

Una Distribución de Probabilidad Predictiva Bayesiana para la Magnitud Máxima Anual de los Sismos con Epicentro en Veracruz

A Bayesian Predictive Probability Distribution for the Annual Maximum Seismic Magnitude with Epicenter in Veracruz

Roberto Carlos Rivera-Flores ^a | Sergio Francisco Juárez-Cerrillo ^{b*}
Francisco Córdoba-Montiel ^c | Silvano Ulises Que-Salinas ^d
Rafael Torres-Orozco ^e

Recibido: 4 de septiembre de 2025.

Aceptado: 23 de marzo de 2026.

^a Universidad, Veracruzana. Xalapa, México. Contacto: zS21013261@estudiantes.uv.mx | ORCID: [0009-0005-2945-9725](https://orcid.org/0009-0005-2945-9725)

^b Observatorio Sismológico y Vulcanológico de Veracruz (OSV), Universidad Veracruzana. Xalapa, México. Contacto: sejuarez@uv.mx | ORCID: [0000-0002-5713-7644](https://orcid.org/0000-0002-5713-7644) *Autor para correspondencia.

^c Observatorio Sismológico y Vulcanológico de Veracruz (OSV), Universidad Veracruzana. Xalapa, México. Contacto: fcordoba@uv.mx | ORCID: [0000-0002-9111-2336](https://orcid.org/0000-0002-9111-2336)

^d Observatorio Sismológico y Vulcanológico de Veracruz (OSV), Universidad Veracruzana. Xalapa, México. Contacto: feanorkdt@gmail.com | ORCID: [0000-0002-1902-960X](https://orcid.org/0000-0002-1902-960X)

^e Observatorio Sismológico y Vulcanológico de Veracruz (OSV), Universidad Veracruzana. Xalapa, México. Contacto: rraftorres@uv.mx | ORCID: [0000-0003-2214-9829](https://orcid.org/0000-0003-2214-9829)

Cómo citar:

Rivera-Flores, R. C., Juárez-Cerrillo, S. F., Córdoba-Montiel, F., Que-Salinas, S. y Torres Orozco, R. (2026). Una Distribución de Probabilidad Predictiva Bayesiana para la Magnitud Máxima Anual de los Sismos con Epicentro en Veracruz. *UVserva*, (21), 65-88. <https://doi.org/10.25009/uvs.vi21.3171>

Resumen: En este trabajo modelamos las magnitudes máximas anuales de los sismos con epicentro en el estado de Veracruz de 1988 a 2024, con la Distribución del Valor Extremo Generalizada (DVEG) bajo un enfoque de inferencia estadística Bayesiana. Con este enfoque inferencial, los parámetros de la DVEG se modelan a su vez con distribuciones de probabilidad las cuales se actualizan con las $n = 37$ magnitudes máximas anuales registradas por el Servicio Sismológico Nacional durante el período de observación. El producto final de nuestro análisis es una distribución de probabilidad predictiva de la magnitud máxima anual. Con nuestra propuesta metodológica esperamos proporcionar un instrumento que apoye a la toma de decisiones sobre riesgo sísmico en la región.

Palabras clave: Distribución del Valor Extremo Generalizada; geovariable extrema; magnitud sísmica; períodos de retorno; predicción Bayesiana de extremos; riesgo sísmico; sismicidad local; Veracruz.

Abstract: *In this study we model the annual maximum magnitudes of those earthquakes with epicenter located in the state of Veracruz, from 1988 to 2024. We utilize the Generalized Extreme Value Distribution (GEVD) to these maximum magnitudes using the Bayesian approach. Within this approach, the GEVD parameters are modeled through probability distributions, which are updated using the $n = 37$ annual maximum magnitudes recorded by the National Seismological Service during the observation period. The final product of our analysis is a predictive probability distribution of the annual maximum magnitude. With this methodological proposal, we aim to provide a reliable tool to support decision-making regarding seismic risk in the region.*

Keywords: *Generalized Extreme Value Distribution; extreme geovariable; seismic magnitude; Bayesian prediction of extremes; return periods; seismic risk; local seismicity; Veracruz.*

Introducción

El estado de Veracruz se encuentra en una zona de alta actividad sísmica la cual se ve influenciada por su ubicación geográfica dentro de la interacción de las placas tectónicas de Norteamérica y del Caribe (Córdoba *et al.*, 2020). Estas condiciones generan un entorno propenso a movimientos sísmicos de diversas magnitudes y eventos históricos de gran magnitud que han causado daños de gran significancia en la región.

Córdoba (2016) explica cómo se combinan estas características en Veracruz para aumentar su vulnerabilidad sísmica. Algunas de estas características son una densidad poblacional significativa en zonas urbanas cercanas a fallas geológicas; infraestructuras que, en muchos casos, no cumplen con estándares sísmicos modernos; así como también condiciones socioeconómicas que limitan las capacidades de respuesta ante sismos extremos.

Las potenciales afectaciones catastróficas que los sismos de magnitud extrema pueden producir en la población y en su entorno hace relevante y vital al estudio de la distribución de probabilidad de la magnitud extrema sísmica. Sin embargo, el análisis estadístico tradicional generalmente se enfoca en tendencias promedio y no en los valores extremos de la magnitud sísmica.

Evaluar el riesgo sísmico plantea el desafío de analizar eventos poco frecuentes. Aquí es donde entran en acción las dos vertientes de nuestro análisis. Por un lado, usamos a la Teoría de Valores Extremos (TVE) ya que es una poderosa herramienta estadística diseñada para entender y modelar la cola superior de distribuciones de probabilidad, (Coles, 2001), Beirlant *et al.* (2006), Held y Sabanés-Bové (2014) y Dutfoy (2021).

Por otro lado, seguimos el enfoque de la inferencia Bayesiana debido a su característica de *aprender* sobre el comportamiento del fenómeno bajo a estudio conforme aumenta la disponibilidad de datos. En el caso de la sismicidad extrema, donde los datos de eventos extremos son escasos, este enfoque resulta muy conveniente. Con los métodos Bayesianos se combina lo que ya se sabe, el conocimiento a-priori, con la evidencia empírica la cual es proporcionada por los datos. El Teorema de Bayes se utiliza para actualizar el conocimiento a-priori con los datos y obtener conocimiento posterior.

En este trabajo construimos un modelo estadístico para evaluar el riesgo de eventos sísmicos extremos en Veracruz. En concreto, la pregunta que se responde en este trabajo es ¿Cuál es la probabilidad de que el sismo de mayor magnitud en cualquier año exceda una magnitud m ? Nuestra propuesta metodológica para estimar estas probabilidades consiste en aplicar la TVE desde una perspectiva Bayesiana.

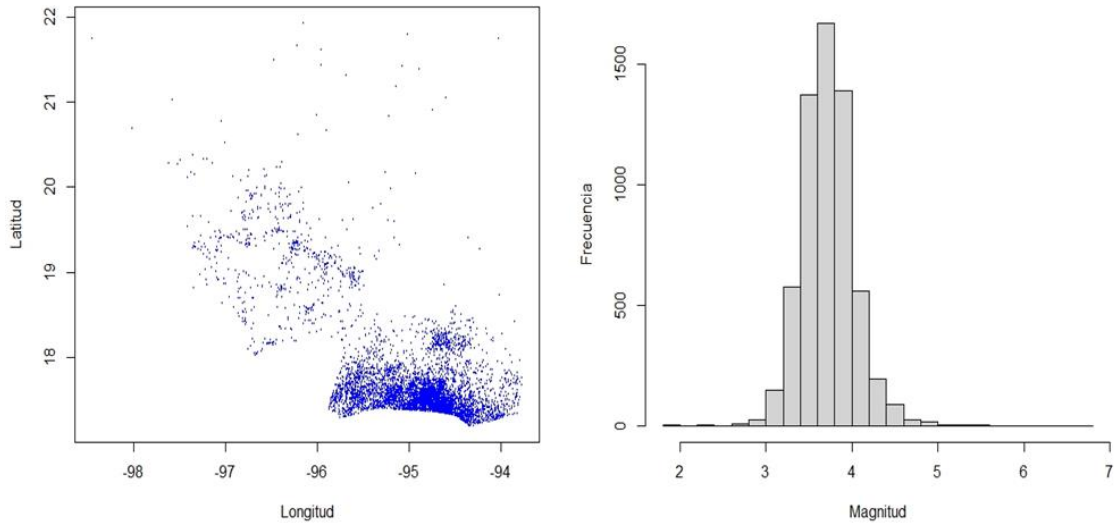
Algunas aplicaciones de la TVE en estudios de sismicidad son el trabajo pionero de Nordquist (1945); posteriormente aparecen los trabajos de Yegulalp y Kuo (1974), Burton (1979), Al-Abbasi y Fahmi (1985). Y más recientemente tenemos a Kijko y Singh (2011), Zöller *et al.* (2014), Pavlenko (2017), Maruyama (2020), García-Bustos *et al.* (2020), Ma *et al.* (2021), Al-Heety *et al.* (2024), Juárez-Cerrillo *et al.* (2024) y Rivera (2025).

El resto del trabajo está estructurado de la siguiente forma. En la Sección 1, presentamos el problema de investigación junto con los datos que se analizan. En la Sección 2 presentamos los aspectos metodológicos que seguimos para construir el modelo predictivo. A continuación, en la Sección 3, presentamos los resultados de la modelación Bayesiana, así como la distribución predictiva que proponemos. Los detalles metodológicos y computacionales en extenso se pueden consultar en Rivera (2025). Finalmente, en la Sección 4, concluimos con algunos comentarios finales.

1. El problema de investigación y los datos

La **Figura 1** muestra las localizaciones geográficas de los 6,088 sismos con epicentro dentro de los límites geográficos del estado de Veracruz, registrados por el Servicio Sismológico Nacional (2024). La distribución de las magnitudes de estos sismos se muestra también en la **Figura 1**.

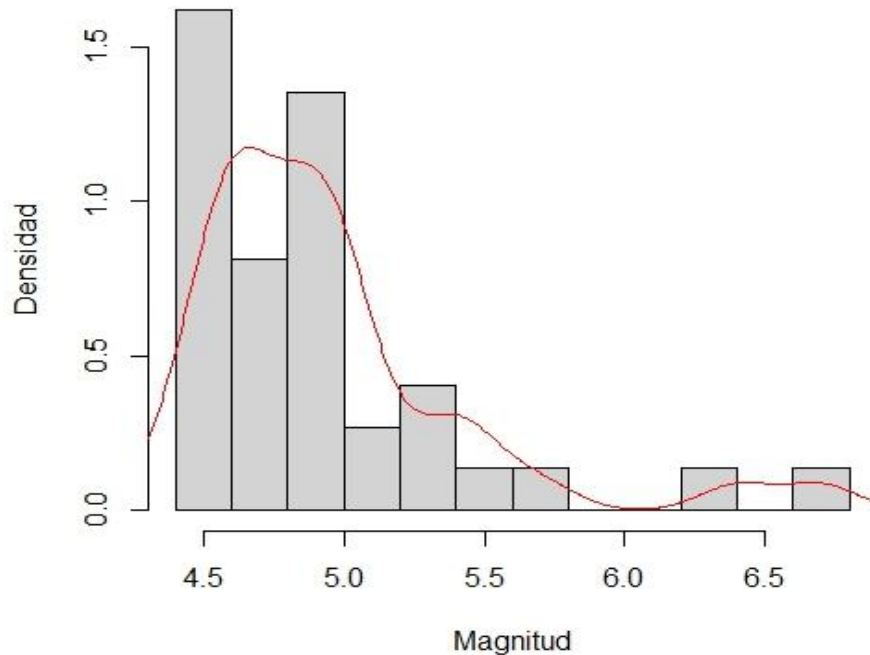
Figura 1
Epicentros y magnitudes de los sismos registrados en Veracruz de 1988 al 2024



Fuente: Elaboración propia con datos del Catálogo del SSN.

Con los datos mostrados en la **Figura 1** nos planteamos el problema de estimar las magnitudes más grandes que se espera ocurran en el estado de Veracruz.

Figura 2
Magnitudes máximas anuales, junto con un estimador no paramétrico de densidad



Fuente: Elaboración propia con datos del Catálogo del SSN.

En la terminología de la TVE, si X_{max} denota a la magnitud máxima de un sismo sobre un período de tiempo fijo (un año, por ejemplo), el nivel de retorno x_p asociado a un período de retorno $1/p$ es la magnitud x_p que se espera se exceda por X_{max} en promedio una vez cada $1/p$ períodos de tiempo. Es decir, el nivel x_p se excede por X_{max} en cualquier período de tiempo con probabilidad p : $P(X_{max} > x_p) = p$.

Para responder al planteamiento anterior, el período total de observación, que va de 1988 a 2024, lo dividimos en períodos de longitud de un año. De cada período (año) se seleccionaron las magnitudes máximas anuales, lo que resulta en 37 observaciones de las magnitudes máximas anuales de los años 1988, 1989, 1990, hasta el 2024. En la **Figura 2** se muestran a estas 37 magnitudes junto con el estimador no paramétrico kernel de la densidad. Véase Lasocki (2021) para un desarrollo detallado de este estimador en estudios de sismicidad. En el Apéndice A se despliega al conjunto de datos.

La **Tabla 1** muestra las estadísticas descriptivas de las magnitudes máximas anuales.

Tabla 1
 Estadísticas descriptivas de las $n = 37$ magnitudes máximas anuales

Mínimo	Q_1	Mediana	Media	Q_3	Máximo
4.4	4.6	4.9	4.95	5.0	6.7

Fuente: Elaboración propia.

2. Metodología

2.1. Predicción Bayesiana

Denotemos con $X = (X_1, X_2, \dots, X_n)$ a una muestra aleatoria generada por el modelo muestral $f(x|\theta)$ y sea $\pi(\theta)$ la distribución de probabilidad a priori del parámetro. Una vez que se tienen a las observaciones $x = (X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$, la distribución $\pi(\theta)$ se actualiza, mediante el Teorema de Bayes, en la distribución posterior $\pi(\theta|x)$

$$\pi(\theta|x) = \frac{L(x|\theta)\pi(\theta)}{\int L(x|\theta)\pi(\theta)d\theta}$$

Donde: $L(x|\theta) = f(x|\theta)$ es la verosimilitud del modelo. La distribución posterior $\pi(\theta|x)$ es el instrumento que se usa para establecer inferencias con respecto al parámetro. Sin embargo, la estadística Bayesiana no se restringe a enunciados sobre el parámetro, también considera el problema fundamental de establecer inferencias sobre observaciones futuras a partir de las observaciones pasadas x . El objetivo de esta tarea predictiva es predecir el valor de una observación futura $Y = X_{n+1}$ proveniente del modelo muestral $f(x|\theta)$. En la estadística Bayesiana se establece que esta

predicción se debe realizar con la distribución de probabilidad condicional de $Y = y|X = x$

$$f(y|x) = \int L(\theta|y)\pi(\theta|x)d\theta$$

Donde: $L(\theta|y) = f(y|\theta)$ es la verosimilitud de $Y = X_{n+1}$. La distribución $f(y|x)$ se llama la distribución predictiva y resulta ser un elemento fundamental de la estadística Bayesiana gracias a sus propiedades de optimalidad. Notemos que $f(y|x)$ promedia a la distribución muestral $L(\theta|y) = f(y|\theta)$ a través de la incertidumbre en el parámetro medida con la distribución posterior $\pi(\theta|x)$. Así $f(y|x)$ combina de forma óptima a dos fuentes de incertidumbre: Por una parte, a la incertidumbre aleatoria descrita en el mecanismo generador de los datos $f(x|\theta)$. Por otra parte, a la incertidumbre epistémica, que es la incertidumbre en los verdaderos parámetros que se cuantifica con la distribución posterior $\pi(\theta|x)$.

2.2. Predicción Bayesiana

La TVE busca describir el comportamiento de procesos naturales con el objetivo de anticipar sus episodios extremos. En este sentido, los problemas de modelación de valores extremos son predictivos.

El método básico y estándar de la TVE para modelar una muestra de máximos muestrales consiste en ajustarles la función de distribución del Valor Extremo Generalizada (DVEG).

$$F(x; \mu, \sigma, \xi) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} & \text{para } \xi \neq 0, \\ \exp \left\{ - \exp \left[- \left(\frac{x - \mu}{\sigma} \right) \right] \right\} & \text{para } \xi = 0. \end{cases}$$

La DVEG tiene un parámetro de localización $-\infty < \mu < \infty$, un parámetro de escala $\sigma > 0$ y un parámetro de forma $-\infty < \xi < \infty$. Cuando $\xi \neq 0$ el soporte de F es $\{x: 1 + \xi(x - \mu)/\sigma > 0\}$, cuando $\xi = 0$ el soporte es $-\infty < x < \infty$, en este caso la DVEG se llama distribución Gumbel.

2.3. Verosimilitud de la DVEG

Sea $X = (X_1, X_2, \dots, X_n)$ una muestra aleatoria de la DVEG $F(x; \mu, \sigma, \xi)$ y sea $x = (x_1, x_2, \dots, x_n)$ una realización de X . La función de verosimilitud de la DVEG es

$$L(\mu, \sigma, \xi) = \prod_{i=1}^n f(x_i; \mu, \sigma, \xi)$$

Donde:

$$f(x; \mu, \sigma, \xi) = \exp \left[-\left(1 + \xi(x - \mu)/\sigma\right)^{-1/\xi} \right] \frac{1}{\sigma} \left\{ 1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right\}^{-\frac{1}{\xi}-1}$$

es la función de densidad de la DVEG. En la inferencia estadística clásica, el instrumento que se utiliza para establecer enunciados inferenciales sobre los parámetros de la DVEG sería su función de log-verosimilitud

$$\log L(\mu, \sigma, \xi) = -n \log(\sigma) - \left(\frac{1}{\xi} + 1 \right) \sum_{i=1}^n \log \left(1 + \xi \frac{x_i - \mu}{\sigma} \right) - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^{-1/\xi}$$

para $1 + \xi[(x_i - \mu)/\sigma] > 0$ y $\xi \neq 0$, véase Coles (2000) o Dutfoy (2021).

2.4. Distribuciones a priori

El problema de la especificación de las distribuciones a priori para las distribuciones de valores extremos lo abordan Burton (1979), Coles y Powell (1996), Northrop y Attalides (2015) y Bousquet y Keller (2017) entre otros. En este trabajo especificamos la distribución a-priori de los parámetros μ, σ y ξ con las distribuciones

$$\begin{aligned} \mu &\sim N(\mu_0 = 5, \sigma_\mu^2 = 1), \\ \sigma &\sim \text{Gamma}(\alpha = 2, \beta = 0.5), \\ \xi &\sim N(\mu_\xi = 0, \sigma_\xi^2 = 0.1). \end{aligned}$$

La elección de los valores de los hiperparámetros la fundamentamos en la inspección de la distribución de las magnitudes máximas observadas. Los detalles que justifican esta elección se pueden consultar en Rivera (2025). Además, hacemos el supuesto de que los parámetros son probabilísticamente independientes, por lo tanto, la distribución a priori $\pi(\mu, \sigma, \xi)$ es el producto de las respectivas densidades de μ, σ y ξ .

2.5. Distribución posterior

La distribución posterior de los parámetros es

$$\pi(\mu, \sigma, \xi|x) = \frac{L(\mu, \sigma, \xi)\pi(\mu, \sigma, \xi)}{\int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty L(\mu, \sigma, \xi)\pi(\mu, \sigma, \xi) d\mu d\xi d\sigma}$$

2.6. Distribución predictiva

Finalmente, para predecir el valor de una observación futura $Y = X_{n+1}$ de $F(x; \mu, \sigma, \xi)$ condicional en los datos X , utilizamos a la distribución predictiva dada por

$$f(y|x) = \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(\mu, \sigma, \xi) \pi(\mu, \sigma, \xi | x) d\mu d\xi d\sigma.$$

Tanto la distribución posterior $\pi(\mu, \sigma, \xi | x)$ como la distribución predictiva $f(y|x)$ no tienen una expresión cerrada. Por lo tanto, recurrimos a las poderosas y modernas herramientas de estadística Bayesiana computacional, ver Albert (2009) o Matsuura (2022), para aproximarlas mediante simulación Monte Carlo. En particular implementamos algoritmos de Cadenas de Markov Monte Carlo para aproximar por simulación a la distribución posterior de los parámetros de la DVEG y a la distribución predictiva. La implementación computacional del análisis lo realizamos con el software estadístico R (R Core Team, 2025). Los paquetes de R que se utilizaron para el análisis están listados en el **Apéndice C**. Con el objetivo de replicabilidad del análisis, hemos puesto el código R usado en el apéndice C.

3. Resultados

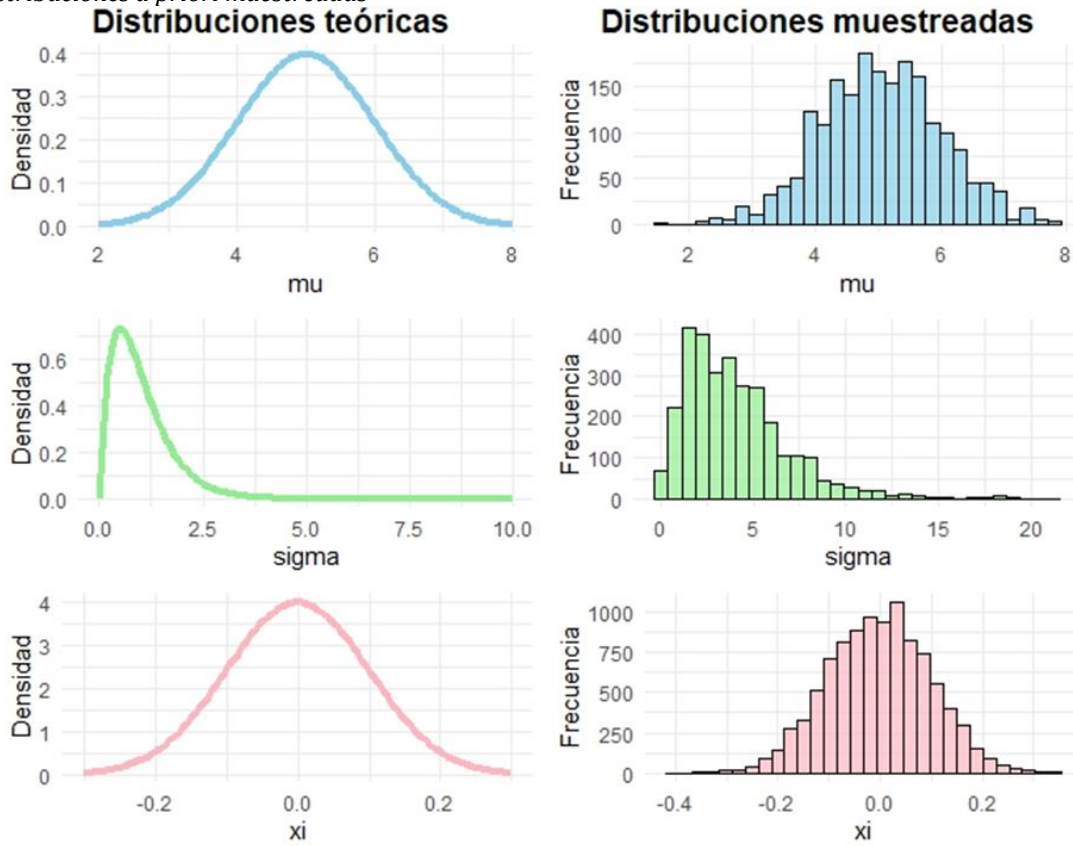
3.1. Distribución posterior

En la **Figura 3** visualizamos los resultados del muestreo de las, *a priori*, teóricas; generado con el algoritmo Metrópolis. Vemos que las distribuciones muestreadas reflejan a las distribuciones a priori teórica.

La **Figura 4** muestra las trazas de las cadenas de Markov simuladas para cada parámetro de la DVEG. Se observa que las cadenas simuladas se mezclan y recorren todo el espacio de alta densidad sin quedar atrapadas en regiones específicas.

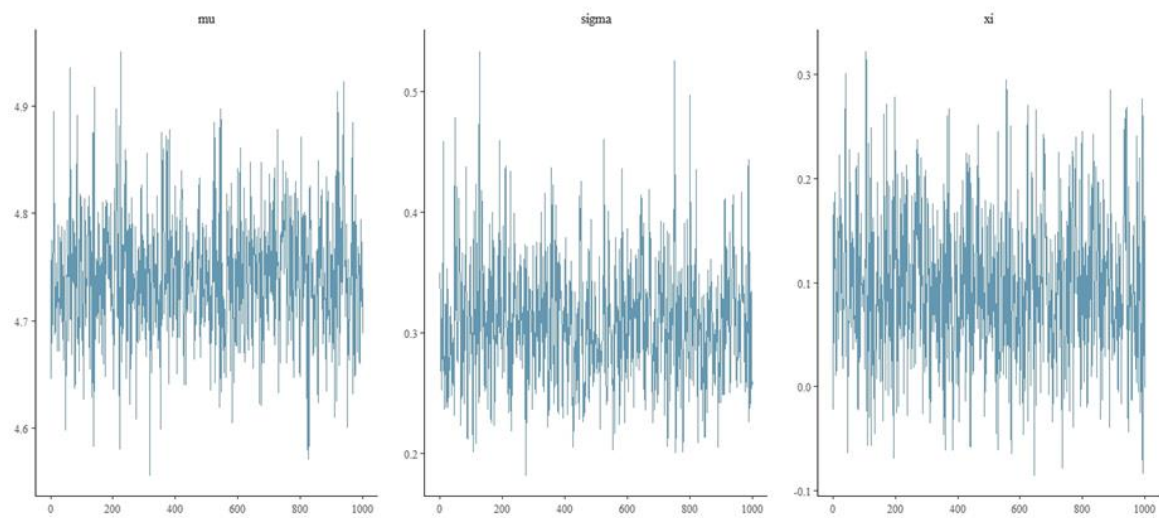
En la **Figura 5** apreciamos una rápida disminución en las autocorrelaciones de las muestras simuladas de las cadenas de Markov para cada parámetro. Esto indica que la aproximación Monte Carlo a la distribución posterior es adecuada en términos de que los valores simulados se pueden considerar muestras aleatorias provenientes de las posteriores.

Figura 3
 Distribuciones a priori muestreadas



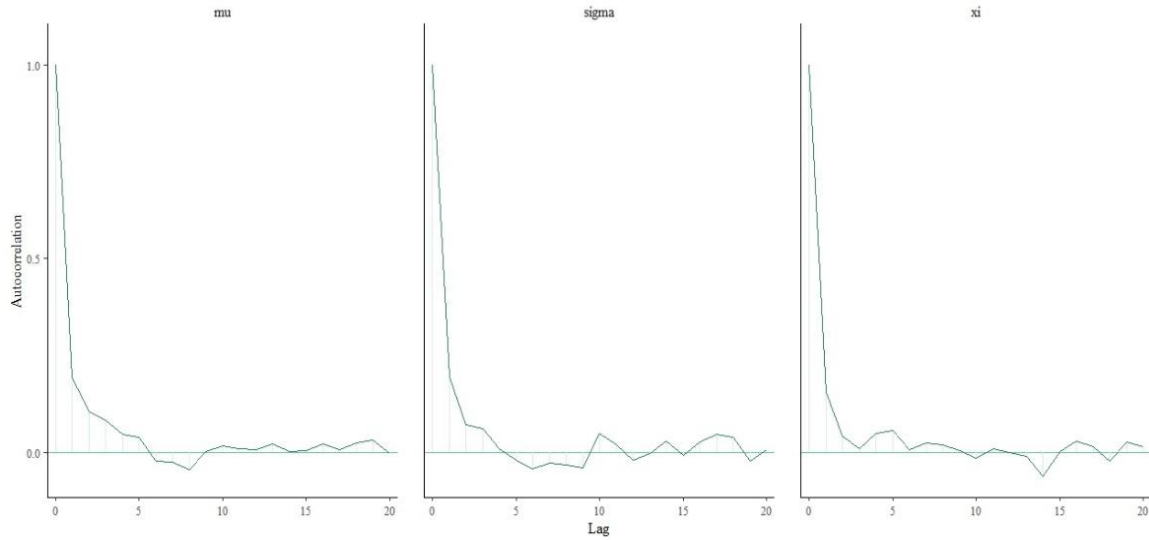
Fuente: Elaboración propia.

Figura 4
 Trazas de las cadenas de Markov Monte Carlo



Fuente: Elaboración propia.

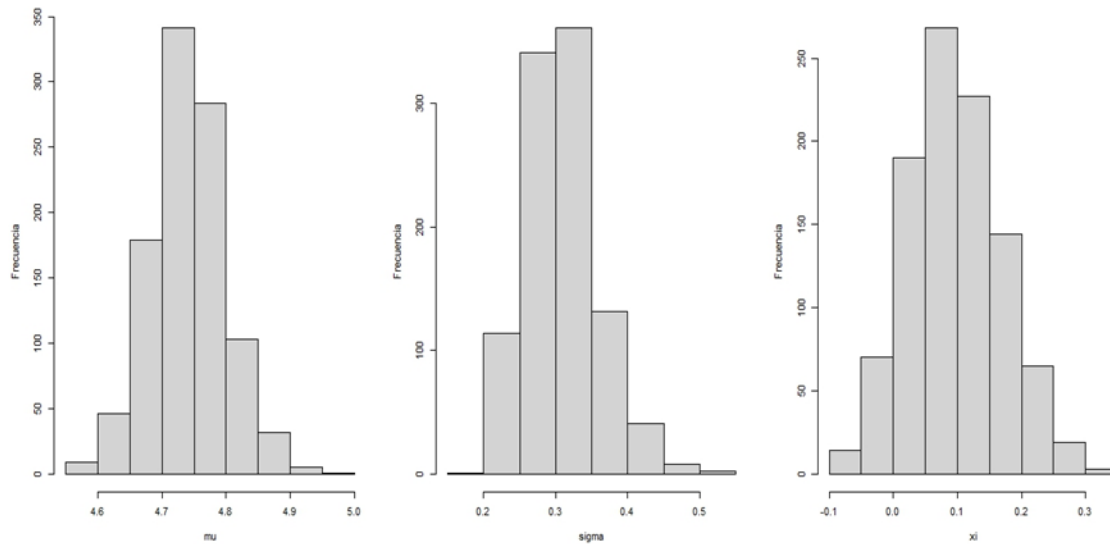
Figura 5
 Funciones de autocorrelación de las cadenas simuladas
 Autocorrelación de los parámetros



Fuente: Elaboración propia.

En la **Figura 6** vemos las distribuciones posteriores de cada parámetro de la DVEG.

Figura 6
 Distribuciones posteriores, de izquierda a derecha: μ , σ , y ξ



Fuente: Elaboración propia.

En la **Tabla 2** mostramos a los intervalos de credibilidad al 95% para cada parámetro.

Tabla 2

Estadísticas descriptivas de las posteriores, e intervalos de credibilidad al 95%

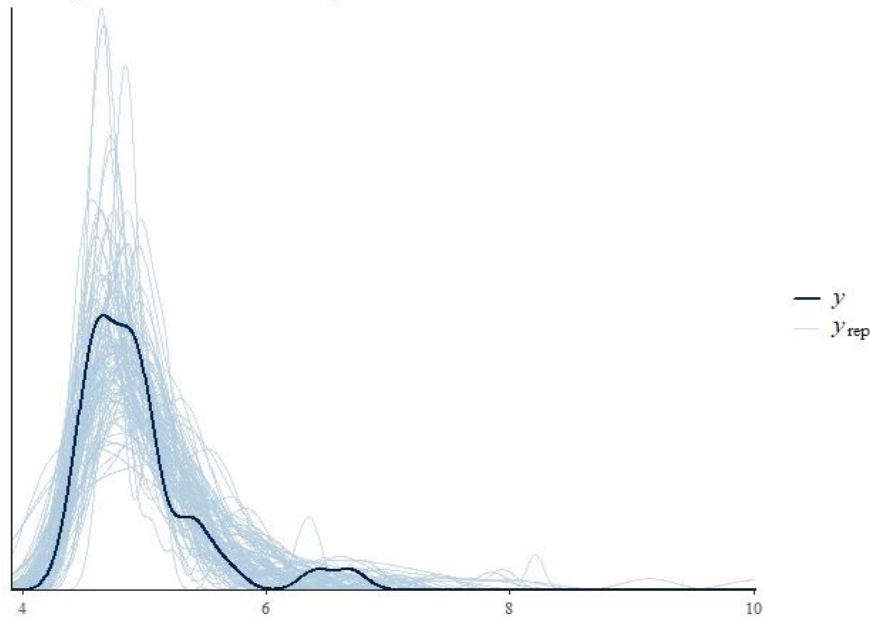
Parámetro	Media	Desv. Est.	Límite inferior	Límite superior
μ	4.74	0.06	4.63	4.87
σ	0.31	0.05	0.22	0.42
ξ	0.10	0.07	-0.03	0.26

Fuente: Elaboración propia.

En la **Figura 7** evaluamos la capacidad de la distribución predictiva para reproducir la forma de la distribución empírica de los datos observados mediante la generación de 100 réplicas de la magnitud máxima anual. Notamos que existe una adecuada correspondencia entre la densidad observada y las densidades de las réplicas simuladas. Esto valida al hecho de que la distribución predictiva captura adecuadamente a la forma de la distribución de las magnitudes máximas anuales.

Figura 7

Densidad no paramétrica ajustada, 100 réplicas de la densidad predictiva
 100 replicas de la distribución predictiva

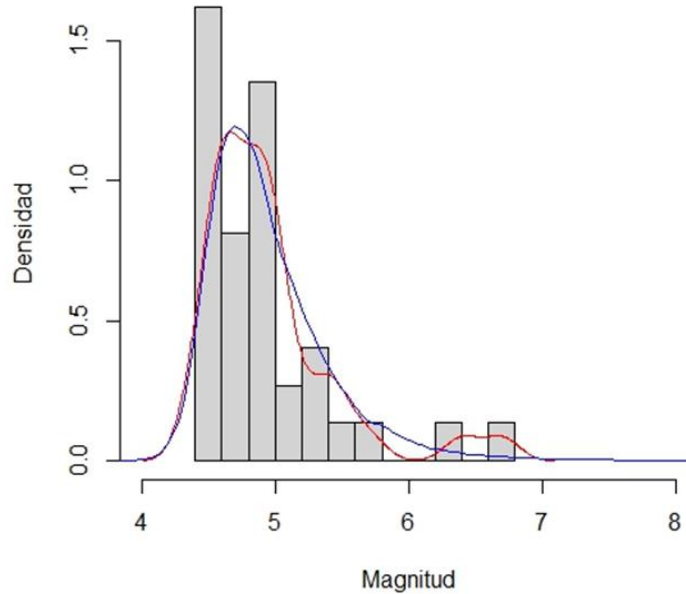


Fuente: Elaboración propia.

3.2. Distribución predictiva

En la **Figura 8** mostramos a la densidad de la distribución predictiva obtenida con el algoritmo de Metrópolis con 37 000 valores simulados. En la **Figura 9** vemos a la función de distribución empírica junto con la función de distribución predictiva Bayesiana.

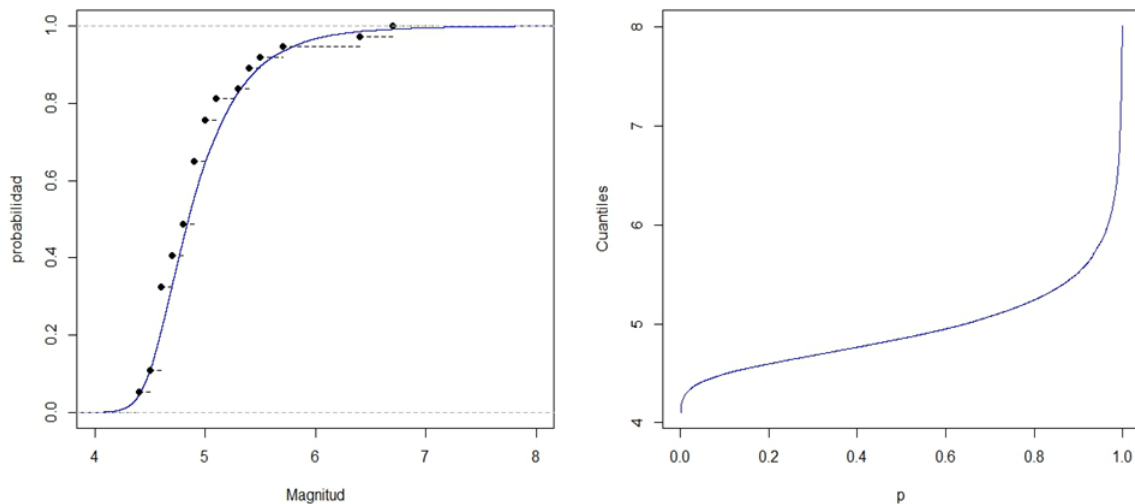
Figura 8
Histograma de las magnitudes máximas anuales



Nota. Línea roja = Densidad no paramétrica de las magnitudes máximas anuales. Línea azul = Distribución Bayesiana predictiva suavizada.

Fuente: Elaboración propia.

Figura 9
Distribución predictiva Bayesiana



Nota. Izquierda: Línea azul continua es la función de distribución predictiva Bayesiana. Línea negra: Función de distribución empírica. Derecha: Función cuantil de la distribución predictiva.

Fuente: Elaboración propia.

Con la distribución predictiva Bayesiana calculamos los niveles de retorno para diferentes períodos de retorno. Presentamos estos niveles de retorno en la **Tabla 3**

junto con la incertidumbre de la predictiva medida con los intervalos de predicción al 95%.

Tabla 3

Niveles de retorno junto con intervalos de predicción al 95%

Cuantil	Período de retorno en años	Nivel de retorno	Límite inferior	Límite superior
0.900	10	5.52	5.47	5.59
0.950	20	5.82	5.75	5.93
0.967	30	6.00	5.88	6.12
0.975	40	6.13	6.00	6.31
0.980	50	6.24	6.08	6.42
0.983	60	6.34	6.16	6.54
0.986	70	6.42	6.22	6.59
0.988	80	6.50	6.31	6.70
0.989	90	6.57	6.34	6.75
0.990	100	6.63	6.37	6.85

Fuente: Elaboración propia.

Como podemos apreciar en las Figuras 8 y 9, la mayoría de los valores se concentran entre 4.85 a 5.15. La densidad de probabilidad por arriba de una magnitud 7 es prácticamente nula, lo cual es consistente con los registros históricos de sismicidad en Veracruz registrados por el SSN. El intervalo de credibilidad central con nivel de probabilidad del 95%, va de 4.35 a 6.11. Los cuantiles superiores de la distribución predictiva son: cuantil 0.9 = 5.52, cuantil 0.95 = 5.81 y cuantil 0.99 = 6.55. Los resultados arrojados por la distribución predictiva en la figura 4 son consistentes con los obtenidos en Juárez *et al.* (2024) pero con un análisis estadístico de la DVEG basado en la verosimilitud. Finalmente, en el Apéndice B desplegamos los cuantiles y_p de la distribución predictiva.

5. Comentarios Finales

En este trabajo hemos estudiado la viabilidad de la TEV combinada con el paradigma Bayesiano para desarrollar un instrumento de predicción de futuros eventos sísmicos extremos con epicentro en el estado de Veracruz. Nuestro producto final es una distribución predictiva resultante del modelo muestral dado por la DVEG con una especificación relativamente simple para las distribuciones a priori de sus parámetros.

Nuestros análisis indican que la distribución predictiva que obtuvimos es efectivamente una herramienta poderosa y útil para capturar la incertidumbre en la predicción de sismos extremos, sin embargo, hay aspectos que merecen investigación futura. Por ejemplo, se puede explorar el otro enfoque de la TVE basado en excedentes sobre umbrales con la distribución Pareto Generalizada, Coles (2001). También se debe considerar incorporar un elemento espacial en la modelación de la sismicidad extrema.

En este trabajo se realiza un análisis basado exclusivamente en eventos sísmicos cuyos epicentros se localizan dentro del territorio del estado de Veracruz, es decir, en la sismicidad local. No obstante, es posible que algunos eventos ocurridos en entidades colindantes, bajo determinadas condiciones, como su magnitud, profundidad y

distancia epicentral, puedan generar efectos significativos en la región de estudio. Por ello, dichos eventos deberán considerarse en análisis posteriores.

En cuanto a la variación en la capacidad de detección de eventos sísmicos, este es un aspecto que debe ser tomado en cuenta. Por ejemplo, a partir de 2013 el estado de Veracruz cuenta con la Red Sísmica de Veracruz, integrada por seis estaciones que se sumaron a las dos estaciones que en ese momento el SSN mantenía en operación en la entidad. Esta red local se complementa además con la cobertura regional del SSN, la cual también se ha incrementado con el tiempo.

Como resultado de esta expansión en la infraestructura de monitoreo, actualmente se registran más eventos sísmicos en comparación con décadas anteriores. Sin embargo, es importante señalar que, aunque la sensibilidad de detección es mayor en la actualidad, la mayoría de los sismos registrados en Veracruz corresponden a eventos de baja magnitud. En este estudio, el interés principal se centra en el caso opuesto, es decir, en el análisis de las magnitudes máximas registradas.

Agradecimientos

Agradecemos a los dos árbitros por las observaciones que nos hicieron y nos llevaron a una mejora sustancial de este trabajo. Los datos utilizados en este estudio son de acceso libre a través del catálogo del Servicio Sismológico Nacional, con la liga de acceso <http://www2.ssn.unam.mx:8080/catalogo/>. El catálogo de sismos del Servicio Sismológico Nacional de México es posible gracias a todo su personal y producto de los cálculos realizados por su área de Análisis e Interpretación de Datos Sísmicos.

Referencias

- Al-Abbasi, J., Fahmi, K. (1985). Estimating maximum magnitude earthquake in Iraq using extreme value statistics. *Geophysical Journal International*, 82(3), 535-548. <https://doi.org/10.1111/j.1365-246x.1985.tb05150.x>
- Albert, J. (2009). *Bayesian Computation with R*. <https://doi.org/10.1007/978-0-387-92298-0>
- Al-Heety, E. A., Rafea, H. F., Mohammad, O. J. (2024). Evaluation of return period and occurrence probability of the maximum magnitude earthquakes in Iraq and surroundings. *IOP Conference Series Earth and Environmental Science*, 1300(1), 012001. <https://doi.org/10.1088/1755-1315/1300/1/012001>
- Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics* (Versión 2.3) [Paquete de R]. <https://CRAN.R-project.org/package=gridExtra>
- Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. L. (2006). *Statistics of Extremes: Theory and Applications*. Wiley. <https://tinyurl.com/4jmskb92>
- Bousquet, N., & Keller, M. (2017). *Bayesian prior elicitation and selection for extreme values* (arXiv:1712.00685). arXiv. <https://arxiv.org/abs/1712.00685>

- Burton, P. W.** (1979). Seismic risk in southern Europe through India examined using Gumbel's third distribution of extreme values. *Geophysical Journal of the Royal Astronomical Society*, 59, 249–280.
- Coles, S.G., and Powell, E.A.** (1996). Bayesian Methods in Extreme Value Modelling: A Review and New Developments. *International Statistical Review*, 64(1), 119–136. <https://doi.org/10.2307/1403426>
- Coles, S.G.** (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer.
- Córdoba Montiel, F., Pérez Campos, X., Sieron, K., y Juárez Cerrillo, S. F.** (2020). Análisis e interpretación de la sismicidad local en Veracruz para la generación de boletines del OSV. *UVserva*, 10, 123-138. <https://doi.org/10.25009/uvs.v0i10.2752>
- Córdoba-Montiel, F.** (2016). La red sísmica de Veracruz. Dirección de Comunicación de la Ciencia. En *La Ciencia y el Hombre*, 3, 24-27. <https://tinyurl.com/3dn457z6>
- Dutfoy, A.** (2021). Univariate extreme value theory: Practice and limitations. En N. Bousquet y P. Bernardara (Eds.), *Extreme value theory with applications to natural hazards* (pp. 83–123). Springer. https://doi.org/10.1007/978-3-030-74942-2_6
- Gabry, J., y Mahr, T.** (2024). *bayesplot: Plotting for Bayesian Models* (Versión 1.11.1) [Paquete de R]. <https://mc-stan.org/bayesplot/>
- García-Bustos, S., Landín, J., Moreno, R., Chong, A., Mulas, M., Mite, M., Cárdenas, N.** (2020). Statistical analysis of the largest Possible earthquake magnitudes on the Ecuadorian coast for selected return periods. *Georisk Assessment And Management Of Risk For Engineered Systems And Geohazards*, 14(1), 56-68. <https://doi.org/10.1080/17499518.2018.1542500>
- Held, L., and Sabanés-Bové, D.** (2014). *Applied Statistical Inference. Likelihood and Bayes*. Springer.
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., y Yasmeen, F.** (2024). *forecast: Forecasting functions for time series and linear models* (Versión 8.23.0) [Paquete de R]. <https://pkg.robjhyndman.com/forecast/>
- Juárez-Cerrillo, S. F., Córdoba-Montiel, F., y Que-Salinas, S. U.** (2024). Análisis estadístico de la sismicidad extrema en Veracruz. En C. M. Welsh y C. A. Ochoa Martínez (Eds.), *Estado del arte de las ciencias de la Tierra y prevención de desastres en Veracruz*. Universidad Veracruzana.
- Kijko, A., and Singh, M.** (2011). Statistical Tools for Maximum Possible Earthquake Magnitude Estimation. *Acta Geophysica*, 59(4), 674-700. <https://doi.org/10.2478/s11600-011-0012-6>
- Lasocki, S.** (2021). Kernel density estimation in seismology. En N. Limnios, E. Papadimitriou, & G. Tsaklidis (Eds.), *Statistical methods and modeling of seismogenesis* (pp. 1–26). Wiley.
- Ma, N., Bai, Y., Meng, S.** (2021). Return period of the largest possible earthquake magnitudes in mainland China based on extreme value theory. *Sensors*, 21(10), 3519. <https://doi.org/10.3390/s21103519>
- Maruyama, F.** (2020). Analyzing the annual maximum magnitude of earthquakes in Japan by extreme value theory. *Open Journal of Applied Sciences*, 10, 817-824. <https://doi.org/10.4236/ojapps.2020.1012057>
- Matsuura, K.** (2022). *Bayesian Statistical Modeling with Stan, R, and Python*. Springer.

- Nordquist**, J. M. (1945). Theory of largest values applied to earthquake magnitudes. *Transactions American Geophysical Union*, 26, 29-31.
- Northrop**, P. J., y **Attalides**, N. (2015). Posterior propriety in Bayesian extreme value analyses using reference priors. *Statistica Sinica*, 26(2), 721-743. <https://doi.org/10.5705/ss.2014.034>
- Pavlenko**, V. (2017). Estimation of the upper bound of seismic hazard curve by using the generalized extreme value distribution. *Nat. Hazards*, 89, 19-33. <https://doi.org/10.1007/s11069-017-2950-z>
- Plummer**, M., **Best**, N., **Cowles**, K., y **Vines**, K. (2006). *coda: Convergence Diagnosis and Output Analysis for MCMC* (Versión actual) [Paquete de R]. <https://journal.r-project.org/archive/>
- R** Core Team. (2025). R: A language and environment for statistical computing [Software]. *R Foundation for Statistical Computing*. <https://www.R-project.org/>
- Rivera-Flores**, R. C. (2025). Distribución del Valor Extremo Generalizada: Análisis bayesiano de datos de sismología en Veracruz. Tesis para obtener el grado de Licenciado en Estadística, Facultad de Estadística e Informática, Universidad Veracruzana.
- Servicio Sismológico Nacional [SSN]**. (2024). *Catálogo de sismos*. Universidad Nacional Autónoma de México. Servicio Sismológico Nacional. <http://www2.ssn.unam.mx:8080/catalogo/>
- Stan** Development Team. (2025). *RStan: The R interface to Stan* (Versión 2.32.7) [Paquete de R]. <https://mc-stan.org/>
- Trapletti**, A., y **Hornik**, K. (2024). *tseries: Time Series Analysis and Computational Finance* (Versión 0.10-58) [Paquete de R]. <https://CRAN.R-project.org/package=tseries>
- Vehtari**, A., **Gabry**, J., **Magnusson**, M., **Yao**, Y., **Bürkner**, P., **Paananen**, T., y **Gelman**, A. (2024). *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models* (Versión 2.8.0) [Paquete de R]. <https://mc-stan.org/loo/>
- Wickham**, H. (2016). *ggplot2: Elegant graphics for data analysis* [Paquete de R]. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wickham**, H., **François**, R., **Henry**, L., **Müller**, K., y **Vaughan**, D. (2023). *dplyr: A Grammar of Data Manipulation* (Versión 1.1.4) [Paquete de R]. <https://CRAN.R-project.org/package=dplyr>
- Yegulalp**, T. M., and **Kuo**, J. T. (1974). Statistical prediction of the occurrence of maximum magnitude earthquakes. *Bulletin of the Seismological Society of America*, 64(2), 393-414. <https://doi.org/10.1785/bssa0640020393>
- Zeileis**, A., y **Grothendieck**, G. (2005). *zoo: S3 Infrastructure for Regular and Irregular Time Series* (Versión actual) [Paquete de R]. <https://doi.org/10.18637/jss.v014.i06>
- Zöller**, G., **Holschneider**, M., **Hainzl**, S., **Zhuang**, J. (2014). The largest expected earthquake magnitudes in Japan: The statistical perspective. *Bulletin Of The Seismological Society Of America*, 104(2), 769-779. <https://doi.org/10.1785/0120130103>

Apéndice A

Datos

Tabla A

Sismos de magnitudes máximas anuales

Fecha	Magnitud	Latitud	Longitud	Referencia	N
16/05/1988	4.8	17.2800	-94.0100	70 km al SUR de LAS CHOAPAS	101
21/02/1989	4.6	17.2700	-94.3700	77 km al SUROESTE de LAS CHOAPAS	97
10/02/1990	4.9	17.8300	-95.0400	10 km al SUROESTE de SAYULA DE ALEMAN	26
25/07/1991	5.3	17.8100	-95.0900	16 km al SUROESTE de SAYULA DE ALEMAN	33
04/06/1992	4.7	18.9700	-95.7200	22 km al NORTE de ALVARADO	33
05/04/1993	4.8	17.5600	-94.6900	45 km al SUR de JALTIPAN DE MORELOS	41
10/11/1994	4.5	19.2900	-95.5300	62 km al NORESTE de ALVARADO	42
03/01/1995	4.4	17.3800	-94.5200	67 km al SUR de MINATITLAN	19
03/06/1996	4.9	17.4800	-94.6500	54 km al SUR de JALTIPAN DE MORELOS	14
06/09/1997	4.7	18.1900	-94.4000	4 km al NORTE de ALLENDE	26
25/04/1998	5.0	17.7200	-94.1400	22 km al SUR de LAS CHOAPAS	15
06/05/1999	4.4	17.3900	-94.4500	67 km al SUR de MINATITLAN	42
20/08/2000	4.6	17.6800	-94.3700	39 km al SUROESTE de LAS CHOAPAS	40
09/10/2001	4.9	18.1900	-95.3400	23 km al NORESTE de J RODRIGUEZ CLARA	37
10/06/2002	4.9	17.5000	-94.7500	48 km al SURESTE de SAYULA DE ALEMAN	35
08/07/2003	4.6	17.5100	-94.8400	43 km al SURESTE de SAYULA DE ALEMAN	43
17/01/2004	5.1	17.6800	-95.6500	41 km al SUROESTE de ISLA	44
14/11/2005	4.7	18.5300	-95.7600	17 km al NORTE de CARLOS A CARRILLO	43
08/05/2006	4.6	17.4600	-94.3900	59 km al SUROESTE de LAS CHOAPAS	44
23/05/2007	5.4	21.9300	-96.1400	169 km al NORESTE de TUXPAM DE RGZ CANO	28
13/01/2008	4.6	17.4800	-94.4200	58 km al SUR de MINATITLAN	48
29/10/2009	5.5	18.9500	-95.6900	21 km al NORESTE de ALVARADO	51
19/08/2010	4.6	17.6500	-94.7700	33 km al SURESTE de SAYULA DE ALEMAN	45
07/04/2011	6.7	17.2200	-94.3200	80 km al SUROESTE de LAS CHOAPAS	67
18/01/2012	4.9	18.2383	-94.4957	13 km al NOROESTE de COATZACOALCOS	53
19/09/2013	4.6	17.7150	-94.6755	28 km al SUR de JALTIPAN DE MORELOS	79
29/07/2014	6.4	17.6988	-95.6373	38 km al SUROESTE de ISLA	111
03/08/2015	4.6	17.7352	-93.8185	35 km al SURESTE de LAS CHOAPAS	234
06/06/2016	5.0	17.3450	-94.3130	67 km al SUROESTE de LAS CHOAPAS	425
25/01/2017	5.1	17.6748	-94.6402	33 km al SUR de JALTIPAN DE MORELOS	432
22/07/2018	5.0	17.3777	-94.6972	62 km al SURESTE de SAYULA DE ALEMAN	354
29/01/2019	5.4	17.3062	-94.1302	67 km al SUR de LAS CHOAPAS	475
11/03/2020	4.9	17.5532	-95.5540	51 km al SUROESTE de J RODRIGUEZ CLARA	595
29/08/2021	5.0	17.8827	-95.0287	7 km al OESTE de SAYULA DE ALEMAN	671
03/03/2022	5.7	18.0908	-95.5727	8 km al NOROESTE de ISLA	629
19/11/2023	4.8	17.4615	-94.8653	48 km al SUR de SAYULA DE ALEMAN	655
16/01/2024	4.5	17.5300	-95.1900	46 km al SUROESTE de SAYULA DE ALEMAN	361

Nota: Variables del catálogo:

1. **Fecha:** Fecha en que sucedió el sismo en tiempo del centro de México.
2. **Magnitud:** Magnitud del sismo.
3. **Latitud:** Latitud geográfica del epicentro del sismo en grados decimales.
4. **Longitud:** Longitud geográfica del epicentro del sismo en grados decimales.
5. **Profundidad:** Profundidad en kilómetros, del epicentro del sismo.
6. **Referencia:** Referencia a una localidad importante en cuanto a número de habitantes y cercana al epicentro del sismo.
7. **N:** Número de sismos en el bloque.

Fuente: Elaboración propia a partir del catálogo del SSN.

Apéndice B

Cuantiles Bayesianos Predictivos

Tabla B
Cuantiles Bayesianos Predictivos

p	x_p	p	x_p	p	x_p	p	x_p
0.005	4.214555	0.255	4.641887	0.505	4.853818	0.755	5.157857
0.010	4.268262	0.260	4.645734	0.510	4.858522	0.760	5.166142
0.015	4.301951	0.265	4.649778	0.515	4.862542	0.765	5.176273
0.020	4.326824	0.270	4.653428	0.520	4.867290	0.770	5.185087
0.025	4.347458	0.275	4.657803	0.525	4.871759	0.775	5.193536
0.030	4.365337	0.280	4.661965	0.530	4.876546	0.780	5.202575
0.035	4.378748	0.285	4.666441	0.535	4.881812	0.785	5.212138
0.040	4.391048	0.290	4.670408	0.540	4.886621	0.790	5.221580
0.045	4.402842	0.295	4.674411	0.545	4.891212	0.795	5.230505
0.050	4.412756	0.300	4.678314	0.550	4.896585	0.800	5.241838
0.055	4.422245	0.305	4.682705	0.555	4.901241	0.805	5.251552
0.060	4.431689	0.310	4.687368	0.560	4.906392	0.810	5.261473
0.065	4.439343	0.315	4.691745	0.565	4.912086	0.815	5.271795
0.070	4.447112	0.320	4.695879	0.570	4.917303	0.820	5.283325
0.075	4.455621	0.325	4.699981	0.575	4.922390	0.825	5.294072
0.080	4.463290	0.330	4.704077	0.580	4.927624	0.830	5.306488
0.085	4.470515	0.335	4.708051	0.585	4.932465	0.835	5.319335
0.090	4.477499	0.340	4.711981	0.590	4.937304	0.840	5.331518
0.095	4.484426	0.345	4.716371	0.595	4.941726	0.845	5.343856
0.100	4.490637	0.350	4.721087	0.600	4.947034	0.850	5.356717
0.105	4.497111	0.355	4.725056	0.605	4.952472	0.855	5.371787
0.110	4.503811	0.360	4.728919	0.610	4.958396	0.860	5.386602
0.115	4.510181	0.365	4.732890	0.615	4.964002	0.865	5.399689
0.120	4.516144	0.370	4.737133	0.620	4.969815	0.870	5.414645
0.125	4.522213	0.375	4.740999	0.625	4.975746	0.875	5.430543
0.130	4.527147	0.380	4.745122	0.630	4.981175	0.880	5.447886
0.135	4.532009	0.385	4.749776	0.635	4.987770	0.885	5.464201
0.140	4.536381	0.390	4.753863	0.640	4.993558	0.890	5.482075
0.145	4.541862	0.395	4.757646	0.645	4.998963	0.895	5.500400
0.150	4.546635	0.400	4.761898	0.650	5.005038	0.900	5.520256
0.155	4.551598	0.405	4.766124	0.655	5.012135	0.905	5.542232
0.160	4.556198	0.410	4.770080	0.660	5.018441	0.910	5.565948
0.165	4.561548	0.415	4.774653	0.665	5.025115	0.915	5.589359
0.170	4.566170	0.420	4.779155	0.670	5.031516	0.920	5.614228
0.175	4.570580	0.425	4.783617	0.675	5.038757	0.925	5.643738
0.180	4.575086	0.430	4.788690	0.680	5.046212	0.930	5.675026
0.185	4.579882	0.435	4.792821	0.685	5.053550	0.935	5.713957
0.190	4.584355	0.440	4.797416	0.690	5.060148	0.940	5.750716
0.195	4.588460	0.445	4.801773	0.695	5.067499	0.945	5.788902
0.200	4.593056	0.450	4.805614	0.700	5.074646	0.950	5.824851
0.205	4.597521	0.455	4.809552	0.705	5.081801	0.955	5.870305
0.210	4.601722	0.460	4.814412	0.710	5.089186	0.960	5.918152
0.215	4.606086	0.465	4.819006	0.715	5.096498	0.965	5.975142
0.220	4.610201	0.470	4.823967	0.720	5.103807	0.970	6.049255
0.225	4.615286	0.475	4.827828	0.725	5.112139	0.975	6.131887
0.230	4.619926	0.480	4.832331	0.730	5.118727	0.980	6.249562
0.235	4.624718	0.485	4.836052	0.735	5.125904	0.985	6.395966
0.240	4.629082	0.490	4.840794	0.740	5.133761	0.990	6.627253
0.245	4.632827	0.495	4.845073	0.745	5.141870	0.995	7.006033
0.250	4.637262	0.500	4.849452	0.750	5.149250	0.999	8.015087

Nota: Cuantiles $Q(p) = x_p$, donde $P(X_{max} \leq x_p) = p$, X_{max} denota la magnitud máxima anual.

Apéndice C

Código R y paquetes usados

```
#####
#
# Código R
#
#####
#
# Carga de librerías
#
library(vctr)
library(rlang)
library(rstan)
library(bayesplot)
library(ggplot2)
library(loo)
library(coda)
library(gridExtra)
library(dplyr)
library(zoo)
library(forecast)
library(tseries)
#
#####
#
# Magnitudes máximas anuales
#
datos <- c(4.8, 4.6, 4.9, 5.3, 4.7, 4.8, 4.5, 4.4, 4.9, 4.7, 5.0,
4.4, 4.6, 4.9, 4.9, 4.6, 5.1, 4.7, 4.6, 5.4, 4.6, 5.5, 4.6, 6.7,
4.9, 4.6, 6.4, 4.6, 5.0, 5.1, 5.0, 5.4, 4.9, 5.0, 5.7, 4.8, 4.5)
#
#####
#
# Análisis descriptivo y exploratorio
#
summary(datos)
# Min. 1st Qu. Median Mean 3rd Qu. Max.
# 4.400 4.600 4.900 4.949 5.000 6.700
#
# Histograma y densidad no paramétrica
kde <- density(x = datos,)
hist(sismos$Magnitud, freq = FALSE, nclass = 16,
xlab = "Magnitud", ylab = "Densidad",main = " ")
points(kde,type = "l", col = "red")
#
# Función de distribución acumulada (CDF)
ecdf_fun <- ecdf(datos)
plot(ecdf_fun,xlab="Magnitudes",ylab="Probabilidad", main=" ")
#
#####
#
# ESTIMACIÓN DE LA DISTRIBUCIÓN PREDICTIVA BAYESIANA
#
# Muestreo de prioris con algoritmo de Metrópolis
# Funciones de densidad logarítmica de prioris
# Por eficiencia y estabilidad computacional usaremos logaritmos

# Localización: Media 5 y varianza 1
log_prior_mu <- function(mu) dnorm(mu, 5, 1, log = TRUE)
```

```

# Escala: Forma 2 y escala 0.5, escala
log_prior_sigma <- function(sigma)
dgamma(sigma, 2, 0.5, log = TRUE)
# Forma: Media 0 y DE 0.1, forma
log_prior_xi <- function(xi)
dnorm(xi, 0, 0.1, log = TRUE)

# Algoritmo Metrópolis
# Muestreo de las a-priori
metropolis <- function(log_target, initial_value,
n_iter, proposal_sd, lower = -Inf, upper = Inf) {
  samples <- numeric(n_iter) # Vector para almacenar muestras
  current <- initial_value
  current_log_dens <- log_target(current) # Evalúa la densidad en el punto
inicial
  accept <- 0 # Contador de aceptaciones
  for (i in 1:n_iter) {
    proposed <- rnorm(1, current, proposal_sd)
    if (proposed < lower || proposed > upper) { # Si está fuera de límites,
rechaza
      samples[i] <- current
      next
    }
    proposed_log_dens <- log_target(proposed)
    log_alpha <- proposed_log_dens - current_log_dens # Ratio de aceptación
    if (log(runif(1)) < log_alpha) { # Acepta o rechaza la propuesta
      current <- proposed
      current_log_dens <- proposed_log_dens
      accept <- accept + 1
    }
    samples[i] <- current # Guarda la muestra
  }
  cat("Tasa de aceptación:", accept / n_iter, "\n")
  return(samples) # Devuelve las muestras generadas
}

# Muestreo de distribuciones a-priori
set.seed(123)
# Generación de muestras con algoritmo Metrópolis
mu_prior_samples <- metropolis(log_prior_mu, 5, 2000, 2)
sigma_prior_samples <- metropolis(log_prior_sigma, 4, 3000, 4, lower = 0)
xi_prior_samples <- metropolis(log_prior_xi, 0, 10000, 0.2)

# Visualización de las a-priori muestreadas
hist_mu <- ggplot(data.frame(mu_prior_samples), aes(x = mu_prior_samples)) +
  geom_histogram(bins = 30, fill = "skyblue", color = "black", alpha = 0.7) +
  labs(x = "mu", y = "Frecuencia") +
  theme_minimal()

hist_sigma <- ggplot(data.frame(sigma_prior_samples), aes(x =
sigma_prior_samples)) +
  geom_histogram(bins = 30, fill = "lightgreen", color = "black", alpha = 0.7) +
  labs(x = "sigma", y = "Frecuencia") +
  theme_minimal()

hist_xi <- ggplot(data.frame(xi_prior_samples), aes(x = xi_prior_samples)) +
  geom_histogram(bins = 30, fill = "lightpink", color = "black", alpha = 0.7) +
  labs(x = "xi", y = "Frecuencia") +
  theme_minimal()

# Visualización de las a-priori teóricas
p_mu <- ggplot(data.frame(x = c(2, 8)), aes(x)) +

```

```

stat_function(fun = dnorm, args = list(mean = 5, sd = 1),
              color = "skyblue", size = 1.5) +
labs(x = "mu", y = "Densidad") +
theme_minimal()

p_sigma <- ggplot(data.frame(x = c(0, 10)), aes(x)) +
  stat_function(fun = dgamma, args = list(shape = 2, scale = 0.5),
              color = "lightgreen", size = 1.5) +
  labs(x = "sigma", y = "Densidad") +
  theme_minimal()

p_xi <- ggplot(data.frame(x = c(-0.3, 0.3)), aes(x)) +
  stat_function(fun = dnorm, args = list(mean = 0, sd = 0.1),
              color = "lightpink", size = 1.5) +
  labs(x = "xi", y = "Densidad") +
  theme_minimal()

# Títulos de la gráfica
titulo_izq <- grid::textGrob("Distribuciones teóricas",
gp = grid::gpar(fontsize = 14, fontface = "bold"))
titulo_der <- grid::textGrob("Distribuciones muestradas",
gp = grid::gpar(fontsize = 14, fontface = "bold"))
grid.arrange(titulo_izq, titulo_der,
              p_mu, hist_mu,
              p_sigma, hist_sigma,
              p_xi, hist_xi,
              ncol = 2, heights = c(0.3, 3, 3, 3))

# Especificación del modelo GEV en Stan

stan_model_code <- "
functions {
  // Función de densidad logarítmica de la distribución GEV
  real gev_lpdf(real y, real mu, real sigma, real xi) {
    real z = 1 + (y - mu) * xi / sigma;
    return -log(sigma) - (1 + 1/xi) * log(z) - pow(z, -1/xi);
  }
  // Función para generar valores aleatorios de la distribución GEV
  real gev_rng(real mu, real sigma, real xi) {
    real u = uniform_rng(0, 1);
    return mu + (sigma / xi) * (pow(-log(u), -xi) - 1);
  }
}
data {
  int<lower=0> N;          // Número de observaciones
  real<lower=0> y[N];     // Datos sísmicos (magnitudes)
}
parameters {
  real mu;               // Parámetro de localización
  real<lower=0> sigma;   // Parámetro de escala
  real xi;               // Parámetro de forma
}
model {
  // Priors
  mu ~ normal(5, 1);     // Prior para magnitudes bajas (~5)
  sigma ~ gamma(2, 0.5);
  xi ~ normal(0, 0.1);   // Prior restrictivo para estabilidad

  // Verosimilitud
  for (n in 1:N)
    target += gev_lpdf(y[n] | mu, sigma, xi);
}
generated quantities {

```

```

real y_rep[N];          // Réplicas para PPC
real y_pred[10];       // Predicciones futuras
vector[N] log_lik;     // Log-verosimilitud para LOO

// Generar réplicas para comprobación predictiva
for (n in 1:N)
  y_rep[n] = gev_rng(mu, sigma, xi);

// Generar 10 predicciones futuras
for (i in 1:10)
  y_pred[i] = gev_rng(mu, sigma, xi);

// Calcular log-verosimilitud para cada observación
for (n in 1:N)
  log_lik[n] = gev_lpdf(y[n] | mu, sigma, xi);
}
"

# Preparación de datos y ejecución del muestreo MCMC
stan_data <- list(N = length(datos), y = datos)
fit <- stan(model_code = stan_model_code, data = stan_data,
           chains = 1, iter = 2000, warmup = 1000,
           control = list(adapt_delta = 0.95))

# Diagnósticos y análisis posterior
# Extraer muestras posteriores y crear matriz para los parámetros
posterior <- extract(fit)
posterior_matrix <- as.matrix(fit, pars = c("mu", "sigma", "xi"))
# Se visualizan los intervalos creíbles de los parámetros
mcmc_intervals(posterior_matrix)

# Percentiles de las posteriores (2.5%, 50%, 97.5%)
print(fit, pars = c("mu", "sigma", "xi"), probs = c(0.025, 0.5, 0.975))

# Inference for Stan model: anon_model.
# 1 chains, each with iter=2000; warmup=1000; thin=1;
# post-warmup draws per chain=1000, total post-warmup draws=1000.
#
#      mean se_mean   sd  2.5%  50%  97.5% n_eff Rhat
# mu    4.74      0 0.06  4.63  4.74  4.87   590    1
# sigma 0.31      0 0.05  0.22  0.30  0.42   470    1
# xi    0.10      0 0.07 -0.03  0.09  0.25   634    1
#
# Graficas de las posteriores'
par(mfrow=c(1,3))
hist(prueba2$mu,xlab="mu",ylab="Frecuencia",main=" ")
hist(prueba2$sigma,xlab="sigma",ylab="Frecuencia",main=" ")
hist(prueba2$xi,xlab="xi",ylab="Frecuencia",main=" ")

# Traza del MCMC para diagnóstico de la convergencia
mcmc_trace(fit, pars = c("mu", "sigma", "xi")) +
ggtitle("Trazas de las cadenas MCMC")
# Las cadenas se deben mezclar bien

# Autocorrelación de las muestras
mcmc_acf(fit, pars = c("mu", "sigma", "xi")) +
ggtitle("Autocorrelación de los parámetros")
# Las autocorrelaciones deben caer rápidamente en el ACF

# Distribución Predictiva Posterior
set.seed(4321)
y_rep_matrix <- as.matrix(fit, pars = "y_rep")

```

```
# Grafica de las primeras 100 predictivas
ppc_dens_overlay(datos, y_rep_matrix[1:100,]) +
ggtitle("100 replicas de la distribución predictiva")
#
# Densidad Predictiva con N = 37,000 observaciones
# simuladas.
#
post_samples <- rstan::extract(fit)
y_rep_total <- as.vector(post_samples$y_rep)
#
# Estadística descriptiva de la predictiva
summary(y_rep_total)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.858  4.637  4.849  4.951  5.149 15.354
boxplot(y_rep_total)
#
# Densidad predictiva suavizada (kernel Density)
densidad_bayesiana <- density(y_rep_total)
# Densidad noparamétrica
kde_magnitudes_max <- density(x = datos)
hist(datos, freq = FALSE, nclass = 16, xlim = c(4,8),
xlab = "Magnitud", ylab = "Densidad", main = " ")
points(kde_magnitudes_max, type = "l", col = "red")
points(densidad_bayesiana, type = "l", col = "blue")
#
# Grafica de:
# Función de distribución acumulada predictiva
# Función de distribución empírica
# Función cuantil
par(mfrow=c(1,2))
# Función de distribución acumulada predictiva
cdf_predictiva_bayesiana <- ecdf(y_rep_total)
cdf_empirica <- ecdf(datos)
plot(cdf_predictiva_bayesiana, xlab="Magnitud",
ylab="probabilidad", xlim=c(4,8), col="blue", main=" ")
lines(cdf_empirica, col = "black", lty = 2)
#
# Gráfica de la función cuantil
pr <- seq(0.001,0.999,0.001)
pr
qqqs <- 0
for(i in 1:length(pr)){
qqqs[i] <- quantile(cdf_predictiva_bayesiana,pr[i])
}
plot(pr,qqqs,type="l",xlab="p",ylab="Cuantiles",col="blue")

# Tabulación de la función cuantil
pr <- c(seq(0.005,0.995,0.005),0.999)
pr
qqqs <- 0
for(i in 1:length(pr)){
qqqs[i] <- quantile(cdf_predictiva_bayesiana,pr[i])
}
data.frame(
  probl=pr[1:50], cuan1=qqqs[1:50],
  prob2=pr[51:100], cuan2=qqqs[51:100],
  prob3=pr[101:150], cuan3=qqqs[101:150],
  prob4=pr[151:200], cuan4=qqqs[151:200])

# Cálculo de los niveles de retorno
retorno <- seq(10,100,10)
p <- 1/retorno
data.frame(m = retorno,
```

```

nivel = quantile(cdf_predictiva_bayesiana,1-p)

      m  nivel
90%   10 5.520256
95%   20 5.824851
96.66667% 30 6.000662
97.5%  40 6.131887
98%   50 6.249562
98.33333% 60 6.342714
98.57143% 70 6.417452
98.75%  80 6.503537
98.88889% 90 6.566392
99%  100 6.627253
#
# Cálculo de los intervalos de predicción
# de los niveles de retorno

zq <- c(0.900,0.950,0.967,0.975,0.980,0.983,0.986,0.988,0.989,0.990)

qs <- matrix(0,nrow=37,ncol=10)
for(j in 1:37) {
cdf <- ecdf(prueba2[,j+3])
qs[j,] <- quantile(cdf,zq)
}

limites <- matrix(0,nrow=10,ncol=2)
for(i in 1:10){
limites[i,] <- quantile(qs[,i],c(0.025,0.975))
}
limites
      [,1]      [,2]
[1,] 5.465444 5.586548
[2,] 5.746483 5.928895
[3,] 5.875959 6.115604
[4,] 6.000379 6.311660
[5,] 6.084417 6.416776
[6,] 6.161104 6.536122
[7,] 6.227682 6.589178
[8,] 6.312203 6.707285
[9,] 6.340043 6.754175
[10,] 6.368076 6.853630

# Cuantiles extremos de la función de distribución
# Bayesiana predictiva
quantile(cdf_predictiva_bayesiana, c(0.9, 0.95, 0.99, 0.999, 0.999))
      90%   95%   99%   99.9%   99.9%
5.520256 5.824851 6.627253 8.015087 8.015087

# FIN
#####

```