

UNIVERSIDAD NACIONAL DEL SANTA
ESCUELA DE POSGRADO
Programa de Doctorado en Estadística Matemática



UNS
ESCUELA DE
POSGRADO

**Modelos Semi-Markovianos para la evaluación del tiempo,
lugar, magnitud y profundidad de los sismos ocurridos
entre Lima - Barranca, Perú, 2022**

**Tesis para optar el grado de
Doctor en Estadística Matemática**

Autor:

Mg. Risco Franco, Carlos Alvaro
Código Orcid: 0000-0003-1323-436X
DNI N° 00419692

Asesor:

Dr. Rubio Jacobo, Luis Alberto
Código ORCID: 0000-0001-5060-9998
DNI N° 18069833

Línea de Investigación
Pronósticos y modelamiento estadístico

Nuevo Chimbote - PERÚ
2024



UNS
ESCUELA DE
POSGRADO

CONSTANCIA DE ASESORAMIENTO DE LA TESIS

Yo, **Luis Alberto, Rubio Jácobo**, mediante la presente certifico mi asesoramiento de la Tesis de doctorado titulada: **Modelos semi-Markovianos para la evaluación del tiempo, lugar, magnitud y profundidad de los sismos ocurridos entre Lima – Barranca Perú 2022**, elaborado por el doctorando **Ms. Risco Franco, Carlos Álvaro**, para obtener el Grado Académico de **Doctor en Estadística Matemática** en la Escuela de Posgrado de la Universidad Nacional del Santa.

Dr. Rubio Jácobo, Luis Alberto

ASESOR

CODIGO ORCID 0000-0001-5060-9998

DNI N° 18069833



UNS
ESCUELA DE
POSGRADO

HOJA DEL AVAL DEL JURADO EVALUADOR

**MODELOS SEMI-MARKOVIANOS PARA LA EVALUACIÓN DEL TIEMPO,
LUGAR, MAGNITUD Y PROFUNDIDAD DE LOS SISMOS OCURRIDOS
ENTRE LIMA - BARRANCA PERU 2022**

TESIS PARA OPTAR EL GRADO DE DOCTOR EN ESTADÍSTICA MATEMÁTICA

Revisado y Aprobado por el Jurado Evaluador:

Dra. América Odar Rosario

PRESIDENTE

CODIGO ORCID 0000-0002-1805-2007

DNI N° 17850879

Dr. Lizandro Baldomero Reyna Zegarra

SECRETARIO

CODIGO ORCID 0000-0002-5474-8423

DNI N° 17850934

Dr. Luis Alberto Rubio Jácono

VOCAL

CODIGO ORCID 0000-0001-5060-9998

DNI N° 18069833



UNS
ESCUELA DE
POSGRADO

ACTA DE EVALUACIÓN DE SUSTENTACIÓN DE TESIS

A los veinticuatro días del mes de octubre del año 2024, siendo las 4.27pm horas, en el aula P-01 de la Escuela de Posgrado de la Universidad Nacional del Santa, se reunieron los miembros del Jurado Evaluador, designados mediante Resolución Directoral N° 339-2024-EPG-UNS de fecha 06.07.2024, conformado por los docentes: Dra. América Odar Rosario (Presidenta), Dr. Lizandro Baldomero Reyna Zegarra (Secretario) y Dr. Luis Alberto Rubio Jacobo (Vocal); con la finalidad de evaluar la tesis titulada **"MODELOS SEMI-MARKOVIANOS PARA LA EVALUACIÓN DEL TIEMPO, LUGAR, MAGNITUD Y PROFUNDIDAD DE LOS SISMOS OCURRIDOS ENTRE LIMA – BARRANCA PERÚ 2022"**; presentado por el tesista Carlos Alvaro Risco Franco, egresado del programa de **Doctorado en Estadística Matemática**.

Sustentación autorizada mediante Resolución Directoral N° 529-2024-EPG-UNS de fecha 22 de octubre de 2024.


La presidenta del jurado autorizó el inicio del acto académico; producido y concluido el acto de sustentación de tesis, los miembros del jurado procedieron a la evaluación respectiva, haciendo una serie de preguntas y recomendaciones al tesista, quien dio respuestas a las interrogantes y observaciones.

El jurado después de deliberar sobre aspectos relacionados con el trabajo, contenido y sustentación del mismo y con las sugerencias pertinentes, declara la sustentación como Aprobado, asignándole la calificación de 19.

Siendo las 5.44pm horas del mismo día se da por finalizado el acto académico, firmando la presente acta en señal de conformidad.


Dra. América Odar Rosario
Presidenta


Dr. Lizandro Baldomero Reyna Zegarra
Secretario


Dr. Luis Alberto Rubio Jacobo
Vocal



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por **Turnitin**. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Francisco Risco
Título del ejercicio: MODELOS SEMI-MARKOVIANOS PARA LA EVALUACION DEL T...
Título de la entrega: TESIS DOCTORADO CARLOS RISCO -OCTUBRE 24-INDICEF2.d...
Nombre del archivo: TESIS_DOCTORADO_CARLOS_RISCO_-OCTUBRE_24-INDICEF2....
Tamaño del archivo: 11.33M
Total páginas: 98
Total de palabras: 20,467
Total de caracteres: 101,664
Fecha de entrega: 05-dic.-2024 02:31p. m. (UTC-0500)
Identificador de la entre... 2542032795

MODELOS SEMI-MÁRKOVIANOS PARA LA
EVALUACIÓN DEL TIEMPO, LUGAR, MAGNITUD Y
PROFUNDIDAD DE LOS SISMOS OCURRIDOS
ENTRE LIMA Y BARRANCA PERÚ, 2022

Informe de Tesis para optar el grado de
Doctor en Estadística Matemática

Autor:

Mg. Carlos Álvaro Risco Franco

Asesor:

Dr. Luis Alberto Rubio Jácome

TESIS DOCTORADO CARLOS RISCO -OCTUBRE 24- INDICEF2.docx

INFORME DE ORIGINALIDAD

16%

INDICE DE SIMILITUD

16%

FUENTES DE INTERNET

1%

PUBLICACIONES

%

TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1	www.utdt.edu Fuente de Internet	3%
2	www.slideshare.net Fuente de Internet	2%
3	doku.pub Fuente de Internet	2%
4	revistas.uni.edu.pe Fuente de Internet	1%
5	www.bdigital.unal.edu.co Fuente de Internet	1%
6	idoc.pub Fuente de Internet	1%
7	alicia.concytec.gob.pe Fuente de Internet	1%
8	qdoc.tips Fuente de Internet	1%
9	edoc.pub Fuente de Internet	

Dedicatoria:
A mis padres, Lorenzo Risco Barrueto y
Eloísa Franco Bouby,
por su amor incondicional.

Agradecimientos:

A mi familia, a mis colegas y administrativos de la UNI y la UNS y a mi asesor el Dr. Luis Alberto Rubio Jacobo por su apoyo constante para la realizaci3n de la presente investigaci3n.

ÍNDICE

CERTIFICACIÓN DEL ASESOR.....	ii
AVAL DEL JURADO.....	iii
RESÚMEN.....	xi
ABSTRACT.....	xii
CAPÍTULO I: INTRODUCCIÓN	01
1.1. DESCRIPCIÓN Y FORMULACIÓN DEL PROBLEMA	03
1.2. FORMULACIÓN DE LOS OBJETIVOS	04
1.3. FORMULACIÓN DE LAS HIPÓTESIS	04
1.4. JUSTIFICACIÓN E IMPORTANCIA	04
CAPÍTULO II: MARCO TEÓRICO	06
2.1. ANTECEDENTES	06
2.2. MARCO CONCEPTUAL	10
2.3. EL MODELO DE RENOVACIÓN	13
2.4. EL MODELO DE LAS CADENAS DE MÁRKOV	18
2.5. EL MODELO SEMI-MÁRKOV	26
2.6. LA DISTRIBUCIÓN DE WEIBULL	28
CAPÍTULO III: METODOLOGÍA	30
3.1. TIPO DE INVESTIGACIÓN	30
3.2. MÉTODO	30
3.3. DISEÑO DE LA INVESTIGACIÓN	30
3.4. OPERACIONALIZACIÓN DE LAS VARIABLES	31
3.5. TÉCNICA E INSTRUMENTO DE RECOLECCIÓN DE DATOS	31
3.6. PROCEDIMIENTO UTILIZADO	32
3.7. TÉCNICAS DE PROCESAMIENTO DE DATOS	32
CAPÍTULO IV: RESULTADOS Y DISCUSIÓN	33

4.1 PRESENTACIÓN DE RESULTADOS	33
4.2 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LAS ZONAS GEOGRÁFICAS DE LOS SISMOS ENTRE LIMA Y BARRANCA	34
4.3 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LA MAGNITUD LOS SISMOS ENTRE LIMA Y BARRANCA	49
4.4 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LA PROFUNDIDAD DE LOS SISMOS	54
4.5 ANÁLISIS DE LOS RESULTADOS Y DISCUSIÓN	60
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES	63
5.1 CONCLUSIONES	63
5.2 RECOMENDACIONES	63
REFERENCIAS BIBLIOGRÁFICAS	64
ANEXO 1 LOS DATOS DE LOS SISMOS DESDE 1961 A 2022	66
ANEXO 2 TIEMPOS DE PERMANENCIA EN LOS ESTADOS DE LAS ZONAS GEOGRÁFICAS, MAGNITUDES Y PROFUNDIDADES	73
ANEXO 3 TABLAS DE RESULTADOS SOBRE LA APLICACIÓN DEL MODELO SEMI-MÁRKOV A LA MAGNITUD Y PROFUNDIDAD DE LOS SISMOS ENTRE LIMA Y BARRANCA	76
ANEXO 4 CÓDIGOS EN EL PROGRAMA R SOBRE EL PROCESAMIENTO DE LOS DATOS DE LOS SISMOS	80

ÍNDICE DE TABLAS

Tabla 1: Operacionalización de las variables que participan en el estudio	31
Tabla 2: Latitud y longitud de las zonas entre Lima y Barranca	34
Tabla 3: Estadísticos descriptivos comparativos de las zonas geográficas	37
Tabla 4: Matriz de frecuencias de las zonas geográficas	37
Tabla 5: Matriz de transición para las zonas entre Lima y Barranca	38
Tabla 6: Ajuste de los tiempos de permanencia para la transición (1,1) a las distribuciones que se presentan	39
Tabla 7: Parámetros de las distribuciones de los tiempos de espera de las transiciones de las zonas geográficas entre Lima y Barranca.	42
Tabla 8: F1 Matriz de probabilidad para $t=1$	43
Tabla 9: F2 Matriz de probabilidad para $t=3$	44
Tabla 10: F3 Matriz de probabilidad para $t=6$	44
Tabla 11: Q1 Matriz de probabilidad interválica para $t=1$	44
Tabla 12: Q2 Matriz de probabilidad interválica para $t=3$	45
Tabla 13: Q3 Matriz de probabilidad interválica para $t=6$	45
Tabla 14 B2 Matriz de probabilidad interválica para t entre 1 y 3	46
Tabla 15: B3 Matriz de probabilidad interválica para t entre 3 y 6	46
Tabla 16: H1 Matriz de probabilidad de ocurrencia de estados, $t=1$	46
Tabla 17 H2 Matriz de probabilidad de ocurrencia de estados, $t=3$	47
Tabla 18: H3 Matriz de probabilidad de ocurrencia de estados, $t=6$	47
Tabla 19: D1 Matriz de probabilidad de supervivencia, $t=1$	48
Tabla 20: D2 Matriz de probabilidad de supervivencia, $t=3$	48
Tabla 21: D3 Matriz de probabilidad de supervivencia, $t=6$	48
Tabla 22: Matriz de frecuencias de los estados de la magnitud	49
Tabla 23: Matriz de transición para los estados de la magnitud de los sismos	50
Tabla 24: Parámetros de las distribuciones para los tiempos de espera para las transiciones de los estados de la magnitud	50
Tabla 25: F2 Matriz de probabilidad para $t=3$	51
Tabla 26: Q2 Matriz de probabilidad interválica para $t=3$	52
Tabla 27: H2 Matriz de probabilidad de ocurrencia de estados para $t=3$	52
Tabla 28: D2 Matriz de probabilidad de supervivencia para $t=3$	53

Tabla 29: Matriz de frecuencias de los estados de la profundidad	55
Tabla 30: Matriz de transición para los estados de profundidad de los sismos	55
Tabla 31: Parámetros de las distribuciones para los tiempos de espera para las transiciones de los estados de la profundidad	56
Tabla 32: F2 Matriz de probabilidad para $t=3$	57
Tabla 33: Q2 Matriz de probabilidad interválica para $t=3$	58
Tabla 34: H2 Matriz de probabilidad de ocurrencia de estados para $t=3$	58
Tabla 35: D2 Matriz de probabilidad de supervivencia para $t=3$	58

ÍNDICE DE FIGURAS

Figura 1: Los paralelos y los meridianos que forman una retícula	13
Figura 2: Las densidades de la distribución Weibull $X \sim WEI(r, s)$, con $r = 1$ y $s = (0.5, 1, 2, 5)$	28
Figura 3: Los sismos de la zona entre Lima y Barranca con datos del IGP	33
Figura 4: Sismos en la zona entre Huacho y Barranca	35
Figura 5: Sismos en la zona entre Lima y Huaral	35
Figura 6: Sismos en la zona entre Chilca y Asia	36
Figura 7: Sismos en la zona Sierra de Lima	36
Figura 8: Gráficos de diagnóstico para la distribución Weibull de los tiempos de espera para la transición t_{11} (zonas geográficas)	40
Figura 9: Gráficos de diagnóstico para la distribución Gamma de los tiempos de espera para la transición t_{11} (zonas geográficas)	41
Figura 10: Frecuencias de la magnitud de los sismos entre Lima y Barranca	49
Figura 11: Gráficos de diagnóstico para la distribución Weibull de los tiempos de espera para la transición t_{11} (Magnitud)	51
Figura 12: Frecuencia de la profundidad de los sismos entre Lima y Barranca	54
Figura 13: Gráficos de diagnóstico para la distribución Weibull de los tiempos de espera para la transición t_{11} (profundidad)	57
Figura 14. La zona de Chilca resaltada con amarillo muestra la probabilidad 0.461 de ocurrencia de un sismo con magnitud mayor a 5.8 Mw y superficial (menor a 60 Km)	61

ÍNDICE DE ANEXOS

Anexo 1: Los datos de tiempo, latitud, longitud, magnitud y profundidad de los sismos utilizados en el presente estudio.	66
Anexo 2: Tiempos de permanencia	73
Anexo 3: Tablas de resultados	76
Anexo 4: Códigos en el programa R	80
Anexo 5: Latitud y longitud de los límites de las zonas geográficas en estudio	84

RESUMEN

El Perú se encuentra en la región sísmica del Pacífico Sur y es frecuentemente afectado por sismos. Terremotos como el ocurrido en Lima en octubre de 1966 y 1974, con cuantiosas pérdidas de vidas humanas y destrucción de viviendas e infraestructura, presentan una necesidad prioritaria de estudios de prevención y evaluación de riesgos de la ocurrencia de sismos de alta magnitud en esta región.

La presente investigación tiene como objetivo identificar zonas con mayor probabilidad de ocurrencia de sismos de alta magnitud en el área de estudio entre Lima y Barranca. Para ello se utilizan herramientas descriptivas, así como el modelo semi-Márkov. Se emplea el modelo semi-Márkov para predecir la magnitud y el lugar de la ocurrencia de eventos sísmicos, asimismo, se amplía el análisis considerando la profundidad de los sismos como una dimensión más. En la investigación se concluye que la zona de Chilca es la que presenta la mayor probabilidad de ocurrencia de sismos de alta magnitud y profundidad superficial.

Palabras clave: Sismos, semi-Márkov, cadenas de Márkov.

ABSTRACT

Peru is located in the South Pacific seismic region and is frequently affected by earthquakes. Earthquakes such as those that occurred in Lima in October 1966 and 1974, with significant loss of human life and destruction of homes and infrastructure, present a priority need for prevention studies and risk assessment of the occurrence of high magnitude earthquakes in this area.

The present research aims to identify areas with a higher probability of occurrence of high magnitude earthquakes in the study area between Lima and Barranca. Descriptive tools are used, as well as the semi-Márkov model. The semi-Márkov model is used to predict the magnitude and location of the occurrence of seismic events, however, here the analysis is extended to the depth of the earthquakes as an additional dimension. The Chilca area has been identified as having the highest probability of occurrence of high magnitude and shallow depth earthquakes.

Keywords: Earthquakes, semi-Márkov, Márkov chains.

CAPÍTULO I: INTRODUCCIÓN

La ocurrencia de sismos de diversa magnitud en el Perú es muy frecuente. Según el IGP, en el año 2022, ocurrieron 807 sismos y, si se toma en cuenta solo los sismos de alta magnitud, esto es sismos con magnitud mayor o igual a 5.0 Mw (*), entre Lima y Barranca, el número promedio es de 4 al año. Estas constantes ocurrencias sísmicas plantean la necesidad de la realización de estudios de prevención, evaluación de riesgos, pronósticos, etc.

En el mundo, el problema del pronóstico de los sismos es un desafío, debido a que es un fenómeno complejo, en el que participan muchas variables, lo cual lo hace impredecible y aleatorio. Sin embargo, con el enfoque probabilístico, algunos han ensayado con modelos llamados sin memoria como el proceso Poisson, para la predicción de dichos fenómenos naturales, Masala (2012).

Se conoce por teoría estadística que, la aplicación del modelo de Poisson tiene como supuesto que los eventos son independientes debido a que cada sismo es independiente en el tiempo y en el espacio de los sismos anteriores. Esta afirmación es controversial puesto que la realidad muestra que este supuesto no se cumple.

Ante esta situación se podría utilizar el proceso de Márkov, que supera las limitaciones del proceso de Poisson. Gracias a la teoría estadística se conoce que las distribuciones de los tiempos de espera en los estados en el proceso de Márkov deben distribuirse como una exponencial, convirtiéndose esta característica en una limitación para su uso en el caso de los sismos.

Masala (2012) agrega que, un nuevo enfoque basado en modelos semi-Márkov permite tomar en consideración el supuesto más realista de la dependencia de los eventos en el tiempo y en el espacio. Estas distribuciones, en un entorno del proceso semi-Márkov, pueden ser de cualquier tipo, refiriéndose a las distribuciones de los tiempos de espera en los estados.

En el presente trabajo, por un lado, se busca identificar la zona geográfica entre Lima y Barranca con mayor probabilidad de ocurrencia de sismos de alta magnitud, aplicando el modelo semi-Márkov en un tiempo determinado; por otro lado, al repetir este procedimiento para la magnitud y también para la profundidad de los sismos, estamos abordando el tema del pronóstico del lugar, magnitud y profundidad en un tiempo determinado.

El presente trabajo se diferencia con la investigación de Masala (2012), el cual utiliza el modelo semi-Márkov solo para determinar la magnitud de los sismos. Asimismo, se diferencia con la investigación de Sadeghian (2012), quien también aplica el modelo semi-Márkov para (*) Mw es la magnitud de momento sísmico, resume en un único número la cantidad de energía liberada por el sismo.

identificar las zonas geográficas y las magnitudes de los sismos, realizando pronósticos de lugar y magnitud. En la presente investigación, en la que también se aplica el modelo semi-Márkov para determinar las zonas geográficas de la ocurrencia de sismos en la región entre Lima y Barranca en el Perú, además de realizar pronósticos sobre la magnitud, se agrega la variable profundidad de los sismos de alta magnitud.

El presente estudio es cuantitativo, descriptivo y correlacional, de corte transversal. Los datos de los sismos con magnitudes mayores o iguales a 5.0 Mw, se obtuvieron de la página web del IGP desde marzo de 1961 a mayo del 2022.

Entre los resultados más importantes del presente trabajo, por un lado, se presenta que entre Lima y Barranca se ha identificado a la zona de Chilca como la que presenta la mayor probabilidad de ocurrencia de un sismo de alta magnitud; por otro lado, se ha determinado el intervalo de magnitud entre 5.4 a 5.7 Mw y a la profundidad menor a 60 Km, con la mayor probabilidad de ocurrencia.

La investigación realizada contiene 5 capítulos. En el capítulo 1, se aborda la formulación del problema, los objetivos, la formulación de la hipótesis, la justificación e importancia del estudio.

En el capítulo 2, se desarrolla el marco teórico, específicamente, los antecedentes y el marco conceptual.

En el Capítulo 3, se presenta la metodología, el diseño de la investigación, la operacionalización de las variables, las técnicas de análisis de datos, entre otros.

En el capítulo 4, se presentan los resultados y discusión sobre la aplicación del modelo semi-Márkov a las zonas geográficas estudiadas entre Lima y Barranca. También se repite la aplicación del modelo semi-Márkov a los estados de la magnitud. Esta aplicación se extiende también a la profundidad de los sismos, lo que constituye nuestro aporte, dado que hasta ahora no se ha encontrado trabajos que consideren las 4 dimensiones: tiempo, lugar, magnitud y profundidad de los sismos.

Finalmente, en el capítulo 5, se presentan las conclusiones y recomendaciones a las cuales se ha llegado.

1.1 DESCRIPCIÓN Y FORMULACIÓN DEL PROBLEMA

Los terremotos, ocurridos en Lima el 17 de octubre de 1966 y el 03 de octubre de 1974, en Ancash en mayo de 1970 y en Ica en agosto del 2007, dejando decenas de miles de muertos, infraestructura dañada y afectando la economía del país, son una realidad lamentable.

Hernando Tavera, jefe del IGP, señala que, en la costa central del Perú, que incluye a Lima, se está acumulando una deformación desde hace más de 275 años y que no solo la costa de Lima sufriría sismos de gran tamaño, sino que también están en riesgo Ancash, Ica, Arequipa, Moquegua y Tacna (El Peruano, 2023). Asimismo “El análisis de los sismos con magnitudes entre 4.0Mw y 7.5Mw (magnitudes intermedias) ha permitido observar, en sus secuencias, determinados patrones sísmicos predecesores (PSP) que emergen y anticipan la aproximación de un gran terremoto” (Flores, Tavera, & Rodriguez, 2012) .

Esta ocurrencia sísmica no solo ocurre en el Perú sino también en distintas partes del mundo. Por ello, diversos investigadores han ensayado la aplicación de modelos, que buscan explicar la magnitud de los sismos y su localización geográfica en función del tiempo. En estos intentos han considerado los modelos de series de tiempo, los modelos de renovación, entre ellos el modelo Poisson y el modelo exponencial han sido los más utilizados. Estos mismos se apoyan en el supuesto de que los sismos son independientes en el tiempo y en el espacio, en particular, con el evento anterior; sin embargo, en la realidad este supuesto no se cumple. Es por ello que los investigadores creen que el modelo semi-Márkov se ajusta mejor a los datos de los sismos.

En un proceso de Márkov, según la teoría estadística, los tiempos de permanencia del proceso en los diferentes estados siguen una distribución exponencial, mientras que en un proceso semi-Márkov, esos tiempos en los estados, pueden tener cualquier distribución. Esta característica lo convierte en un modelo más flexible al ser utilizado.

Debido al problema que aborda el presente estudio es **la predicción de sismos de gran magnitud**, el cual es de gran complejidad, en el presente trabajo se pretende identificar las zonas, magnitudes, profundidad y tiempos, con mayores probabilidades de ocurrencia de sismos, en la región entre Lima y Barranca. Por tanto, nuestra pregunta de investigación es la siguiente:

¿Cuáles son las áreas, magnitudes, profundidades, y tiempos con mayores probabilidades de ocurrencia de sismos de gran magnitud en la región entre Lima y Barranca?

1.2 FORMULACIÓN DE LOS OBJETIVOS

1.2.1 OBJETIVO GENERAL

Evaluar el tiempo, magnitud y profundidad, así como identificar zonas de mayor riesgo de sismos entre Lima y Barranca, ocurridos entre 1961 y 2022, usando los modelos semi-Márkov.

1.2.2 OBJETIVOS ESPECÍFICOS

- a. Identificar una zona geográfica con mayor probabilidad de ocurrencia de sismos de alta magnitud entre Lima y Barranca.
- b. Identificar un intervalo de la magnitud con mayor probabilidad de ocurrencia en los sismos entre Lima y Barranca.
- c. Identificar un intervalo de la profundidad con mayor probabilidad de ocurrencia en los sismos entre Lima y Barranca.
- d. Realizar predicciones en el tiempo, de las zonas, magnitud y profundidad de la ocurrencia de sismos de alta magnitud entre Lima y Barranca.

1.3 FORMULACIÓN DE LA HIPÓTESIS

Existe una zona entre Lima y Barranca, que presenta la mayor probabilidad de ocurrencia de sismos de alta magnitud.

1.4 JUSTIFICACIÓN E IMPORTANCIA

La investigación desde el punto de vista **teórico – metodológico** es importante porque permite poner en práctica la metodología del modelo semi-Márkov, en situaciones concretas del mundo real; en este caso abordar el estudio de variables referidas a los sismos en el Perú como la magnitud, profundidad y localización geográfica.

Desde el punto de vista **práctico** se justifica porque permite abordar el conocimiento de realidades específicas, que permitirá agrupar, relacionar, etc., y sacar conclusiones para la prevención de los sismos, causantes de desastres y grandes daños a la población.

En cuanto a la **contribución social**, la presente investigación se justifica, dado el interés público del tema sísmico, el mismo que puede servir para evaluar mejor el fenómeno y sus tendencias. Y tiene como principal beneficiario a la ciudadanía peruana, dado que las autoridades al contar con más información y comprender mejor este fenómeno, podrían establecer adecuadas medidas de prevención que mitiguen los efectos de los sismos.

Por otro lado, las **instituciones académicas**, serán también beneficiadas porque el presente trabajo, contribuye a la discusión sobre modelos estadísticos aplicados al tema sísmico en el Perú, motivando al desarrollo de investigaciones sobre este tema. También beneficia a quienes estén interesados en el fenómeno de los sismos en el Perú.

CAPÍTULO II: MARCO TEÓRICO

2.1 ANTECEDENTES

En el estudio de la aplicación de modelos estadísticos para identificar áreas con alta probabilidad de ocurrencia de sismos de gran magnitud en la región entre Lima y Barranca, varios autores han explorado diferentes enfoques y modelos.

En primer lugar, Flores et al. (2012) realizan un estudio de toda la costa y mar del Perú, el cual incluye a la zona entre Lima y Barranca. En este estudio los autores, aplicaron el método de predicción de terremotos a plazo intermedio, conocido como M8 (Borok, 2002), al borde occidental del Perú a fin de predecir la posible ocurrencia de sismos de gran magnitud en el futuro. Para su aplicación utilizaron entre otros los catálogos sísmicos del IGP (Instituto Geofísico del Perú), previamente evaluados para constituir una base homogénea de datos. En el análisis retroactivo confirmaron la formación de TIPs (Tiempos de incremento de probabilidad) antes de los terremotos ocurridos en el Perú en los años 1970, 1974, 1996, 2001 y 2007. En el análisis prospectivo se identificó la presencia de dos CTIPs correspondientes a áreas circulares de 300 km con centro, para el primero en el punto $-11,15^{\circ}\text{S}$ y $-78,37^{\circ}\text{O}$ ubicado al Nor-Oeste de la ciudad de Lima, y para el segundo en $-19,55^{\circ}\text{S}$ y $-69,97^{\circ}\text{O}$ ubicado en el límite Perú-Chile. Ellos concluyeron que, en ambas áreas, es posible la ocurrencia de un gran sismo dentro del periodo de tiempo 2008-2013.

A continuación, se exponen otros antecedentes relacionados con la determinación del modelo más adecuado para ajustar los datos de sismos de alta magnitud. En este contexto, se mencionan los modelos sin memoria, como el modelo Poisson, el proceso de Márkov y el proceso semi-Márkov.

Masala (2012) señala:

Que es ya conocido que el modelo de Poisson proporciona un enfoque típico para una serie aleatoria de eventos. Este modelo asume que las ocurrencias de terremotos son eventos independientes en el espacio y el tiempo. Sin embargo, esta característica "sin memoria" no es adecuada para este tipo de eventos.

También que, en un entorno del Proceso de Márkov, la función de distribución del tiempo de espera en cada estado debe ser exponencial, mientras que en un entorno del proceso semi-Márkov, estas distribuciones pueden ser de cualquier tipo. Y en este último enfoque permite tomar en consideración el supuesto más realista de la dependencia de los eventos en el tiempo y en el espacio. Por ejemplo, la probabilidad

de que ocurra un terremoto en este caso es dependiente de la magnitud del evento anterior, así como del tiempo transcurrido.

También se presenta el tema de la conveniencia de incluir una dimensión más, puesto que Masala (2012) aplica el modelo semi-Márkov solo a la magnitud de los sismos, en cambio Sadeghian (2012) utiliza una dimensión más, en este caso la magnitud y las zonas geográficas o regiones.

Sadeghian (2012), también aplica el modelo semi-Márkov a datos de sismos y señala:

Que en su trabajo tiene por objetivo de pronosticar el lugar y tiempo de terremotos, usando el modelo semi-Márkov, para ello divide la provincia de Teherán en seis regiones, y por otro lado aplica el modelo semi-Márkov a intervalos de la magnitud, agrupando los terremotos según su magnitud en tres clases. Finalmente procede a predecir la probabilidad del lugar de ocurrencia de sismos en la provincia de estudio.

Alvarez (2005) que aplica un modelo de renovación de Márkov:

El objetivo es aplicar el proceso de renovación de Márkov estacionario a tres tipos de magnitud de los terremotos para pronosticar los sismos en Turquía. Él utiliza a la distribución de Weibull para modelar los tiempos de permanencia en los estados de la magnitud. El autor realiza pronósticos de la magnitud de los sismos en el área de estudio.

Pavani (2011) señala:

Que, el objetivo de los autores es modelar grandes sismos, acorde a la geofísica inmersa. Los modelos disponibles en la literatura pueden dividirse en algunas pocas categorías. Esencialmente se pueden dividir en modelos Poisson y no Poisson. Los modelos Poisson se usan ampliamente en ingeniería de análisis de riesgo sísmico, pero pueden no ser apropiados para grandes terremotos, en el caso de falla específica, porque cualquier proceso de Poisson posee la propiedad de falta de memoria. Propone un nuevo método para evaluar el riesgo sísmico, esto es, un modelo de recurrencia no-Poisson, el cual toma en consideración la ocurrencia histórica, mejorando con algunas restricciones de orden de la física.

Las autoras proponen una mixtura de exponencial y Weibull para el modelo de la distribución de los tiempos de espera (aplican el modelo semi-Márkov). Los parámetros requeridos son estimados por el método de máxima verosimilitud. Ellas validan su

modelo, usando datos de terremotos de alta severidad ocurridos en Turquía, durante el siglo XX.

Las autoras concluyen que, con los resultados obtenidos, revelan una buena aproximación a los datos reales.

Votsi et al. (2012):

Tienen como objetivo, introducir el modelo semi-Márkov en el tiempo continuo, como un modelo candidato para la descripción de patrones de sismicidad en el dominio del tiempo, en el norte del mar Egeo de Grecia. Estimadores de los núcleos del modelo semi-Márkov, funciones de renovación de Márkov y funciones de transición son calculados a través de un método no paramétrico. Por otra parte, ellos estiman los tiempos de impacto para la ocurrencia espacial de los terremotos más fuertes y los intervalos de confianza de ciertos indicadores importantes. La clasificación de los estados del modelo se basa en la magnitud de los terremotos. Ellos calculan la tasa de ocurrencia de terremotos instantáneos entre los estados del modelo, así como la tasa de ocurrencia total de terremotos. Para aumentar la consistencia entre el modelo y el proceso de generación de terremotos, incluyen características sismo tectónicas como un componente importante en el modelo.

Como conclusión, ellos proponen una nueva clasificación de estados que combina los estados de orientación de magnitud y falla. Este modelo tiene en cuenta las características sismo tectónicas y contribuye significativamente a la evaluación del riesgo sísmico en la región en estudio.

Orfanogiannaki, Karlis, & Papadopoulos (2010) señalan que:

Los autores aplican el modelo de Poisson y el modelo oculto de Márkov para explicar la frecuencia de sismos en el área de Killini, mar Jónico de Grecia, en el periodo de 1990-2006.

Este modelo asume que la secuencia de estados no observables representa una cadena de Márkov de estados finito y que la distribución de las observaciones en cualquier tiempo sigue una distribución de Poisson, cuya tasa depende solamente del estado actual de la cadena. Consecuentemente, el modelo permite a una región tener tasa de sismicidad variable.

Los autores concluyen que, el modelo usado es mejor que el tradicional, dado que una transición de un estado a otro no solamente depende del número de eventos, sino también del estado del sistema.

Altinok & Kolcak (1999) en su investigación:

Usan la teoría semi-Márkov (método no paramétrico), sobre sismos ocurridos en el norte de la falla Anatolia entre 1902 y 1992, para pronosticar regiones y magnitud de los sismos en el tiempo. Ellos asumen que dos sismos consecutivos ocurridos en la misma zona de discontinuidad estructural no son eventos independientes, pero que dependen del intervalo de tiempo entre ellos. Además, los autores obtienen las probabilidades de ocurrencia, determinando las probabilidades de transición de región a región y transiciones de magnitud a magnitud de esos sismos.

Ellos concluyen que, la determinación de probabilidades conjuntas demuestra que la ocurrencia de sismos puede ser investigado en las tres dimensiones: espacio, tiempo y magnitud.

Risco (2018) señala que:

En este trabajo exploratorio, tiene por objetivo ajustar la magnitud de los sismos en el 2017, utilizando modelos de series de tiempo, en particular incluye a la zona entre Lima e Ica, resultando el ajuste significativo para el modelo ARIMA (1,1,0), el cual evidencia una dependencia de la magnitud de los sismos respecto al tiempo.

La discusión expuesta anteriormente evidencia que el modelo de Poisson ya no es adecuado para modelar la magnitud de los datos de sismos, ya que estos presentan dependencia temporal y espacial respecto al evento anterior. Por ello, en el presente trabajo se emplean los modelos semi-Márkov para ajustar la magnitud y ubicación espacial de los datos de los sismos ocurridos entre Lima y Barranca, incorporando además una dimensión adicional al análisis, es decir, se incluye la aplicación del modelo semi-Márkov a la profundidad de los sismos.

2.2 MARCO CONCEPTUAL

El estudio de los sismos de gran magnitud en zonas geográficas específicas de Lima y Barranca en el Perú es crucial debido a la frecuencia de estos eventos y potencialidad para generar graves impactos en la población y su infraestructura. En este contexto, la aplicación de modelos probabilísticos, como los modelos semi-Márkov se presentan como herramientas clave para la identificación y análisis de la ocurrencia de sismos de alta magnitud.

IGP (2022) define:

Sismo: Se define al proceso de generación y liberación de energía para posteriormente propagarse en forma de ondas por el interior de la tierra. Al llegar a la superficie, estas ondas son registradas por las estaciones sísmicas y el sacudimiento del suelo por ellas producidas es percibido por la población.

Réplicas: Después de la ocurrencia de un sismo de gran magnitud, es posible la ocurrencia de muchos sismos de menor magnitud entorno al hipocentro del sismo principal. A estos sismos se les conoce como réplicas. Algunas series de réplicas duran largo tiempo, incluso, pueden prolongarse por años. El área que encierra a los epicentros de las réplicas se llama área de réplicas.

Sismología: Es una rama de la geofísica que estudia los sismos y las propiedades elásticas de la tierra. Entre sus campos de investigación se considera: a) el estudio de la propagación de las ondas sísmicas en el interior de la tierra a fin de conocer su estructura interna, b) el estudio de las causas que dan origen a eventos sísmicos y c) la prevención de daños. La sismología incluye, entre otros fenómenos, el estudio de tsunamis, el comportamiento dinámico del suelo y las vibraciones asociadas a erupciones volcánicas.

Límite de placa: Dícese del lugar en donde dos o más placas están en contacto. Existen tres límites de placa: límites divergentes, límites convergentes y límites transformantes. En el caso del Perú, las Placas de Nazca y Sudamérica son del tipo convergentes. En el presente trabajo de investigación se utilizaron los datos de las variables de sismos registrados por el Instituto Geofísico del Perú.

Tiempo de realización del sismo: Es el registro de la fecha, hora y minutos de duración del sismo.

Escala de magnitud: Representa a la escala que mide el total de la energía liberada en el foco sísmico y originalmente corresponde a la escala de Richter, propuesta en el año 1935. Es una escala logarítmica, lo que hace que los niveles asignados no tengan un comportamiento lineal y permiten medir sismos muy pequeños hasta los que alcanzarían valores en magnitud del orden de 6.5 MI (llamada también escala de magnitud local, de ahí sus siglas “MI”). En la actualidad, la escala de magnitud más acertada y utilizada es la **escala de magnitud de momento (Mw)** en razón que permite medir sin restricción sismos pequeños y grandes como el ocurrido en Japón en el año 2011.

Hora local y universal (UTC): Indica la hora o el tiempo que corresponde a una determinada región en el globo terrestre de acuerdo con su ubicación y longitud geográfica con respecto al meridiano estándar de referencia; es decir, el meridiano de Greenwich o París. Según el estándar, cada 5° de longitud corresponden a una hora de tiempo. Para el caso de Perú la diferencia horaria es de menos 5 horas con respecto a la hora universal (UTC).

Escala de magnitud de momento: Puerto Rico (2021)

A diferencia de otras mediciones de magnitud, la magnitud de momento se relaciona directamente a las propiedades de la falla. En forma de ecuación se tiene:

$$M_w = \log M_0 / 1.5 - 10.73$$

M_0 es el momento sísmico en unidades dyne x cm y puede ser definido por las propiedades físicas de la falla:

$$M_0 = \hat{\mu} \frac{1}{4} DS = \hat{\mu} \frac{1}{4} DfL^2$$

$\hat{\mu}$: rigidez en la profundidad de la fuente

D: promedio del desliz en la falla

S: el área de la falla

$$S = fL^2$$

f: relación de anchura a la longitud

L: longitud de la falla

En términos simples, la magnitud de momento se relaciona directamente con el área de la falla que se desplazó durante el terremoto, cuán lejos la falla se deslizó y cuánto el material de la falla se resistió al desplazamiento. Debido a que esta magnitud se vincula directamente a las propiedades de falla es considerada más útil para los científicos. Las ventajas de magnitud de momento son las siguientes:

Directamente vinculada al proceso de falla del terremoto No se satura: para magnitudes grandes (>6.5) mb y Ms puede subestimar el tamaño del evento. Cuando este sucede los sismólogos lo llaman magnitud de saturación. Conserva la simplicidad de las otras escalas de magnitud (orden de 1 es comparable a las otras magnitudes).

Sistema de coordenadas geográficas

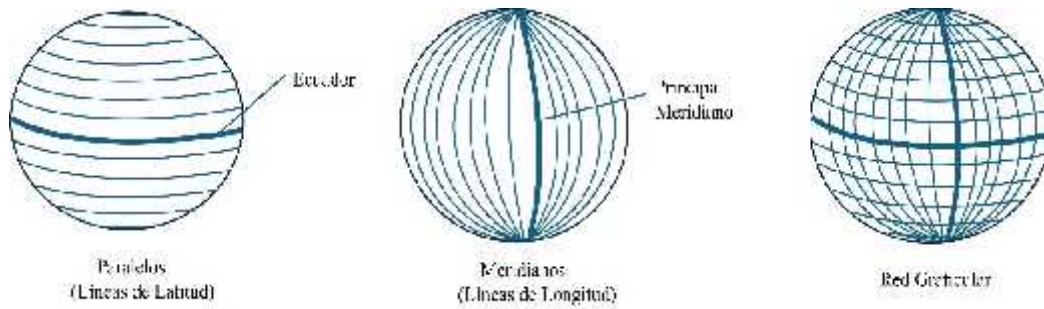
Esri (2023) establece que:

Un sistema de coordenadas geográficas (GCS) utiliza una superficie esférica tridimensional para definir ubicaciones en la Tierra. Con frecuencia, a los GCS, Geographic Coordinate System (sistema de coordenadas geográficas) se los llama incorrectamente datum, pero un datum es solo una parte de un GCS. Un GCS incluye una unidad angular de medida, un meridiano base y un datum (basado en un esferoide). Para hacer referencia a un punto se utilizan los valores de latitud y longitud. La longitud y la latitud son ángulos medidos desde el centro de la Tierra hasta un punto de superficie. Los ángulos se suelen medir en grados (o en grados centesimales).

Latitud: Los valores de latitud se miden respecto al ecuador y van desde -90° en el Polo Sur hasta $+90^\circ$ en el Polo Norte.

Longitud: Mide el ángulo a lo largo del ecuador desde cualquier punto de la Tierra. Se acepta que Greenwich en Londres es la longitud 0 en la mayoría de las sociedades modernas. Las líneas de longitud son círculos máximos que pasan por los polos y se llaman meridianos.

Figura 1: Los paralelos y los meridianos que forman una retícula



Elaboración propia

La figura 1 nos muestra la línea ecuatorial, cuya distancia a un punto en el hemisferio Sur, es la latitud y por convención es negativo. En forma similar, también se puede observar el meridiano principal, cuya distancia hacia el Oeste es la longitud y por convención también es negativo.

Para identificar las zonas de mayor riesgo sísmico entre Lima y Barranca, surge la necesidad de presentar brevemente los modelos de Renovación, Cadenas de Márkov y Semi-Márkov.

2.3 EL MODELO DE RENOVACIÓN

Aquí la teoría del modelo de renovación ha sido tomada de Ross (2007).

En un proceso de Poisson, los tiempos de Inter Llegadas se distribuyen como una exponencial y son además independientes e idénticamente distribuidos.

Una natural generalización del modelo de Poisson es considerar un proceso de conteo para el cual los tiempos de Inter Llegadas son independientes e idénticamente distribuidas, con una distribución arbitraria F . Tal proceso de conteo es llamado un proceso de renovación.

Formalmente, sea $\{X_n, n = 1, 2, \dots\}$ una secuencia de variables aleatorias independientes no negativas con distribución común F .

Para evitar trivialidades, suponemos que $P\{X_n = 0\} < 1$.

De la no-negatividad de X_n se sigue que $E[X_n]$ existe, incluso puede ser ∞ , y denotamos como

$$u = \int_0^{\infty} x dF(x)$$

Dejamos que $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, $n \geq 1$, y definimos $N(t) = \sup\{n : S_n \leq t\}$.

Se sigue que de la ley fuerte de los grandes números que $\frac{S_n}{n} \rightarrow u$, con probabilidad 1.

Definición 2.1

$\{N(t); t \geq 0\}$ es un proceso de renovación si los tiempos de Inter llegadas son independientes e idénticamente distribuidas con distribución arbitraria $F(t)$.

Tenemos la siguiente propiedad,

$$S_n \leq t \Leftrightarrow N(t) \geq n$$

$$\text{O } N(t) \geq n \Leftrightarrow S_n \leq t \tag{2.1}$$

De (1) Obtenemos

$$P\{N(t) = n\} = P\{N(t) \geq n\} - P\{N(t) \geq n + 1\}$$

$$= P\{S_n \leq t\} - P\{S_{n+1} \leq t\} \tag{2.2}$$

$$= F_n(t) - F_{n+1}(t)$$

Sea $m(t) = EN(t)$

$m(t)$ es llamada la función de renovación y la teoría de renovación está relacionada con sus propiedades.

Proposición 2.1

$$m(t) = \sum_{n=1}^{\infty} F_n(t) \tag{2.3}$$

Prueba

Sea $N(t) = \sum_{n=1}^{\infty} A_n$

donde $A_n = \begin{cases} 1 & \text{si la n-esima renovacion} \\ & \text{ha ocurrido en } [0,t] \\ 0 & \text{en otro caso} \end{cases}$

De aquí,

$$\begin{aligned}
E[N(t)] &= E\sum_{n=1}^{\infty} A_n \\
&= \sum_{n=1}^{\infty} E A_n \\
&= \sum_{n=1}^{\infty} P\{A_n = 1\} \\
&= \sum_{n=1}^{\infty} P\{S_n \leq t\} \\
&= \sum_{n=1}^{\infty} F_n(t)
\end{aligned}$$

Donde el intercambio del esperado y la suma es por la no negatividad de las A_n

Proposición 2.2. $m(t) < \infty \quad \forall t \geq 0$.

También $E(N(t))^r$, $m(t) < \infty \quad \forall t \geq 0, r \geq 0$

Proposición 2.3. Existe una correspondencia uno a uno entre la distribución de interlegadas F y la función de renovación $m(t)$.

Prueba. Tomando transformada de Laplace a ambos lados de

$$\begin{aligned}
m(t) &= \sum_{n=1}^{\infty} F_n(t), \text{ tenemos:} \\
\tilde{m}(s) &= \left\{ \sum_{n=1}^{\infty} F_n(s) \right\} \\
&= \sum_{n=1}^{\infty} [\tilde{F}(s)]^n \\
&= \frac{\tilde{F}(s)}{1 - \tilde{F}(s)}
\end{aligned} \tag{2.4}$$

O equivalente $\tilde{F}(s) = \frac{\tilde{m}(s)}{1 + \tilde{m}(s)}$ de aquí \tilde{F} es determinado por $\tilde{m}(t)$ y como la transformada de Laplace determina la distribución, se tiene que F también es determinado por $m(t)$.

Ecuación de renovación

Una ecuación integral para $m(t)$ puede ser obtenida condicionando sobre el tiempo de la primera renovación.

Haciendo esto, obtenemos,

$$m(t) = \int_0^{\infty} E [N(t) / X_1 = x] dF(x) \quad (2.5)$$

Tenemos que

$$E [N(t) / X_1 = x] = \begin{cases} 0 & x > t \\ 1 + m(t - x) & x \leq t \end{cases} \quad (2.6)$$

Poniendo (2.6) en (2.5) obtenemos,

$$m(t) = \int_0^t [1 + m(t - x)] dF(x) \quad (2.7)$$

$$m(t) = F(t) + \int_0^t m(t - x) dF(x)$$

La ecuación (7) es conocida como la ecuación de renovación y puede ser resuelta para $m(t)$

Una generalización de la ecuación de renovación es la siguiente:

$$g(t) = h(t) + \int_0^t g(t - x) dF(x) \quad (t \geq 0) \quad (2.8)$$

Donde h y F son conocidas y g es una función desconocida a ser determinada como la solución de la integral (8).

La ecuación integral (8) se dice que es una ecuación de tipo renovación y su solución está dada por a siguiente:

Proposición 2.4

Si $g(t) = h(t) + \int_0^t g(t - x) dF(x) \quad (t \geq 0)$

Entonces $g(t) = h(t) + \int_0^t h(t - x) dm(x)$

Donde:

$$m(x) = \sum_{n=1}^{\infty} F_n(x)$$

Prueba. La ecuación (8) establece que,

$$g = h + g * F$$

Y tomando transformada de Laplace produce,

$$\tilde{g}(s) = \tilde{h}(s) + \tilde{g}(s) \tilde{F}(s)$$

$$O \quad \tilde{g}(s) = \frac{\tilde{h}(s)}{1 - \tilde{F}(s)}$$

El cual es equivalente a,

$$\begin{aligned} \tilde{g}(s) &= \tilde{h}(s) \left[1 + \frac{\tilde{F}(s)}{1 - \tilde{F}(s)} \right] \\ &= \tilde{h}(s) + \tilde{h}(s) \frac{\tilde{F}(s)}{1 - \tilde{F}(s)} \end{aligned}$$

Pero sabemos $\tilde{m}(s) = \frac{\tilde{F}(s)}{1 - \tilde{F}(s)}$, obtenemos

$$\begin{aligned} \tilde{g}(s) &= \tilde{h}(s) + \tilde{h}(s)\tilde{m}(s) \\ \tilde{g}(s) &= \{h(s) + h(s)m(s)\} \\ g(s) &= h + h * m \end{aligned}$$

Puesto que la transformada de Laplace determinada la función, esto produce el resultado deseado, esto es

$$g(t) = h(t) + \int_0^t h(t-x) dm(x)$$

donde $m(x) = \sum_{n=1}^{\infty} F_n(x)$

Teoremas límites

$\frac{1}{u} = \frac{1}{\int_0^{\infty} x dF(x)}$ es frecuentemente llamada la tasa del proceso, donde $\frac{1}{\infty} = 0$.

La justificación teórica para esto es dada por el siguiente teorema.

Teorema 2.1 Con probabilidad 1,

$$\frac{N(t)}{t} \rightarrow \frac{1}{u} \text{ cuando } t \rightarrow \infty$$

Corolario 2.1.

Si $u < \infty$, entonces $E[S_{N(t)+1}] = u \cdot (m(t) + 1)$ (2.9)

Teorema 2.2 (El teorema elemental de renovación)

$$\frac{m(t)}{t} \rightarrow \frac{1}{u} \text{ cuando } t \rightarrow \infty$$

Teorema 2.3 (Teorema lavelave de renovación)

Si F no es discreta, y si $h(t)$ es directamente Riemann integrable, entonces,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_0^n h(t-x) dm(x) = \frac{1}{u} \int_0^{\infty} h(t) dt$$

Teorema 2.4 Teorema del límite central para el proceso de renovación

Sean u y \dagger^2 que se asumen finitos, representan la media y la varianza de la distribución de los tiempos de interllegadas. Entonces,

$$P \left\{ \frac{N(t) - t/u}{\dagger \sqrt{t/u^3}} < y \right\} \rightarrow \frac{1}{\sqrt{2f}} \int_{-\infty}^y e^{-x^2/2} dx \text{ cuando } t \rightarrow \infty. \quad (2.10)$$

Se puede demostrar que $\frac{Var(N(t))}{t}$ converge a $\frac{\dagger^2}{u^3}$, esto es

$$\lim_{t \rightarrow \infty} \frac{V(N(t))}{t} = \frac{\dagger^2}{u^3} \quad (2.11)$$

2.4 EL MODELO DE LAS CADENAS DE MÁRKOV

Aquí la teoría de las cadenas de Márkov ha sido tomada de Ross (2007).

Definición 2.2

Un modelo probabilístico para un experimento aleatorio es un espacio de probabilidad dada por una terna (Ω, A, P) donde Ω es el espacio muestral, A es una sigma-álgebra de subconjuntos (eventos) de Ω y P es una medida de probabilidad definida en A (para $A \in A$, $P(A)$ es la probabilidad de A).

Una medida de probabilidad es un mapeo $P: A \rightarrow [0,1]$ tal que

(i) $P(\Omega) = 1$ y

(ii) Para cualquier secuencia $A_n \in A$, $n \geq 1$, Donde las A_n son disjuntos por pares,

$$P \left(\bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} P(A_n)$$

El par (Ω, A) se llama espacio medible.

Definición

Un proceso estocástico $\{X_n, n = 0, 1, 2, \dots\}$ con un espacio de estados finito o contable, se dice que es una cadena de Márkov si para todos los estados $i_0, i_1, \dots, i_{n-1}, i_n, j$ y todo $n \geq 0$,

$$\begin{aligned} P\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i\} \\ = P\{X_{n+1} = j | X_n = i\} \end{aligned} \quad (2.12)$$

Si $P\{X_{n+1} = j | X_n = i\}$ es independiente de n , entonces se dice que la cadena de Márkov posee probabilidades de transición estacionarias. Solo consideramos cadenas de Márkov con esta propiedad y denotamos por,

$$P_{ij} = P\{X_{n+1} = j | X_n = i\}$$

El valor P_{ij} representa la probabilidad de que el proceso, cuando esté en el estado i , haga una transición al estado j . Dado que las probabilidades no son negativas y dado que el proceso debe hacer una transición a algún estado, tenemos que

$$P_{ij} \geq 0, \quad i, j \geq 0; \quad \sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, \dots$$

Sea P la matriz de probabilidades de transición de un paso P_{ij} , de modo que P se puede representar por

$$P = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \dots \\ P_{10} & P_{11} & P_{12} & \dots \\ \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

2.4.1 ECUACIONES DE CHAPMAN KOLMOGOROV

Ya hemos definido las probabilidades de transición de un paso P_{ij} ; ahora definimos las probabilidades de transición de n pasos P_{ij}^n como la probabilidad de que un proceso en el estado i esté en el estado j después de n transiciones adicionales. Es decir,

$$P_{ij}^n = P\{X_{n+m} = j | X_m = i\}, \quad n \geq 0, i, j \geq 0$$

Por supuesto $P_{ij}^1 = P_{ij}$, las ecuaciones de Chapman-Kolmogorov proporcionan un método para calcular estas probabilidades de transición de n pasos.

Estas ecuaciones son:

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad \text{para todo } n, m \geq 0 \text{ y toda } i, j \quad (2.13)$$

La suma de todos los estados intermedios k produce la probabilidad de que el proceso esté en el estado j después de $n + m$ transiciones. Formalmente, tenemos:

$$\begin{aligned} P_{ij}^{n+m} &= P\{X_{n+m} = j | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P\{X_{n+m} = j, X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P\{X_{n+m} = j | X_n = k, X_0 = i\} P\{X_n = k | X_0 = i\} \\ &= \sum_{k=0}^{\infty} P_{kj}^m P_{ik}^n \end{aligned}$$

Si dejamos que $\mathbf{P}^{(n)}$ denote la matriz de probabilidades de transición de n pasos P_{ij}^n , entonces, la ecuación (2.13) afirma que:

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)} \cdot \mathbf{P}^{(m)}$$

donde el punto representa la multiplicación de matrices. Por lo tanto, en particular,

$$\mathbf{P}^{(2)} = \mathbf{P}^{(1+1)} = \mathbf{P} \cdot \mathbf{P} = \mathbf{P}^2$$

y por inducción

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{n-1} \cdot \mathbf{P} = \mathbf{P}^n$$

Es decir, la matriz de transición de n pasos se puede obtener multiplicando la matriz \mathbf{P} por sí misma n veces.

Hasta ahora, todas las probabilidades que hemos considerado son probabilidades condicionales. Por ejemplo, P_{ij}^n es la probabilidad de que el estado en el tiempo n sea j dado que el estado inicial en el tiempo 0 sea i . Si se desea la distribución incondicional del estado en el tiempo n , es necesario especificar la distribución de probabilidad del estado inicial. Denotemos esto por

$$r_i \equiv P\{X_0 = i\}, \quad i \geq 0 \quad \left(\sum_{i=0}^{\infty} r_i = 1 \right)$$

Todas las probabilidades incondicionales pueden calcularse condicionando el estado inicial. Es decir,

$$\begin{aligned} P\{X_n = j\} &= \sum_{i=0}^{\infty} P\{X_n = j | X_0 = i\} P\{X_0 = i\} \\ &= \sum_{i=0}^{\infty} P_{ij}^n r_i \end{aligned}$$

2.4.2 CLASIFICACIÓN DE ESTADOS

Se dice que el estado j es accesible desde el estado i si $P_{ij}^n > 0$ para algún $n \geq 0$.

Nótese que esto implica que el estado j es accesible desde el estado i sí y solo si, comenzando en i , es posible que el proceso alguna vez al entrar en el estado j . Esto es cierto ya que si j no es accesible desde i , entonces:

$$\begin{aligned} P\{\text{alguna vez entrar a } j | \text{empezando en } i\} &= P\left\{ \bigcup_{n=0}^{\infty} \{X_n = j\} | X_0 = i \right\} \\ &\geq \sum_{n=0}^{\infty} P\{X_n = j | X_0 = i\} \\ &= \sum_{n=0}^{\infty} P_{ij}^n \\ &= 0 \end{aligned}$$

Se dice que dos estados i y j , que son accesibles entre sí, se comunican y escribimos

$$i \leftrightarrow j.$$

Tenga en cuenta que cualquier estado se comunica consigo mismo ya que, por definición,

$$P_{ii}^0 = P\{X_0 = i | X_0 = i\} = 1$$

La relación de comunicación satisface las siguientes tres propiedades:

- (i) El estado i se comunica con el estado i , para todo $i \geq 0$.
- (ii) Si el estado i se comunica con el estado j , entonces el estado j se comunica con el estado i .
- (iii) Si el estado i se comunica con el estado j , y el estado j se comunica con el estado k , entonces el estado i se comunica con el estado k .

Las propiedades (i) y (ii) se derivan inmediatamente de la definición de comunicación.

Para probar (iii) suponga que i se comunica con j y j se comunica con k ; por lo tanto,

existen enteros n y m tales que $P_{ij}^n > 0$, $P_{jk}^m > 0$. Ahora, por las ecuaciones de

Chapman-Kolmogorov, tenemos que:

$$P_{ik}^{n+m} = \sum_{r=0}^{\infty} P_{ir}^n P_{rk}^m \geq P_{ij}^n P_{jk}^m > 0$$

Por tanto, el estado k es accesible desde el estado i . De manera similar, podemos mostrar que el estado i es accesible desde el estado k . Por tanto, los estados i y k se comunican.

Se dice que dos estados que se comunican están en la misma clase. Una consecuencia fácil de, (i), (ii) y (iii) es que dos clases de estados cualesquiera sean idénticas o disjuntas. En otras palabras, el concepto de comunicación divide el espacio de estados en varias clases separadas de una cadena de Márkov

Se dice que la cadena es irreductible si solo hay una clase, es decir, si todos los estados se comunican entre sí.

Para cualquier estado i , dejamos que f_i denote la probabilidad de que, comenzando en el estado i , el proceso alguna vez vuelva a entrar en el estado i , se dice que el estado i es recurrente si $f_i = 1$ y transitorio si $f_i < 1$.

Si el estado i es recurrente, entonces, comenzando en el estado i , el proceso volverá a entrar en el estado i una y otra vez, así, infinitas veces.

Por otro lado, si el estado i es transitorio, cada vez que el proceso entra en el estado i , habrá una probabilidad positiva, es decir, $1 - f_i$ que nunca volverá a entrar en este estado. Por lo tanto, comenzando en el estado i , la probabilidad de que el proceso esté en el estado i durante exactamente n períodos de tiempo es igual a $f_i^{n-1}(1 - f_i)$, $n \geq 1$. En otras palabras, si el estado i es transitorio, entonces, comenzando en el estado i , el número de períodos de tiempo que el proceso estará en el estado i tiene una distribución geométrica con media finita $1/(1 - f_i)$.

Proposición 2.5

El estado i es

$$\begin{aligned} \text{recurrente si } \sum_{n=1}^{\infty} P_{ii}^n &= \infty, \\ \text{transitorio si } \sum_{n=1}^{\infty} P_{ii}^n &< \infty \end{aligned}$$

Corolario 2.2

Si el estado i es recurrente y el estado j se comunica con el estado i , entonces el estado j es recurrente.

Prueba. Para probar esto, primero notamos que, dado que el estado i se comunica con el estado j , existen números enteros k y m tales que $P_{ij}^k > 0$, $P_{ji}^m > 0$. Ahora, para cualquier número entero n

$$P_{jj}^{m+n+k} \geq P_{ji}^m P_{ii}^n P_{ij}^k$$

De lo anterior, se obtiene, sumando sobre n , que

$$\sum_{n=1}^{\infty} P_{jj}^{m+n+k} \geq P_{ji}^m P_{ij}^k \sum_{n=1}^{\infty} P_{ii}^n = \infty$$

ya que $P_{ji}^m P_{ij}^k > 0$ y $\sum_{n=1}^{\infty} P_{ii}^n$ es infinito ya que el estado i es recurrente.

Por lo tanto, por la proposición anterior se sigue que el estado j también es recurrente.

2.4.3 PROBABILIDADES LÍMITES

Hay dos propiedades adicionales de los estados de una cadena de Márkov que debemos considerar. Se dice que el estado i tiene un período d si $P_{ii}^n = 0$ siempre que n no sea divisible por d , y d es el número entero más grande con esta propiedad. Se dice que un estado con período 1 es aperiódico.

Si el estado i es recurrente, entonces se dice que es positivo recurrente sí, comenzando en i , el tiempo esperado hasta que el proceso vuelve al estado i es finito. Se puede demostrar que la recurrencia positiva es una propiedad de clase. Si bien existen estados recurrentes que no son recurrentes positivos, se puede demostrar que en una cadena de Márkov de estados finitos todos los estados recurrentes son recurrentes positivos. Los estados aperiódicos recurrentes positivos se denominan ergódicos.

Observaciones

- a. Dado que $f_j = \lim_{n \rightarrow \infty} P_{ij}^n$ existe y es independiente del estado inicial i , no es

difícil ver (heurísticamente) que las f 's debe satisfacer la ecuación anterior. Pues, se deriva una expresión para $P\{X_{n+1} = j\}$ condicionando el estado en el tiempo n ; es decir,

$$\begin{aligned} P\{X_{n+1} = j\} &= \sum_{i=0}^{\infty} P\{X_{n+1} = j | X_n = i\} P\{X_n = i\} \\ &= \sum_{i=0}^{\infty} P_{ij} P\{X_n = i\} \end{aligned}$$

Dejando $n \rightarrow \infty$, y asumiendo que podemos poner el límite dentro de la suma, conduce a

$$f_j = \sum_{i=0}^{\infty} P_{ij} f_i$$

- b. Se puede demostrar que f_j , la probabilidad límite de que el proceso esté en el estado j en el tiempo n , también es igual a la proporción de tiempo a largo plazo en que el proceso estará en el estado j .
- c. En el caso periódico recurrente positivo irreducible, todavía tenemos que los $f_j, j \geq 0$, son la única solución no negativa de

$$f_j = \sum_i f_i P_{ij}, \quad \sum_j f_j = 1 \quad (2.14)$$

Pero ahora f_j debe interpretarse como la proporción de tiempo a largo plazo que la cadena de Márkov está en el estado j .

Esta parte ha sido tomada de Taha (2012).

El tiempo medio del primer retorno o tiempo medio de recurrencia, y se calcula en una cadena de Márkov de n estados como

$$\mu_{jj} = \frac{1}{f_j}, j = 1, 2, \dots, n$$

El tiempo medio del primer paso $\tilde{\mu}_{ij}$, definido como el número esperado de transiciones para llegar por primera vez al estado j desde el estado i .

Si suponemos que la matriz de transición $\mathbf{P} = \left\| P_{ij} \right\|$ tiene m estados.

- Si $f_{ij} < 1$ no es seguro que el sistema pase alguna vez del estado i al estado j y $\tilde{\mu}_{ij} = \infty$.
- Si $f_{ij} = 1$ la cadena de Márkov es ergódica y el tiempo medio del primer paso del estado i al estado j se calcula como:

$$\tilde{\mu}_{ij} = \sum_{n=1}^{\infty} n f_{ij}^n$$

Una forma más simple de determinar el tiempo medio del primer paso de todos los estados en una matriz de n transiciones, \mathbf{P} , es utilizar la siguiente fórmula basada en una matriz:

$$\left\| \tilde{\mu}_{ij} \right\| = (\mathbf{I} - \mathbf{N}_j)^{-1} \mathbf{1}, j \neq i \quad (2.15)$$

Donde: \mathbf{I} = Matriz de identidad de orden $(m - 1)$.

\mathbf{N}_j = Matriz de transición \mathbf{P} sin su fila j -ésima fila y su j -ésima columna del estado destino j .

$\mathbf{1}$ = Vector columna con $(m - 1)$ elementos con todos los elementos iguales a 1.

La operación matricial $(\mathbf{I} - \mathbf{N}_j)^{-1} \mathbf{1}$ en esencia, es la suma de las columnas de $(\mathbf{I} - \mathbf{N}_j)^{-1}$

2.5 EL MODELO SEMI-MÁRKOV

Masala (2012), en relación al modelo semi-Márkov, postula que:

Sea Ω un espacio muestral, F una sigma álgebra de conjuntos de Ω y P una medida de probabilidad sobre ella.

Denote por $E = \{I, II, III, \dots\}$ el espacio de estados y $\{\Omega, A, P\}$ representa el espacio de probabilidad. Se denota por $m = \text{card } E$ y se introduce las siguientes variables aleatorias:

$$J_n : \Omega \rightarrow E \quad \text{y} \quad S_n : \Omega \rightarrow [0, +\infty),$$

donde J_n representa el estado de la n -ésima transición y S_n representa el tiempo

cronológico de la n -ésima transición. $N(t)$ es el proceso de conteo $\{N(t), t \geq 0\}$ el

cual está asociada al proceso puntual $(S_n)_{n \in \mathbb{N}}$, definido como:

$N(t) = \sup \{n, \in \mathbb{N} : S_n \leq t\}$ para $t \geq 0$, el cual representa el número de transiciones en el horizonte $(0, t)$.

De la definición de S_n , se introduce “la duración del proceso” $(X_n)_{n \in \mathbb{N}}$ como la siguiente familia de variables aleatorias:

$$X_0 = 0$$

$$X_{n+1} = S_{n+1} - S_n$$

Aquí X_{n+1} representa la duración del tiempo gastado en el estado J_n .

Definición: Dadas las variables aleatorias

$$J_n : \Omega \rightarrow E \quad S_n : \Omega \rightarrow \mathbb{R}$$

el proceso (J_n, S_n) es un proceso semi-Márkov de tiempo discreto si el núcleo Q asociado con el proceso está dado para el caso homogéneo:

$$Q = [Q_{ij}(t)] = [P(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = i)] \quad \text{con } i, j \in E; t \in \mathbb{R}^+$$

Se introduce la matriz de transición P de este proceso:

$$P = [p_{ij}] = \left[\lim_{t \rightarrow \infty} Q_{ij}(t) \right] \quad \text{con } i, j \in E; t \in \mathbb{R}^+$$

Se define la probabilidad de que el proceso abandone el estado i dentro de un tiempo t como:

$$H_i(t) = [P(S_{n+1} - S_n \leq t | J_n = i)]$$

También esta relación es válida:

$$H_i(t) = \sum_{j=1}^m Q_{ij}(t) \quad (2.16)$$

Se introduce las siguientes probabilidades que son específicas para el caso discreto homogéneo:

$$B = [B_{ij}(t)] = [P(J_{n+1} = j, S_{n+1} - S_n = t | J_n = i)]$$

De la definición de Q, se obtiene:

$$b_{ij}(t) = \begin{cases} Q_{ij}(0) = 0 & \text{si } t = 0 \\ Q_{ij}(t) - Q_{ij}(t-1) & \text{si } t > 0 \end{cases}$$

La función de distribución condicional de los tiempos de espera, dado el estado actual y el estado siguiente, viene dada por:

$$F = [F_{ij}(t)] = [P(S_{n+1} - S_n \leq t | J_n = i, J_{n+1} = j)] \quad (2.17)$$

Las siguientes relaciones se pueden deducir fácilmente:

$$F_{ij}(t) = \begin{cases} Q_{ij}(t)/p_{ij} & \text{si } p_{ij} \neq 0 \\ 1 & \text{si } p_{ij} = 0 \end{cases} \quad (2.18)$$

De aquí resulta que $Q_{ij}(t) = p_{ij} Q_{ij}(t)$

2.6 LA DISTRIBUCIÓN DE WEIBULL

La distribución de Weibull, según Osaki (1992), es ideal para modelar fenómenos en los que la mayor concentración de eventos ocurre en los primeros intervalos de tiempo, con una disminución gradual en los intervalos posteriores. Esto da lugar a una cola asimétrica y alargada hacia la derecha, reflejando una mayor probabilidad de eventos tempranos y una menor probabilidad de eventos a medida que el tiempo avanza.

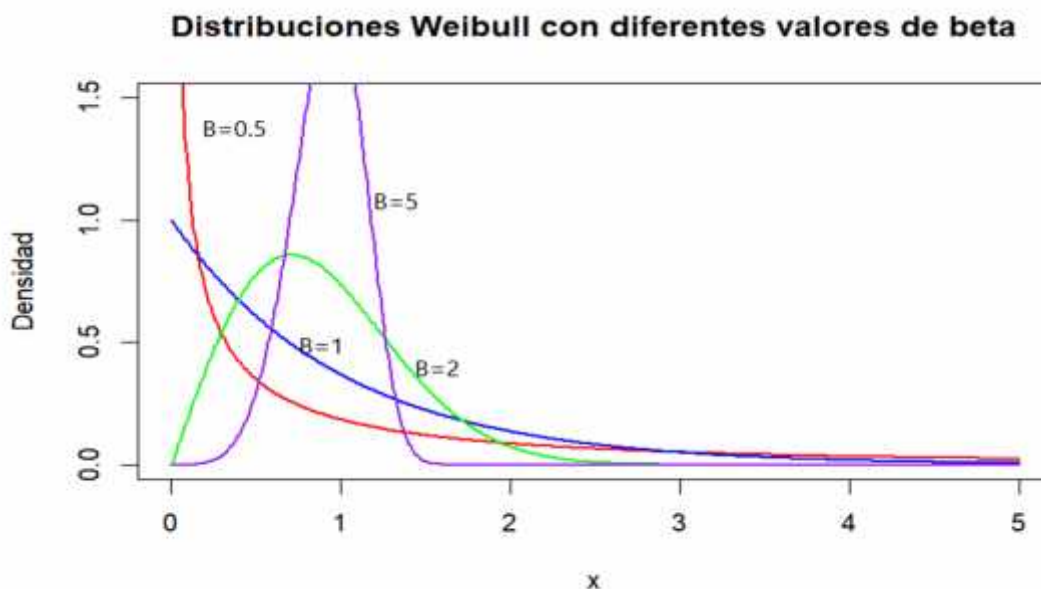
La densidad de la distribución Weibull es dado por

$$f_X(x) = r s (r x)^{s-1} e^{-(r x)^s} \quad (x \geq 0),$$

Y es denotado por $X \sim WEI(r, s)$, donde r ($r > 0$) y s ($s > 0$) son parámetros. La correspondiente distribución es dada por

$$F_X(x) = 1 - e^{-(r x)^s}.$$

Figura 2. Las densidades de la distribución $X \sim WEI(r, s)$, con $r = 1$ y $s = (0.5, 1, 2, 5)$



La media y la varianza pueden ser expresados en función de la función Gamma.

$$E[X] = \int_0^{\infty} x r s (r x)^{s-1} e^{-(r x)^s} dx = \frac{1}{r} \Gamma\left(1 + \frac{1}{s}\right),$$

$$Var(X) = E[X^2] - E[X]^2 = \left(\frac{1}{r}\right)^2 \left\{ \Gamma\left(1 + \frac{2}{s}\right) - \Gamma\left(1 + \frac{1}{s}\right)^2 \right\}$$

En el presente trabajo se ha seleccionado la distribución de Weibull porque es la que mejor se ajusta a los datos de los tiempos de espera en los estados de las zonas geográficas, en los estados de las magnitudes y en los estados de las profundidades de los sismos en estudio.

CAPÍTULO III: METODOLOGÍA

3.1 TIPO DE INVESTIGACIÓN

El presente implica un estudio correlacional de corte transversal, de enfoque cuantitativo y explicativo.

Es correlacional, porque se investiga la relación entre variables, como la magnitud y el tiempo, lo que implica una dependencia entre ellas.

Es de corte transversal, porque se está utilizando datos de sismos en un periodo determinado, lo cual caracteriza un estudio de este tipo, ya que no sigue los eventos a lo largo del tiempo.

Es de enfoque cuantitativo, porque se utiliza variables cuantitativas (magnitudes de los sismos, tiempos, profundidad y localizaciones) y técnicas estadísticas como los modelos de Márkov, lo que define claramente un enfoque cuantitativo.

Y es explicativo, porque se busca explicar por qué se dan ciertos patrones o relaciones entre los eventos sísmicos, más allá de solo describirlos.

3.2. MÉTODO

En esta investigación se ha empleado el método inductivo, que parte de la observación de información específica para llegar a conclusiones generales. Este enfoque implica la observación de la ocurrencia de sismos, el análisis de los registros proporcionados por el Instituto Geofísico del Perú (IGP), la validación de los datos mediante modelos probabilísticos y la formulación de conclusiones generales. La metodología se basa en principios probabilísticos, por lo que los resultados obtenidos pueden ser precisos o estar sujetos a error.

3.3 DISEÑO DE LA INVESTIGACIÓN

3.3.1 Población

La población está conformada por todas las posibles realizaciones de sismos que podrían haber ocurrido entre Lima y Barranca de marzo de 1961 a mayo del 2022 registrados por el Instituto Geofísico del Perú.

3.3.2 Muestra

Dado que en este caso se trata de un proceso estocástico, estamos limitados a trabajar con una realización del proceso, no corresponden los cálculos para la obtención de una muestra. Y dado que se está estudiando solo a los sismos de alta magnitud, la realización del proceso estará conformada por todos los sismos con una magnitud mayor o igual a 5

Mw, que ocurren en el área de Lima y Barranca, desde marzo de 1961 a mayo del 2022 registrados por el Instituto Geofísico del Perú.

Se ha realizado la delimitación del estudio a los sismos que ocurren entre Lima y Barranca, por ser esta el área de estudio la que incluye a Lima y que tengan una magnitud mayor o igual a 5 Mw porque se está estudiando solo a los sismos de alta magnitud.

3.4 OPERACIONALIZACIÓN DE LAS VARIABLES

Las variables consideradas en este estudio incluyen el tiempo, que se registra mediante la fecha, la hora, los minutos y los segundos de ocurrencia del sismo. También se toma en cuenta la localización, la cual se determina a partir de las coordenadas de latitud y longitud del sismo. Además, se incluyen la magnitud y la profundidad del evento sísmico.

Tabla 1: Operacionalización de las variables que participan en el estudio

Variables	Dimensión	Indicador	Escala
Tiempo	Tiempo	Hora, Min, Seg.	Numérica
Localización	Espacio	Latitud, Longitud	Numérica
Magnitud	Onda sísmica	Valor numérico 0-10	Mw, Escala Richter
Profundidad	Profundidad	Valor numérico Km	Numérica

Elaboración propia

3.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Los datos utilizados en esta investigación provienen del Instituto Geofísico del Perú (IGP), que se encarga de la recolección de información sísmica mediante un riguroso proceso de cálculo. Este proceso incluye el registro detallado del tiempo (fecha, hora, minutos y segundos), la localización a través de coordenadas de latitud y longitud, así como la magnitud y profundidad de los sismos.

Los datos proporcionados por el IGP son altamente precisos y su validez para investigaciones científicas está ampliamente respaldada. En este estudio, los datos fueron obtenidos de la página web del IGP, la cual ofrece acceso público a esta información. Por lo tanto, no ha sido necesario desarrollar un instrumento propio para la recolección de los datos.

3.6 PROCEDIMIENTO UTILIZADO

Para ejecutar el presente trabajo y aplicar el modelo semi-Márkov se precedió de la siguiente manera.

- A. En el área geográfica entre Lima y Barranca, con el uso del software Google Earth Pro, se establecieron 4 zonas geográficas, (ver anexo 5) que corresponden a 4 estados, que permitieron evaluar el tiempo de los sismos en dichas zonas. Luego se estimó la matriz de transición P para los estados de las zonas geográficas.

Se halló la distribución de los tiempos de espera de cada transición.

Se extrajeron de la base de datos, los tiempos de espera empíricos; en seguida, se estimaron los parámetros de las distribuciones Weibull de cada transición.

Luego, se estimaron las matrices relacionadas con las probabilidades de intervalo.

- B. En la segunda parte se consideró la magnitud de los sismos, para ello se procedió a clasificar la magnitud de sismos en tres estados, permitiendo 9 posibles transiciones. Luego se estimó la matriz de transición P.

En el siguiente paso, se halló la distribución de los tiempos de espera de cada transición. Se extrajo de la base de datos, los tiempos de espera empíricos y se seleccionó a la distribución de Weibull para estimar los parámetros respectivos.

En similar forma, al caso anterior, se estimaron las matrices relacionadas con las probabilidades de intervalo, de las magnitudes.

- C. En forma similar a los casos anteriores, se ha aplicado el modelo semi-Márkov a los intervalos (estados) de la profundidad de los sismos.

3.7 TÉCNICAS DE PROCESAMIENTO DE DATOS

Para el procesamiento de los datos, se utilizaron las técnicas estadísticas descriptivas, así como las técnicas que corresponden al modelo semi-Márkov.

Para el procesamiento de los datos se utilizó el software R versión 4.3.0, y SPSS v-21 y Excel.

CAPÍTULO IV: RESULTADOS Y DISCUSIÓN

4.1 PRESENTACIÓN DE RESULTADOS

En la primera parte se presentan los resultados de la aplicación del modelo semi -Márkov en las zonas geográficas de los sismos. Para ello se han usado los datos que publica el IGP, en su página web, los cuales corresponden a una base de datos de los sismos desde el 17 de marzo de 1961 hasta el 12 de Mayo de 2022, para magnitudes mayores o iguales a 5.0 Mw. Ver Anexo 1.

Figura 3: Los sismos de la zona entre Lima y Barranca con datos del IGP.



En la figura 3, se presenta con puntos los sismos ocurridos entre Lima y Barranca desde 1961 y 2022 en las zonas que fueron delimitadas aproximadamente, con la ayuda del software GoogleEarth Pro.

4.2 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LAS ZONAS GEOGRÁFICAS DE LOS SISMOS ENTRE LIMA Y BARRANCA

Para ajustar el modelo semi-Márkov a los datos de los sismos de alta magnitud a las zonas de estudio, se sigue la siguiente metodología:

El área de estudio, la que se encuentra entre Lima y Barranca, se ha dividido en 4 zonas geográficas, (ver anexo 5, las latitudes y longitudes de los límites de las zonas geográficas) las cuales en el modelo semi-Márkov constituyen 4 estados, permitiendo 16 posibles transiciones, como se verá más adelante.

Tabla 2: Latitud y longitud de las zonas entre Lima y Barranca

Zona geográfica	Variable	N	Mínimo	Máximo
Huacho - Barranca	Latitud	80	-11.502	-10.78
	Longitud	80	-79.4115	-77.0409
Lima - Huaral	Latitud	54	-12.338	-11.5097
	Longitud	54	-78.56	-76.76
Chilca - Asia	Latitud	98	-13.03	-12.36
	Longitud	98	-78.194	-76.34
Sierra de Lima	Latitud	35	-13.01	-10.85
	Longitud	35	-77.18	-75.95

Elaboración propia

En la tabla 2, se muestran los valores mínimo y máximo para las variables latitud y longitud, las cuales fueron obtenidas con la ayuda del software GoogleEarth Pro.

Se ha delimitado las 4 zonas, el área geográfica entre Huacho y Barranca, el área entre Lima y Huaral, entre Chilca y Asia y la última área geográfica la Sierra de Lima.

Figura 4: Sismos en la zona entre Huacho y Barranca



En la figura 4, se observan, en puntos graficados, los sismos ocurridos en esta zona entre Huacho y Barranca.

Figura 5: Sismos en la zona entre Lima y Huaral



Elaboración propia

En la figura 5, se presentan, en puntos graficados, los sismos ocurridos en esta zona de Lima y Huaral.

Figura 6: Sismos en la zona entre Chilca y Asia



Elaboración propia

En la figura 6, se representan, en puntos graficados, los sismos ocurridos en esta zona de Chilca y Asia.

Figura 7: Sismos en la zona de la Sierra de Lima.



Elaboración propia

En la figura 7, se presentan, en puntos graficados, los sismos ocurridos en esta zona de la sierra de Lima.

Nota. Con motivos de simplificación, a los estados de las zonas geográficas en estudio, en adelante las llamaremos: Huacho, Lima, Chilca y Sierra Lima.

Tabla 3: Estadísticos descriptivos comparativos de las zonas geográficas

Zonas geo	Variables	N	Mínimo	Máximo	Media	Desv. típ.	Asimet	Curtosis
Huacho	Magnitud	80	5	8	5.35	0.43	3.6	18.85
	Profundidad	80	9	113	44.45	19.02	1.17	1.8
	Tiempo inter	80	0.06	595.45	78.45	91.57	2.72	11.96
Lima	Magnitud	54	5	7.7	5.34	0.42	3.5	17.78
	Profundidad	54	10	123	53.22	23.62	0.75	0.81
	Tiempo inter	54	0.4	935.92	88.16	144.03	4.36	23.22
Chilca	Magnitud	98	5	7.2	5.32	0.33	2.26	9.31
	Profundidad	98	11	97	41.46	16.87	0.73	0.43
	Tiempo inter	97	0.02	424.49	54.18	73.38	2.95	10.97
Sierra	Magnitud	35	5	6.6	5.27	0.39	2.38	5.38
Lima	Profundidad	35	11	142	91.37	35.08	-1.05	0.28
	Tiempo inter	35	0.92	431.58	118.82	106.51	1.13	0.84

Elaboración propia

Como se puede observar en la tabla 3, las medias de las magnitudes son parecidas. Sin embargo, se puede observar que la zona de Chilca tiene en la variable profundidad del sismo, la que presenta, respecto a las otras zonas, el menor valor medio (41km) y también el menor valor medio del tiempo de interllegada (54 días). Este comportamiento significa un mayor peligro para Lima por su cercanía y poca profundidad, si ocurre un sismo en esta zona.

En el caso de los tiempos de interllegada, también la zona Sierra de Lima se diferencia de las zonas de la costa con promedio 118 días y 106 días de desviación estándar, lo cual significa poca frecuencia de sismos en esta zona.

Tabla 4: Matriz de frecuencias de las zonas geográficas

Estados	Huacho	Lima	Chilca	Sierra Lima	Total
Huacho	29	18	27	6	80
Lima	14	10	22	8	54
Chilca	27	20	34	16	97
Sierra Lima	9	6	15	5	35
Total	79	54	98	35	266

Elaboración propia

En la tabla 4, se presenta el número de sismos de alta magnitud ocurridos en las respectivas zonas geográficas. Por ejemplo, el valor 22 significa que han ocurrido en Chilca, dado que el sismo anterior ocurrió en Lima.

Para hallar las probabilidades de transición se usa la siguiente relación.

$$\hat{p}_{ij} = \frac{a_{ij}}{\sum_{k=1}^m a_{ik}}$$

Masala (2012) señala que estas estimaciones gozan de las propiedades de los estimadores de máxima verosimilitud. Luego se estimó la matriz de transición P. Esta matriz es constante, pues, en este estudio se considera el caso homogéneo.

Tabla 5: Matriz de transición para las zonas entre Lima y Barranca

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.3625	0.2250	0.3375	0.0750
Lima	0.2593	0.1852	0.4074	0.1481
Chilca	0.2784	0.2062	0.3505	0.1649
Sierra Lima	0.2571	0.1714	0.4286	0.1429

Elaboración propia

En la matriz de transición presentada en la tabla 5, se observa al estado de Huacho con 0.3625 y al estado Chilca con 0.3505, las probabilidades mayores para repetir los sismos en sus estados respectivos.

En el siguiente paso, se halla la distribución de los tiempos de espera de cada transición. Se extrae de la base de datos, los tiempos de espera empíricos y se utiliza la distribución que mejor se ajuste a los datos de los sismos. Se ha utilizado la herramienta `riskDistribution` del software estadístico R. Se realiza el ajuste de varias distribuciones clásicas que se muestran en la tabla 6.

Tabla 6: Ajuste de los tiempos de permanencia para la transición (1,1) a las distribuciones que se presentan.

Distrib.	logL	AIC	BIC	Chisq	Chi(p)	AD(v)	H(AD)	KS(v)	H(KS)
Normal	-165.99	335.97	338.71	24.05	0	1.67	rechaz	0.18	Nulo
Cauchy	-167.92	339.84	342.57	14.01	0.01	2.4	no rechaz	0.27	Nulo
Logisti	-166.03	336.06	338.79	22.19	0	1.45	rechaz	0.2	Nulo
Expon	-151.67	305.34	306.7	40.89	0	3.4	rechaz	0.19	Nulo
Chi-sq	-807.51	1617	1618.4	Inf	0	Inf	Nulo	0.56	Nulo
Unifor	NULL	NULL	NULL	23.91	0	Inf	Nulo	0.23	Nulo
Gamma	-144.75	293.5	296.24	2.86	0.58	0.47	NA	0.11	Nulo
Logn	-150.7	305.4	308.13	7.16	0.13	1.42	rechaz	0.19	Nulo
Weibull	-146.05	296.09	298.83	4.09	0.39	0.7	no rechaz	0.12	Nulo
F	-161.37	326.74	329.48	30.27	0	4.29	Nulo	0.32	Nulo
Student	-181.97	365.94	367.3	87.52	0	18.16	Nulo	0.55	Nulo
Gomp.	-151.67	307.34	310.07	40.89	0	3.4	Nulo	0.19	Nulo

Elaboración propia.

A continuación, se explica la tabla 6:

- Logaritmo de la verosimilitud (LogL)

Distribución Weibull: - 146.05

El LogL mide la probabilidad de que el modelo ajustado haya generado los datos observados. Un valor más alto (menos negativo) indica un mejor ajuste a los datos.

Interpretación: El valor de -146.05 es el segundo más alto (la distribución Gamma tiene -144.75) que el de las demás distribuciones, por tanto, la distribución Gamma y Weibull son candidatas al mejor ajuste a los datos, en comparación con las demás distribuciones.

- Criterio de Información de Akaike (AIC)

Distribución Weibull: 296.09

El AIC mide el ajuste del modelo penalizando la complejidad. Un AIC más bajo indica un mejor ajuste. En nuestro caso, solo el AIC de la distribución Gamma con 293.5 supera este valor, siguiendo ambas distribuciones como candidatas a ser escogidas para el ajuste de los datos.

- Criterio de Información Bayesiana (BIC)

El BIC penaliza modelos complejos más severamente que el AIC. Un BIC más bajo, indica un mejor ajuste. En nuestro caso, nuevamente el BIC de la distribución Gamma 296.24 supera ligeramente al BIC de la distribución Weibull, 298.83, aquí nuevamente indica que la Gamma se ajusta mejor a los datos.

- La prueba de Anderson – Darling (AD)

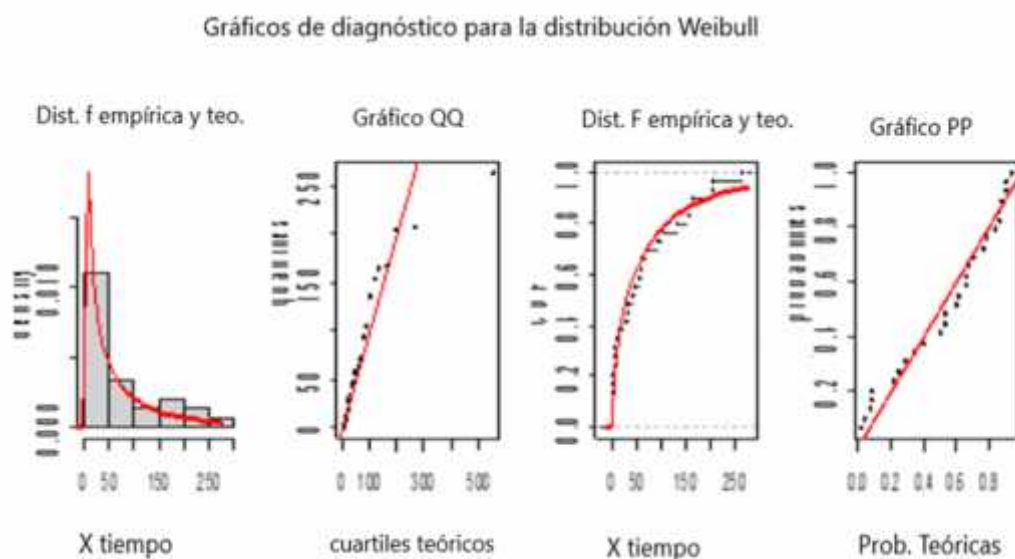
La prueba de Anderson – Darling es más sensible en las colas de la distribución. Un valor más bajo indica un mejor ajuste. Aquí la Gamma tiene un valor más bajo (0.47) que la Weibull (0.7), lo que sugiere que la distribución Gamma se ajusta mejor en términos de la prueba AD.

- La prueba de Kolmogorov – Smirnov (KS)

La prueba KS evalúa la mayor diferencia teórica y la empírica. Ambos modelos tienen valores muy cercanos (0.11 para la Gamma y 0.12 para Weibull) y, en ambos casos, no se rechaza la hipótesis de ajuste. Sin embargo, la Gamma es ligeramente menor, lo que indica un mejor ajuste.

Si comparamos con el diagnóstico para la distribución Gamma, el cual se presenta a continuación, se observa una gran similitud con el de la distribución de Weibull.

Figura 8: Gráficos de diagnóstico para la distribución de Weibull de los tiempos de espera para la transición t11 (Zonas geográficas)

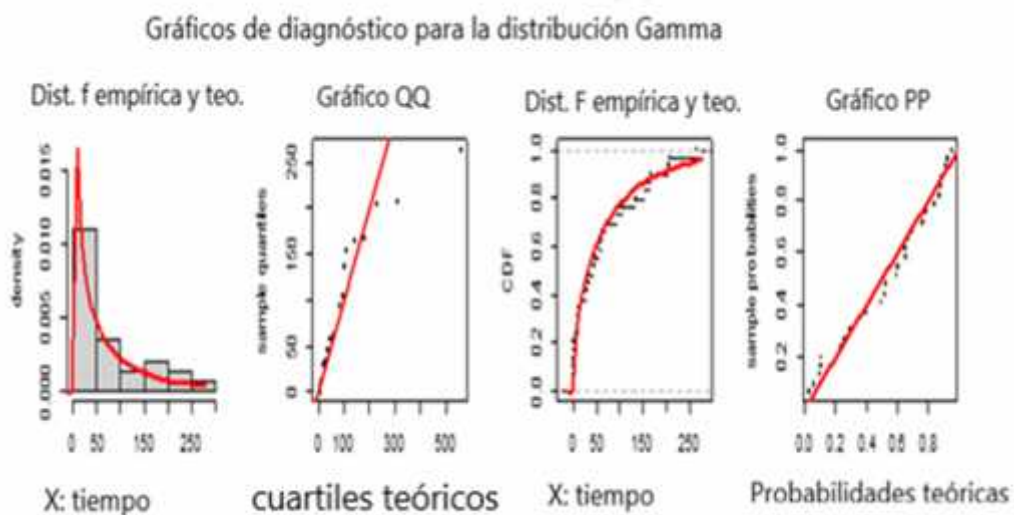


Elaboración propia

En la figura 8, se presentan los gráficos de diagnóstico para la distribución Weibull para la primera transición de la zona 1 a la zona 1, llamada t11.

El primer gráfico nos muestra una buena aproximación entre la distribución empírica f y la distribución f de Weibull (en rojo). En el segundo gráfico llamado gráfico QQ de los cuartiles teóricos y los cuartiles de los datos representados con los puntos negros, se nota una buena aproximación. En el tercer gráfico, se nota una buena aproximación entre la distribución empírica F , los puntos negros y la distribución F de Weibull. Y, en el último gráfico, la línea roja corresponde a las probabilidades teóricas de la distribución Weibull en una escala logarítmica y las probabilidades empíricas representadas con los puntos negros se nota una buena aproximación.

Figura 9: Gráficos de diagnóstico para la distribución Gamma de los tiempos de espera para la transición t11 (Zonas geográficas)



Elaboración propia

En la figura 9, un análisis similar al gráfico 7, se puede observar una buena aproximación entre la distribución Gamma y los datos de los tiempos de permanencia que corresponden a la transición t11.

-Por tanto, la distribución Gamma como la Weibull se ajustan adecuadamente a los datos según las pruebas de bondad de ajuste, pese a que la distribución Gamma tiene valores mejores en LogL, AIC, BIC, Chi-Square, AD, y un resultado ligeramente mejor en el test de Kolmogorov – Smirnov.

Si comparamos los gráficos de diagnóstico de ambas distribuciones, veremos que hay gran similitud en ambos gráficos, por tanto, nos quedamos con la distribución de Weibull para el ajuste de los datos.

Como la distribución elegida es la Weibull, se hallan sus parámetros respectivos, resultando, el parámetro de forma 0.6196 y el parámetro de escala 51.9981.

Algunas tablas para las demás transiciones se presentan en el Anexo 2.

De esta forma se ha procedido con el ajuste de los tiempos de permanencia para las demás transiciones (1,2), (1,3), etc., los cuales se presentan en la siguiente tabla.

Tabla 7: Parámetros de las distribuciones de los tiempos de espera de las transiciones de las zonas geográficas de Lima – Barranca.

Transición	Distribución	Param. Forma	Param. Escala
t11	Weibull	0.6196	51.9981
t12	Weibull	1.1721	64.4789
t13	Weibull	0.7524	86.1233
t14	Weibull	1.1461	82.8814
t21	Weibull	0.9694	84.0880
t22	Weibull	1.4164	99.4614
t23	Weibull	0.5772	55.6904
t24	Weibull	0.6618	60.2558
t31	Weibull	0.9441	60.3321
t32	Weibull	0.7477	40.4532
t33	Weibull	0.6260	25.6440
t34	Weibull	0.8185	77.8140
t41	Weibull	2.2874	70.9320
t42	Weibull	2.4586	153.9122
t43	Weibull	0.9625	136.7479
t44	Weibull	0.5827	92.5135

Elaboración propia

En la tabla 7, se muestra los parámetros de las distribuciones de los tiempos de espera para las transiciones de los estados de las zonas geográficas. Por ejemplo, para la transición del primer estado, que corresponde a Huacho al mismo estado Huacho; esto es t11, luego de las pruebas de ajuste, ha resultado la distribución de Weibull con los parámetros de forma y escala con valores 0.6196 y 51.9981, respectivamente.

Para el cálculo de los parámetros de las distribuciones de los tiempos de espera en los estados de las zonas geográficas, los datos se presentan en el Anexo 2, Tabla B1.

En esta parte vamos a presentar los resultados obtenidos para obtener las probabilidades de intervalo según el modelo semi-Márkov

De acuerdo con Másala (2012):

En un algoritmo para hallar las probabilidades de intervalo del modelo semi-Márkov homogéneo se fija el horizonte en el tiempo $[0, T]$ a utilizar y una subdivisión finita con K intervalos de longitud h .

Los datos de entrada son la matriz de transición P y las distribuciones de los tiempos de espera $F_{ij}(t)$.

Las matrices F son de $m \times m$ (m denota el número de estados) y ellos deben ser evaluados para $t = 1, \dots, k$ usando las distribuciones previamente determinadas.

Se deja en evidencia que $F(0) = 0$. Esto es la probabilidad que ocurra un sismo en el tiempo $t = 0$ es cero.

Se calcula para el caso, $t = 1, 3$ y 6 meses (3 valores en el tiempo, en este caso $k = 3$).

Tabla 8: F1, Matriz de probabilidad para $t=1$ mes

Estados	Huacho	Lima	Chilca	Sierra de Lima
Huacho	0.5089548	0.3349283	0.3638232	0.2680286
Lima	0.3080156	0.1673202	0.5032858	0.4675855
Chilca	0.4037158	0.5505393	0.6681936	0.3676598
Sierra de Lima	0.1303655	0.0177877	0.2072339	0.4047849

Elaboración propia

Estas probabilidades han sido obtenidas para cada transición, utilizando la distribución de Weibull presentada en la tabla 6 con sus parámetros respectivos, para el periodo de tiempo de $t = 1$ mes. Estas probabilidades son independientes de las probabilidades de transición de los estados dados en la tabla 5.

Por ejemplo, en la tabla 8, el valor 0.3349283 es la probabilidad Weibull (1.1721, 64.4789, $t = 1$), la cual significa la probabilidad que ocurra un sismo en el intervalo de tiempo (0, 1 mes) simplemente. Luego más adelante cuando consideremos a la matriz de transición respectiva, tendremos la probabilidad condicional que ocurra en Lima, dado que el anterior ocurrió en Huacho.

Tabla 9: F2, Matriz de probabilidad para t=3 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.7545869	0.7719657	0.6443051	0.6668104
Lima	0.6563344	0.5802017	0.7326577	0.7285792
Chilca	0.7674817	0.8376902	0.8885977	0.6758241
Sierra Lima	0.8216371	0.23458919	0.4875511	0.6262165

Elaboración propia

En la tabla 9, el valor 0.8376902 es la probabilidad Weibull (0.7477, 40.4532, $t = 3$), la cual significa la probabilidad que ocurra un sismo en el intervalo de tiempo (0, 3 meses) simplemente. Luego, más adelante, cuando se considere la matriz de transición respectiva, se tendrá la probabilidad condicional de que ocurra en Lima dado que el último ocurrió en Chilca (matriz Q2).

Y así se ha calculado para F3, considerando $t = 6$ meses.

Tabla 10: F3, Matriz de probabilidad para t=6 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.8844895	0.9642472	0.8247112	0.9121733
Lima	0.8764827	0.901416	0.8602897	0.8729362
Chilca	0.9396	0.9527807	0.9661903	0.8628479
Sierra Lima	0.9997787	0.76997082	0.7282277	0.7709399

Elaboración propia

En la tabla 10, el valor 0.9661903 es la probabilidad Weibull (0.7477, 40.4532, $t = 6$), la cual significa la probabilidad que ocurra un sismo en el intervalo de tiempo (0, 6 meses).

Las matrices Q son definidas como $Q(t) = F(t) \times P$ (producto de elemento por elemento).

Tabla 11: Q1, Matriz de probabilidad interválica para t=1 mes

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.1845	0.0754	0.1228	0.0201
Lima	0.0799	0.0310	0.2050	0.0693
Chilca	0.1124	0.1135	0.2342	0.0606
Sierra Lima	0.0335	0.0030	0.0888	0.0578

Elaboración propia

En la tabla 11, por ejemplo, para la transición Chilca – Lima, el valor 0.1135 se interpreta como la probabilidad que ocurra un sismo de alta magnitud en la ciudad de Lima, en un periodo de 1 mes, dado que el último ocurrió en Chilca.

Tabla 12: Q2, Matriz de probabilidad interválica para t=3 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.2735	0.1737	0.2175	0.0500
Lima	0.1702	0.1074	0.2985	0.1079
Chilca	0.2136	0.1727	0.3115	0.1115
Sierra Lima	0.2113	0.0402	0.2090	0.0895

Elaboración propia

En la tabla 12, por ejemplo, para la transición Chilca – Lima, el valor 0.1727, se interpreta como la probabilidad que ocurra un sismo de alta magnitud en la ciudad de Lima, en un periodo de 3 meses, dado que el último ocurrió en Chilca.

Tabla 13: Q3, Matriz de probabilidad interválica para t=6 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.3206	0.2170	0.2783	0.0684
Lima	0.2272	0.1669	0.3505	0.1293
Chilca	0.2615	0.1964	0.3387	0.1423
Sierra Lima	0.2571	0.1320	0.3121	0.1101

Elaboración propia

En igual forma, en la tabla 13, para la transición Chilca – Chilca, el valor 0.3387 se interpreta como la probabilidad que ocurra un sismo de alta magnitud en la ciudad de Chilca, en un periodo de 6 meses, dado que el último ocurrió en Chilca.

A continuación, se calculan las probabilidades interválicas, con las matrices B, las cuales están dadas por:

$$B(t) = 0 \quad \text{si } t = 0$$

$$B(t) = Q(t) - Q(t-1) \quad \text{si } t > 0$$

Se observa que $B1 = Q1$

Los elementos de la matriz B1 se interpretan como la probabilidad que ocurra un sismo en la zona i, dado que ha estado en la zona j, en el tiempo de 1 mes.

Tabla 14: B2, Matriz de probabilidad interváltica para t entre 1 y 3 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.0890	0.0983	0.0947	0.0299
Lima	0.0903	0.0765	0.0934	0.0387
Chilca	0.1013	0.0592	0.0773	0.0508
Sierra Lima	0.1778	0.0372	0.1201	0.0316

Elaboración propia

Los valores de la tabla 14, matriz B2, son las probabilidades que ocurra un sismo en la zona j dado que el último ha estado en la zona i, en el intervalo de tiempo entre 1 y 3 meses.

Por ejemplo, el valor 0.1201 es la probabilidad que ocurra un sismo de alta magnitud en Chilca, dado que el último ocurrió en la Sierra de Lima, en un periodo entre 1 y 3 meses. En forma similar se interpreta la matriz B3, cambiando el intervalo de tiempo entre 1 y 6 meses.

Tabla 15: B3, Matriz de probabilidad interváltica para t entre 3 y 6 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.0471	0.0433	0.0609	0.0184
Lima	0.0571	0.0595	0.0520	0.0214
Chilca	0.0479	0.0237	0.0272	0.0308
Sierra Lima	0.0458	0.0918	0.1031	0.0207

Elaboración propia

En la tabla 15, los valores de la matriz B3 corresponden a las probabilidades que ocurra un sismo en la zona j, dado que ha estado en la zona i, en el intervalo de tiempo entre 3 a 6 meses.

Se deduce, entonces, las matrices H de la siguiente manera:

$$H_{ij}(t) = \begin{cases} 0 & \text{si } i \neq j \\ \sum_{k=1}^m Q_{ik}(t) & \text{si } i = j \end{cases}$$

Estas matrices representan la probabilidad de salir del estado i en el tiempo t. Esto es la probabilidad que ocurra un sismo de alta magnitud en la zona de estudio entre Lima y Barranca, dado que el último sismo ocurrió en la zona i.

Tabla 16: H1, Matriz de probabilidad de ocurrencia para t=1 mes

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.4027	0	0	0
Lima	0	0.3852	0	0
Chilca	0	0	0.5207	0
Sierra Lima	0	0	0	0.1832

Elaboración propia

Los valores de la tabla 16 se interpretan como las probabilidades que el proceso salga del respectivo estado en el tiempo de un mes. Esto es, que ocurra un sismo de alta magnitud, dado que el último sismo ocurrió en dicho estado.

Tabla 17: H2, Matriz de probabilidad de ocurrencia para t=3 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.7147	0	0	0
Lima	0	0.6840	0	0
Chilca	0	0	0.8093	0
Sierra Lima	0	0	0	0.5499

Elaboración propia

Por ejemplo, en la tabla 17, el valor 0.8093 es la probabilidad que ocurra un sismo de alta magnitud en la zona de estudio, entre Lima y Barranca, dado que el último sismo de alta magnitud ocurrió en Chilca, considerando que fue evaluado en el tiempo de 3 meses. Las matrices Q y H son las más importantes.

Tabla 18: H3, Matriz de probabilidad de ocurrencia de estados, t=6

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.8843	0	0	0
Lima	0	0.8740	0	0
Chilca	0	0	0.9390	0
Sierra Lima	0	0	0	0.8113

Elaboración propia

En forma similar en la tabla 18, el valor 0.8740 es la probabilidad que ocurra un sismo de alta magnitud en la zona de estudio entre Lima y Barranca, dado que el último sismo de alta magnitud ocurrió en Lima, considerando que fue evaluado en el tiempo de 6 meses.

Finalmente, se define las matrices D como $D(t) = I - H(t)$, cuyos valores corresponden a las probabilidades de permanecer en los estados respectivos, en los tiempos de 1, 3 y 6 meses.

Tabla 19: D1, Matriz de probabilidad de sobrevivencia, t=1

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.5973	0	0	0
Lima	0	0.6148	0	0
Chilca	0	0	0.4793	0
Sierra Lima	0	0	0	0.8168

Elaboración propia

Tabla 20: D2, Matriz de probabilidad de sobrevivencia t=3 meses

Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.2853	0	0	0
Lima	0	0.3160	0	0
Chilca	0	0	0.1907	0
Sierra Lima	0	0	0	0.4501

Elaboración propia

Tabla 21: D3, Matriz de probabilidad de sobrevivencia t=6 meses

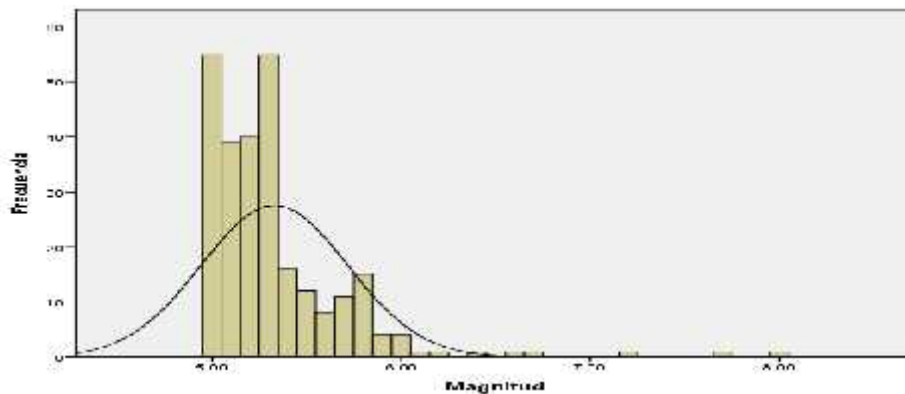
Estados	Huacho	Lima	Chilca	Sierra Lima
Huacho	0.1157	0	0	0
Lima	0	0.1260	0	0
Chilca	0	0	0.0610	0
Sierra Lima	0	0	0	0.1887

Elaboración propia

En las tablas 19, 20 y 21 se observa que, para un intervalo de tiempo de 1, 3 y 6 meses, respectivamente, es la sierra de Lima la que presenta la mayor probabilidad de sobrevivencia de los sismos de alta magnitud, esto es, que no ocurra un sismo en dicha zona.

4.3 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LA MAGNITUD LOS SISMOS ENTRE LIMA Y BARRANCA

Figura 10: Frecuencias de la magnitud de los sismos entre Lima y Barranca



Elaboración propia

En la figura 10, se muestra que la mayor frecuencia de los sismos está en el intervalo de 5.0 a 5.3 Mw., esto es en el primer intervalo de la magnitud.

Para ajustar el modelo semi-Márkov a los datos de la magnitud de los sismos en estudio, se sigue el siguiente proceso metodológico:

Primero debemos hallar la matriz de transición para los 3 intervalos de la magnitud de los sismos de la zona entre Lima y Barranca.

Másala (2016) usó los intervalos I ($M < 4.7$), estado II ($4.7 < M < 5$) y el III ($M \geq 5$).

En este caso, el comportamiento de los sismos en el Perú es diferente al caso de Italia.

Para hallar esta matriz se ha dividido la magnitud de los sismos en 3 estados:

Mag 1: Magnitud de los sismos entre 5.0 y 5.3

Mag 2: Magnitud de los sismos entre 5.4 y 5.7

Mag 3: Magnitud de los sismos mayor o igual a 5.8

Los criterios para esta elección están relacionados a las características de las magnitudes de los sismos que se presentan en esta zona entre Lima y Barranca, como a la cantidad de sismos en cada intervalo de la magnitud en el tiempo desde 1961 a 2022.

A continuación, se obtiene la matriz de frecuencias de los estados de la magnitud.

Tabla 22: Matriz de frecuencias de los estados de la magnitud.

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más	Total
De 5.0 a 5.3	137	31	21	189
De 5.4 a 5.7	28	12	6	46
De 5.8 a más	23	4	4	31
Total	188	47	31	266

Elaboración propia

En la tabla 22, se observa al primer intervalo cuya frecuencia es mayor, ha resultado 137 sismos que han ocurrido en el primer intervalo y han regresado al mismo intervalo. Luego se estimó la matriz de transición P como en el caso de las zonas geográficas.

Tabla 23: Matriz de transición para los estados de la magnitud

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.7249	0.164	0.1111
De 5.4 a 5.7	0.6087	0.2609	0.1304
De 5.8 a más	0.7419	0.129	0.129

Elaboración propia

La tabla 23 muestra que la mayor probabilidad de transición es para el primer intervalo, esto es, si ocurre un sismo de magnitud en el intervalo de 5.0 a 5.3 Mw, hay una probabilidad de 0.7249 que el siguiente sismo también sea de una magnitud del intervalo 5.0 a 5.3 Mw.

En el siguiente paso, se hallan la distribución de los tiempos de espera de cada transición. Se extrae de nuestra base de datos los tiempos de espera empíricos y utilizaremos la distribución de Weibull, utilizando la herramienta `riskDistribution` de R, los cuales se muestran en la tabla 24.

Tabla 24: Parámetros de las distribuciones para los tiempos de espera para las transiciones de los estados de la magnitud

Transición	Distribución	P. forma	P. escala
t11	Weibull	0.7085	68.9513
t12	Weibull	0.851	61.661
t13	Weibull	0.9716	83.1488
t21	Weibull	0.8129	56.5256
t22	Weibull	0.4766	45.1662
t23	Weibull	1.0732	24.7894
t31	Weibull	0.7799	58.5485
t32	Weibull	1.6067	101.8433
t33	Weibull	2.6411	52.9517

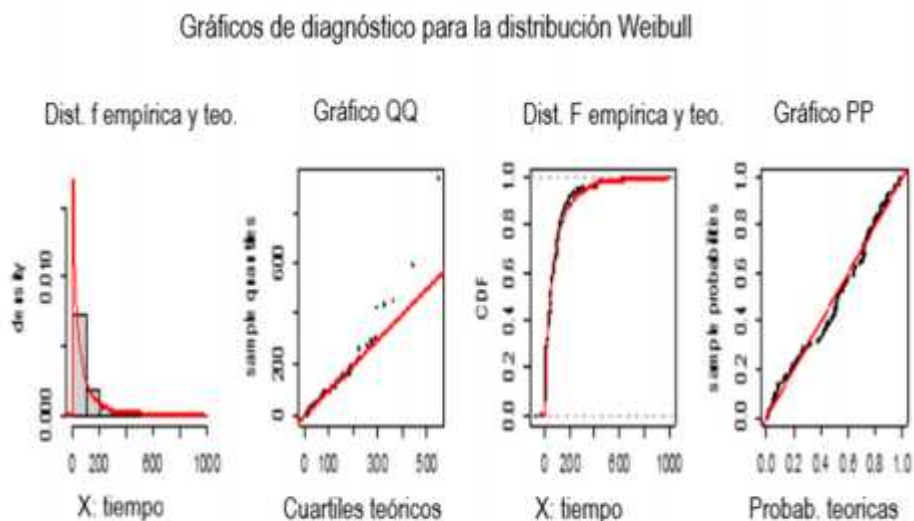
Elaboración propia

En la tabla 24, se muestra los parámetros de las distribuciones de los tiempos de espera para las transiciones de los estados de la magnitud. Por ejemplo, para la transición del

primer intervalo de 5.0 a 5.3 Mw al mismo intervalo 5.0 a 5.3 Mw. Esto es t_{11} , luego de las pruebas de ajuste, ha resultado la distribución de Weibull con los parámetros de forma y escala con valores 0.7085 y 68.9513, respectivamente.

Para el cálculo de los parámetros de las distribuciones de los tiempos de espera en los estados de la magnitud, los datos se presentan en el Anexo 2, Tabla B2.

Figura 11: Gráficos de diagnóstico para la distribución Weibull de los tiempos de espera para la transición t_{11} (Magnitud)



Elaboración propia

En la figura 11, se observa una buena aproximación de la distribución Weibull a los datos de los tiempos de permanencia de la primera transición t_{11} . En esta parte se va a presentar los resultados para el ajuste del modelo semi-Márkov para la magnitud de los sismos, similar al caso de las zonas geográficas.

Se presentan las matrices de probabilidades, (similar al caso de las zonas geográficas presentadas en el apartado anterior) solo para el caso en que $t = 3$ meses. (Para el caso $t=1$ y $t=3$ se presentan en el Anexo 3, Tablas C1 a C7).

Tabla 25: F2, Matriz de probabilidad para $t=3$ meses

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.7011	0.7483	0.6604
De 5.4 a 5.7	0.7676	0.7507	0.9815
De 5.8 a más	0.753	0.5595	0.9827

Elaboración propia

En este caso, en la tabla 25 el valor 0.7507 es la probabilidad de la transición (2,2) y corresponde al cálculo de la probabilidad Weibull (0.4766, 45.1662, $t = 3$), (ver tabla 23), la cual significa la probabilidad que ocurra un sismo en el intervalo de tiempo (0, 3 meses) en la transición de magnitud del intervalo 5.4 a 5.7 a 5.4 a 5.7, respectivamente.

Las matrices Q son definidas como $Q(t) = F(t) \times P$ (producto de elemento por elemento). A continuación, se presenta Q2, para $t=3$. (Para Q1 y Q3 para $t=1$ y para $t=6$ meses se presenta en el Anexo 3).

Tabla 26: Q2, Matriz de probabilidad interválica para $t=3$

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.5082	0.1227	0.0734
De 5.4 a 5.7	0.4673	0.1958	0.128
De 5.8 a más	0.5587	0.0722	0.1268

Elaboración propia

En esta tabla 26, la interpretación del valor 0.1958 es la probabilidad que ocurra un sismo en el intervalo de tiempo (0, 3 meses), en el intervalo de la magnitud del intervalo 5.4 a 5.7, dado que el último sismo ocurrió en la magnitud del intervalo de 5.4 a 5.7.

Similar al caso de las zonas geográficas, se deducen las matrices H.

Los valores de estas matrices H_i representan la probabilidad de salir del estado i en el tiempo t .

Calculamos para $t = 3$ meses. (Para H_1 y H_3 para $t=1$ y $t=6$ respectivamente, se presenta en el anexo 3)

Tabla 27: H2, Matriz probabilidad de ocurrencia, $t=3$

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.7043	0	0
De 5.4 a 5.7	0	0.7911	0
De 5.8 a más	0	0	0.7577

Elaboración propia

En la tabla 27, el valor 0.7577 es la probabilidad que ocurra un sismo de alta magnitud de 5.0 Mw a más, dado que el último sismo ocurrió en el intervalo de 5.8 Mw a más, evaluado en el tiempo de 3 meses.

Finalmente, se define las matrices D como $D(t) = I - H(t)$, las cuales corresponden a las probabilidades de permanecer en los estados respectivos en el tiempo de 3 meses.

(Para $t = 1$ y $t = 6$ los cálculos están en el anexo 3).

Tabla 28: D2, Matriz de probabilidad de sobrevivencia $t=3$

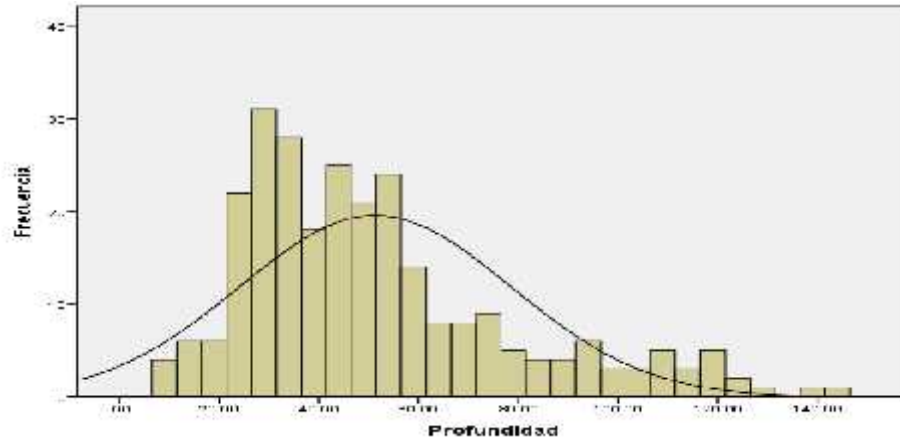
Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.2957	0	0
De 5.4 a 5.7	0	0.2089	0
De 5.8 a más	0	0	0.2423

Elaboración propia

Los valores de la tabla 28 corresponden a las probabilidades de no ocurrencia de los sismos en los intervalos de la magnitud respectiva. Por ejemplo, el valor 0.2089 es la probabilidad que no ocurra un sismo de alta magnitud, dado que el último sismo ocurrió en el intervalo de magnitud de 5.4 a 5.7 Mw.

4.4 AJUSTE DEL MODELO SEMI-MÁRKOV A LOS DATOS DE LA PROFUNDIDAD DE LOS SISMOS ENTRE LIMA Y BARRANCA

Figura 12. Frecuencia de la profundidad de los sismos entre Lima y Barranca.



Elaboración propia

En la figura 12, se observa que la mayor parte de los sismos ocurren dentro de la profundidad menor a 60 Km, los cuales son considerados de tipo superficial y se caracterizan por ser más destructivos que los que ocurren a mayor profundidad.

También se ha seguido la siguiente metodología:

Matriz de transición para los 3 intervalos de la profundidad de los sismos de la zona entre Lima y Barranca.

Para hallar matriz de transición se ha dividido la profundidad de los sismos en 3 estados:

Prof. 1: Profundidad de los sismos menor o igual a 60 Km (grave)

Prof. 2: Profundidad de los sismos entre 60 a 100 Km (intermedio)

Prof. 3: Profundidad de los sismos mayor a 100 Km (leve)

Los criterios para esta elección, por un lado, están relacionados a las características de las profundidades de los sismos que se presentan, en esta zona, entre Lima y Barranca y, por otro lado, a la cantidad de sismos en cada intervalo.

Tabla 29: Matriz de frecuencias de los estados de la profundidad

Estados	Hasta 60 Km	De 60 a 100Km	Más de 100 Km	Total
Hasta 60 Km	143	36	15	194
De 60 a 100 Km	37	12	2	51
Más de 100 Km	14	3	4	21
Total	194	51	21	266

Elaboración propia

Luego se estimó la matriz de transición P como en los casos anteriores.

En la matriz de frecuencias de la tabla 29, de los estados de la profundidad, se observa que lo más frecuente corresponde a sismos con una profundidad menor o igual a 60 Km.

Tabla 30: Matriz de transición para los estados de la profundidad de los sismos

Estados	Hasta 60 Km	De 60 a 100 Km	Más de 100 Km
Hasta 60 Km	0.7371	0.1856	0.0773
De 60 a 100 Km	0.7255	0.2353	0.0392
Más de 100 Km	0.6667	0.1429	0.1905

Elaboración propia

En la tabla 30, se puede observar que la probabilidad 0.7255 indica que, si ocurre un sismo con una profundidad en el intervalo entre 60 y 100 Km hay una probabilidad de 0.7255 que el siguiente sismo tenga una profundidad menor o igual a 60 Km.

En el siguiente paso, se obtiene la distribución de los tiempos de espera de cada transición. Se extrae de nuestra base de datos, los tiempos de espera empíricos y luego de las pruebas de ajuste respectivos, similar al caso de las zonas geográficas, se usa la distribución de Weibull, como la distribución que mejor se ajusta a los datos.

Tabla 31: Parámetros de las distribuciones para los tiempos de espera para las transiciones de los estados de la profundidad

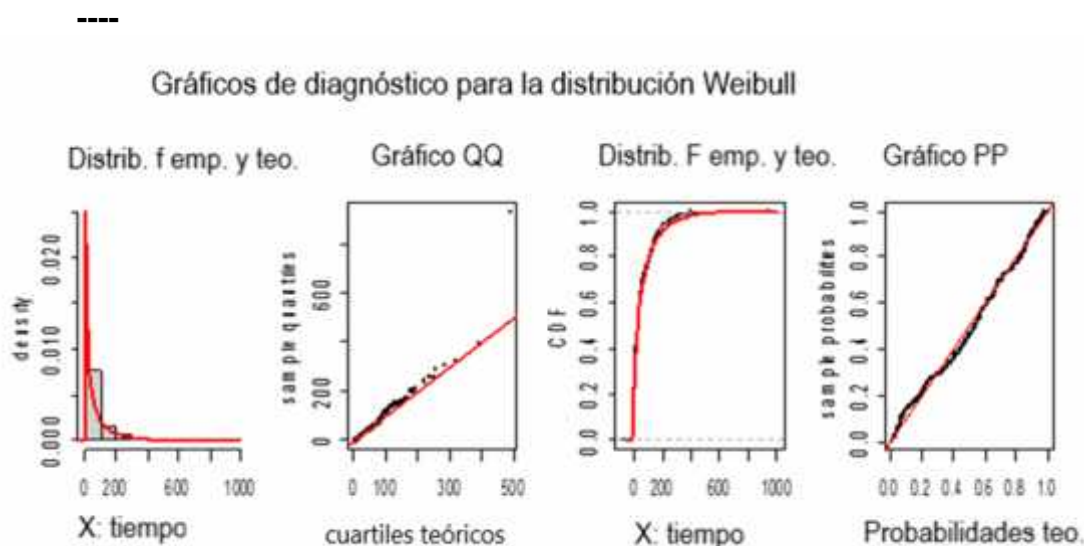
Transición	Distribución	P. forma	P. escala
t11	Weibull	0.6590	52.6067
t12	Weibull	1.1558	70.8590
t13	Weibull	0.6966	61.5722
t21	Weibull	0.8993	77.0484
t22	Weibull	0.7165	45.6066
t23	Weibull	2.8357	73.9258
t31	Weibull	1.1996	154.1307
t32	Weibull	1.5722	74.5467
t33	Weibull	1.1252	223.8478

Elaboración propia

En la tabla 31, se muestra los parámetros de las distribuciones de los tiempos de espera para las transiciones de los estados de la profundidad. Por ejemplo, para la transición del primer intervalo menor a 60 Km al mismo intervalo menor a 60 Km. Esto es t11, luego de las pruebas de ajuste, ha resultado la distribución de Weibull con los parámetros de forma y escala con valores 0.6590 y 52.6067, respectivamente.

Para el cálculo de los parámetros de las distribuciones de los tiempos de espera en los estados de la profundidad, los datos se presentan en el Anexo 2, Tabla B3.

Figura 13. Gráficos de diagnóstico para la distribución Weibull de los tiempos de espera para la transición t11 (Profundidad)



Elaboración propia

La figura 13 muestra los gráficos de diagnóstico del ajuste de la distribución de Weibull, en el cual se observa un buen ajuste, tanto de la distribución f de la distribución F, así como de los gráficos de cuartiles QQ y del gráfico de probabilidad PP.

En esta parte se va a presentar los resultados del ajuste del modelo semi-Márkov, a los datos de la profundidad de los sismos similar al caso de las zonas geográficas y la magnitud de los sismos.

En este caso $t = 3$ meses. (para $t=1$ y $t=6$ se presenta en el Anexo 3 Tablas C8 a C15)

Tabla 32: F2 Matriz de probabilidad para $t=3$ meses

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.7594	0.7324	0.7282
De 60 a 100Km	0.6834	0.8036	0.8257
Más de 100Km	0.4081	0.7394	0.3014

Elaboración propia

En la tabla 32, el valor 0.7594 es la probabilidad que ocurra un sismo de profundidad hasta 60Km, evaluado en la distribución Weibull, cuyos parámetros corresponden a la transición t11, evaluado en el tiempo de 3 meses.

Las matrices Q son definidas como $Q(t) = F(t) \times P$ (Producto de elemento por elemento).

A continuación, presentamos Q2 para t=3 meses. (Para t=1 y para t=6 se presenta en el anexo 3).

Tabla 33: Q2 Matriz de probabilidad interválica para t=3

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.5598	0.1359	0.0563
De 60 a 100Km	0.4958	0.1891	0.0324
Más de 100Km	0.2721	0.1056	0.0574

Elaboración propia

En la tabla 33, se puede observar el valor 0.5598 es la probabilidad que ocurra un sismo de profundidad menor o igual a 60 Km, dado que el último sismo fue también de esta profundidad en el tiempo de 3 meses.

Ahora se calculan las matrices H, similar al caso de las zonas geográficas:

A continuación se presenta H2 para t=3 meses. (Para t=1 y para t=6 se presenta en el Anexo 3).

Tabla 34: H2 Matriz de probabilidad de ocurrencia t=3

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.752	0	0
De 60 a 100Km	0	0.7172	0
Más de 100Km	0	0	0.4351

Elaboración propia

En la tabla 34, se observa que es el primer estado de la profundidad (Hasta 60 Km), resulta con la más alta probabilidad de ocurrencia (0.752) y esta característica corresponde a los sismos superficiales, los cuales son los más destructivos.

Finalmente, se calculan las matrices D como $D(t) = I - H(t)$, las cuales corresponden a las probabilidades de permanecer en los estados respectivos en los tiempos de 3 meses. (Para t=1 y t=6 esto es D1 y D3, se presenta en el Anexo 3).

Tabla 35: D2 Matriz de probabilidad de sobrevivencia t=3

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.248	0	0
De 60 a 100Km	0	0.2828	0
Más de 100Km	0	0	0.5649

Elaboración propia

En la tabla 35, se muestra las probabilidades de no ocurrencia de los estados de la profundidad. A manera de ejemplo, el valor 0.248 es la probabilidad que no ocurra un sismo de profundidad menor o igual a 60 Km, dado que el último sismo fue justamente de esta profundidad en un tiempo de 3 meses.

4.5 ANÁLISIS DE LOS RESULTADOS Y DISCUSIÓN

Se han escogido 4 zonas geográficas con mayor frecuencia de sismos de magnitud mayor o igual a 5.0 Mw, ubicadas entre Lima y Barranca, las cuales se delimitaron en forma arbitraria utilizando el GoogleEarthPro (ver Anexo 5). Las zonas geográficas son: Zona 1: Huacho-Barranca; Zona 2: Lima- Huaral; Zona 3: Chilca-Asia y Zona 4: Sierra de Lima.

En el tema del estudio de las dimensiones de la ocurrencia de los sismos, Masala (2012), aplica el modelo semi-Márkov solo para el estudio de la magnitud de los sismos. Sadeghian (2012), aplica el modelo semi-Márkov en el estudio de la magnitud de los sismos y a las zonas geográficas. Producto de su investigación, divide la provincia de Teherán en 6 regiones y encuentra probabilidades de ocurrencia para cada región, combinando con los resultados de los intervalos de la magnitud.

En el presente trabajo, se aplica el modelo semi-Márkov a las zonas geográficas (se ha dividido el área entre Lima y Barranca en 4 zonas), a la magnitud (se han considerado tres intervalos) y a la profundidad de los sismos (se ha dividido en tres intervalos). En comparación con Masala (2012) y Sadeghian (2012), se ha incluido una dimensión más: la profundidad de los sismos.

En el Perú, no hay ningún trabajo que aplique el modelo semi-Márkov, que incluya además del tiempo, las tres dimensiones, lugar, magnitud y profundidad de los sismos, siendo esto un aporte del presente trabajo.

De acuerdo con la metodología del modelo semi-Márkov, el cual se resume en la matriz de transición de los estados o zonas geográficas, en la matriz Q_i y en la matriz H_i , para $i=1,3,6$ meses, de probabilidades de ocurrencia de sismos en dichas zonas, se puede afirmar que son las zonas de Chilca y Huacho, las que presentan las mayores probabilidades de ocurrencia de sismos de gran magnitud, con 0.3115 y 0.2735, respectivamente. (ver Tabla 12, Matriz Q2 para los estados de las zonas geográficas).

Para el caso de ocurrencia de sismos en los estados (intervalos) de las magnitudes, podemos afirmar que para el primer intervalo de 5.0 a 5.3 presenta la mayor probabilidad de ocurrencia con 0.5082, seguido del intervalo de 5.4 a 5.7 Mw, el cual presenta una probabilidad de ocurrencia de 0.1958, evaluado en el tiempo de 3 meses. (ver Tabla 26, Matriz Q2).

Para el caso de los estados de la profundidad, ha resultado el primer estado (hasta 60 Km), con la mayor probabilidad de ocurrencia 0.5598, es decir, sismos con profundidad superficial, lo cual es preocupante, debido a que estos sismos son los más destructivos. (ver Tabla 33, Matriz Q2).

Se tomó en cuenta la afirmación teórica de Sadeghian (2012) quien sostiene que “Las probabilidades de transición de los estados de las zonas geográficas son independientes de las probabilidades de transición de los estados de la magnitud” Página 6.

En el presente trabajo se extiende esta idea, a la independencia de las probabilidades de transición de los estados de las zonas geográficas, de los estados de la magnitud y de los estados de la profundidad de los sismos.

Combinando los resultados anteriores de las dos dimensiones (zonas geográficas y profundidad) para sismos de profundidad superficial (hasta 60 Km), resultó que en Chilca la probabilidad de ocurrencia es de $0.3115 \times 0.5598 = 0.17$, evaluado en el tiempo de 3 meses.

Figura 14: La zona de Chilca presenta una probabilidad 0.17 de ocurrencia de un sismo de alta magnitud y superficial (profundidad menor a 60 Km.) en un periodo de 3 meses.



Elaboración propia

Entre las limitaciones del presente trabajo, se puede notar que el número de sismos con magnitudes mayores o iguales a 5.8 Mw es relativamente pequeño, lo cual es una limitación para las estimaciones de los parámetros de los tiempos de espera de las transiciones de los estados de la magnitud.

Se agradece a todas las personas que hicieron posible la realización del presente trabajo de investigación.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

Se ha estudiado las 4 zonas de mayor concentración de sismos en el departamento de Lima: Zona 1: Huacho-Barranca; Zona 2: Lima- Huaral; Zona 3: Chilca y Zona 4: Sierra de Lima, llegando a las siguientes conclusiones:

1. Al aplicarse la metodología del modelo semi-Márkov, a las zonas geográficas en estudio, Chilca presenta la mayor probabilidad de ocurrencia de sismos de gran magnitud.
2. Al aplicarse el modelo semi-Márkov a los intervalos de las magnitudes, el primer intervalo de 5.0 a 5.3 Mw, presenta la mayor probabilidad de ocurrencia.
3. Al aplicarse el modelo semi-Márkov a los intervalos de la profundidad, el primer intervalo (hasta 60 Km) presenta la mayor probabilidad de ocurrencia, esto es los sismos superficiales.
4. Además, al combinar el lugar y la profundidad de los sismos, en un periodo de 3 meses y considerando una profundidad de hasta 60 Km, Chilca presenta una probabilidad de 0.17 de ocurrencia de sismos de alta magnitud.

5.2. RECOMENDACIONES

Se recomienda incorporar más información de los sismos, ampliando el rango de la magnitud, es decir, se podrían considerar los sismos a partir de una magnitud de 4.0 ó 4.5 Mw a más.

Se recomienda aplicar el modelo semi-Márkov en un periodo de tiempo de 20 años para ver su evolución.

Se recomienda tomar las previsiones del caso, por el riesgo de ocurrencia de un sismo de alta magnitud y superficial en la zona de Chilca, la cual está relativamente cerca de Lima y podría ser muy destructivo para la ciudad de Lima.

REFERENCIAS BIBLIOGRÁFICAS

- Flores, C., Tavera, H., & Rodriguez, L. (2012). Aplicación del Algoritmo M8 en el borde occidental del Perú: Incrementos de probabilidad para la ocurrencia de grandes sismos EMOTOS. *Bol. Soc. geol. Perú* 106, 11-21.
- Votsi, I., Limnios, N., Tsaklidis, G., & Papadimitriou, E. (2012). Estimation of the Expected Number of Earthquake Occurrences Based on Semi-Markov Models. *Methodol Comput Appl Probab*, 685–703 .
- Altinok, Y., & Kolcak, D. (1999). An application of the semi-Markov model for earthquake occurrences in North Anatolia, Turkey. *J. Balkan Geophys. Soc.* 2 , 90–99.
- Alvarez, E. (2005). Estimation in Stationary Markov Renewal Processes with Application to Earthquake Forecasting in Turkey. *Methodology and Computing in Applied Probability*(7), 119–130.
- Borok, K. (2002). Earthquake prediction: State-of-the-Art and Emerging Possibilities. *International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Warshavskoe Moscow, Russia*, 79.
- Esri. (2023). Obtenido de <https://www.esri.com/es-es/arcgis/products/arcgis-desktop/resources>
- IGP, I. G. (2022). *IGP*. Obtenido de <https://www.igp.gob.pe/servicios/centro-sismologico-nacional/glosario>
- Masala, G. (2012). Earthquakes occurrences estimation through a parametric semi-Markov approach. *Journal of Applied Statistics* , 81-96.
- Orfanogiannaki, K., Karlis, D., & Papadopoulos, G. A. (2010). Identifying Seismicity Levels via Poisson Hidden Markov Models. *Pure Appl. Geophys.* 167, 919–931.
- Osaki, S. (1992). *Applied Stochastic System Modeling*. Berlin: © Springer-Verlag.
doi:10.1007/978-3-642-84681-6
- Pavani, E. G. (2011). About Earthquake Forecasting by Markov. *Methodol Comput Appl Probab* (2011) (13), 155-169. doi:10.1007/s11009-009-9137-3
- Puerto Rico, U. (Octubre de 2021). *Puerto Rico Red Sísmica*. Obtenido de <http://www.prsn.uprm.edu/Spanish/educacion/terremotos/tamano.php>
- Risco, C. (2018). Modelos ARIMA para el análisis de datos de sismos en el Perú en el 2017. *Revista IECOS*(18), 143-157.
- Ross, S. (2007). Introduction to Probability Models. *Elsevier Inc*, 433-442.

- Sadeghian, R. (2012). Forecasting time and place of earthquakes using a Semi-Markov model (with case study in Tehran province). *Journal of Industrial Engineering International*, 8-20. Obtenido de <http://www.jiei-tsb.com/content/8/1/20>
- Taha, H. (2012). *Investigación de operaciones*. Mexico: Pearson. Obtenido de ISBN: 978-607-32-0796-6
- Tavera, H., & Bendezu, L. (15 de Agosto de 2017). El Comercio - Perú. *El comercio*.

Anexo 1 Los datos de tiempo, latitud, longitud, magnitud y profundidad de los sismos utilizados en el presente estudio.

Tabla A1. Los datos de los sismos desde 1961 al 2022, magnitud ≥ 5 Mw.

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
1	17.03.1961	10:06:28	-11.1	-78.8	60	5
2	31.03.1963	05:51:00	-10.8	-78.5	32	5.8
3	28.04.1963	05:22:05	-11.7	-78.1	10	5.3
4	05.07.1963	05:48:13	-11.6	-77.5	55	6
5	17.09.1963	05:54:36	-10.78	-78.27	61	6.7
6	07.10.1963	12:43:53	-12.9	-76.8	69	5.7
7	17.10.1966	21:41:58	-10.83	-78.65	37	8
8	19.11.1966	18:20:30	-10.79	-79.13	22	5.3
9	18.12.1966	23:59:32	-10.82	-79	20	5.1
10	22.03.1967	05:05:19	-11.05	-79.15	26	5.1
11	03.06.1967	06:11:08	-10.98	-79.21	28	5.1
12	01.08.1967	01:13:43	-12.99	-76.72	52	5.8
13	22.08.1967	07:42:46	-10.99	-78.35	47	5.3
14	22.08.1967	09:59:36	-12.64	-77.02	59	5.2
15	16.10.1967	01:31:48	-11.5	-78.4	83	5.1
16	03.02.1968	11:07:50	-12.6	-76.59	68	5.3
17	03.02.1968	11:43:07	-12.71	-76.59	67	5.3
18	03.04.1969	03:47:18	-11.4	-76.6	33	5.2
19	09.02.1970	16:15:55	-12.72	-77.33	31	5.1
20	29.06.1970	03:02:33	-11.12	-79.05	44	5.3
21	14.08.1970	05:41:42	-10.93	-79.27	53	5.1
22	23.09.1970	08:32:00	-11.78	-77.64	75	5
23	27.10.1970	09:37:41	-12.93	-76.55	77	5
24	30.01.1971	10:19:57	-11	-78.04	41	5.1
25	28.05.1971	06:50:02	-12.58	-77	53	5.3
26	22.09.1971	22:15:01	-10.97	-78.16	56	5.6
27	01.10.1971	21:08:35	-12.74	-77.16	26	5.1
28	03.10.1971	10:29:27	-12.75	-77.36	24	5
29	03.10.1971	14:58:27	-12.78	-77.28	31	5.5
30	09.10.1971	21:54:32	-12.9	-76.9	33	5.3
31	29.11.1971	20:14:14	-11.34	-77.79	54	5.6
32	03.04.1972	01:29:50	-11.64	-76.6	96	5.3
33	19.06.1972	15:50:29	-12.12	-77.66	42	5.4
34	25.04.1973	16:51:13	-10.9	-76.67	117	5
35	24.12.1973	18:18:12	-12.71	-77.54	38	5.7
36	05.01.1974	08:33:51	-12.39	-76.29	92	6.6
37	27.09.1974	16:09:03	-12.6	-77.8	25	5.3
38	03.10.1974	14:21:34	-12.28	-77.54	21	7.7
39	03.10.1974	15:28:19	-12.8	-76.37	25	5
40	03.10.1974	20:39:27	-12.52	-77.73	34	5.2

Fuente Instituto Geofísico del Perú IGP

Continuación...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
41	03.10.1974	0.867847	-12.64	-77.67	97	5
42	03.10.1974	0.926539	-12.27	-77.86	69	5.2
43	05.10.1974	0.35412	-12.39	-77.61	79	5.1
44	07.10.1974	0.206863	-12.22	-77.92	28	5.3
45	09.10.1974	0.743009	-12.59	-77.92	30	5.2
46	09.10.1974	0.912951	-12.59	-76.8	30	5.3
47	10.10.1974	0.828484	-12.5	-77.68	31	5.6
48	02.11.1974	0.040023	-12.73	-78	27	5.3
49	09.11.1974	0.541609	-12.52	-77.59	13	7.2
50	11.11.1974	0.307894	-11.37	-79.1	33	5.1
51	14.11.1974	0.803912	-12.88	-77.08	42	5.7
52	24.12.1974	0.037558	-12.82	-77.78	48	5.1
53	23.05.1975	0.809884	-10.84	-78.73	64	5.3
54	11.08.1975	0.443981	-11.84	-77.74	74	5.8
55	09.10.1975	0.565463	-12.26	-77.97	22	5.7
56	29.10.1975	0.263009	-11.04	-78.06	60	5.8
57	18.01.1976	0.978449	-11.7	-76.68	113	5.2
58	24.07.1976	0.2425	-11.92	-76.79	123	5.2
59	01.11.1976	0.126319	-12.41	-77.82	38	5.5
60	26.01.1977	0.63625	-12.19	-77.93	23	5.4
61	13.10.1977	0.96978	-12.3	-77.94	31	5.5
62	22.01.1978	0.700023	-11.71	-77.79	50	5.4
63	29.07.1978	0.523611	-11.59	-76.48	94	5.3
64	14.11.1978	0.278495	-11.05	-78.42	33	5.2
65	26.12.1978	0.184363	-12.04	-77.82	33	5.2
66	15.06.1979	0.37059	-12.24	-76.76	74	5.7
67	03.07.1979	0.66985	-12.8	-77.63	33	5.3
68	15.07.1979	0.380938	-12.4	-77.12	49	5.3
69	01.08.1979	0.695093	-11.2	-77.95	75	5.5
70	02.01.1980	0.935567	-10.98	-79.24	25	5.5
71	02.01.1980	0.937847	-10.83	-79.01	48	5.6
72	02.01.1980	0.964213	-11	-79.14	29	5.4
73	03.01.1980	0.879676	-10.95	-78.9	46	5.6
74	04.01.1980	0.314502	-11.01	-79.18	45	5.4
75	05.01.1980	0.217708	-11.22	-79.22	33	5.3
76	10.04.1980	0.031991	-11.62	-77.46	67	5.3
77	04.07.1981	0.634977	-11.11	-78.47	32	5.2
78	22.07.1981	0.320093	-11.14	-78.19	39	5.3
79	13.02.1982	0.375718	-10.84	-78.74	33	5.2
80	23.03.1982	0.215394	-12.34	-77.6	58	5.4

Continuación ...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
81	28.03.1982	0.97559	-12.77	-75.98	88	6.2
82	25.04.1982	0.610579	-11.17	-78.57	33	5.3
83	16.02.1983	0.581528	-12.36	-78.07	11	5.3
84	02.03.1983	0.297002	-11.53	-77.8	57	5.8
85	25.03.1983	0.084734	-11.4	-77.3	35	5
86	08.04.1983	0.304907	-10.82	-78.29	33	5.2
87	23.07.1983	0.126285	-11.4	-78.8	90	5.1
88	24.10.1983	0.025081	-12.85	-76.66	48	5.8
89	01.11.1983	0.41287	-11.32	-77.99	81	5.3
90	25.12.1983	0.017245	-11.63	-78.56	85	5
91	26.01.1984	0.813183	-12.36	-76.93	53	5.6
92	09.02.1984	0.054977	-12.92	-76.34	63	5.7
93	24.03.1984	0.894236	-10.99	-78.3	44	5.2
94	05.05.1984	0.950382	-12.54	-78.17	30	5.3
95	08.05.1984	0.677072	-11.42	-76.78	112	5.1
96	17.11.1984	0.228356	-12.69	-77.01	39	5
97	13.01.1985	0.555139	-12.14	-77.83	55	5.1
98	24.02.1985	0.940509	-12.62	-77.19	52	5
99	01.04.1985	0.344699	-12.78	-76.68	43	5.4
100	19.05.1985	0.780949	-12.28	-77.03	67	5.2
101	30.06.1985	0.716609	-11.33	-78.51	57	5.3
102	11.10.1985	0.130544	-12.43	-77.15	74	5.3
103	25.10.1985	0.432315	-10.97	-79.03	19	5.3
104	31.10.1985	0.004375	-10.85	-79.04	40	5.3
105	04.11.1985	0.942442	-12.46	-77.98	22	5.5
106	17.01.1986	0.176343	-10.95	-78.71	32	5.8
107	14.02.1986	0.060486	-11.85	-77.88	31	5
108	17.03.1986	0.502917	-12.76	-76.79	26	5.3
109	05.05.1986	0.084653	-12.26	-76.46	44	5.1
110	05.05.1986	0.551528	-12.56	-76.57	29	5
111	22.05.1986	0.109919	-12.4	-76.97	30	5.1
112	05.07.1986	0.454144	-10.98	-76.79	117	5.2
113	04.10.1986	0.892419	-12.9	-76.7	50	5.1
114	12.10.1986	0.618854	-12.86	-76.89	46	5.1
115	06.11.1986	0.048947	-12.57	-76.98	36	5
116	14.11.1986	0.288264	-11.96	-77.08	53	5
117	16.11.1986	0.182257	-12.67	-78.14	27	5.2
118	23.12.1986	0.482292	-12.82	-77.27	24	5.2
119	12.05.1987	0.679572	-11.64	-78.03	52	5.2
120	19.05.1987	0.462025	-10.85	-76.79	106	5

Continuación...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
121	09.06.1987	0.779606	-12.83	-76.14	72	5.4
122	15.06.1987	0.266065	-12.64	-77.02	42	5.8
123	26.08.1987	0.450868	-10.9	-77.18	79	5.2
124	27.11.1987	0.34765	-12.88	-76.76	50	5.3
125	25.01.1988	0.302801	-11.45	-77.87	69	5.8
126	02.07.1988	0.270926	-12.65	-76.54	83	5.3
127	07.07.1988	0.159954	-12.58	-77.35	53	5.1
128	20.08.1988	0.260775	-12.32	-76.41	76	5
129	01.01.1989	0.427998	-12	-77.41	46	5.5
130	08.06.1989	0.782731	-12.57	-78.01	25	5.1
131	08.07.1989	0.50375	-13.01	-75.99	92	5.1
132	15.10.1989	0.697616	-12.82	-76.72	65	5.3
133	07.12.1989	0.227581	-12.52	-76.99	43	5
134	16.02.1990	0.635069	-11.73	-77.15	69	5.3
135	13.03.1990	0.821146	-12	-76.86	47	5.5
136	16.03.1990	0.084421	-12.89	-76.66	29	5.3
137	20.03.1990	0.776377	-11.78	-76.16	80	5.2
138	03.05.1990	0.559502	-12.09	-77.21	61	5
139	14.06.1990	0.333831	-11.59	-76.95	42	5.2
140	14.06.1990	0.546007	-12.74	-76.6	15	5.4
141	15.07.1990	0.425347	-12.11	-76.96	34	5
142	08.10.1990	0.876574	-12.44	-76.83	40	5.2
143	29.04.1991	0.758634	-11.3	-77.5	60	6.1
144	09.10.1991	0.72412	-12.45	-77.45	35	5.4
145	22.12.1991	0.819051	-11.78	-77.38	73	5.2
146	27.04.1992	0.089676	-13.03	-76.36	37	5
147	03.05.1992	0.236192	-12.61	-75.95	23	5.4
148	27.11.1992	0.867465	-13.01	-76.81	21	5.2
149	18.04.1993	0.386377	-11.65	-76.84	107	6.4
150	22.08.1993	0.622975	-12.64	-76.96	61	5.5
151	07.09.1993	0.893275	-10.78	-78.09	36	5.1
152	04.03.1994	0.297662	-12.79	-77.04	42	5.1
153	15.03.1994	0.315046	-12.81	-76.97	29	5
154	26.05.1994	0.490012	-11.36	-76.81	94	5
155	04.06.1994	0.764826	-12.69	-76.63	32	5
156	20.06.1994	0.438657	-11.67	-77.98	42	5
157	22.06.1994	0.857083	-11.26	-78.24	32	5.2
158	05.07.1994	0.082766	-11.64	-76.2	142	5
159	10.02.1995	0.721944	-11.9	-77.56	41	5.1
160	02.06.1995	0.971088	-11.11	-79.03	22	5.1

Continuación...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
161	21.06.1995	0.68963	-11.8	-77.87	54	5.8
162	31.07.1995	0.366991	-10.78	-78.83	40	5.9
163	17.09.1995	0.782303	-10.91	-78.94	38	5
164	24.09.1995	0.054016	-10.98	-78.98	28	5.1
165	27.09.1995	0.348414	-11.71	-78.11	24	5
166	31.12.1995	0.407396	-11.44	-78.58	9	5
167	13.02.1996	0.054005	-11.73	-77.65	31	5.2
168	20.05.1996	0.79485	-11.08	-77.33	65	5
169	03.08.1996	0.97272	-12.54	-76.93	45	5.3
170	07.08.1996	0.965	-12.71	-77.08	55	6
171	01.10.1996	0.961215	-12.85	-77.29	42	5.8
172	01.11.1997	0.553079	-12.49	-77.08	59	5.3
173	30.01.1998	0.741308	-12.64	-76.95	43	5.4
174	20.03.1998	0.45963	-12.42	-78.19	26	5.5
175	20.03.1998	0.547662	-11.25	-78.46	18	5
176	16.04.1998	0.284688	-12.9	-77.09	36	5
177	19.05.1998	0.195023	-10.97	-77.37	53	5
178	02.10.1998	0.234965	-11.49	-77.91	84	5.1
179	05.11.1998	0.1564	-10.8	-79.09	37	5.7
180	25.12.1998	0.601586	-11.87	-77.67	30	5.7
181	06.05.1999	0.121875	-12.51	-77.24	42	5.3
182	18.08.1999	0.075961	-11.73	-76.74	98	5.2
183	10.12.1999	0.929132	-11.15	-78.63	49	5
184	09.05.2000	0.46015	-12.98	-76.74	59	5.3
185	06.12.2000	0.283137	-11.1	-78.33	56	5.8
186	08.02.2001	0.316053	-11.26	-79.32	29	5.1
187	31.10.2001	0.926678	-11.02	-78.93	36	5.3
188	17.04.2002	0.891111	-11.18	-78.03	30	5
189	11.05.2002	0.446759	-11.53	-78.01	42	5.9
190	14.06.2002	0.314676	-11.52	-76.41	16	5
191	08.05.2003	0.689595	-12.97	-77.39	50	5.9
192	28.05.2003	0.893576	-12.6	-77.37	20	5.8
193	02.07.2003	0.85463	-11.94	-76.9	90	5
194	06.10.2003	0.310405	-11.75	-78.24	13	5.3
195	14.03.2004	0.28706	-12.99	-77.48	54	5.3
196	02.04.2004	0.704606	-12.96	-77.36	44	5.3
197	29.07.2004	0.416817	-11.38	-78.79	25	5.7
198	17.02.2005	0.133866	-11.3	-76.32	121	5.7
199	02.03.2005	0.575347	-11.83	-76.14	120	5.3
200	08.05.2006	0.060324	-11.34	-76.28	105	5

Continuación...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
201	21.07.2006	0.546505	-11.34	-79.41	50	5.3
202	02.01.2007	0.738785	-11.4	-79.08	31	5.3
203	28.07.2007	0.59162	-11.15	-78.01	49	5.3
204	29.03.2008	0.535683	-12.24	-77.29	50	5.6
205	07.06.2008	0.546169	-12.47	-77.4	66	5.3
206	07.09.2008	0.451319	-10.88	-77.81	113	5.1
207	26.04.2010	0.74294	-12.73	-76.98	36	5
208	01.02.2011	0.559769	-10.89	-76.6	126	5.1
209	23.09.2011	0.427199	-12.12	-76.07	110	5
210	26.11.2011	0.291609	-10.89	-78.49	54	5.4
211	19.12.2011	0.234688	-12.25	-77.38	44	5.2
212	29.12.2011	0.573391	-12.63	-76.89	72	5.2
213	14.02.2012	0.196007	-12.31	-77.39	47	5.2
214	19.03.2012	0.973264	-11.22	-78.18	51	5.2
215	25.03.2012	0.958993	-12.68	-77.83	35	5
216	08.06.2012	0.53772	-12.57	-77.19	54	5
217	21.06.2012	0.928588	-11.71	-76.96	96	5.1
218	29.07.2012	0.733218	-12.01	-77.2	61	5
219	01.11.2012	0.06765	-11.96	-76.01	22	5.1
220	04.11.2012	0.077928	-12.13	-76.03	110	5.3
221	29.12.2012	0.081713	-11.44	-77.04	97	5
222	31.12.2012	0.365486	-11.44	-77.85	51	5.3
223	04.05.2013	0.170972	-12.86	-77.92	30	5.1
224	18.06.2013	0.777789	-11.8	-77.4	62	5
225	12.07.2013	0.81081	-12.63	-77.19	38	5
226	25.11.2013	0.838137	-12.78	-76.49	62	5.8
227	02.01.2014	0.935938	-11.2	-78.43	28	5
228	03.06.2014	0.898762	-12.52	-77.14	33	5.5
229	17.06.2014	0.84338	-11.27	-77.22	49	5.1
230	15.11.2014	0.012928	-12.76	-77.14	23	6
231	05.04.2015	0.248194	-11.67	-78.05	39	5
232	04.06.2015	0.899236	-12.46	-77.68	15	5.1
233	28.08.2015	0.777222	-12.35	-76.28	88	5.4
234	27.10.2015	0.683252	-10.79	-78.6	56	5.2
235	25.12.2015	0.290359	-10.81	-78.46	31	5
236	06.05.2016	0.619711	-11.8	-77.76	46	5.2
237	18.07.2016	0.066539	-11.35	-78.18	59	5.7
238	17.08.2016	0.525266	-11.6	-76.93	108	5.1
239	14.10.2016	0.906331	-11.02	-79.3	37	5.2
240	15.10.2016	0.572303	-12.44	-78.07	23	5.1

Continuación...

Nro	FechaUTC	HoraUTC	Latitud	Longitud	Profund	Magnit
241	19.10.2016	0.277882	-12.44	-78.09	14	5.3
242	27.10.2016	0.021933	-12.89	-76.99	52	5
243	28.10.2016	0.157975	-12.03	-77.66	52	5.3
244	09.11.2016	0.09397	-11.5	-76.2	137	5
245	08.12.2016	0.796678	-10.94	-78.32	41	5.2
246	17.12.2016	0.307905	-11.01	-79.25	35	5.9
247	22.12.2016	0.697118	-12.73	-77.8	26	5.3
248	06.01.2017	0.955174	-11.49	-78.58	28	5.4
249	01.02.2017	0.485035	-11.18	-76.75	129	5
250	09.07.2017	0.046806	-11.82	-77.57	54	5.2
251	14.07.2017	0.737338	-11.04	-77.05	118	5
252	19.08.2017	0.140509	-10.98	-79.33	31	5.4
253	14.09.2017	0.346794	-11.9	-76.33	11	5.2
254	23.10.2017	0.006377	-12.41	-77.45	34	5.1
255	24.11.2017	0.677685	-12.16	-76.25	110	5.1
256	05.02.2018	0.810486	-12.75	-78.19	31	5.2
257	07.04.2018	0.843808	-10.91	-78.53	51	5.3
258	17.05.2018	0.463333	-12.79	-76.87	49	5.8
259	22.05.2018	0.547639	-11.26	-78.15	48	5.3
260	27.09.2018	0.916748	-11.51	-78.13	52	5.3
261	20.04.2021	0.203808	-12.65	-76.88	44	5
262	23.06.2021	0.12103	-12.77	-77.01	45	6
263	25.08.2021	0.302188	-12.41	-77.57	31	5
264	28.11.2021	0.272569	-12.06	-77.54	64	5.2
265	07.01.2022	0.435475	-11.97	-76.84	105	5.6
266	31.03.2022	0.008125	-11.29	-78.41	27	5.2
267	12.05.2022	0.913738	-12.48	-77.07	41	5.5

Anexo 2 Tiempos de permanencia

Tabla B1. Tiempos de permanencia (días) en los estados de las zonas geográficas

t11	t12	t13	t14	t21	t22	t23	t24	t31	t32	t33	t34	t41	t42	t43	t44
59.0	27.2	19.2	125.8	73.2	67.2	33.4	309.7	29.0	5.7	0.5	424.5	107.5	76.1	311.2	20.5
32.9	39.2	58.3	80.3	19.6	58.4	0.4	187.7	20.1	0.1	1.9	11.8	28.0	188.0	242.7	12.1
28.8	79.8	0.7	12.9	449.0	259.6	1.9	4.2	54.4	1.4	0.6	3.0	113.1	133.3	264.4	431.6
94.0	41.3	109.1	202.4	22.3	101.0	1.2	6.7	139.7	85.1	5.6	48.5	73.1	43.8	192.7	233.6
72.2	95.2	117.4	26.0	41.8	170.2	99.2	33.4	94.4	13.6	0.4	43.1	63.4	219.1	0.9	2.1
45.1	37.4	8.9	25.1	1.4	24.6	17.4	94.7	116.3	56.2	0.1	71.3	54.1	157.5	90.5	
153.7	53.4	2.3		111.7	41.6	31.0	11.2	50.9	47.3	0.3	43.2	59.8		5.8	
0.1	27.2	296.6		39.7	95.9	41.6	4.0	1.5	7.0	0.9	29.8	28.1		92.5	
0.1	19.0	92.1		94.3	37.9	30.1		149.0	139.5	22.8	3.1	35.7		98.5	
1.0	2.1	41.9		96.1	39.3	1.3		16.4	70.2	6.0	5.1			207.2	
0.9	43.4	102.7		33.2		157.4		7.0	30.5	39.8	141.9			125.4	
0.3	49.2	3.0		72.6		2.8		43.1	73.7	11.7	71.3			8.5	
17.6	23.9	158.3		57.5		0.7		13.1	15.8	13.8	103.1			327.3	
205.3	244.6	162.8		82.4		84.4		73.9	34.9	35.9	280.7			38.3	
13.1	22.3	177.9				126.8		58.3	46.6	16.6	84.9			72.7	
105.3	132.3	74.8				131.6		202.9	12.5	7.9	31.0				
5.4	29.1	26.5				159.3		15.6	44.2	24.6					
47.4	127.5	150.9				69.5		0.5	140.0	36.2					
6.8		595.5				9.2		32.3	0.0	4.3					
135.2		6.0				23.8		210.5	94.3	52.7					
33.2		123.4				59.2		117.7		10.3					
63.3		151.9				935.9		91.5		4.0					
264.3		150.8						37.8		55.0					
167.9		0.9						13.9		396.0					
164.5		4.3						14.7		89.6					
206.7		39.8						60.8		48.7					
1.1		41.0						4.5		19.7					
58.7										18.3					
8.8										75.0					
										135.8					
										3.6					
										7.3					
										63.2					
										62.1					

Elaboración propia – Datos IGP

Tabla B2. Los tiempos de permanencia (días) en los estados de la magnitud de los sismos

t11				t12	t13	t21	t22	t23	t31	t32	t33
28.76	56.23	111.72	23.78	116.28	59.00	8.93	85.13	29.00	27.24	19.25	73.24
94.00	41.56	6.78	59.25	.56	67.22	5.62	259.64	11.76	32.90	58.44	39.69
72.21	41.78	2.05	58.68	50.91	58.26	125.84	100.97	19.57	20.05	162.76	54.97
.68	102.72	94.35	132.29	76.06	5.67	309.66	153.70	4.22	264.36	125.39	19.69
54.42	13.13	43.41	57.53	242.70	6.04	22.83	.06	73.94	.40		
109.06	5.43	96.05	.91	.91	79.81	39.80	.06	5.78	1.54		
.54	30.06	74.79	3.57	2.31	13.58	187.70	.96		80.26		
424.49	48.50	26.55	7.28	99.24	92.13	17.37	.88		27.98		
311.16	.92	32.28	.02	170.18	58.35	.31	13.81		22.30		
139.68	16.55	135.20	11.16	16.38	202.88	43.05	48.74		7.03		
45.13	43.11	103.12	28.09	37.38	141.87	47.34	49.16		27.18		
39.24	90.45	113.08	157.49	31.02	18.97	157.43	202.42		71.27		
33.36	7.89	150.93	4.05	35.94	3.97	2.82			158.30		
94.40	24.62	264.32	38.35	3.00	210.46	30.55			47.37		
117.43	7.05	167.93	31.01	20.46	23.89	73.72			395.96		
1.88	1.29	95.85	72.68	133.26	327.31	207.24			63.28		
.36	36.18	159.31	60.81	24.64	135.81	15.62			33.45		
.14	139.48	18.29	127.55	.67	150.84	.54			34.89		
.13	6.68	431.58	935.92	5.09	8.80	131.60			37.84		
1.93	92.45	73.06	94.30	89.55	39.84	12.13			140.01		
1.35	4.27	164.55		33.23	63.20	69.54			4.31		
1.21	43.16	206.74		117.70		22.29			4.46		
.26	29.78	91.55		244.59		13.90			62.12		
149.04	98.50	595.45		63.43		59.78					
187.98	52.70	280.74		151.94		29.07					
107.52	70.23	233.56		84.90		25.96					
41.28	3.08	9.23		72.62		25.14					
11.67	43.78	46.57		14.70		82.44					
95.22	41.56	33.20		35.74							
449.03	84.43	5.97		39.27							
17.63	126.82	74.96		41.01							
205.32	177.89	12.54									
296.61	10.30	37.93									
13.08	71.32	94.73									
105.30	8.49	2.07									
53.41	15.76	54.08									
41.89	1.44	1.08									
2.95	12.86	123.37									
192.68	219.08	44.17									

Elaboración propia – Datos IGP

Tabla B3. Los tiempos de permanencia (días) en los estados de la profundidad de los sismos

t11			t12	t13	t21	t22	t23	t31	t32	t33	
59.00	11.67	202.88	18.29	73.24	309.66	29.00	19.25	91.55	242.70	20.46	187.98
27.24	.06	162.76	117.70	54.42	80.26	424.49	109.06	39.27	99.24	125.39	12.13
67.22	.06	5.09	164.55	39.24	2.95	94.40	.54		192.68	54.08	431.58
32.90	.96	207.24	206.74	125.84	43.11	76.06	33.36		90.45		233.56
28.76	.88	177.89	244.59	11.76	6.68	264.36	.13		219.08		
94.00	.31	10.30	22.29	.14	141.87	1.35	1.93		73.06		
72.21	17.63	15.76	33.20	149.04	12.86	58.44	79.81		595.45		
58.26	205.32	1.44	5.97	187.70	202.42	107.52	53.41		63.43		
20.05	37.38	111.72	74.96	170.18	280.74	17.37	158.30		57.53		
.68	296.61	18.97	123.37	16.38	2.07	153.70	98.50		28.09		
311.16	13.58	39.69	151.94	95.22	29.07	449.03	43.78		157.49		
139.68	22.30	47.37	13.90	4.22	11.16	27.98	37.93		35.74		
45.13	13.08	6.78	150.84	105.30	25.96	92.13			72.68		
117.43	41.89	2.05	140.01	7.03	4.05	31.02			82.44		
116.28	56.23	94.35	59.25	13.81	31.01	43.05					
8.93	41.56	43.41	58.68	47.34		41.78					
1.88	35.94	3.97	132.29	102.72		13.13					
.56	5.43	54.97	72.62	71.27		5.78					
5.62	3.00	395.96	.91	58.35		92.45					
50.91	73.94	89.55	3.57	43.16		4.27					
5.67	27.18	48.74	7.28	29.78		133.26					
.40	30.06	.54	.02	70.23		52.70					
.36	48.50	26.55	8.80	3.08		24.64					
1.21	.92	32.28	4.31	73.72		41.56					
.26	16.55	49.16	14.70	71.32		126.82					
.91	7.89	131.60	25.14	96.05		15.62					
22.83	24.62	150.93	38.35	135.20		8.49					
6.04	7.05	210.46	60.81	103.12		74.79					
1.54	1.29	63.28	39.84	34.89		33.23					
2.31	36.18	264.32	4.46	69.54		113.08					
39.80	139.48	167.93	127.55	9.23		95.85					
19.57	157.43	23.89	935.92	12.54		46.57					
85.13	2.82	33.45	63.20	44.17		94.73					
259.64	.67	327.31	62.12	135.81		1.08					
100.97	30.55	19.69	41.01	84.90		23.78					
41.28	84.43	159.31		94.30		37.84					
						59.78					

Elaboración propia – Datos IGP

Anexo 3 Tablas de resultados

Tablas de resultados sobre la aplicación del modelo semi-Márkov a la magnitud y profundidad de los sismos entre Lima y Barranca:

Probabilidades de ocurrencia de sismos, en las respectivas transiciones de los intervalos de la magnitud de los sismos, en el tiempo de 1 mes.

Tabla C1. F1 Matriz de probabilidad para t=1 mes

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.4257	0.4182	0.3102
De 5.4 a 5.7	0.4498	0.5608	0.7069
De 5.8 a más	0.4477	0.1309	0.1999

Elaboración propia

Tabla C2. F3 Matriz de probabilidad para t=6 meses

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.8610	0.9170	0.8797
De 5.4 a 5.7	0.9230	0.8553	0.9998
De 5.8 a más	0.9094	0.9177	1.0000

Elaboración propia

Las matrices Q son definidas como $Q(t) = F(t) \times P$ (producto de “elemento por elemento”).

Tabla C3. Q3 Matriz de probabilidad interválica para t=6 meses

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.6241	0.1504	0.0977
De 5.4 a 5.7	0.5618	0.2231	0.1304
De 5.8 a más	0.6747	0.1184	0.1290

Elaboración propia

Deducimos entonces las matrices H de la siguiente manera (1):

$$H_{ij}(t) = \begin{cases} 0 & \text{si } i \neq j \\ \sum_{k=1}^m Q_{ik}(t) & \text{si } i = j \end{cases}$$

Estas matrices representan la probabilidad de salir del estado i en el tiempo t.

Tabla C4. H1, Matriz probabilidad de ocurrencia, t=1 mes

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.4116	0	0
De 5.4 a 5.7	0	0.5123	0
De 5.8 a más	0	0	0.3748

Elaboración propia

Tabla C5. H3, Matriz probabilidad de ocurrencia, t=6 meses

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.8723	0	0
De 5.4 a 5.7	0	0.9153	0
De 5.8 a más	0	0	0.9221

Elaboración propia

Finalmente, se define las matrices D como $D(t) = I - H(t)$, las cuales corresponden a las probabilidades de permanecer en los estados respectivos en los tiempos de 1, 3 y 6 meses.

Tabla C6. D1, Matriz de probabilidad de sobrevivencia, t=1 mes

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.5884	0	0
De 5.4 a 5.7	0	0.4877	0
De 5.8 a más	0	0	0.6252

Elaboración propia

Tabla C7. D3, Matriz de probabilidad de sobrevivencia, t=6 meses

Estados	De 5.0 a 5.3	De 5.4 a 5.7	De 5.8 a más
De 5.0 a 5.3	0.1277	0	0
De 5.4 a 5.7	0	0.0847	0
De 5.8 a más	0	0	0.0779

Elaboración propia

Sobre la aplicación del modelo semi-Márkov a la profundidad de los sismos

En esta parte se presentan los resultados del ajuste del modelo semi-Márkov a los datos de la profundidad de los sismos similar al caso de las zonas geográficas y la magnitud de los sismos.

Se deja en evidencia que $F(0) = 0$.

En este caso, $t = 1, 3, 6$ meses, $k = 3$.

Tabla C8. F1, Matriz de probabilidad para $t=1$ mes

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.4987	0.3095	0.4545
De 60 a 100Km	0.3483	0.5232	0.0746
Más de 100Km	0.1310	0.2126	0.0990

Elaboración propia

Tabla C9. F3, Matriz de probabilidad para $t=6$ meses

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.8945	0.9470	0.8789
De 60 a 100Km	0.8829	0.9310	1.0000
Más de 100Km	0.7002	0.9817	0.5427

Elaboración propia

Las matrices Q son definidas como $Q(t) = F(t) \times P$ (producto de elemento por elemento).

Tabla C10. Q1, Matriz de probabilidad interválica para $t=1$ mes

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.3676	0.0574	0.0351
De 60 a 100Km	0.2527	0.1231	0.0029
Más de 100Km	0.0873	0.0304	0.0188

Elaboración propia

Tabla C11. Q3, Matriz de probabilidad interválica para $t=6$ mes

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.6594	0.1757	0.0680
De 60 a 100Km	0.6406	0.2191	0.0392
Más de 100Km	0.4668	0.1402	0.1034

Elaboración propia

Deducimos entonces las matrices H de la misma manera como en el caso de las zonas geográficas y las magnitudes.

Estas matrices representan la probabilidad de salir del estado i en el tiempo t .

Tabla C12. H1, Matriz de probabilidad de ocurrencia, $t=1$ mes

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.4602	0	0
De 60 a 100Km	0	0.3787	0
Más de 100Km	0	0	0.1365

Elaboración propia

Tabla C13. H3, Matriz de probabilidad de ocurrencia, $t=6$ meses

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.9031	0	0
De 60 a 100Km	0	0.8988	0
Más de 100Km	0	0	0.7104

Elaboración propia

Finalmente, se define las matrices D como $D(t) = I - H(t)$, las cuales corresponden a las probabilidades de permanecer en los estados respectivos en los tiempos de 1, 2 y 3 meses.

Tabla C14. D1, Matriz de probabilidad de sobrevivencia, $t=1$ mes

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.5398	0	0
De 60 a 100Km	0	0.6213	0
Más de 100Km	0	0	0.8635

Elaboración propia

Tabla C15. D3, Matriz de probabilidad de sobrevivencia, $t=6$ meses

Estados	Hasta 60Km	De 60 a 100Km	Más de 100Km
Hasta 60Km	0.0969	0	0
De 60 a 100Km	0	0.1012	0
Más de 100Km	0	0	0.2896

Elaboración propia

Anexo 4 Códigos en el programa R

Códigos en el programa R para el procesamiento de los datos de los sismos.

```
library(readxl)
> resuns51 <- read_excel("~/Directorio de trabajo/archivo.xlsx",
+   sheet = "Hoja4", range = "A1:D5")
> View(resuns51)
> T<- as.matrix(resuns51)
> T
      Huacho   Lima   Chilca Sierra de Lima
[1,] 0.3625000 0.2250000 0.3375000   0.0750000
[2,] 0.2592593 0.1851852 0.4074074   0.1481481
[3,] 0.2783505 0.2061856 0.3505155   0.1649485
[4,] 0.2571429 0.1714286 0.4285714   0.1428571
> library(readxl)
> resuns51 <- read_excel("~/Directorio de trabajo/archivo.xlsx",
+   sheet = "Hoja1", range = "A1:D5")
> View(resuns51)
> F1<- as.matrix(resuns51)
> F1
      Huacho   Lima   Chilca Sierra de Lima
[1,] 0.5089548 0.33492830 0.3638232   0.2680286
[2,] 0.3080156 0.16732020 0.5032858   0.4675855
[3,] 0.4037158 0.55053930 0.6681936   0.3676598
[4,] 0.1303655 0.01778773 0.2072339   0.4047849
> library(readxl)
> resuns51 <- read_excel("~/Directorio de trabajo/archivo.xlsx",
+   sheet = "Hoja2", range = "A1:D5")
> View(resuns51)
> F2<- as.matrix(resuns51)
> F2
      Huacho   Lima   Chilca Sierra de Lima
[1,] 0.7545869 0.7719657 0.6443051   0.6668104
[2,] 0.6563344 0.5802017 0.7326577   0.7285792
[3,] 0.7674817 0.8376902 0.8885977   0.6758241
[4,] 0.8216371 0.2345892 0.4875511   0.6262165
> library(readxl)
> resuns51 <- read_excel("~/Directorio de trabajo/archivo.xlsx",
+   sheet = "Hoja3", range = "A1:D5")
> View(resuns51)
> F3<- as.matrix(resuns51)
> F3
      Huacho   Lima   Chilca Sierra de Lima
[1,] 0.8844895 0.9642472 0.8247112   0.9121733
[2,] 0.8764827 0.9014160 0.8602897   0.8729362
[3,] 0.9396000 0.9527807 0.9661903   0.8628479
[4,] 0.9997787 0.7699708 0.7282277   0.7709399
> Q1<- F1*T
> Q1
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.18449612 0.075358868 0.12279033 0.02010214
[2,] 0.07985590 0.030985222 0.20504236 0.06927193
[3,] 0.11237450 0.113513258 0.23421219 0.06064492
[4,] 0.03352256 0.003049325 0.08881453 0.05782641

```

```
> Q2<- F2*T
```

```
> Q2
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.2735378 0.17369228 0.2174530 0.05001078
[2,] 0.1701608 0.10744476 0.2984902 0.10793766
[3,] 0.2136289 0.17271963 0.3114672 0.11147614
[4,] 0.2112781 0.04021529 0.2089505 0.08945950

```

```
> Q3<- F3*T
```

```
> Q3
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.3206274 0.2169556 0.2783400 0.0684130
[2,] 0.2272363 0.1669289 0.3504884 0.1293239
[3,] 0.2615381 0.1964496 0.3386646 0.1423254
[4,] 0.2570860 0.1319950 0.3120976 0.1101343

```

```
> #para sumar las filas de una matriz
```

```
> suma_filas1 <- rowSums(Q1)
```

```
> suma_filas1
```

```
[1] 0.4027475 0.3851554 0.5207449 0.1832128
```

```
> #para poner las sumás en la diagonal de la matriz
```

```
> H1<- diag(suma_filas1)
```

```
> H1
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.4027475 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.3851554 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.5207449 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.1832128

```

```
> #para los demás Qs
```

```
>
```

```
> suma_filas2 <- rowSums(Q2)
```

```
> suma_filas2
```

```
[1] 0.7146938 0.6840334 0.8092919 0.5499034
```

```
> H2<- diag(suma_filas2)
```

```
>
```

```
>
```

```
> suma_filas3 <- rowSums(Q3)
```

```
> suma_filas3
```

```
[1] 0.8843361 0.8739774 0.9389778 0.8113128
```

```
> H3<- diag(suma_filas3)
```

```
> H3
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.8843361 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.8739774 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.9389778 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.8113128

```

```
> H2
```



```

Huacho Lima Chilca Sierra de Lima
[1,] 0.7146938 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.6840334 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.8092919 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.5499034

```

Para el analisis de sobrevivencia

```
I4<- diag(4)
```

```
> I4
```

```

[,1] [,2] [,3] [,4]
[1,] 1 0 0 0
[2,] 0 1 0 0
[3,] 0 0 1 0
[4,] 0 0 0 1

```

```
> D1<- I4-H1
```

```
> D1
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.5972525 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.6148446 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.4792551 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.8167872

```

```
> D2<- I4-H2
```

```
> D2
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.2853062 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.3159666 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.1907081 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.4500966

```

```
> D3<- I4-H3
```

```
> D3
```

```

Huacho Lima Chilca Sierra de Lima
[1,] 0.1156639 0.0000000 0.0000000 0.0000000
[2,] 0.0000000 0.1260226 0.0000000 0.0000000
[3,] 0.0000000 0.0000000 0.06102216 0.0000000
[4,] 0.0000000 0.0000000 0.0000000 0.1886872

```

Las probabilidades de intervalo

B1=Q1

	Huacho	Lima	Chilca	Sierra de Lima
[1,]	0.18449612	0.075358868	0.12279033	0.02010214
[2,]	0.07985590	0.030985222	0.20504236	0.06927193
[3,]	0.11237450	0.113513258	0.23421219	0.06064492
[4,]	0.03352256	0.003049325	0.08881453	0.05782641

B2<- Q2-Q1

> B2

	Huacho	Lima	Chilca	Sierra de Lima
[1,]	0.08904164	0.09833341	0.09466264	0.02990863
[2,]	0.09030487	0.07645954	0.09344781	0.03866573
[3,]	0.10125443	0.05920637	0.07725505	0.05083122
[4,]	0.17775555	0.03716596	0.12013594	0.03163309

> B3<- Q3-Q2

> B3

	Huacho	Lima	Chilca	Sierra de Lima
[1,]	0.04708969	0.04326334	0.06088706	0.01840222
[2,]	0.05707549	0.05948413	0.05199822	0.02138622
[3,]	0.04790922	0.02373000	0.02719741	0.03084929
[4,]	0.04580784	0.09177971	0.10314711	0.02067477

Anexo 5. Latitud y longitud de los límites de las zonas geográficas en estudio

