STAT 501 PROJECT

Stochastic Modeling of Seismic Events: A Poisson Process Analysis of the Gulf of California Region

Abdhamid Mubarak, Owoyemi Qudus, Feri Syamsudin Pakiah

Department of Mathematics and Applied Statistics
King Fahd University of Petroleum & Minerals

Tuesday, 9 December 2025

Table of Contents I

Table of Contents II

Introduction

Earthquakes are among the most destructive natural phenomena, characterized by the sudden release of energy in the Earth's crust. Understanding the temporal distribution of these events is crucial for seismic risk assessment and hazard mitigation.

Historically, the Poisson Process has been the standard model for estimating earthquake recurrence rates.

- Seminal work by Gardner and Knopoff (1974) established that while main shocks often follow a Poisson distribution, the presence of aftershocks introduces clustering that violates the independence assumption.

Introduction

Earthquakes are among the most destructive natural phenomena, characterized by the sudden release of energy in the Earth's crust. Understanding the temporal distribution of these events is crucial for seismic risk assessment and hazard mitigation.

Historically, the Poisson Process has been the standard model for estimating earthquake recurrence rates.

- Seminal work by Gardner and Knopoff (1974) established that while main shocks often follow a Poisson distribution, the presence of aftershocks introduces clustering that violates the independence assumption.

Introduction

Earthquakes are among the most destructive natural phenomena, characterized by the sudden release of energy in the Earth's crust. Understanding the temporal distribution of these events is crucial for seismic risk assessment and hazard mitigation.

Historically, the Poisson Process has been the standard model for estimating earthquake recurrence rates.

- Seminal work by Gardner and Knopoff (1974) established that while main shocks often follow a Poisson distribution, the presence of aftershocks introduces clustering that violates the independence assumption.

Aim & Objective of the Study

Aim

The primary aim of this project is to statistically evaluate whether the temporal occurrence of major earthquakes in the Gulf of California region can be modeled as a memoryless Poisson process.

Objectives

- To analyze inter-arrival times of independent events and check for an **Exponential Distribution**.
- To test the **memoryless property** using ACF analysis.
- To perform a Kolmogorov–Smirnov test and Split-Half Stationarity test to verify the model at a 5% significance level.

Aim & Objective of the Study

Aim

The primary aim of this project is to statistically evaluate whether the temporal occurrence of major earthquakes in the Gulf of California region can be modeled as a memoryless Poisson process.

Objectives

- To analyze inter-arrival times of independent events and check for an **Exponential Distribution**.
- To test the **memoryless property** using ACF analysis.
- To perform a Kolmogorov–Smirnov test and Split-Half Stationarity test to verify the model at a 5% significance level.

Aim & Objective of the Study

Aim

The primary aim of this project is to statistically evaluate whether the temporal occurrence of major earthquakes in the Gulf of California region can be modeled as a memoryless Poisson process.

Objectives

- To analyze inter-arrival times of independent events and check for an **Exponential Distribution**.
- To test the **memoryless property** using ACF analysis.
- To perform a Kolmogorov–Smirnov test and Split-Half Stationarity test to verify the model at a 5% significance level.

Literature Review

- **Vere-Jones (1970)**
  Developed foundational stochastic models for earthquake occurrences, distinguishing between independent background events and triggered (aftershock) sequences. This work established the conceptual basis for applying Poisson processes in seismology.
- **Gardner & Knopoff (1974)**
  Demonstrated that earthquake catalogs contain dependent aftershocks that violate Poisson assumptions. Introduced declustering methods to isolate *main shocks*, a necessary preprocessing step before statistical modeling.

Literature Review

- **Vere-Jones (1970)**
  Developed foundational stochastic models for earthquake occurrences, distinguishing between independent background events and triggered (aftershock) sequences. This work established the conceptual basis for applying Poisson processes in seismology.
- **Gardner & Knopoff (1974)**
  Demonstrated that earthquake catalogs contain dependent aftershocks that violate Poisson assumptions. Introduced declustering methods to isolate *main shocks*, a necessary preprocessing step before statistical modeling.

Raw Earthquake Catalog

ID	Time (UTC)	Mag	Lat	Lon	Depth (km)
us6000j1ab	2023-01-04 13:22:10	4.9	29.45	-113.22	10.2
us6000j2df	2023-01-05 05:41:52	5.1	31.12	-115.04	12.5
us6000j3xy	2023-01-11 16:03:29	4.9	26.84	-110.75	9.8
us6000j4bb	2023-01-14 09:18:47	5.3	33.92	-118.51	14.0
us6000j5cc	2023-01-17 22:40:31	5.7	27.34	-111.22	11.0

Table 1: Snapshot of first 5 records of the dataset

Research Methodology I

This section presents the procedures used to construct, estimate, and validate the Poisson Process (PP) model for the seismic events of the Gulf of California (GoC) region compiled from January 2023 to November 2025.

The study area is geographically bound to encompass the major transformation systems of the GoC. The raw seismic catalog must be fully specified, including the source, time, and geographic coordinates used.

The analysis is restricted to events reliably recorded above a minimum magnitude of 4.9 (Mag). It is determined using the Maximum Likelihood Estimate (MLE) method by assessing the

Research Methodology II

fitness of the observed frequency–magnitude distribution to the Gutenberg–Richter Law:

$$\log_{10}(N) = a - bM \tag{1}$$

Only events with $M > Mag$ were retained, ensuring the calculated rate λ is not biased by detection limits.

To satisfy the assumption of event independence, dependent events (foreshocks and aftershocks) are removed using a declustering algorithm which isolates the background seismicity (mainshocks).

Research Methodology III

The procedure uses magnitude-dependent space (Δd) and time (Δt):

$$\log_{10}(\Delta t) = p_1 + p_2 M \quad \text{(days)} \tag{2}$$

$$\log_{10}(\Delta d) = q_1 + q_2 M \quad \text{(km)} \tag{3}$$

The de-clustered sequence of mainshocks was modeled as a Poisson process with constant rate parameter λ.
Under the Poisson process, the probability of observing exactly n events in a fixed time interval t is:

$$P(N(t) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \tag{4}$$

Research Methodology IV

Also, the defining property is that the time T elapsed between consecutive independent events follows an exponential distribution:

$$f(T \mid \lambda) = \lambda e^{-\lambda T}, \quad T > 0 \tag{5}$$

The mean occurrence rate λ is estimated using the Maximum Likelihood Estimator (MLE). Given N_{obs} total declustered events observed over the total period T_{obs}:

$$\lambda = \frac{N_{\text{obs}}}{T_{\text{obs}}} \quad \text{(events per time unit)} \tag{6}$$

Research Methodology V

The K–S test was used to verify the Poisson process assumption that the inter-arrival times (T_i) between consecutive mainshocks follow the exponential distribution with the estimated rate λ. The test statistic D_i was calculated as:

$$D = \max_i \left| F_{\text{emp}}(s) - F_{\text{th}}(s) \right|, \quad F_{\text{th}}(s) = 1 - e^{-s} \tag{7}$$

The Autocorrelation Test was applied directly to the sequence of inter-arrival times $(T_1, T_2, \ldots)$ to ensure that the procedure

Research Methodology VI

successfully removed all temporal dependencies. It was calculated for various time lags using the formula:

$$\widetilde{\rho}_k = \frac{\sum_{i=1}^{N-k}(T_i - \overline{T})(T_{i+k} - \overline{T})}{\sum_{i=1}^{N-k}(T_i - \overline{T})^2} \tag{8}$$

The Split-Half Test was used to assess the rate stability of the Poisson process by comparing the estimated event rates from two distinct, equal-length sub-periods of the catalog. The rates were estimated for the first half $(\widehat{\lambda}_1)$ and the second half $(\widehat{\lambda}_2)$, and the

Research Methodology VII

difference between the two rates was evaluated using the Standardized Z-Score:

$$Z = \frac{|\widehat{\lambda}_1 - \widehat{\lambda}_2|}{\sqrt{\frac{\widehat{\lambda}_1}{T_1} + \frac{\widehat{\lambda}_2}{T_2}}} \tag{9}$$

Geographic Distribution of Main Shocks

Spatial Clustering Along the Plate Boundary

map_mainshocks.png

- **High Concentration**: Events cluster primarily along the transform fault system of the Gulf of California.
- **Catalog Validation**: This confirms that the de-clustered catalog accurately captures the active plate-boundary corridor expected in the study area.

Geographic Distribution of Main Shocks

Spatial Clustering Along the Plate Boundary

map_mainshocks.png

- **High Concentration**: Events cluster primarily along the transform fault system of the Gulf of California.
- **Catalog Validation**: This confirms that the de-clustered catalog accurately captures the active plate-boundary corridor expected in the study area.

Inter-arrival Time Distribution

Validation of the Exponential Distribution

hist_exp.png

- **Short Waiting Times**: Most waiting times are short, with frequency decreasing exponentially as the gap between earthquakes increases.
- **Poisson Support**: The fitted exponential curve closely follows the empirical histogram, which strongly supports the Poisson-process assumption for the de-clustered data.

Inter-arrival Time Distribution

Validation of the Exponential Distribution

hist_exp.png

- **Short Waiting Times**: Most waiting times are short, with frequency decreasing exponentially as the gap between earthquakes increases.
- **Poisson Support**: The fitted exponential curve closely follows the empirical histogram, which strongly supports the Poisson-process assumption for the de-clustered data.

Independence of Inter-arrival Times

Autocorrelation Test (3.6)

acf_plot.png

- **Independence Confirmed**: Autocorrelation values for all lags $k \geq 1$ lie within the approximate 95% confidence bounds.
- **No Clustering**: There is no clear pattern of positive or negative correlation, indicating that successive waiting times are approximately independent.
- **Result**: The de-clustering algorithm was successful in removing temporal dependencies.

Independence of Inter-arrival Times

Autocorrelation Test (3.6)

acf_plot.png

- **Independence Confirmed**: Autocorrelation values for all lags $k \geq 1$ lie within the approximate 95% confidence bounds.
- **No Clustering**: There is no clear pattern of positive or negative correlation, indicating that successive waiting times are approximately independent.
- **Result**: The de-clustering algorithm was successful in removing temporal dependencies.

Independence of Inter-arrival Times

Autocorrelation Test (3.6)

acf_plot.png

- **Independence Confirmed**: Autocorrelation values for all lags $k \geq 1$ lie within the approximate 95% confidence bounds.
- **No Clustering**: There is no clear pattern of positive or negative correlation, indicating that successive waiting times are approximately independent.
- **Result**: The de-clustering algorithm was successful in removing temporal dependencies.

Discussion of Findings

Summary of Model Validation

- **Spatial Validation**: Patterns are consistent with the known tectonic structure of the Gulf of California, validating the choice of the study region.
- **Temporal Fit**: The close exponential fit to inter-arrival times and the lack of significant autocorrelation confirm the de-clustered main shocks satisfy the independence and distribution requirements of a Poisson process.
- **Stationarity**: Significant evidence supports the use of a stationary Poisson process to describe the occurrence of moderate-to-large earthquakes in this region over the study period.
- **Practical Application**: Under this model, the estimated rate parameter λ provides a practical tool for quantifying the probability of at least one event in any future time window.

Discussion of Findings

Summary of Model Validation

- **Spatial Validation**: Patterns are consistent with the known tectonic structure of the Gulf of California, validating the choice of the study region.
- **Temporal Fit**: The close exponential fit to inter-arrival times and the lack of significant autocorrelation confirm the de-clustered main shocks satisfy the independence and distribution requirements of a Poisson process.
- **Stationarity**: Significant evidence supports the use of a stationary Poisson process to describe the occurrence of moderate-to-large earthquakes in this region over the study period.
- **Practical Application**: Under this model, the estimated rate parameter λ provides a practical tool for quantifying the probability of at least one event in any future time window.

Discussion of Findings

Summary of Model Validation

- **Spatial Validation**: Patterns are consistent with the known tectonic structure of the Gulf of California, validating the choice of the study region.
- **Temporal Fit**: The close exponential fit to inter-arrival times and the lack of significant autocorrelation confirm the de-clustered main shocks satisfy the independence and distribution requirements of a Poisson process.
- **Stationarity**: Significant evidence supports the use of a stationary Poisson process to describe the occurrence of moderate-to-large earthquakes in this region over the study period.
- **Practical Application**: Under this model, the estimated rate parameter λ provides a practical tool for quantifying the probability of at least one event in any future time window.

Discussion of Findings

Summary of Model Validation

- **Spatial Validation**: Patterns are consistent with the known tectonic structure of the Gulf of California, validating the choice of the study region.
- **Temporal Fit**: The close exponential fit to inter-arrival times and the lack of significant autocorrelation confirm the de-clustered main shocks satisfy the independence and distribution requirements of a Poisson process.
- **Stationarity**: Significant evidence supports the use of a stationary Poisson process to describe the occurrence of moderate-to-large earthquakes in this region over the study period.
- **Practical Application**: Under this model, the estimated rate parameter λ provides a practical tool for quantifying the probability of at least one event in any future time window.

Application of the Poisson Model

Forecasting Seismic Activity

- After validating that the earthquake occurrences in the Gulf of California follow a stationary Poisson process, the model can be used for **forecasting**.
- The key parameter is the estimated rate λ:

$$\lambda = \frac{N_{\text{events}}}{T_{\text{total}}}$$

  - This rate represents the **average number of main shocks per unit time** (e.g., per hour).
- Once λ is known, we can compute crucial probabilities for hazard assessment:
  - The probability of observing at least one event in a given window,
  - The probability of multiple events,
  - The expected waiting time before the next event.

Application of the Poisson Model

Forecasting Seismic Activity

- After validating that the earthquake occurrences in the Gulf of California follow a stationary Poisson process, the model can be used for **forecasting**.
- The key parameter is the estimated rate λ:

$$\lambda = \frac{N_{\text{events}}}{T_{\text{total}}}$$

  - This rate represents the **average number of main shocks per unit time** (e.g., per hour).
- Once λ is known, we can compute crucial probabilities for hazard assessment:
  - The probability of observing at least one event in a given window,
  - The probability of multiple events,
  - The expected waiting time before the next event.

Application of the Poisson Model

Forecasting Seismic Activity

- After validating that the earthquake occurrences in the Gulf of California follow a stationary Poisson process, the model can be used for **forecasting**.
- The key parameter is the estimated rate λ:

$$\lambda = \frac{N_{\text{events}}}{T_{\text{total}}}$$

  - This rate represents the **average number of main shocks per unit time** (e.g., per hour).
- Once λ is known, we can compute crucial probabilities for hazard assessment:
  - The probability of observing at least one event in a given window,
  - The probability of multiple events,
  - The expected waiting time before the next event.

Prediction: Waiting Time Distribution

- Under the Poisson process, the waiting time between independent events follows an exponential distribution:

$$P(T \leq t) = 1 - e^{-\lambda t}$$

- This allows us to answer questions such as:
  - *"What is the probability that an earthquake occurs within the next day?"*
  - *"How likely is another event within a week or a month?"*

Time Window	t **(hours)**	**Probability** $P(T \leq t)$
1 day (24 h)	24	0.030
7 days (168 h)	168	0.194
30 days (720 h)	720	0.602

These probabilities are computed using the estimated rate λ.

Prediction: Number of Earthquakes in a Time Window

- The number of earthquakes expected in time interval t follows:

$$P(N(t) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

- Examples of practical use:
  - Probability of **no** earthquake occurring in the next 7 days:

$$P(N(7 \text{ days}) = 0) = e^{-\lambda t}$$

  - Probability of **one or more** events:

$$P(N(t) \geq 1) = 1 - e^{-\lambda t}$$

- This helps estimate short-term and medium-term seismic hazard.

Prediction: Number of Earthquakes in a Time Window

- The number of earthquakes expected in time interval t follows:

$$P(N(t) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

- Examples of practical use:
  - Probability of **no** earthquake occurring in the next 7 days:

$$P(N(7 \text{ days}) = 0) = e^{-\lambda t}$$

  - Probability of **one or more** events:

$$P(N(t) \geq 1) = 1 - e^{-\lambda t}$$

- This helps estimate short-term and medium-term seismic hazard.

Prediction: Number of Earthquakes in a Time Window

- The number of earthquakes expected in time interval t follows:

$$P(N(t) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

- Examples of practical use:
  - Probability of **no** earthquake occurring in the next 7 days:

$$P(N(7 \text{ days}) = 0) = e^{-\lambda t}$$

  - Probability of **one or more** events:

$$P(N(t) \geq 1) = 1 - e^{-\lambda t}$$

- This helps estimate short-term and medium-term seismic hazard.

Interpretation for the Gulf of California

- The forecast shows a **low but non-zero daily probability** of a magnitude ≥ 4.9 event.
- Over longer windows (weeks to months), the likelihood increases substantially due to the constant rate.
- The Poisson model provides:
  - A **quantitative** estimate of near-future earthquake likelihood,
  - A **simple** yet effective tool for risk communication,
  - A **baseline model** for more complex hazard modeling.
- These predictions can assist researchers, planners, and civil protection agencies in short-term preparedness decisions.

Interpretation for the Gulf of California

- The forecast shows a **low but non-zero daily probability** of a magnitude ≥ 4.9 event.
- Over longer windows (weeks to months), the likelihood increases substantially due to the constant rate.
- The Poisson model provides:
  - A **quantitative** estimate of near-future earthquake likelihood,
  - A **simple** yet effective tool for risk communication,
  - A **baseline model** for more complex hazard modeling.
- These predictions can assist researchers, planners, and civil protection agencies in short-term preparedness decisions.

Interpretation for the Gulf of California

- The forecast shows a **low but non-zero daily probability** of a magnitude ≥ 4.9 event.
- Over longer windows (weeks to months), the likelihood increases substantially due to the constant rate.
- The Poisson model provides:
  - A **quantitative** estimate of near-future earthquake likelihood,
  - A **simple** yet effective tool for risk communication,
  - A **baseline model** for more complex hazard modeling.
- These predictions can assist researchers, planners, and civil protection agencies in short-term preparedness decisions.

Interpretation for the Gulf of California

- The forecast shows a **low but non-zero daily probability** of a magnitude ≥ 4.9 event.
- Over longer windows (weeks to months), the likelihood increases substantially due to the constant rate.
- The Poisson model provides:
  - A **quantitative** estimate of near-future earthquake likelihood,
  - A **simple** yet effective tool for risk communication,
  - A **baseline model** for more complex hazard modeling.
- These predictions can assist researchers, planners, and civil protection agencies in short-term preparedness decisions.

Summary of Key Findings I

- The earthquake catalog was filtered and declustered so that only **independent main shocks** with magnitude ≥ 4.9 remained.
- When looking at the timing between earthquakes, the gaps (inter-arrival times) follow a pattern that decreases smoothly, i.e. meaning short waiting times are common and long waiting times are rare and matches the shape of an **exponential distribution**, which is exactly what a Poisson process predicts.

Summary of Key Findings II

- The independence test (ACF) showed no meaningful correlation between one earthquake and the next — **earthquakes do not "remember" each other** in this dataset.
- The split-half test showed the average rate of earthquakes is basically the same in the first and second half of the study period, meaning **the activity level is steady and not changing over time**.

Conclusion

- The data behaves exactly as expected from a **stationary Poisson process**.
- This means:
  - Earthquakes occur randomly in time,
  - Each earthquake happens independently of the previous one,
  - The average rate of earthquakes remains stable,
  - The waiting time between events follows a predictable exponential pattern.
- In practical terms, the Poisson model allows us to estimate the probability of at least one earthquake happening within a chosen time window (day, week, month, etc.).

Conclusion

- The data behaves exactly as expected from a **stationary Poisson process**.
- This means:
  - Earthquakes occur randomly in time,
  - Each earthquake happens independently of the previous one,
  - The average rate of earthquakes remains stable,
  - The waiting time between events follows a predictable exponential pattern.
- In practical terms, the Poisson model allows us to estimate the probability of at least one earthquake happening within a chosen time window (day, week, month, etc.).

Conclusion

- The data behaves exactly as expected from a **stationary Poisson process**.
- This means:
  - Earthquakes occur randomly in time,
  - Each earthquake happens independently of the previous one,
  - The average rate of earthquakes remains stable,
  - The waiting time between events follows a predictable exponential pattern.
- In practical terms, the Poisson model allows us to estimate the probability of at least one earthquake happening within a chosen time window (day, week, month, etc.).

References I

- Gardner, J. K., & Knopoff, L. (1974). *Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?* Bulletin of the Seismological Society of America, 64(5), 1363–1367.
- Vere-Jones, D. (1970). *Stochastic models for earthquake occurrence.* Journal of the Royal Statistical Society: Series B, 32(1), 1–62.
- Reiter, L. (1990). *Earthquake Hazard Analysis: Issues and Insights.* New York: Columbia University Press.
- United States Geological Survey (USGS). (2025). *Earthquake Catalog.* Retrieved from `https://earthquake.usgs.gov/earthquakes/search/`

Thank You!

Stochastic Modeling of Seismic Events

Gulf of California Region Analysis

Questions are welcome.