pesadas que la lognormal, lo que la hace adecuada para representar situaciones en las que las pérdidas de gran magnitud tienen una probabilidad relativamente más alta que la que sugeriría un modelo lognormal (Klugman et al., 2012).

Para la cola de la distribución se emplearon distribuciones específicamente diseñadas para capturar comportamientos extremos: Pareto máximo, Pareto clásica, distribución generalizada de Pareto (GPD), Burr y Weibull, además de versiones de los estimadores kernel enfocadas en la región de mayores pérdidas. Las distribuciones de tipo Pareto y Burr se caracterizan por colas pesadas, en las que la probabilidad de observar pérdidas muy grandes decrece lentamente, propiedad coherente con la naturaleza catastrófica de los eventos analizados (Embrechts et al., 1997; McNeil et al., 2015). La distribución GPD, por su parte, surge como modelo límite para los excedentes sobre un umbral alto en el marco de la teoría de valores extremos, lo que la convierte en una opción natural para describir la cola de la distribución de pérdidas una vez fijado el umbral (Embrechts et al., 1997). El uso de la Weibull en la cola permite contrastar escenarios de cola menos pesada, mientras que los kernels proporcionan una referencia no paramétrica que ayuda a evaluar la adecuación de los supuestos paramétricos en la región extrema.

La estimación de los parámetros de cada distribución se realizó mediante un enfoque bayesiano usando el algoritmo de Metropolis–Hastings, lo que permitió obtener distribuciones posteriores para los parámetros de severidad en cada combinación de provincia–sector y de provincia–categoría (Robert & Casella, 2004). Este procedimiento se aplicó de manera sistemática a todas las familias consideradas, y la elección del modelo final para el cuerpo y la cola de cada marginal se basó en un proceso de selección en dos etapas: en primer lugar, se verificó la adecuación de los ajustes mediante pruebas de bondad de ajuste por medio de la prueba de Kolmogorov–Smirnov y, en segundo lugar, se compararon únicamente los modelos aceptados mediante criterios de información como AIC y BIC, seleccionando en cada caso la combinación de cuerpo

y cola que proporcionó el mejor compromiso entre calidad de ajuste y complejidad del modelo (Burnham & Anderson, 2002).

Modelado de dependencia mediante cópulas

Las distribuciones exploradas en la sección anterior motivan el uso de cópulas para modelar la dependencia entre dimensiones clave del riesgo, como provincia, sector y categoría de desastre. Las diferencias observadas en las colas de las distribuciones por provincia y sector sugieren que la correlación lineal resulta insuficiente para describir la estructura conjunta de las pérdidas, especialmente cuando interesa la ocurrencia simultánea de eventos extremos. En este contexto, el marco de cópulas permite separar el modelado de las distribuciones marginales de pérdidas de la estructura de dependencia que las vincula, proporcionando una descripción más flexible y realista del riesgo agregado (Joe, 2014; Nelsen, 2006).

A partir de las pérdidas estandarizadas por combinaciones de provincia-sector y provincia-categoría, se ajustaron diversas familias de cópulas bivariadas con el objetivo de capturar patrones de dependencia no lineal y, en particular, la dependencia en cola. En específico, se consideraron cópulas de tipo Clayton, Gumbel, Frank, Gauss y t-Student, las cuales permiten representar asimetrías en las colas inferiores o superiores, así como estructuras de dependencia más cercanas a la correlación elíptica clásica (Joe, 2014; Nelsen, 2006). Para cada par de variables se seleccionó la familia y los parámetros que minimizan criterios de información como el AIC, una vez verificada la adecuación básica del ajuste mediante herramientas gráficas y medidas de concordancia, tales como los coeficientes de Kendall y Spearman (Patton, 2012).

La estructura de dependencia estimada se utilizó posteriormente para generar simulaciones conjuntas de pérdidas, combinando las distribuciones marginales ajustadas con las cópulas seleccionadas. Sobre estas simulaciones se calcularon medidas de riesgo como el Valor en Riesgo (VaR) y el Valor en Riesgo Condicional (CVaR) a niveles de confianza del 95 % y 99 % para las combinaciones provincia-sector, provincia-categoría y sector-categoría, siguiendo las recomendaciones habituales en gestión de riesgo financiero y catastrófico (McNeil et al., 2015). Estas métricas permiten cuantificar el tamaño de las pérdidas potenciales bajo escenarios extremos, incorporando tanto la severidad marginal como la dependencia entre componentes del sistema, y ofrecen así insumos cuantitativos para la planificación de medidas de mitigación y de financiamiento del riesgo de desastres en Costa Rica.

Resultados y Discusión

Resultados del ajuste marginal

En esta sección se presentan los resultados del ajuste de las distribuciones marginales de severidad para algunas combinaciones representativas de provincia, sector y categoría de desastre. Para cada caso se evaluaron múltiples modelos para el cuerpo (Weibull, gamma, lognormal, Fisk y estimadores kernel) y para la cola (GPD, Burr, Pareto, Pareto máximo, lognormal en cola y kernels), seleccionando el modelo final con base en la prueba de Kolmogorov-Smirnov y en criterios de información como el AIC y el BIC.

A nivel de categorías, para los eventos hidrometeorológicos el cuerpo de la distribución fue mejor representado por un modelo Weibull, mientras que la cola se ajustó mediante una distribución de Burr. Este resultado es coherente con la presencia de una parte central relativamente flexible y una cola pesada, donde la probabilidad de observar pérdidas muy grandes decrece lentamente (Embrechts et al., 1997; McNeil et al., 2015).

En el ámbito espacial, los resultados por provincia muestran cierta heterogeneidad en la forma del cuerpo, pero un patrón relativamente consistente en la cola. Para Limón, el cuerpo se describió mejor con un estimador kernel de tipo tophat y la cola con una distribución GPD, mientras que para Puntarenas y San José el cuerpo fue capturado por modelos Weibull y kernel Epanechnikov, respectivamente, y la cola

nuevamente por distribuciones de tipo Burr o GPD. Estos hallazgos refuerzan la idea de que, aunque la parte central de la distribución puede variar entre provincias, la región extrema presenta sistemáticamente colas pesadas, consistentes con la naturaleza catastrófica de las pérdidas.

Desde la perspectiva sectorial, el sector infraestructura mostró un cuerpo ajustado por un estimador kernel tophat y una cola descrita por una distribución Burr, mientras que en el sector social el cuerpo se representó adecuadamente con un kernel gaussiano y la cola con una GPD. En ambos casos, la selección de modelos con colas pesadas indica que una fracción importante del riesgo económico se concentra en unos pocos eventos de gran magnitud, en línea con la evidencia documentada en la literatura de riesgo operativo y catastrófico (Gómez & Pérez, 2009; Mora & Ramírez, 2013). En conjunto, estos resultados justifican el uso de técnicas de teoría de valores extremos y un tratamiento explícito de la cola en la cuantificación del riesgo asociado a desastres naturales.

Tabla 1: Resumen de modelos ajustados para el cuerpo y la cola en diferentes categorías y provincias.

Caso	Umbral óptimo	Proporción cola	Modelo cuerpo	Modelo cola
Categoría hidrometeorológica	21.52	0.076	Weibull	Burr
Provincia Limón	20.98	0.167	KDE-tophat	GPD
Provincia Puntarenas	21.11	0.185	Weibull	Burr
Provincia San José	20.40	0.194	KDE-epanechnikov	GPD
Sector infraestructura	20.99	0.159	KDE-tophat	Burr
Sector social	20.13	0.183	KDE-gaussian	GPD

Fuente: Elaboración propia con datos del CNE 2005–2023.

La tabla 1 resume los principales hallazgos del ajuste marginal para las combinaciones representativas de provincias, sectores y categorías de desastre. En primer lugar, los umbrales óptimos seleccionados para separar el cuerpo y la cola se ubican en un rango estrecho —aproximadamente entre 20 y 21.5 en escala logarítmica— lo que sugiere que las pérdidas comienzan a exhibir comportamiento extremo a partir de montos comparables en todos los grupos analizados. Estos valores corresponden a pérdidas superiores al orden de cientos de millones de colones, magnitud consistente con los eventos de mayor severidad registrados por la CNE.

En segundo lugar, la proporción de observaciones que pertenece a la cola oscila entre 7 % y 19 %, siendo particularmente elevada en las provincias de San José, Puntarenas y Limón, así como en los sectores infraestructura y social. Estas proporciones relativamente altas reflejan la presencia de colas pesadas y la importancia de un número reducido de eventos de gran magnitud en la explicación del riesgo total. En términos prácticos, estos resultados indican que una fracción no despreciable del riesgo económico se concentra en eventos extremos, lo que refuerza la necesidad de utilizar herramientas basadas en teoría de valores extremos para su adecuada caracterización.

En cuanto a la forma del cuerpo de la distribución, se observa una mayor heterogeneidad: modelos paramétricos como la Weibull y la lognormal resultaron adecuados en algunos casos, mientras que en otros fue necesario recurrir a métodos no paramétricos basados en estimadores kernel. Esta variabilidad sugiere que la parte central de la distribución de pérdidas depende de las características propias de cada provincia o sector, tales como su exposición física, densidad poblacional o infraestructura existente. Sin embargo, en la cola se aprecia un patrón consistente: las distribuciones Burr y GPD fueron sistemáticamente seleccionadas como modelos óptimos, lo cual es coherente con la presencia de colas pesadas y con la evidencia empírica de eventos catastróficos de gran magnitud.

En conjunto, los resultados de la tabla 1 muestran que, aunque el comportamiento central de las pérdidas puede variar entre regiones, sectores y categorías, la estructura extrema es notablemente estable y está dominada por colas pesadas. Esto tiene implicaciones directas para la cuantificación del riesgo de desastres naturales, dado que los modelos que no incorporen explícitamente la cola tenderían a subestimar la

probabilidad y la severidad de eventos catastróficos. Por esta razón, el enfoque semiparamétrico cuerpo–cola y el uso de distribuciones como Burr y GPD se justifican plenamente dentro del marco de análisis adoptado en este estudio.

Resultados del ajuste de cópulas y medidas de riesgo conjuntas

Para cada tipo de relación (provincia–sector, provincia–categoría y sector–categoría) se ajustaron cópulas bivariadas y, para la cópula seleccionada en cada pareja, se estimó la correlación de Kendall τ junto con las medidas de riesgo conjunto VaR y CVaR al 95 % y 99 %. En términos de dependencia, la mayoría de las parejas presentan coeficientes de Kendall τ positivos de magnitud moderada, lo que indica que las pérdidas tienden a ser simultáneamente altas en las dimensiones analizadas. Las combinaciones que involucran la categoría hidrometeorológica y los sectores infraestructura y social muestran, en general, los valores más elevados de τ , evidenciando una estructura de riesgo fuertemente interrelacionada en estos componentes. Además, las familias de cópulas seleccionadas se distribuyen principalmente entre modelos con capacidad para capturar dependencia en cola (como Clayton o Gumbel) y modelos elípticos (como la cópula Gaussiana), lo cual sugiere que en varios casos la dependencia se intensifica cuando las pérdidas son extremas en ambas variables, un comportamiento coherente con la ocurrencia de eventos climatológicos severos que afectan simultáneamente a múltiples provincias y sectores (Joe, 2014; Nelsen, 2006).

A partir de las cópulas ajustadas se calcularon el Valor en Riesgo (VaR) y el Valor en Riesgo Condicional (CVaR) conjuntos al 95 % y 99 % para cada pareja. En las tablas completas, los escenarios que combinan la categoría hidrometeorológica con los sectores infraestructura y social, así como ciertas parejas provincia–sector (por ejemplo, provincias costeras altamente expuestas), presentan los niveles más elevados de VaR y CVaR. En muchos de estos casos se observa que el CVaR supera de manera significativa al VaR, especialmente al 99 %, lo que indica que, una vez superado el cuantil correspondiente, las pérdidas esperadas en la cola se incrementan de forma pronunciada. Este patrón es típico de distribuciones con colas pesadas y refuerza la importancia de considerar explícitamente tanto la estructura de dependencia como el comportamiento extremo en la evaluación del riesgo de desastres naturales (6; 5; McNeil et al., 2015).

Tabla 2: Resumen de cópulas seleccionadas y dependencia (coeficiente de Kendall τ).

Caso	Cópula elegida	τ
Alajuela – Hidrometeorológico	SC	0.37
Limón – Hidrometeorológico	C	0.20
San José – Hídrico	SC	0.26
Limón – Infraestructura	C270	-0.18
Hídrico – Hidrometeorológico	SC	0.44
Social – Geológico	N	0.28

Fuente: Elaboración propia con datos del CNE 2005–2023.

Nota: SC denota una cópula t–Student, C denota una cópula de Clayton, N denota una cópula Normal o Gaussiana y C270 corresponde a una rotación de 270 grados de la cópula de Clayton.

En la tabla 2 se observa que la mayoría de las parejas presentan coeficientes de Kendall τ positivos de magnitud baja a moderada, lo que indica una dependencia positiva entre las pérdidas conjuntas. Por ejemplo, las parejas Alajuela–Hidrometeorológico y Hídrico–Hidrometeorológico muestran valores de τ en torno a 0.37 y 0.44, respectivamente, lo que sugiere que los episodios de pérdidas elevadas en la provincia o el sector tienden a coincidir con pérdidas elevadas en la categoría de desastre correspondiente.

En contraste, algunas parejas presentan valores negativos de τ , como el caso Limón–Infraestructura, asociado a una cópula de tipo C270 (rotación de Clayton), donde $\tau \approx -0.18$. Esto indica una dependencia

negativa débil, en el sentido de que niveles altos de pérdida en una variable tienden a coincidir con niveles relativamente más bajos en la otra, al menos dentro del rango observado. La presencia simultánea de dependencias positivas y negativas pone de manifiesto la heterogeneidad en la forma en que distintos territorios y sectores responden a los eventos extremos.

En conjunto, estos resultados muestran que la estructura de dependencia no es homogénea: existe una agrupación de pares con dependencia positiva moderada, especialmente en combinaciones relacionadas con la categoría hidrometeorológica y los sectores infraestructura y social, y algunos casos puntuales de dependencia negativa. El uso de cópulas permite capturar estas relaciones de manera más adecuada que una simple correlación lineal, especialmente cuando interesa el comportamiento conjunto en las colas de la distribución de pérdidas (Joe, 2014; Nelsen, 2006).

Medidas de riesgo conjunto: VaR y CVaR

Tabla 3: VaR y CVaR conjuntos para parejas seleccionadas (escala logarítmica).

Caso	VaR ₉₅	CVaR ₉₅	VaR ₉₉	CVaR ₉₉
Limón – Hidrometeorológico	26.02	27.07	27.74	28.66
San José – Hídrico	25.22	26.39	27.11	28.16
Limón – Infraestructura	25.97	26.98	27.61	28.54
Social – Geológico	26.17	27.88	28.84	30.60

Fuente: Elaboración propia con datos del CNE 2005–2023.

En la tabla 3 se aprecia que, para todos los casos, el CVaR supera de manera sistemática al VaR, tanto al 95 % como al 99 %, lo cual refleja que las pérdidas esperadas condicionadas a superar el cuantil correspondiente son sustancialmente mayores que el propio umbral de VaR. Este efecto es especialmente notable en parejas asociadas a la categoría hidrometeorológica y al sector social, como SOCIAL–Geológico y Limón–Hidromet., donde las diferencias entre VaR y CVaR al 99 % son particularmente amplias.

En conjunto, estos resultados confirman la presencia de colas pesadas y ponen de manifiesto que una parte importante del riesgo económico se concentra en un reducido número de eventos extremos. Además, la combinación de colas marginales pesadas con dependencia positiva en las colas implica que los escenarios conjuntos de pérdida severa no son eventos excepcionales bajo el modelo, sino componentes relevantes del perfil de riesgo. Esto justifica el uso conjunto de VaR y CVaR en la evaluación del riesgo de desastres naturales y refuerza la necesidad de incorporar explícitamente la estructura de dependencia en los análisis de solvencia y planificación financiera (Gómez & Pérez, 2009; McNeil et al., 2015; Mora & Ramírez, 2013).

Conclusiones

Este estudio desarrolló un marco probabilístico multivariado para analizar la severidad y la dependencia espacial y sectorial de las pérdidas económicas asociadas a desastres naturales en Costa Rica, utilizando información de la CNE para el período 2005–2023. A partir de un proceso de depuración y reclasificación de la base de datos, se caracterizó la distribución de las pérdidas por provincia, sector y categoría de desastre, identificando patrones de asimetría marcada y colas derechas largas en todos los grupos analizados.

Los resultados del ajuste marginal mostraron que, si bien el comportamiento central de las pérdidas presenta cierta heterogeneidad entre provincias y sectores —reflejada en la necesidad de combinar modelos paramétricos y no paramétricos en el cuerpo de la distribución—, la región extrema está dominada por colas pesadas, adecuadamente descritas por distribuciones de tipo Burr y GPD. Los umbrales que delimitan la cola se concentraron en un rango estrecho en escala logarítmica, indicando que el comportamiento extremo se activa a partir de montos de pérdida comparables entre grupos. Asimismo, la proporción de

observaciones en la cola resultó relativamente alta en provincias y sectores clave como San José, Puntarenas, Limón, infraestructura y social, lo que evidencia la importancia de unos pocos eventos de gran magnitud en la explicación del riesgo total.

En cuanto a la dependencia, el análisis mediante cópulas bivariadas reveló coeficientes de Kendall τ positivos de magnitud baja a moderada en la mayoría de las parejas, especialmente en aquellas vinculadas con la categoría hidrometeorológica y los sectores infraestructura y social. Esto sugiere que las pérdidas elevadas tienden a ocurrir de manera conjunta en estas dimensiones, y que los eventos extremos pueden generar impactos simultáneos en múltiples provincias y sectores. Algunos casos puntuales de dependencia negativa ponen de manifiesto la heterogeneidad en la respuesta de diferentes territorios y sectores frente a los eventos extremos, aspecto que merece exploraciones futuras más detalladas.

Las medidas de riesgo conjunto VaR y CVaR al 95 % y 99 % confirmaron que una fracción importante del riesgo agregado se concentra en escenarios extremos conjuntos, con diferencias significativas entre VaR y CVaR, especialmente al 99 %. Esto es consistente con la presencia de colas pesadas y con la intensificación de la dependencia en las colas. Desde una perspectiva actuarial y de política pública, estos hallazgos resaltan la necesidad de incorporar explícitamente el comportamiento extremo y la dependencia espacial y sectorial en el diseño de instrumentos de aseguramiento, mecanismos de financiamiento del riesgo y estrategias de inversión en infraestructura resiliente.

Como líneas de trabajo futuro, sería relevante extender el análisis a modelos de cópulas de dimensión superior que permitan capturar de manera conjunta la dependencia entre múltiples provincias y sectores, así como integrar información de frecuencia de eventos para construir modelos de riesgo agregados más completos. Asimismo, la incorporación de escenarios de cambio climático y de proyecciones socioeconómicas podría contribuir a evaluar la evolución futura del riesgo de desastres y a diseñar estrategias de adaptación más robustas.

Agradecimiento

Al grupo le gustaría expresar su gratitud al profesor Maikol Solis y Esteban Bermudez por su colaboración durante el trabajo. Así como al asistente del curso y los compañeros.

Bibliografía

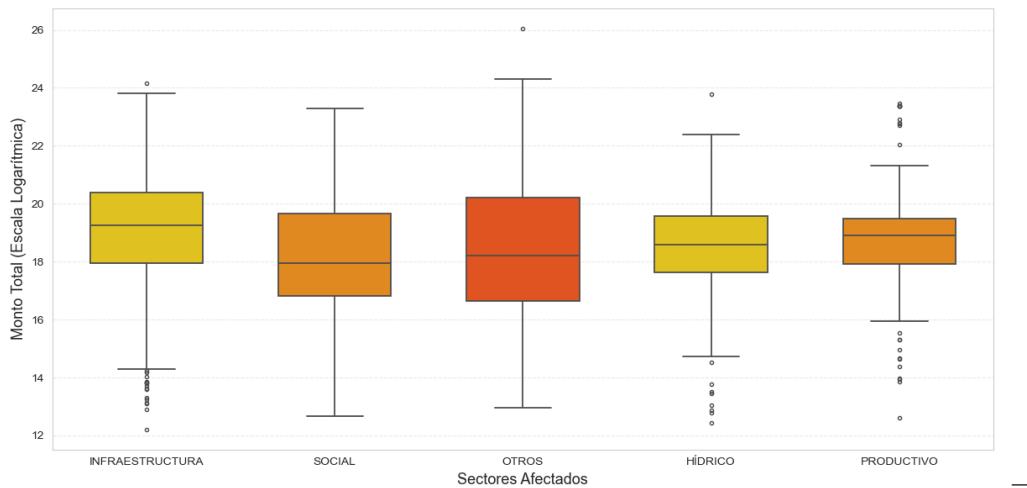
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2.^a ed.). Springer.
- de Emergencias (CNE), C. N. (2025). *Datos Abiertos* [Sitio oficial de transparencia — sección “Datos Abiertos”]. Consultado el 28 de noviembre de 2025, desde https://www.cne.go.cr/transparencia/datos_abiertos.aspx
- Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.
- Gómez, J., & Pérez, C. (2009). *Teoría del riesgo y aplicaciones actuariales*. Editorial Académica.
- Joe, H. (2014). *Dependence Modeling with Copulas*. CRC Press.
- Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss Models: From Data to Decisions* (4.^a ed.). Wiley.
- McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools* (2.^a ed.). Princeton University Press.
- MIDEPLAN, S., MAG. (2019, noviembre). *Costa Rica: Impacto de los Fenómenos Naturales para el período 1988-2018, por sectores, provincias, cantones y distritos: Compendio*. Ministerio de Agricultura y Ganadería. Consultado el 28 de noviembre de 2025, desde https://documentos.mideplan.go.cr/share/s/MQOaFN8jQTGMfuNRKR_IOg

- Ministerio de Hacienda. (2020, diciembre). *Informe de Riesgos: Desastres Naturales* (Informe de riesgos). Ministerio de Hacienda, Dirección de Crédito Público, Departamento de Política de Endeudamiento Público. Consultado el 28 de noviembre de 2025, desde <https://www.hacienda.go.cr/docs/InformeRiesgosFiscalesDesastresNaturales.pdf>
- Mora, L., & Ramírez, A. (2013). *Construcción de modelos de riesgo extremo*. Editorial Universitaria.
- Nelsen, R. B. (2006). *An Introduction to Copulas* (2.^a ed.). Springer.
- Patton, A. J. (2012). Copula-Based Models for Financial Time Series. En T. G. Andersen, R. A. Davis, J.-P. Kreiss & T. Mikosch (Eds.), *Handbook of Financial Time Series and Applications* (pp. 767-785). Springer.
- Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods* (2.^a ed.). Springer.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman; Hall.

Anexos

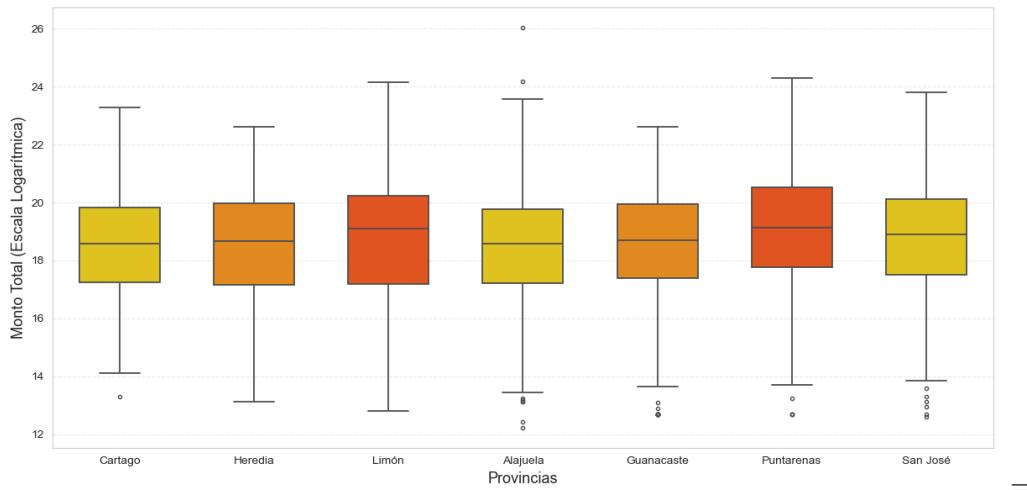
Enlace al repositorio de Github : <https://github.com/CA0307-II-2025/grupo-2>

Figura 1: Distribución del total en montos económicos de los sectores afectados.



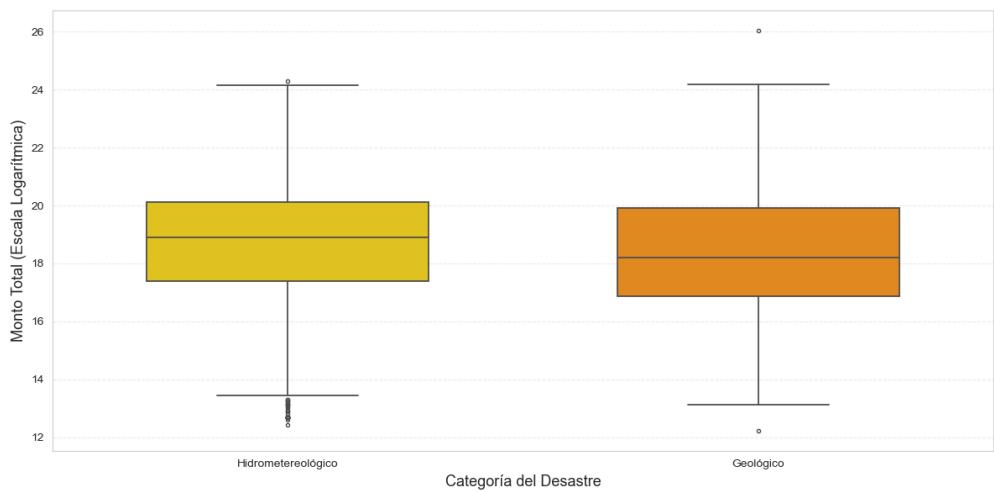
Fuente: Elaboración propia con datos del CNE 2005–2023.

Figura 2: Distribución del total en montos económicos de las provincias afectados.



Fuente: Elaboración propia con datos del CNE 2005–2023.

Figura 3: Distribución del total en montos económicos de las categorías de desastres naturales.



Fuente: Elaboración propia con datos del CNE 2005–2023.