

CUANTIFICACIÓN DE INCERTIDUMBRE BAYESIANA APROXIMADA EN PROBLEMAS INVERSOS DE ODE

T E S I S

Que para obtener el grado de
Maestro en Probabilidad y Estadística

Presenta

César Isaí García Cornejo

Director de Tesis:

Dr. José Andrés Christen Gracia

Autorización de la versión final

Dedicatoria . . .

Resumen

El estudio del fenómeno de la violencia en México que acontece durante la conocida “Guerra contra las drogas” ha sido abordado desde diferentes perspectivas y mediante enfoques tanto cualitativos como cuantitativos. En referencia al último enfoque, aunque ya existen trabajos que analizan este problema a través de modelos estadísticos, aún existen aspectos por cubrir.

A partir de un análisis exploratorio de los datos seleccionamos el conjunto de covariables más relevantes y delimitamos una región para la modelación. Ya que los procesos de este tipo inducen correlación espacial, proponemos una serie de parametrizaciones de un modelo espacio-temporal jerárquico bayesiano con efectos de espacio y tiempo separables.

Palabras clave: modelación espacio-temporal, modelo jerárquico bayesiano, ciencia de datos, cómputo estadístico, inla, datos de áreas

Agradecimientos

A mis padres ...

Índice general

Resumen	III
Agradecimientos	V
Índice de figuras	IX
1. Introducción	1
1.1. Proceso de modelación	1
2. Antecedentes	3
2.1. Problema inverso	3
2.1.1. Modelos descritos por ODEs	4
2.2. Solución bayesiana a problemas inversos	5
2.3. Forward map aproximado y consistencia	5
2.4. MCMC	5
2.4.1. Metropolis-Hastings	5
2.5. Necesidad de trabajar con aproximaciones del Forward map	5
3. Implementación	7
3.1. Modelos de trabajo	7
3.1.1. Modelo Gravitatorio	8
3.1.2. Modelo Logístico	11
3.1.3. Modelo SIR	13
3.2. Enfoque bayesiano al problema inverso	14
3.2.1. Distribución posterior	15
3.2.2. Simulación con MCMC	18
3.2.3. Simulación del modelo SIR	32
3.3. Inferencia con aproximación.	35
3.3.1. Regla de interpolación(aproximación de forward map)	35
3.3.2. Simulación MCMC	36
3.3.3. Experimentación con la convergencia.	37
4. Conclusiones y trabajo futuro	49
Referencias	51

Índice de figuras

3.1. Varias trayectorias con la dinámica gravitación sujeta a fricción para varios valores de los parámetros g, b	11
3.2. Varias trayectorias con la dinámica de crecimiento logístico para varios valores de los parámetros K, r	13
3.3. Trayectoria de cada grupo del modelo SIR con $\beta = 0.009$ y $\gamma = 0.5$	14
3.4. Distribuciones a priori para $\theta_1 = g$ y $\theta_2 = b$	20
3.5. Muestra \mathbf{y} del modelo gravitatorio.	20
3.6. Cadena obtenida por MCMC para la distribución posterior conjunta.	21
3.7. Distribución posterior para el parámetros g	21
3.8. Distribución posterior para el parámetro b	22
3.9. Distribución predictiva para el modelo gravitatorio con σ conocido.	22
3.10. Distribuciones a priori para $\theta_1 = g$ y $\theta_2 = b$	23
3.11. Muestra \mathbf{y} del modelo gravitatorio.	24
3.12. Cadena obtenida por MCMC para la distribución posterior conjunta.	24
3.13. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	36
3.14. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	37
3.15. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	38
3.16. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	39
3.17. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	40
3.18. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	41
3.19. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	42
3.20. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	43
3.21. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	44
3.22. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	45

3.23. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	46
3.24. Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.	47

Capítulo 1

Introducción

En una amplia gama de disciplinas se utilizan modelos que pretenden describir algún fenómeno. La inmensa mayoría de estos son modelos matemáticos, teniendo relaciones entre variables expresadas por ecuaciones dependientes de ciertos parámetros. El tipo de modelos considerados son aquellos cuyas variables se pueden expresar ya sea en una forma funcional o en forma de ecuación diferencial ordinaria (ODE).

1.1. Proceso de modelación

El proceso de modelación para cierto fenómeno involucra una estructura cohesionada en tres principales sectores. Al investigar un fenómeno se plantea, propone y prueban las variables de las cuales el fenómeno responde a cambios. Dicho de otra forma, se seleccionan las variables que influyen en la descripción del modelo. Dicho sector se conoce como **parametrización**.

Posteriormente, una vez postulado las variables del modelo, sigue estructurar ya se explícita o implícitamente las relaciones entre las mencionadas variables. Se puede proponer modelos lineales generalizados, series temporales, ecuaciones diferenciales o simplemente una ecuación algebraica con las variables involucradas y además parámetros de ajuste que en adelante conoceremos solamente como parámetros. Este sector de la modelación es conocido como **modelación directa**.

El último gran sector de la modelación es **el problema inverso**. Como su nombre

sugiere, se refiere al problema de abordar cierta inferencia sobre los parámetros del modelo dada una muestra propia del fenómeno modelado. Que es lo contrario al problema directo, el cual dado unos parámetros fijos busca encontrar explícitamente la dinámica del fenómeno.

El proceso de modelación no sigue una línea recta. Es por ello que se optó por llamar a sus partes como sectores de modelación en lugar de etapas de modelación, ya que todos estos sectores siguen un camino cíclico según las virtudes de cada modelo en particular.

Una vez aclarado que los modelos considerados son aquellos que pueden expresarse como una ODE dependientes de parámetros, la tesis se centra en el problema inverso desde el enfoque bayesiano. En el capítulo 2 se describe el paradigma bayesiano aplicado al problema inverso. Además de los antecedentes necesarios para hacer de este texto autocontenido.

En el capítulo 3 se describe las implementaciones realizadas para la inferencia bayesiana de los parámetros de tres modelos particulares según el paradigma convencional. Además, se propone una forma aproximada para realizar el mismo procedimiento, donde se cuestiona la factibilidad de la propuesta como sustituto del paradigma original.

Capítulo 2

Antecedentes

En el capítulo 1 se habla a grandes rasgos de la estructura implícita de la modelación. Un sector importante es el del problema inverso. Aquí formalizaremos el concepto y detallaremos sus lineamientos. Se describe el enfoque bayesiano del problema inverso, que es un paradigma ampliamente estudiado. Otro enfoque al problema inverso es el operativo, el cual se aborde desde el problema directo como un operador actuando en el modelo, y el problema inverso como un operador inverso de este. Este último enfoque sale de los límites de estudio del texto.

2.1. Problema inverso

El estudio del problema inverso se remonta a mediados del siglo 17. Los físicos contemporáneos interesados en la relación causa-efecto de diferentes fenómenos físicos lograron predecir el comportamiento en la dinámica subyacente. Así, para la ley de gravitacional de newton, una vez obtenidas las masas de los cuerpos celestes (causas) es posible predecir las trayectorias que siguen en el tiempo (efecto). Estudiar los efectos de un fenómeno dada ciertas causas es el problema directo, se requiere la metodología descrita en la introducción. El problema inverso toma la dirección opuesta, dados los efectos, es de interés investigar las causas que lo ocasionaron. Retomando un ejemplo newtoniano, dado un campo gravitatorio circundando cierto cuerpo celeste, es de interés conocer la masa de dicho cuerpo. El problema inverso padece de no ser único,

suele ser el caso que a ciertos efectos se correspondan varias causas. el hecho de que no exista una mapeo uno a uno entre causa y efecto complica el problema inverso. Sin embargo, la información que se tenga a priori del fenómeno es crucial para determinar con una menor ambigüedad las causas.

2.1.1. Modelos descritos por ODEs

Los modelos matemáticos para describir fenómenos pueden venir de una amplia variedad de modelos. Tomese que se desea describir los costos de la renta de inmuebles en cierta ciudad Y (efecto). Supongase que de análisis previos se concluyo que dicho fenómeno esta vinculado al tamaño del inmueble X_1 , número de habitaciones X_2 , ingreso medio por habitante X_3 (causas). Un modelo plausible es por modelos lineales generalizados

$$Y = G(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3) + \varepsilon$$

donde G es la función liga, $\varepsilon \sim N(0, \sigma^2)$ un error aleatorio y $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ son los parámetros del modelo. El problema directo considera predecir el valor de Y conociendo $X = (X_1, X_2, X_3)$ y los parámetros del modelo β . Para esta clase de modelos el problema directo no representa complejidad. En contraste, el problema inverso precisa estimaciones de β y σ dada muestras de Y y X , siendo un problema no trivial.

2.2. Solución bayesiana a problemas inversos

Tratar a fondo sobre la teoría de problemas inversos. Cómo se define. Que es el Forward map. Métodos convencionales (superficialmente para resolución)

Método para problemas inversos (lineales)

Método por estadística bayesiana

Explicar como abordar el problema inverso con estadística bayesiana

2.3. Forward map aproximado y consistencia

Teoría

Aquí explicar la teoría clásica para la estadística bayesiana.

Ejemplos

2.4. MCMC

Introducción a la simulación por Monte Carlo.

2.4.1. Metropolis-Hastings

:

Explicar teoría de MCMC, teoremas de cadenas de Markov(apendice)

2.5. Necesidad de trabajar con aproximaciones del Forward map

Esbozar el algoritmo y explicar la practicidad de este. Poner un ejemplo ilustrativo.

Capítulo 3

Implementación

Debido a que nos encontramos en un terreno carente de fundamentos teóricos, nuestra mejor arma para probar que grado de aproximación de los problemas inversos son consistentes es por medio de experimentación metodológica. Esto nos permite revelar obstáculos del nuevo método así como también las grandes ventajas.

Gran parte del trabajo de tesis se centra primordialmente en la experimentación con la propuesta innovadora para la solución bayesiana del problema inverso. Se hacen los experimentos con una implementación en Python cuyo código se encuentra adjunto.

3.1. Modelos de trabajo

Existen una variedad de modelos a los que se puede aplicar la metodología descrita para problemas inversos. De hecho se puede implementar para cualquier modelo descrito por una ecuación diferencial ordinaria. Sin embargo, cada modelo tiene sus bemoles, lo que dificulta calibrar (en general), para cada aspecto en la distribución de los parámetros, como puede ser el caso de tomar las distribuciones a priori adecuadas, la variación de la muestra, los dominios adecuados para graficar las distribuciones a posterior, entre otras. Por ello, nos permitiremos enfocarnos en solo un par de modelos que son el modelo gravitatorio sujeto a fricción para un partícula en caída y el modelo de crecimiento logístico para una población en crecimiento.

3.1.1. Modelo Gravitatorio

Consideremos una partícula puntual en un campo gravitatorio cercano a la superficie terrestre. Por medio de la dinámica clásica podemos modelar dicha caída con las leyes de Newton. Sea $x(t)$ la distancia recorrida (unidimensional) por la partícula puntual en el tiempo t . De la segunda ley de Newton sabemos que

$$\sum F_i = m\ddot{x}(t) \quad (3.1)$$

donde $\sum F_i$ es la suma de fuerzas ejercida sobre la partícula con masa m .

Para el modelo de caída libre nos dice que la fuerza ejercida en la partícula es constante y se le conoce como constante de aceleración gravitacional g . De forma que la trayectoria se rige de (3.1) con únicamente la fuerza gravitacional $F = mg$. La ecuación de la dinámica es

$$\ddot{x}(t) = g$$

bajo las condiciones iniciales $x(0) = x_0, \dot{x}(0) = v_0$. La solución a la ecuación dinámica es [Sears y cols. \(1986\)](#)

$$x(t) = \frac{1}{2}gt^2 + v_0t + x_0$$

Sin embargo, dentro del marco establecido para el modelo gravitacional, no es un modelo realista ya que este no está considerando la desaceleración producto de la fricción con el medio que interactúa (aire). Por ello es necesario modelar la fuerza de fricción de forma adecuada. Un modelo físico plausible para la fricción es considerar una fuerza opuesta a la fuerza de gravedad y que está depende de la velocidad de la partícula. Consideremos dicha fuerza de fricción como $F_f = -bv(x)$. Podemos pensar que el modelo para la fuerza de fricción es una aproximación de los primeros polinomios de Taylor a primer grado, donde se descarta el termino constante ya que no tiene sentido físico que la fricción obedezca a marcos de referencia.

El modelo que estamos interesados es entonces el **modelo gravitatorio sujeto**

a **fricción**, cuya ecuación dinámica es

$$m\ddot{x} = mg - b\dot{x} \quad (3.2)$$

que representa la trayectoria de la partícula sujeta a dos fuerzas opuestas, la gravitatoria y la fricción con el medio. Además con las condiciones iniciales $x(0) = x_0, \dot{x}(0) = v_0$. Notemos que dados los parámetros $\theta = (g, b)$ podemos determinar unívocamente la trayectoria de interés.

Para poder hacer inferencia bayesiana en el problema inverso, es necesario construir el forward map $F(\theta)$. Esto puede ser de dos maneras, podemos resolver la ecuación diferencial y obtener $x(t)$ en términos de $\theta = (g, b)$ en caso de que exista dicha solución explícita. Otra forma de construir el forward map es utilizando métodos numéricos para resolver la EDO.

En este modelo sí tenemos solución explícita a la ecuación de la dinámica. Para poder resolverla expresamos la EDO en términos de la velocidad $v(t) = \dot{x}(t)$. Teniendo una EDO lineal de primer orden.

La solución a la ecuación dinámica

$$m \frac{dv}{dt} = mg - bv \quad (3.3)$$

se obtiene del siguiente análisis. Primeramente notemos que la aceleración de la partícula se anula. Es decir, la velocidad tiene una asíntota. Esto corresponde al caso cuando la fuerza gravitacional es igual y opuesta a la fuerza de fricción. Por tanto podemos definir una velocidad terminal v_T de la relación

$$mg - bv_T = 0 \quad \Rightarrow \quad v_T = \frac{mg}{b} \quad (3.4)$$

de reordenar e integrar (3.3)

$$\frac{dv}{dt} = -\frac{b}{m}(v - v_T) \quad (3.5)$$

e integrando con la condición inicial $v_0 = 0$

$$\int_0^v \frac{dv}{v - v_T} = -\frac{b}{m} \int_0^t dt$$

obtenemos

$$\log \left(\frac{v_t - v}{v_t} \right) = -\frac{b}{m} t$$

despejando para v

$$v = v_T \left(1 - \exp \left\{ -\frac{b}{m} t \right\} \right) \quad (3.6)$$

donde vemos que en efecto la velocidad terminal v_T es una asíntota debido a que $v(t)$ ya que se aproxima a v_T a medida que pasa el tiempo.

Para obtener la función de la trayectoria, simplemente integramos una vez más obteniendo

$$x(t) = \int_0^t v_T \left(1 - \exp \left\{ -\frac{b}{m} t \right\} \right) dt + x_0$$

tomando $x_0 = 0$, finalmente nos queda que la trayectoria de la partícula es

$$x(t) = v_T \left[t - \frac{m}{b} \left(1 - \exp \left\{ -\frac{b}{m} t \right\} \right) \right] \quad (3.7)$$

Por tanto, podemos tomar el forward map $F(g, b) = x(t)$, prescindiendo de m al considerar $m = 1$.

Por otro lado, recordemos que la idea principal es resolver el problema inverso. Es decir, dada una muestra de la trayectoria $x(t_i)$ como podemos inferir los parámetros $\theta = (g, b)$. La ciencia actual ya es conocida con buen grado de exactitud la aceleración de la gravedad g . Sin embargo, vamos a considerar que desconocemos su valor, lo que puede aplicar para constantes en alturas distintas a la superficie terrestre o constantes gravitacionales en otro planeta. Con fines ilustrativos esbozamos en la Fig 3.1 varias trayectorias según sus parámetros (g, b) y notamos en todas un carácter creciente en

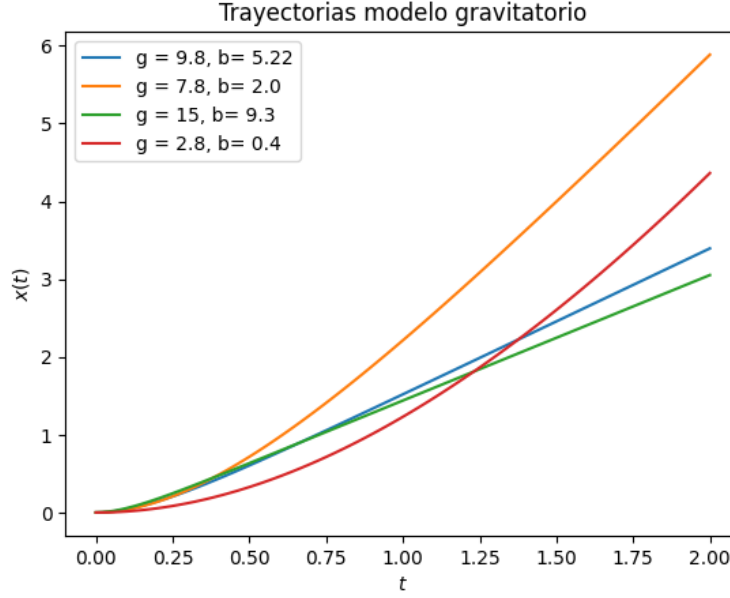


Figura 3.1: Varias trayectorias con la dinámica gravitación sujeta a fricción para varias valores de los parámetros g, b .

el tiempo.

3.1.2. Modelo Logístico

Un modelo de crecimiento poblacional simple caracterizado por una ecuación diferencial es aquel que la tasa de cambio de la población es proporcional al tamaño actual de la población. Sea $P(t)$ el tamaño de la población al tiempo t . El modelo descrito es

$$\frac{dP}{dt} = \lambda P(t) \quad (3.8)$$

donde λ es la tasa de crecimiento y con la condición $P(0) = P_0$.

La solución explícita al modelo en (3.8) es

$$P(t) = P_0 e^{\lambda t} \quad (3.9)$$

que es una función creciente en el tiempo. Sin embargo, este modelo solo es valioso para pequeños intervalos de tiempo, ya que el crecimiento en una la tasa de cambio

de la población tiende a decaer a medida que se consumen los recursos.

Para saldar con la dificultad planteada, podemos modelar la tasa de crecimiento de una población considerando tanto al tamaño de la población como a los recursos aún presentes. Consideremos a K como la capacidad de sustento, que se puede interpretar como la cantidad máxima de población que los actuales recursos permiten mantener, entonces esperamos que a medida que crece la población, la tasa de crecimiento de la misma decrezca.

El modelo que se propone es conocido como el modelo de crecimiento logístico. Este nos dice que el cambio de la población es proporcional al tamaño de la población misma así como a la proporción de recursos disponibles. Dicho modelo se puede escribir con la ecuación de Verhulst

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right) \quad (3.10)$$

donde r es la tasa de crecimiento y K la capacidad de sustento.

A pesar de que la ecuación diferencial del modelo de crecimiento logístico es no lineal tiene solución explícita. Notemos que la EDO pertenece a la familia de las EDO de Bernoulli, esto es tiene la forma

$$\frac{dy}{dt} + P(x)y = Q(x)y^n$$

y además puede transformarse a una EDO lineal de primer orden con $u(x) = y^{1-n}$.

Por tanto, del cambio $u(t) = P^{-1}$ obtenemos la solución explícita a (3.8)

$$P(t) = \frac{KP_0 e^{rt}}{K + P_0 (e^{rt} - 1)} \quad (3.11)$$

donde verificamos que $\lim_{t \rightarrow \infty} = K$.

Podemos ver en la Fig 3.2 que sus trayectorias son no decrecientes y tiene asintóticamente al parámetro K .

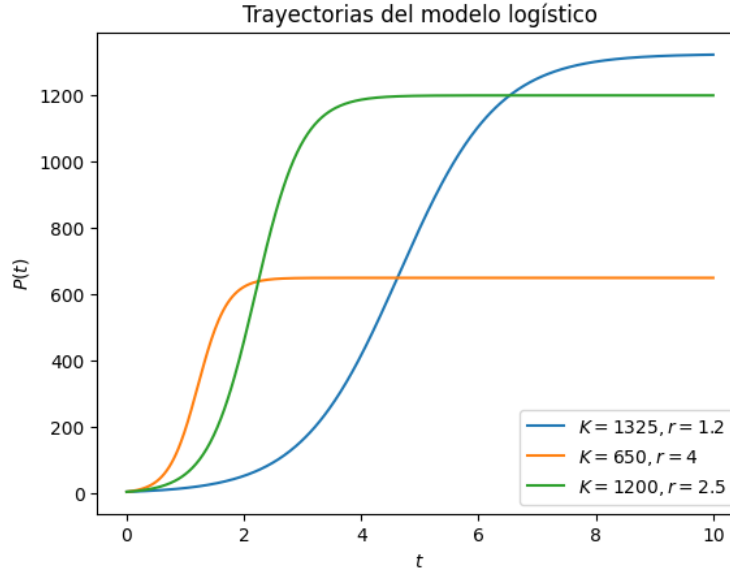


Figura 3.2: Varias trayectorias con la dinámica de crecimiento logístico para varios valores de los parámetros K, r .

3.1.3. Modelo SIR

Los modelos epidemiológicos clasifican a la población en clases que distinguen si han tenido o no cierta enfermedad. Un modelo que captura con buena aproximación a varias enfermedades se es el modelo SIR. Este propone dividir a las N personas de una población en tres grupos:

1. Susceptible (S): Todas las personas que no han tenido la enfermedad ni están enfermas del brote de interés.
2. Infectado (I): Son las personas que actualmente portan el patógeno y son además un vector de transmisión.
3. Recuperado (R): Para aquellas personas que se han recuperado de la enfermedad y ya tienen inmunidad o también para aquellas que perecieron.

Así, el modelo propuesto supone que la tasa de cambios de las personas susceptibles decrece proporcionalmente a la cantidad de infectados y la cantidad de susceptibles restantes. De igual forma, la cantidad de recuperados crece con una tasa proporcional a la cantidad de infectados. La dinámica del modelo SIR se da con el siguiente sistema

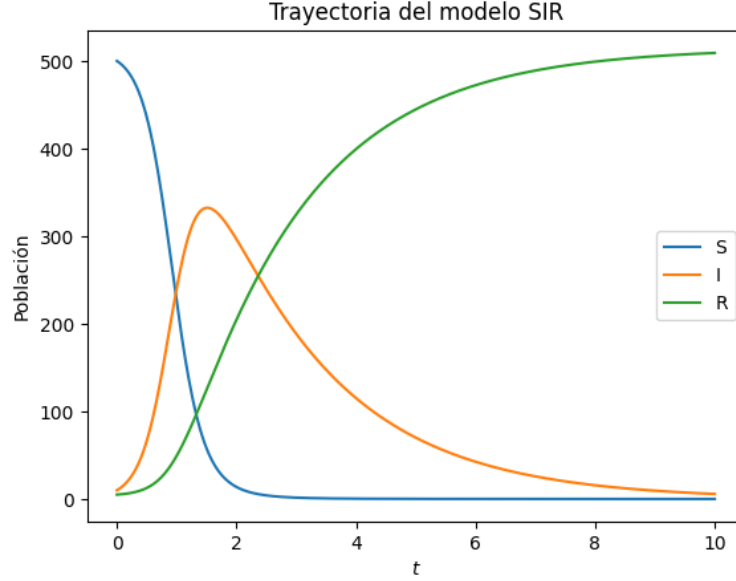


Figura 3.3: Trayectoria de cada grupo del modelo SIR con $\beta = 0.009$ y $\gamma = 0.5$.

de EDO's

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}\tag{3.12}$$

donde β es la tasa de infección, γ es la tasa de recuperación y $S + I + R = N$.

La solución explícita para $S(t)$, $I(t)$, $R(t)$ no existe en general, por ello se requieren métodos numéricos para resolver numéricamente para cada grupo y poder construir el forward map. En la Fig. 3.3 tenemos un ejemplo para un caso del modelo SIR con tasa de infección $\beta = 0.009$ y tasa de recuperación $\gamma = 0.5$

3.2. Enfoque bayesiano al problema inverso

Una vez establecidos los modelos sobre los que trabajaremos, nos enfocamos en abordar el problema inverso para los mismos. Para ello, es necesario tener a disposición una muestra de la trayectoria del modelo al cual deseamos inferir sus parámetros.

Según sea el caso, podemos tomar mediciones de caída de un objeto para ciertos tiempo, o podemos contar la población para cada intervalo de tiempo, observar los grupos de personas en un brote de cierto patógeno. Denotemos estas observaciones por y_1, y_2, \dots, y_n .

Sin embargo, no es fácil tener acceso a dicha muestra, por lo que optaremos por simular los datos directamente del modelo seleccionando los parámetros convenientemente y además agregar un ruido gaussiano. Es decir, tomamos $\theta \in \mathbb{R}^d$ fijo, luego consideramos los tiempos a los cuales corresponderá la muestra, denotados por t_1, t_2, \dots, t_n . Posteriormente, con el forward map F_θ generado según el modelo (ya sea analítico o numérico) evaluamos para cada tiempo (denotado por $F_\theta(t_i)$) y agregamos un ruido $\varepsilon_i \sim N(0, \sigma^2)$. Así la muestra considerada se constituye por

$$y_i = F_\theta(t_i) + \varepsilon_i \quad (3.13)$$

para $i = 1, 2, \dots, n$. Dicho de otra forma la muestra contemplada

$$\{(y_1, t_1), (y_2, t_2), \dots, (y_n, t_n)\} \quad (3.14)$$

un conjunto de n tuplas.

Una vez establecida la muestra y_i , ahora nuestro propósito es abordar el problema inverso, esto es buscamos θ tal que $y_i \approx F_\theta(t_i)$. Del capítulo 2 sabemos que hay varias maneras de abordar el problema, en adelante solamente se usará el enfoque bayesiano del problema inverso.

3.2.1. Distribución posterior

En la teoría de la estadística bayesiana, el paradigma central se basa en obtener la distribución posterior de los parámetros de interés dada ciertas observaciones a las que denotamos por $\mathbf{y} = (y_1, y_2, \dots, y_n)$. Luego, sabemos que salvo el caso de análisis conjugado podemos acceder directamente a dicha distribución posterior. En el caso del problema inverso no es posible determinar la distribución por medio de análisis

conjugado en caso general, debido a que el forward map somete a la distribución posterior a una familia diferente a la familia de la distribución a priori.

Para el vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_d) \in \mathbb{R}^d$, recordemos la ecuación (ecu)

$$\pi(\theta|\mathbf{y}) \propto f(\mathbf{y}|\theta)\pi(\theta) \quad (3.15)$$

que nos da la distribución posterior salvo una constante en términos de la verosimilitud y la distribución a priori. Para obtener la verosimilitud, observemos de la ecuación (3.13) que

$$\mathbf{y}_i \sim N(F_\theta(t_i), \sigma^2)$$

por lo que la verosimilitud para θ es la correspondiente al modelo normal, es decir

$$\mathcal{L}(\theta, \sigma) = f(\mathbf{y}|\theta) = \prod_{i=1}^n \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - F_\theta(t_i))^2 \right\}, \quad (3.16)$$

simplificando

$$\mathcal{L}(\theta, \sigma) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - F_\theta(t_i))^2 \right\}, \quad (3.17)$$

que es la verosimilitud en función de los parámetros θ y la desviación estándar σ .

Por otro lado, la distribución a priori representa el conocimiento previo a observar los datos. Para nuestro caso particular, los modelos propuestos en la sección (citar sección) dependen de parámetros no negativos, es decir, $\theta_i \geq 0$ para toda $i = 1, \dots, d$. Incluso algunos de ellos se tiene bastante certeza de su valor verdadero, que podría pensarse en ajustar una distribución a priori normal. Sin embargo, ajustar una normal conlleva a tener problemas en el soporte, por lo que el siguiente paso es considerar una normal truncada.

Finalmente, se optó por proponer distribuciones a priori gamma, debido a que existe una parametrización en la cual se asemeja a una distribución normal pero con

el soporte adecuado. La distribución gamma que tiene por máximo a θ_i^* y asemeja una normal es

$$\theta_i \sim \text{Gamma}\left(\alpha_i, \frac{\theta_i^*}{\alpha_i}\right) \quad (3.18)$$

con θ_i^* y α_i parámetros conocidos.

La parametrización de la $\text{Gamma}(\alpha, \beta)$ es tal que su función de densidad sea

$$f(\theta|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp\{-\beta\theta\}$$

con α, β conocidos.

Como θ_i se proponen independientes, entonces la distribución a priori para θ es

$$\begin{aligned} f(\theta|\alpha) &= f(\theta_1|\alpha_1) \cdots f(\theta_d|\alpha_d) \\ &= \prod_{i=1}^n \left[\frac{1}{\Gamma(\alpha_i)} \left(\frac{\theta_i^*}{\alpha_i} \right)^{\alpha_i} \theta_i^{\alpha_i-1} \exp\left\{-\frac{\theta_i^*}{\alpha_i} \theta_i\right\} \right] \end{aligned} \quad (3.19)$$

desarrollando

$$\begin{aligned} f(\theta|\alpha) &= \\ &= \frac{1}{\Gamma(\alpha_1)} \left(\frac{\theta_1^*}{\alpha_1} \right)^{\alpha_1} \theta_1^{\alpha_1-1} \exp\left\{-\frac{\theta_1^*}{\alpha_1} \theta_1\right\} \cdots \frac{1}{\Gamma(\alpha_d)} \left(\frac{\theta_d^*}{\alpha_d} \right)^{\alpha_d} \theta_d^{\alpha_d-1} \exp\left\{-\frac{\theta_d^*}{\alpha_d} \theta_d\right\} \end{aligned} \quad (3.20)$$

con $\alpha = (\alpha_1, \dots, \alpha_p)$ y $\theta^* = (\theta_1^*, \dots, \theta_d^*)$ conocidos.

Sustituyendo (3.17) y (3.19) en (3.15) se tiene la forma funcional de la distribución posterior

$$\pi(\theta, \sigma|\mathbf{y}) \propto \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - F_\theta(t_i))^2\right\} \prod_{i=1}^n \left[\frac{1}{\Gamma(\alpha_i)} \left(\frac{\theta_i^*}{\alpha_i} \right)^{\alpha_i} \theta_i^{\alpha_i-1} \exp\left\{-\frac{\theta_i^*}{\alpha_i} \theta_i\right\} \right] \quad (3.21)$$

Notemos que la distribución en (3.21) no pertenece a una familia conocida. Para lidiar con este problema usamos métodos Monte Carlo. Más precisamente, podemos

simular de la distribución posterior (3.21).

3.2.2. Simulación con MCMC

Al incursionar a la estadística bayesiana se aprende que se puede obtener la distribución posterior por medio del famoso análisis conjugado. Es decir se propone una distribución a priori convenientemente para que la distribución posterior pertenezca a la misma familia. Sin embargo, el análisis conjugado puede ser no trivial o incluso no existir. Otra forma es usando métodos numéricos para calcular la integral del denominador, sin embargo no es optimo para funciones de varias variables. Por ello es que simular de la distribución posterior con MCMC (Markov Chain Monte Carlo) es la vía optima para tener una muestra de dicha distribución, con dicha muestra podemos estimar cualquier momento de la distribución. Si estamos interesados en obtener la media de X , en beneficio de la ley de grandes números, podemos estimar como

$$\mathbb{E}[X] \approx \frac{1}{n} \sum_{i=1}^n X_i \quad (3.22)$$

De igual forma si estamos interesados en la media del estadístico $h(X)$ para un función h arbitraria, se estima por

$$\mathbb{E}[h(X)] \approx \frac{1}{n} \sum_{i=1}^n h(X_i) \quad (3.23)$$

donde X_i es la simulación de la distribución posterior.

La distribución objetivo que se implementa en el MCMC es la distribución posterior (3.21). Recordemos que la distribución posterior depende del forward map $F_\theta(t_i)$ obtenido de evaluar los tiempos de la muestra. Recordemos que $F_\theta(t_i)$ es la version discretizada de F_θ . Sea $\theta \in \mathbb{R}^d$ la función

$$\theta \mapsto F_\theta = x(t) \quad (3.24)$$

tal que $x(t)$ es la solución de la ecuación diferencial ordinaria $x(t) = G(t, x(t), x'(t), x''(t) \dots)$

del modelo bajo consideración.

Simulación del modelo gravitatorio

Para poder hacer simulaciones, es necesario determinar completamente la distribución posterior. El Forward map para el modelo gravitatorio es $F_\theta = F(g, b) = x(t)$ con la ecuación diferencial de la dinámica gravitacional dada en (3.2) o equivalentemente en (3.7). Además, es necesario saber como abordar el caso de la desviación estándar σ . Existen dos formas, pensar en σ como parámetro conocido, por lo que no es necesario hacer inferencia, el otro caso es para σ desconocida y veremos los detalles después de explicar el primer caso.

Para el caso σ **conocido**, la distribución posterior (3.21) tendrá como variables únicamente al vector de parámetros θ , al ser una EDO dependiente solamente de g y b se sigue que $\theta = (g, b)$ siendo un vector en \mathbb{R}^d con $d = 2$. Además es necesario definir las constantes para las distribuciones a priori gamma.

Consideremos las distribuciones a priori (3.18) como

$$g \sim \text{Gamma}\left(\alpha, \frac{\theta_1^*}{\alpha}\right) \quad (3.25)$$

$$b \sim \text{Gamma}\left(\beta, \frac{\theta_2^*}{\beta}\right) \quad (3.26)$$

con $\alpha = 10$, $\beta = 1.1$ con máximo en $\theta_1^* = 10$ y $\theta_2^* = 2$. Podemos observar las distribuciones a priori en la Fig. 3.4.

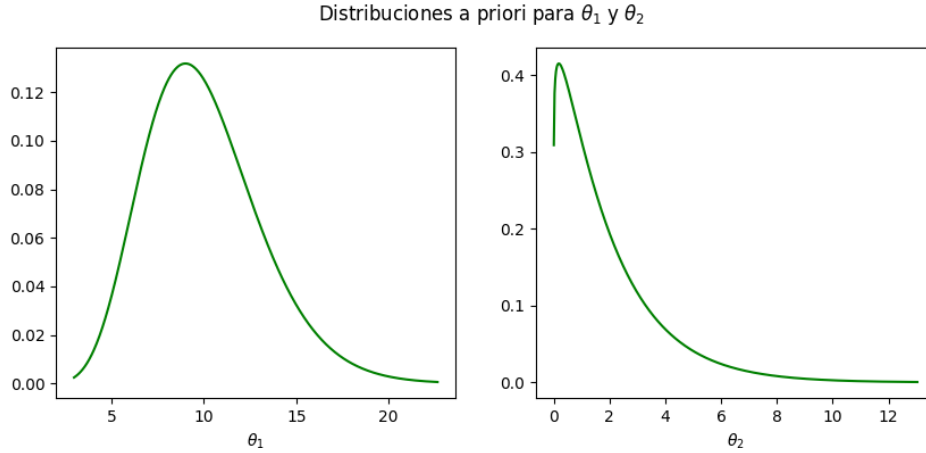


Figura 3.4: Distribuciones a priori para $\theta_1 = g$ y $\theta_2 = b$.

Ahora, generamos una muestra del forward map gravitatorio y agregamos un ruido gaussiano con $\sigma = 0.1$ obteniendo la muestra como en la Fig. 3.5.

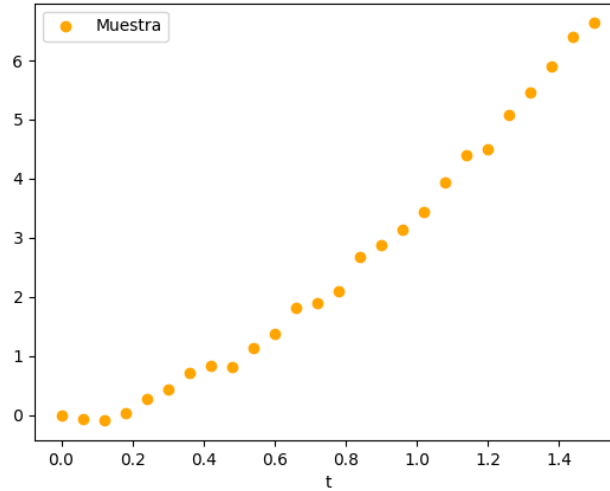


Figura 3.5: Muestra y del modelo gravitatorio.

Luego, con la metodología propuesta, usando el algoritmo MCMC Metropolis-Hastings para una cadena de tamaño $T = 500,000$ y un burn in de 20,000 tenemos que la trayectoria seguida en el espacio de parámetros $\theta = (g, b)$ para la distribución posterior es el mostrado en la Fig. 3.6. Notemos que dicha posterior tiene alta correlación entre parámetros. Para valores altos de g es necesario tener más fricción b con el medio para la misma trayectoria y .

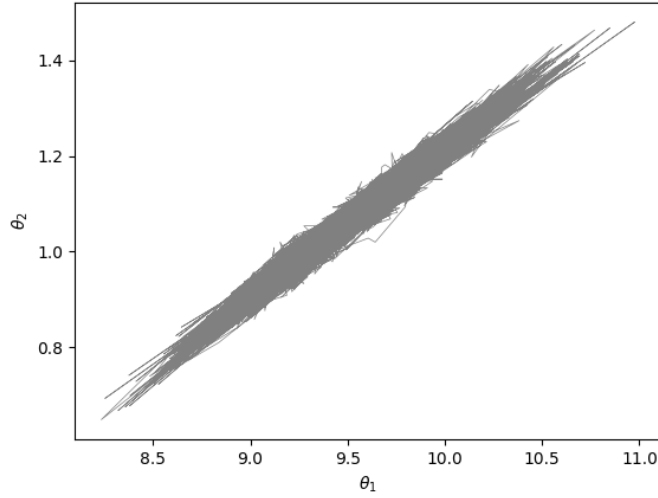


Figura 3.6: Cadena obtenida por MCMC para la distribución posterior conjunta.

Nos es de interés las marginales de la distribución posterior. Dado que la simulación de la cadena es bidimensional, de (3.23) tomamos $h_1(X_1, X_2) = 1_{\{X_1\}}$ y $h_2(X_1, X_2) = 1_{\{X_2\}}$, dicho de otra forma, simplemente tomamos el histograma sobre una componente de la cadena dada por MCMC. Podemos ver las estimaciones a las distribuciones marginales posteriores para los parámetros g y b en la Fig.

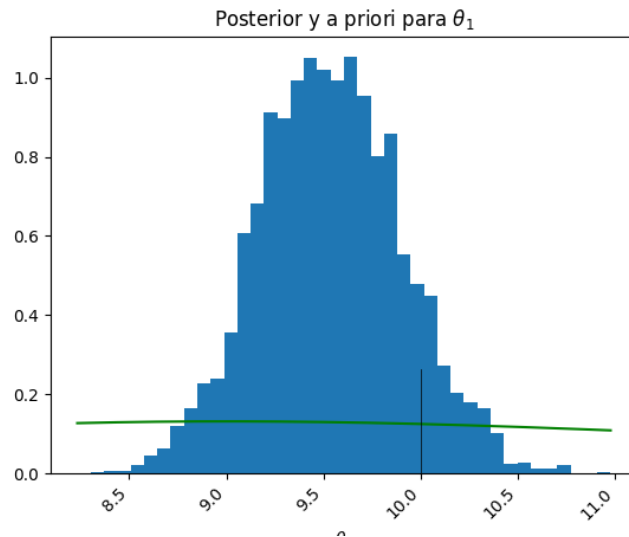


Figura 3.7: Distribución posterior para el parámetros g .

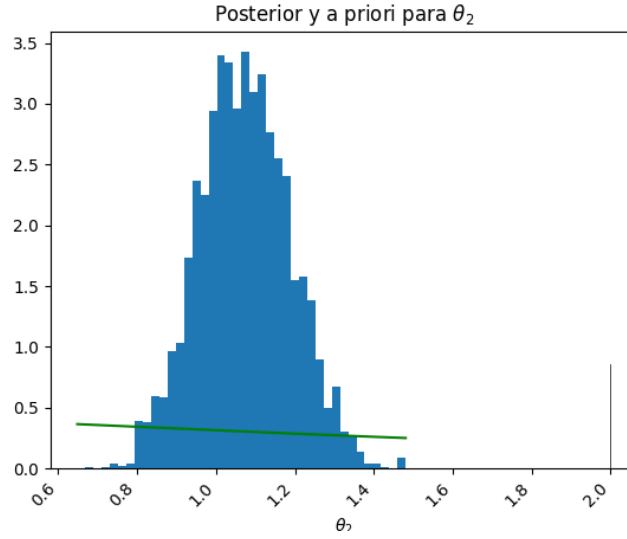


Figura 3.8: Distribución posterior para el parámetro b .

Finalmente, es de interés observar la distribución que puede generar las curvas predictivas. Es decir, de la cadena X_t fruto de MCMC, tomamos submuestreos en intervalos $X_{[iN, (i+1)N]}$ para tomar estimaciones de g y b y así aplicar el forward map para cada estimación. En la Fig.

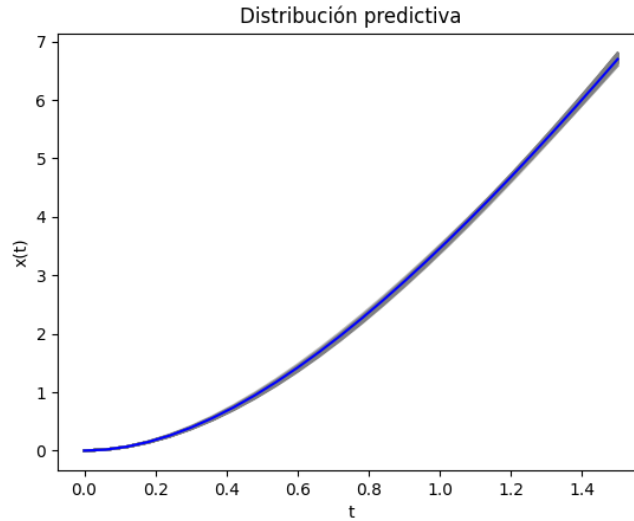


Figura 3.9: Distribución predictiva para el modelo gravitatorio con σ conocido.

Por otro lado, el análisis previo se dio con σ conocido. Sin embargo, este supuesto es irreal. Pues a pesar que el supuesto de normalidad de los errores es plausible, no

lo es considerar la desviación estándar como parámetro conocido.

El caso de σ **desconocido** se considera a dicho parámetro como uno más de la distribución posterior dada en (3.21).

En el estudio de la distribución posterior para este caso, análogo al estudio previo con σ conocido, proponemos las siguiente distribuciones a priori para g y b mostradas en la Fig. 3.10

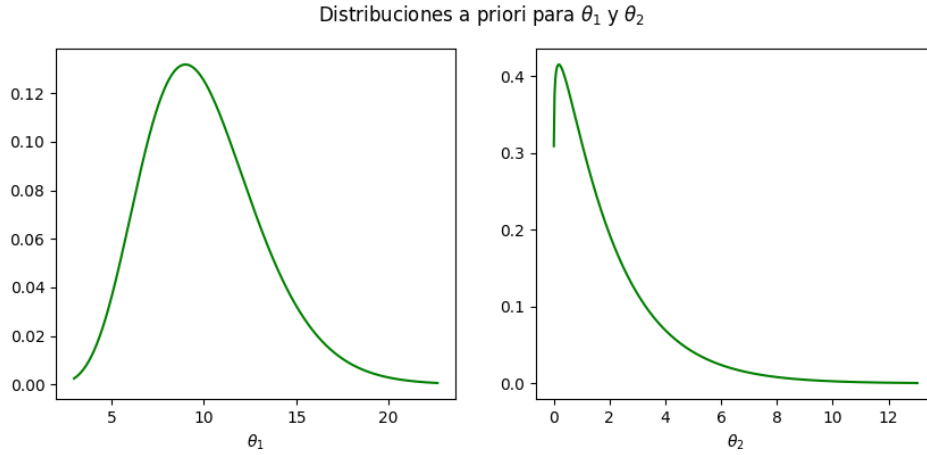


Figura 3.10: Distribuciones a priori para $\theta_1 = g$ y $\theta_2 = b$.

En la Fig. (3.11) tenemos una muestra del modelo gravitatorio. Así, con el algoritmo de MCMC Metropolis-Hastings, la trayectoria de la cadena para la distribución posterior conjunta es 3.12

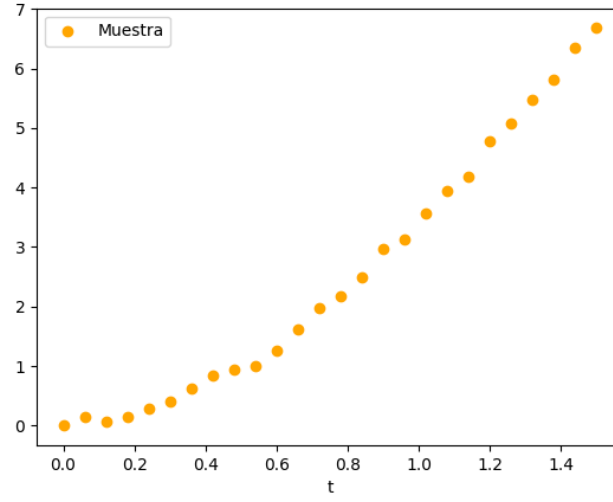


Figura 3.11: Muestra y del modelo gravitatorio.

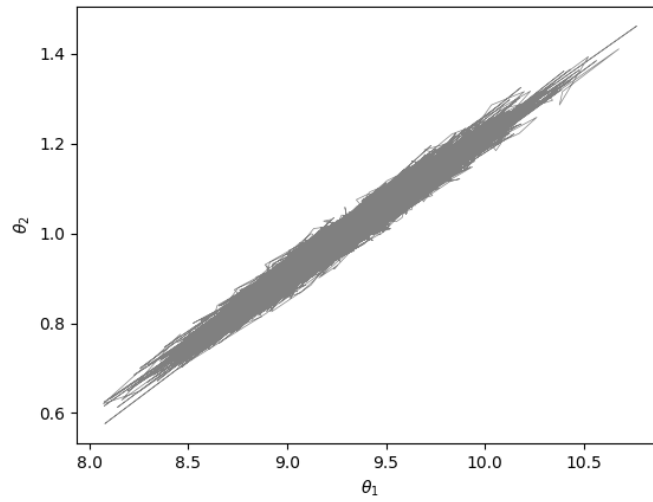
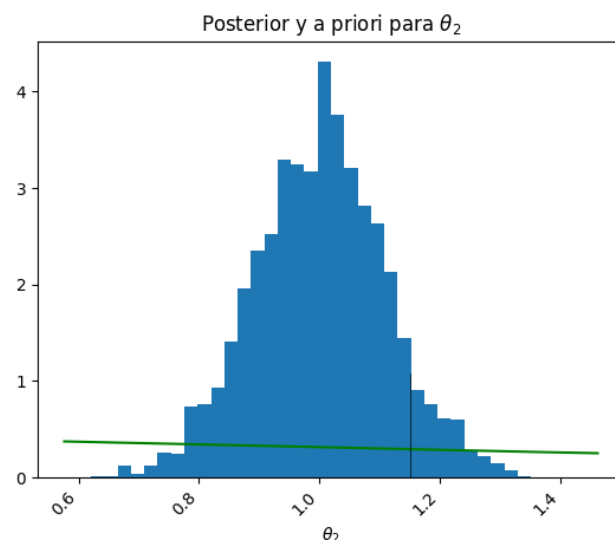
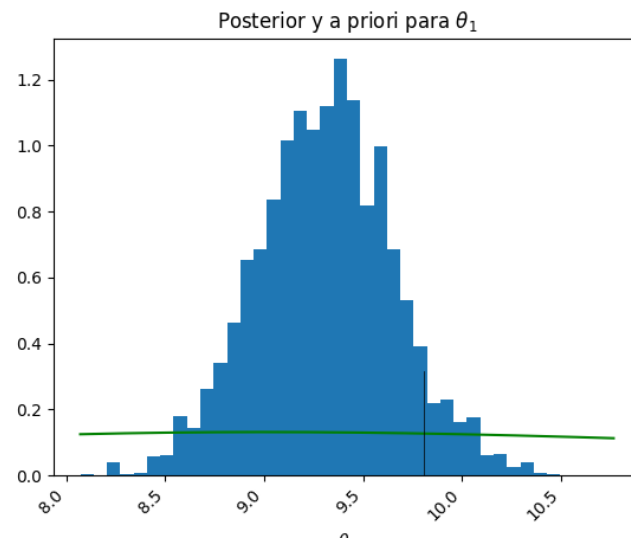
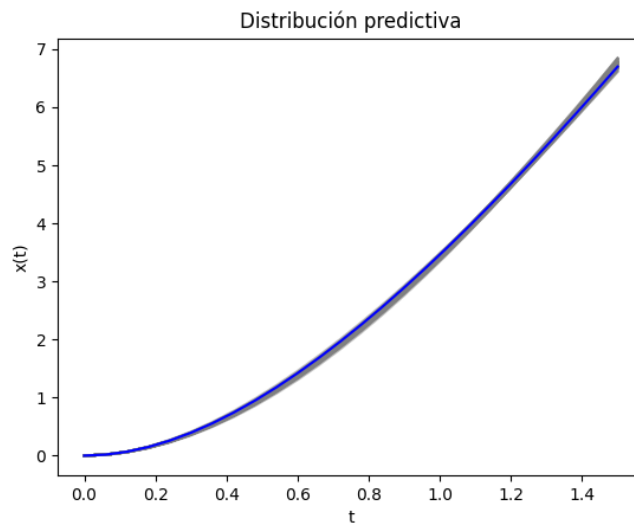
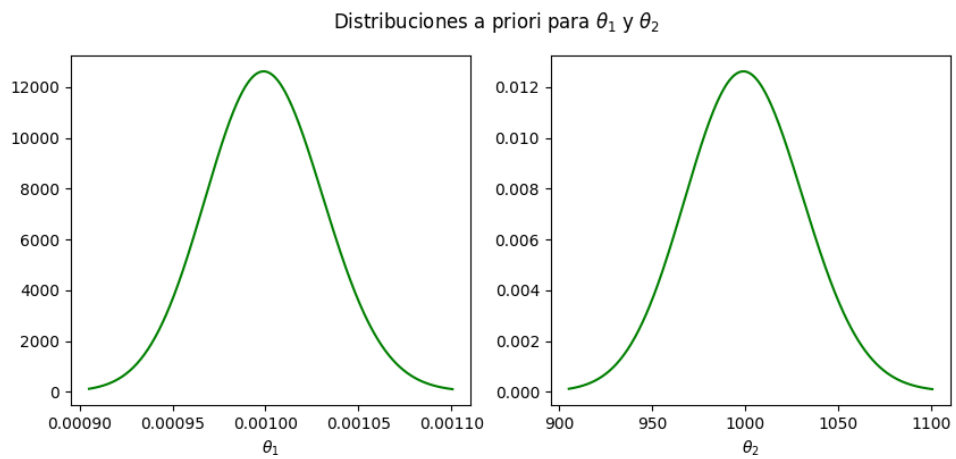


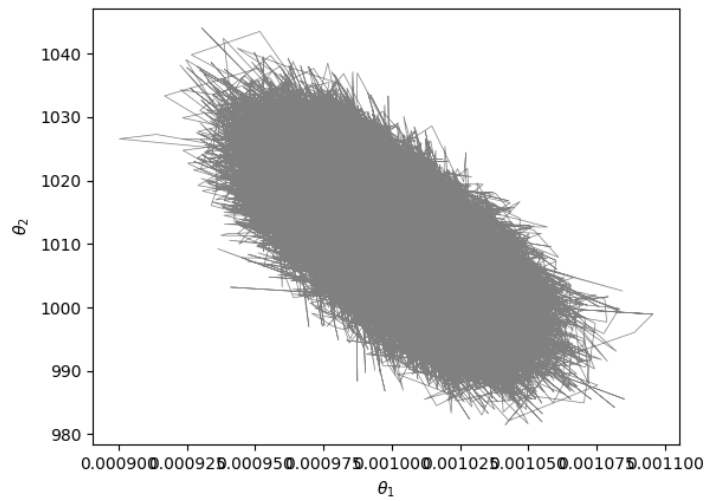
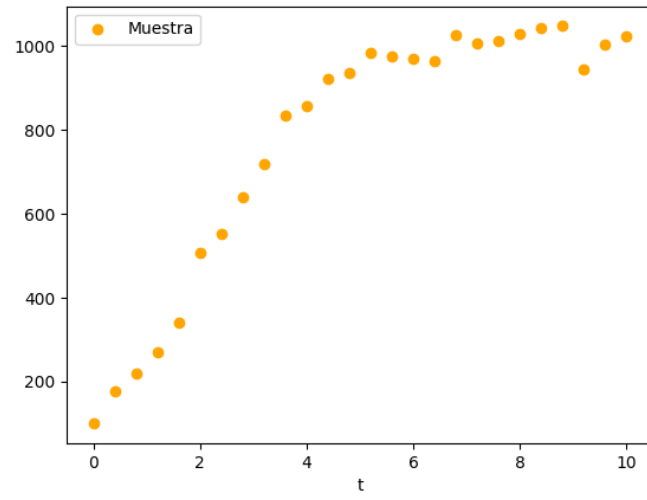
Figura 3.12: Cadena obtenida por MCMC para la distribución posterior conjunta.

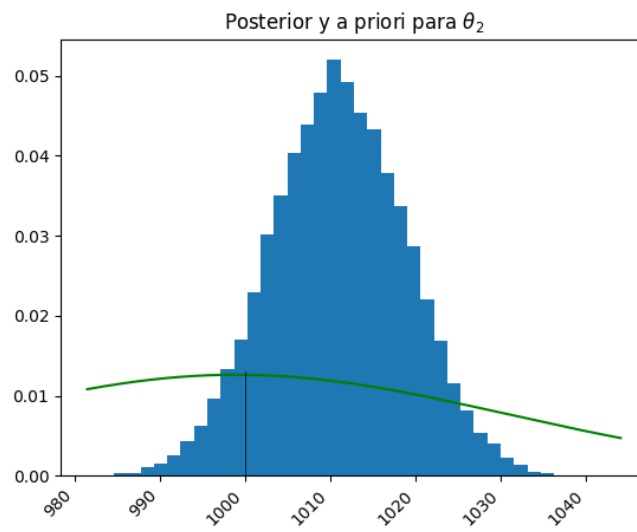
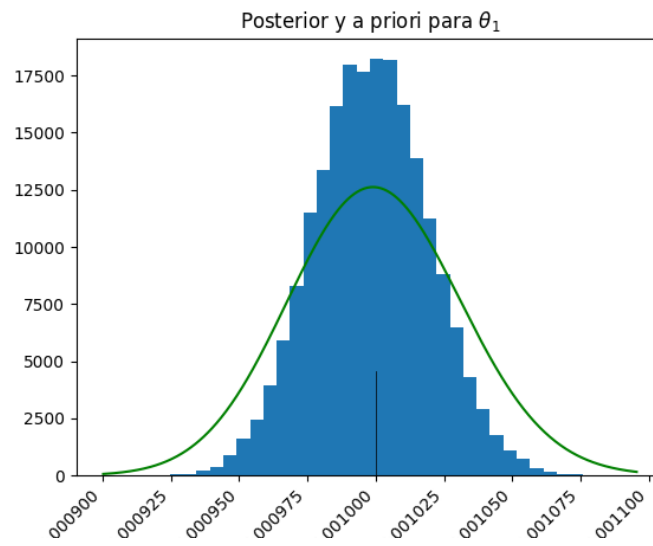


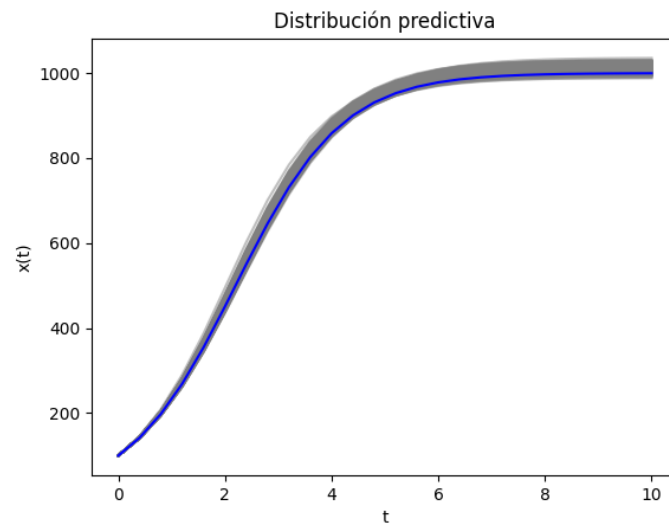


Modelo logístico

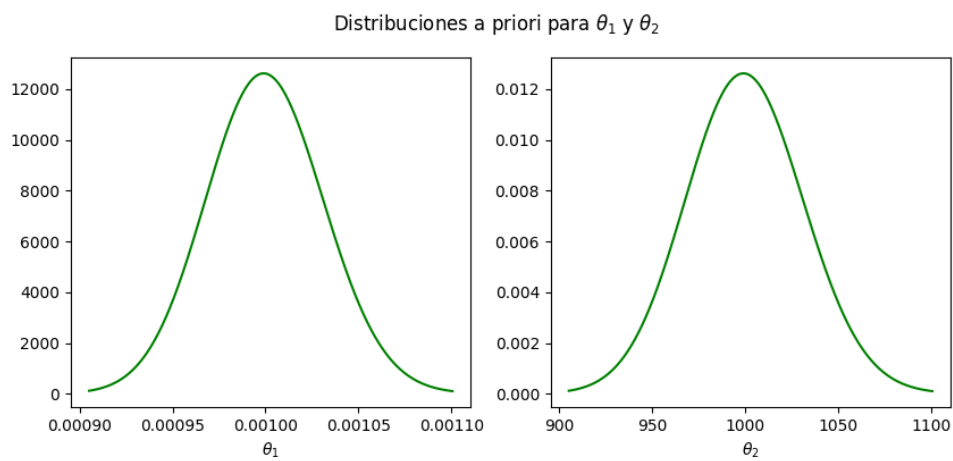


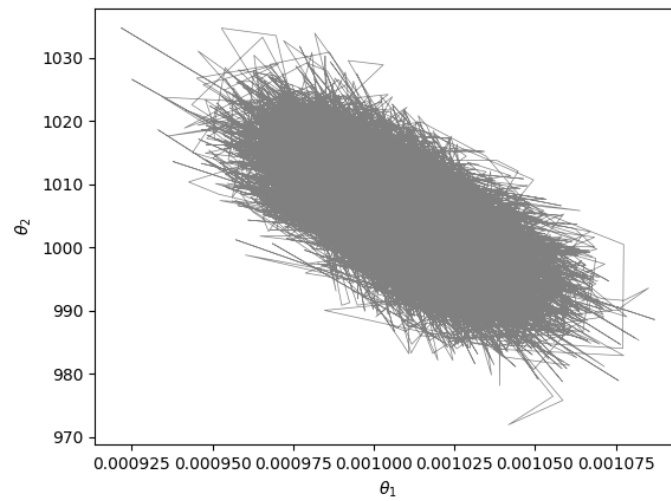
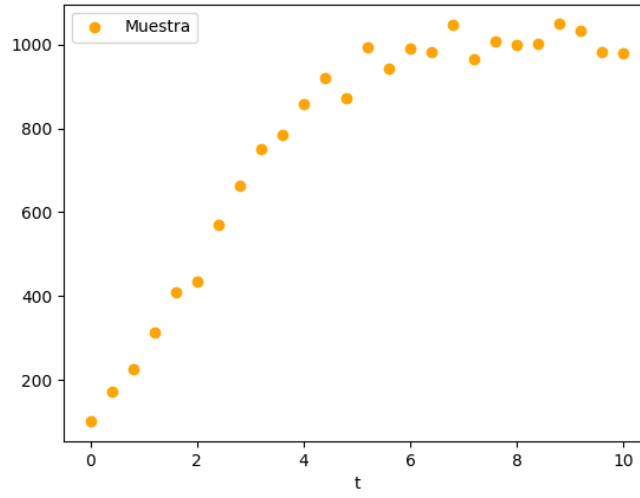


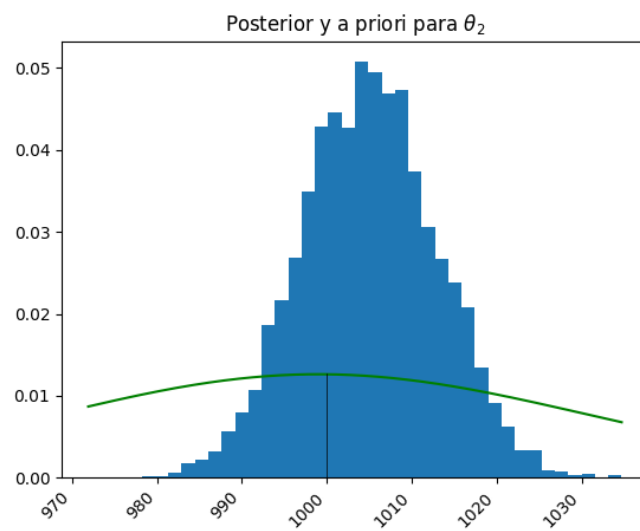
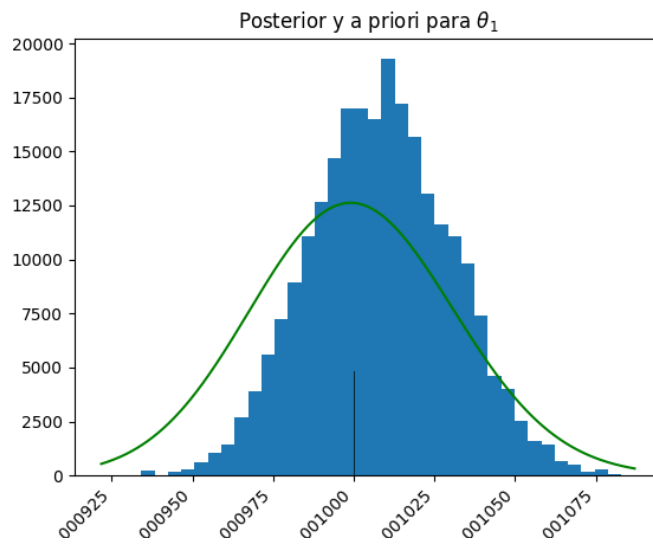


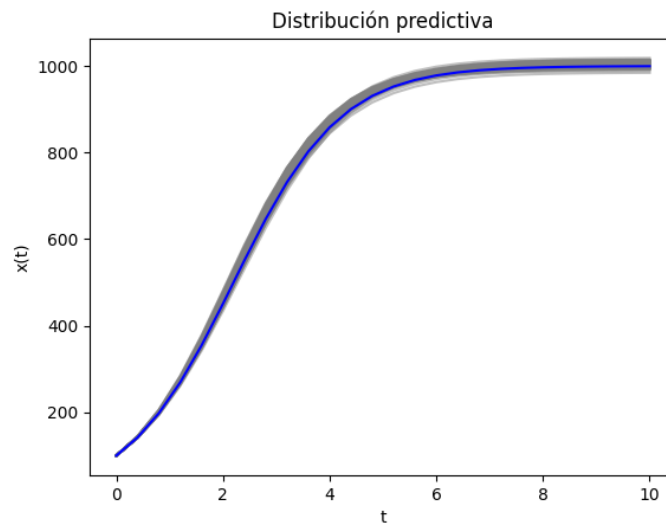


Sigma desconocido

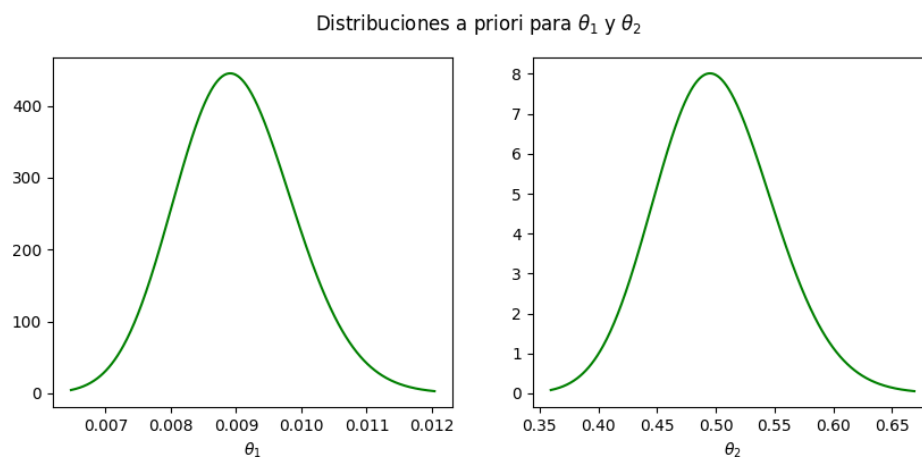


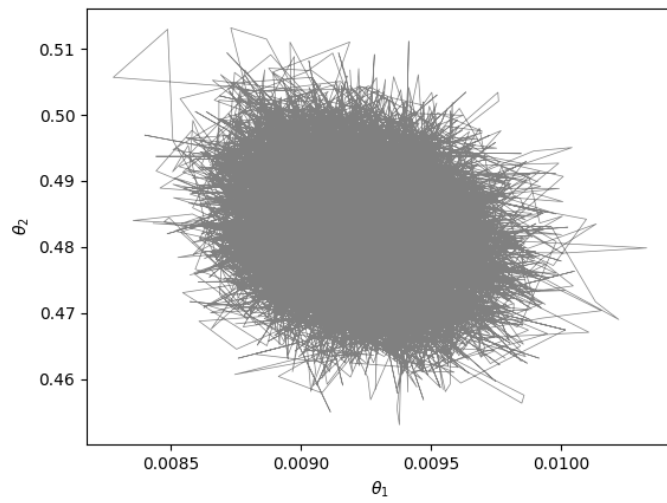
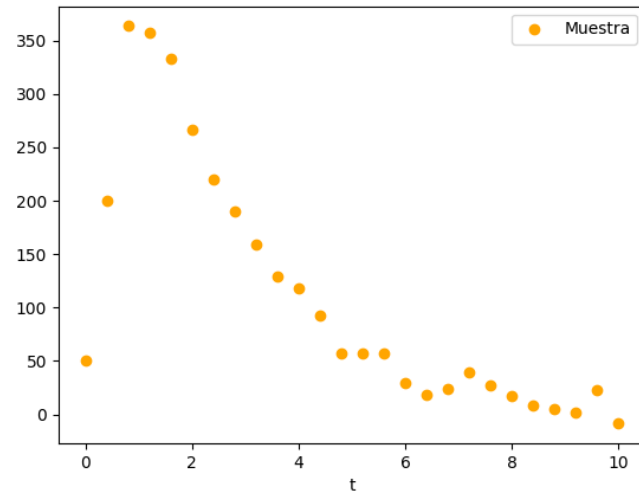


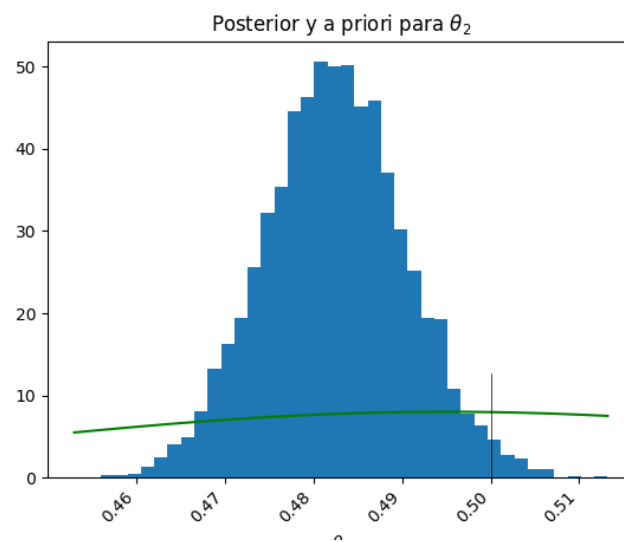
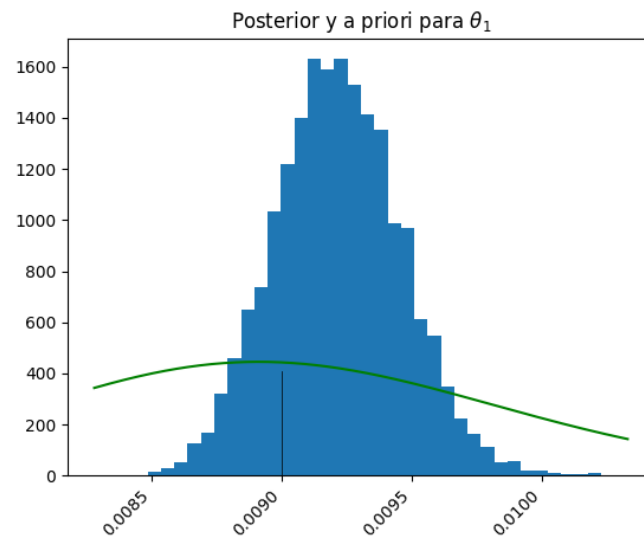


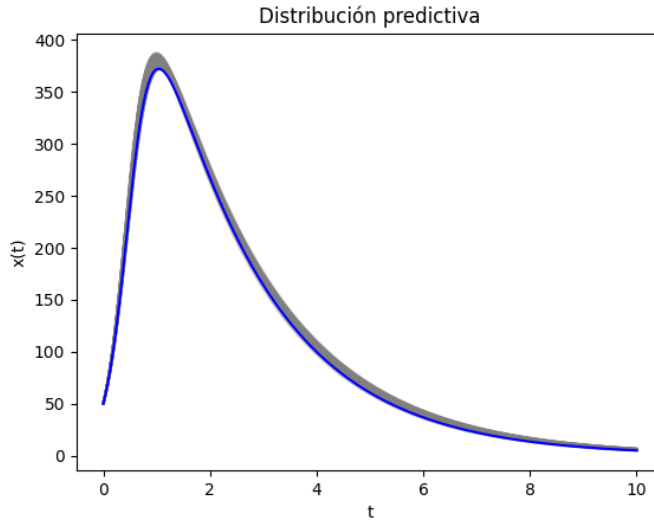


3.2.3. Simulación del modelo SIR









3.3. Inferencia con aproximación.

Se propone una nueva forma de hacer el análisis bayesiano para el problema inverso. Ya hemos visto que la metodología desarrollada que la distribución posterior depende del forward map, y por ello el algoritmo de Metropolis-Hastings requiere que a cada paso en el espacio de estados Θ en la cadena X_t , se solucione la ecuación diferencial del modelo para los parámetros propuestos por MCMC. Sin embargo, tenemos cadenas de medio millón de pasos, lo que complica el análisis bayesiano. En esta ocasión discutiremos una manera más óptima de hacer el estudio.

3.3.1. Regla de interpolación(aproximación de forward map)

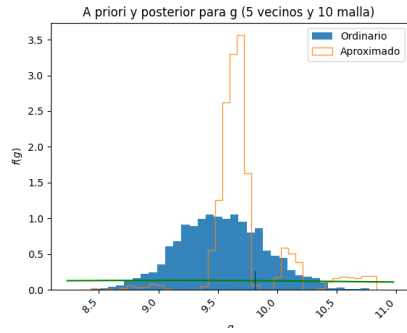
Consideremos un enmallado donde para cada θ_i se particione en intervalos discretos de $[a_i, b_i]$ en m intervalos equidistantes para toda $i \in \{1, 2, \dots, d\}$.

[explicar detalles partición](#)

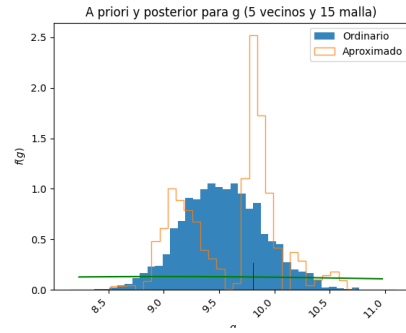
3.3.2. Simulación MCMC

Forward aproximado para el modelo gravitatorio

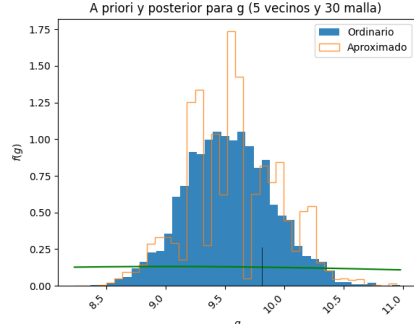
En el análisis bayesiano sobre el modelo gravitatorio obtuvimos la distribución posterior con el forward map ordinario. En este estudio se proponen hacer el mismo análisis con mallas cada vez más finas. Luego graficamos cada posterior con el forward map aproximado para estudiar su convergencia.



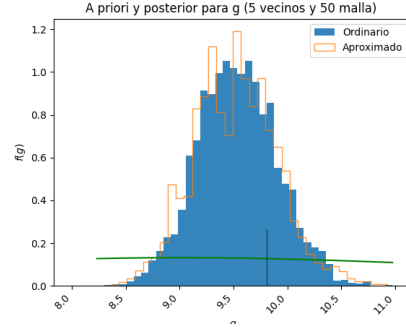
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15



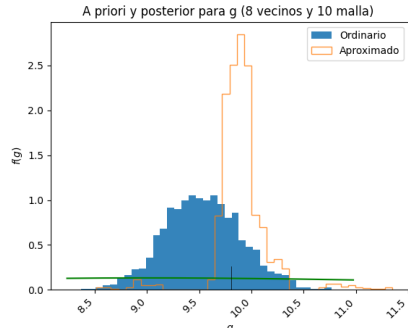
(c) Posterior con Forward aproximado en malla de 30x30



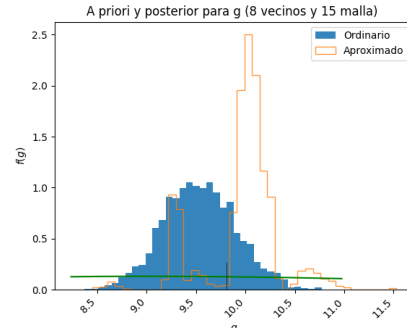
(d) Posterior con Forward aproximado en malla de 50x50

Figura 3.13: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.

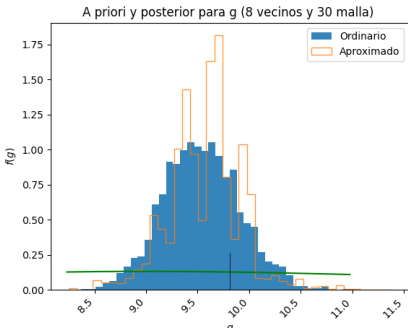
Modelo logístico



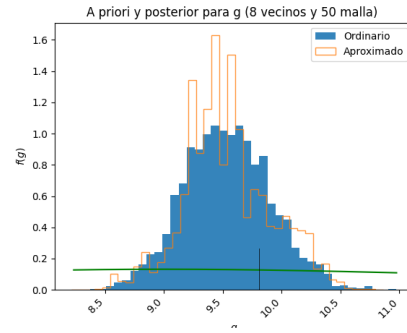
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15



(c) Posterior con Forward aproximado en malla de 30x30



(d) Posterior con Forward aproximado en malla de 50x50

Figura 3.14: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.

3.3.3. Experimentación con la convergencia.

En esta sección se abordan las diferentes simulaciones para cada modelo. Empezando por el modelo de gravedad sujeto a fricción, nos interesa la distribución posterior de los parámetros g y b . Con la metodología establecida y usando MCMC con la a priori [Grafico de la priori](#) anteriormente mencionada obtenemos una simulación de la distribución posterior conjunta. Tomando la marginal de cada parámetro y haciendo el histograma para la estimación de la distribución marginal tenemos las siguientes distribuciones posteriores marginales.

Poner la distribución posterior marginal con Forward ordinario

Consideremos ahora la distribución posterior conjunta obtenida con el Forward map aproximado. En este caso, se hace el experimento para la cantidad de 5,8,16 vecinos cercanos y a su vez cada uno de ello haciendo la malla cada vez más fina.

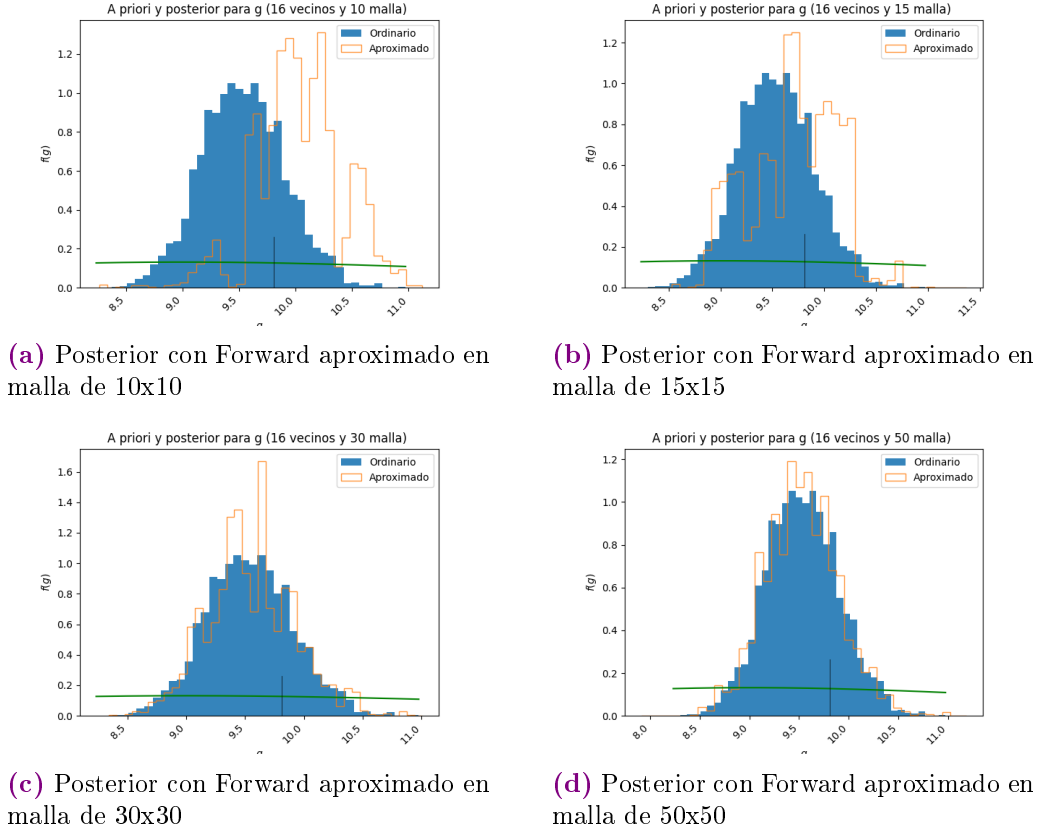


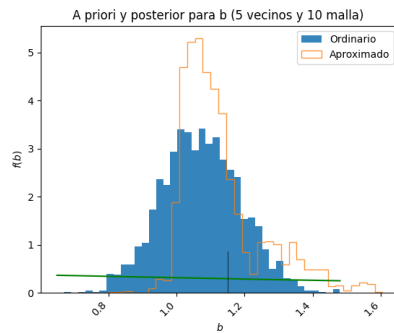
Figura 3.15: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.

Para cada Forward con la cantidad dada de vecinos cercanos se propone una malla en el espacio de parámetros $(g, b) \in \Theta$ de tamaño 10x10, 15x15, 30x30, 50x50.

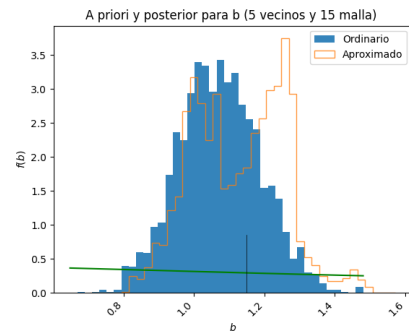
La muestra inicial de la trayectoria que sigue el modelo se hace mediante simulación desde la misma ecuación diferencial dada por el modelo. Es de nuestra elección la cantidad de muestra n que queremos para la distribución posterior conjunta de los parámetros. Sin embargo, para dicha simulación es necesario fijar dos parámetros, que llamaremos parámetros verdadero y obtener la trayectoria $x_t|\theta$.

Para la simulación de los datos consideramos una muestra de 30 datos del modelo de gravedad en el intervalo de tiempo $[0, 1.5]$.

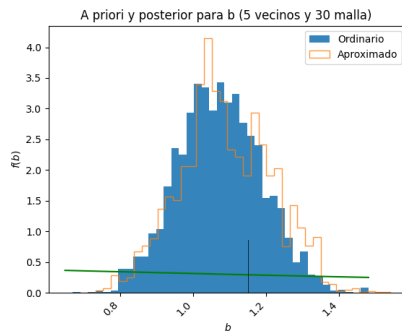
Las distribuciones posteriores marginales de g se muestran en las figuras 3.24 ??
??



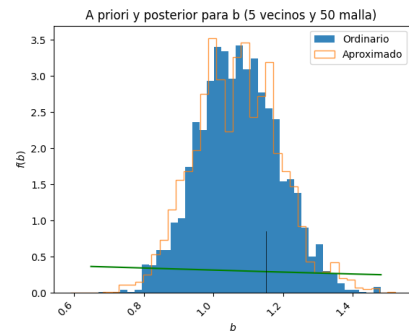
(a) Posterior con Forward aproximado en
malla de 10x10



(b) Posterior con Forward aproximado en
malla de 15x15

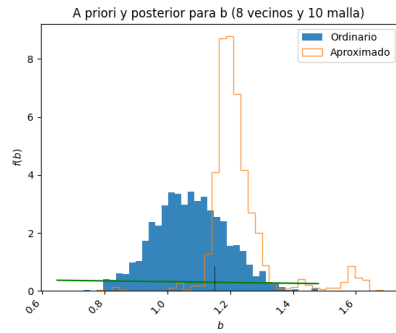


(c) Posterior con Forward aproximado en
malla de 30x30

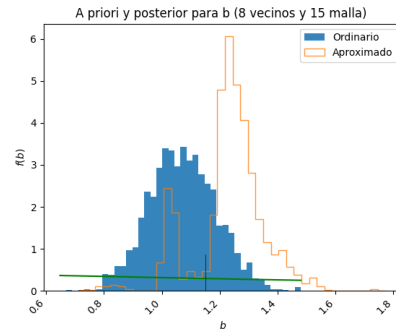


(d) Posterior con Forward aproximado en
malla de 50x50

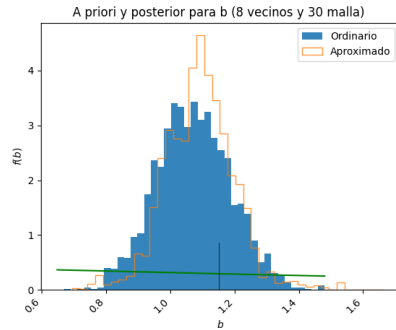
Figura 3.16: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



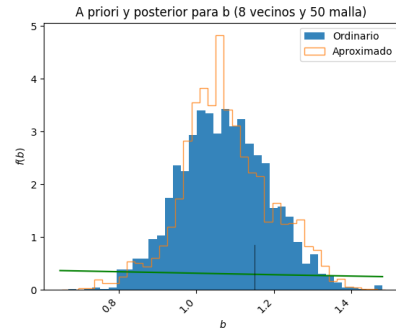
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15

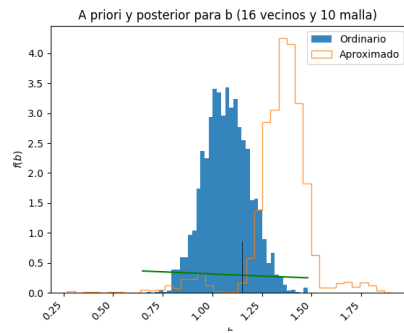


(c) Posterior con Forward aproximado en malla de 30x30

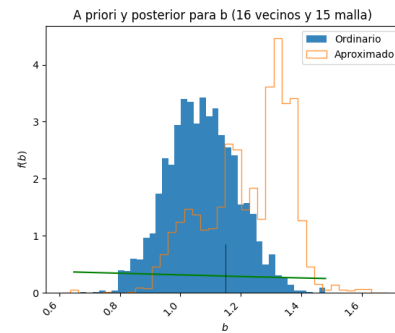


(d) Posterior con Forward aproximado en malla de 50x50

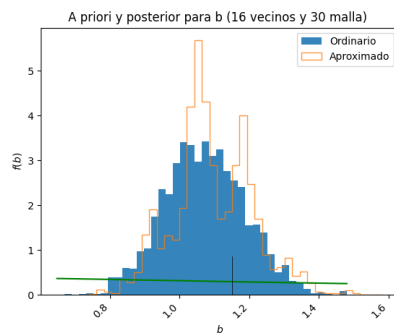
Figura 3.17: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



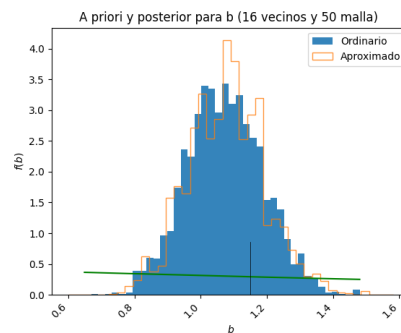
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15

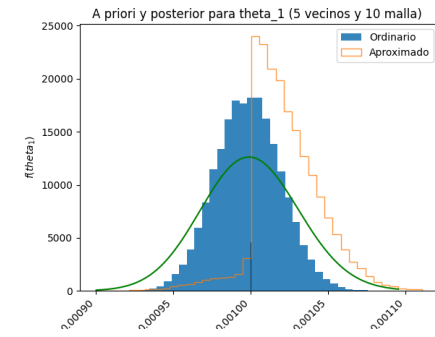


(c) Posterior con Forward aproximado en malla de 30x30

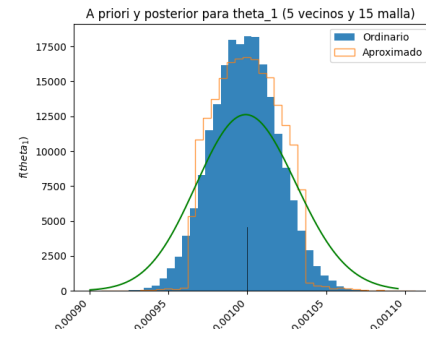


(d) Posterior con Forward aproximado en malla de 50x50

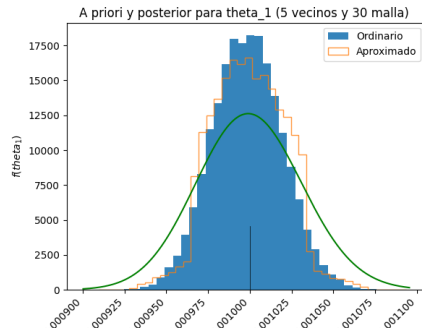
Figura 3.18: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



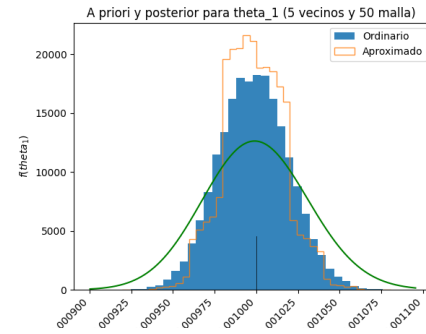
(a) Posterior con Forward aproximado en mall de 10x10



(b) Posterior con Forward aproximado en mall de 15x15

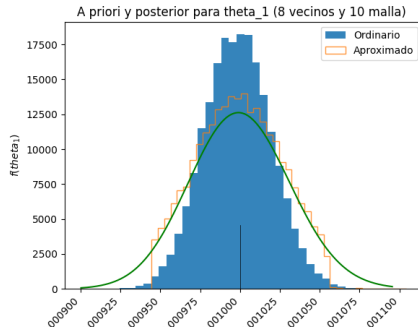


(c) Posterior con Forward aproximado en mall de 30x30

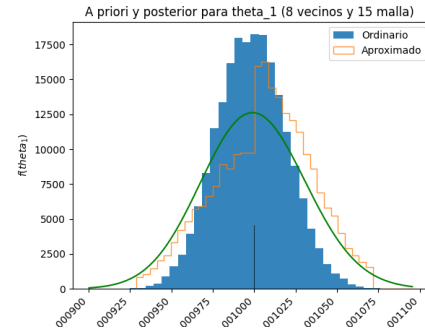


(d) Posterior con Forward aproximado en mall de 50x50

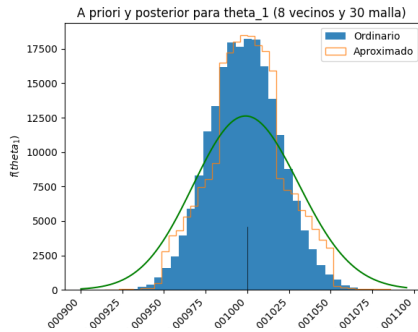
Figura 3.19: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



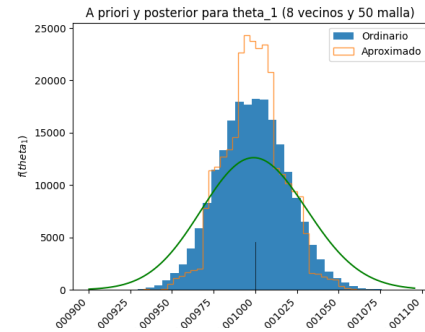
(a) Posterior con Forward aproximado en
malla de 10x10



(b) Posterior con Forward aproximado en
malla de 15x15

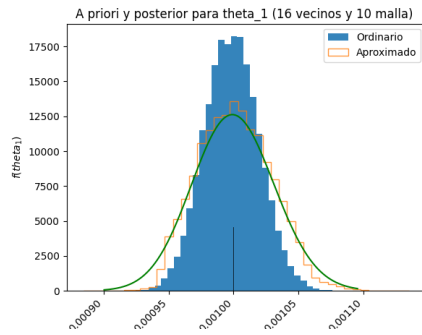


(c) Posterior con Forward aproximado en
malla de 30x30

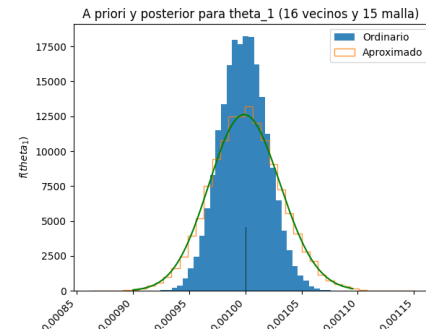


(d) Posterior con Forward aproximado en
malla de 50x50

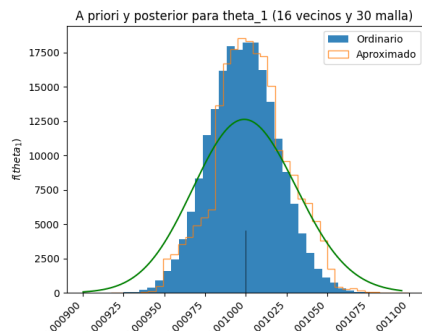
Figura 3.20: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



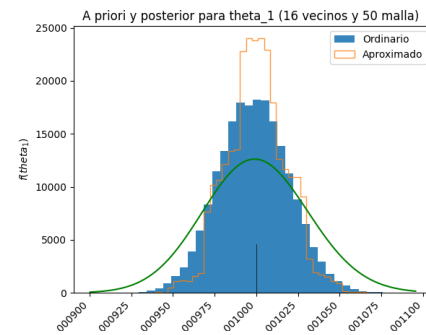
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15

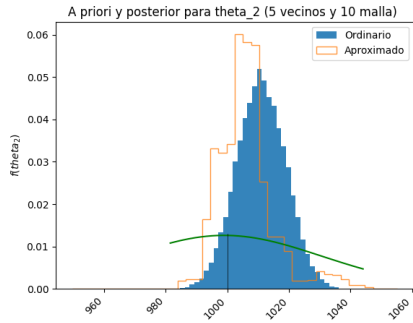


(c) Posterior con Forward aproximado en malla de 30x30

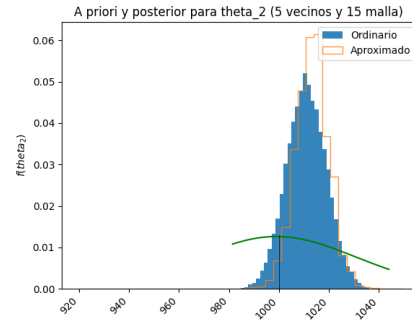


(d) Posterior con Forward aproximado en malla de 50x50

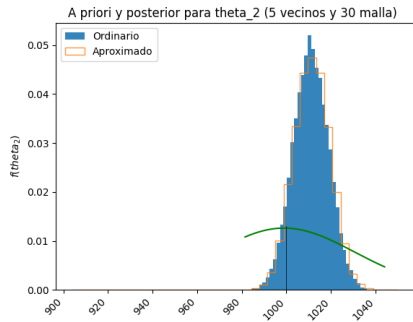
Figura 3.21: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



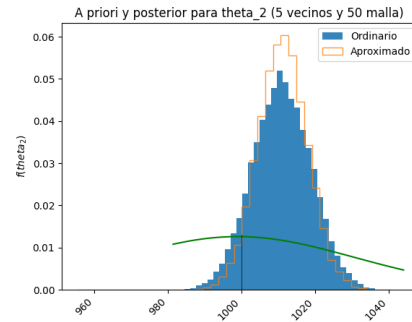
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15

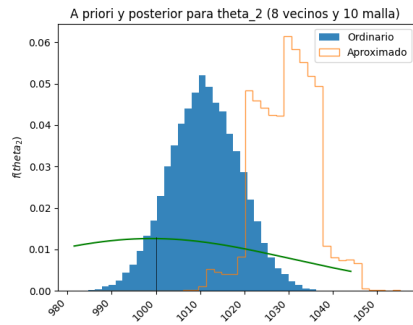


(c) Posterior con Forward aproximado en malla de 30x30

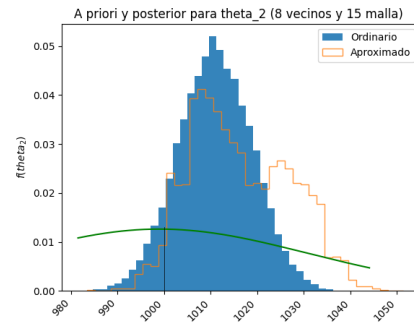


(d) Posterior con Forward aproximado en malla de 50x50

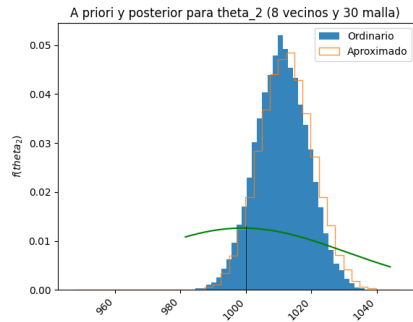
Figura 3.22: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



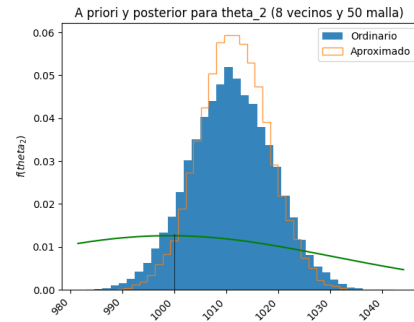
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15

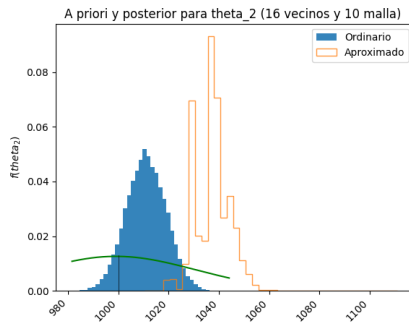


(c) Posterior con Forward aproximado en malla de 30x30

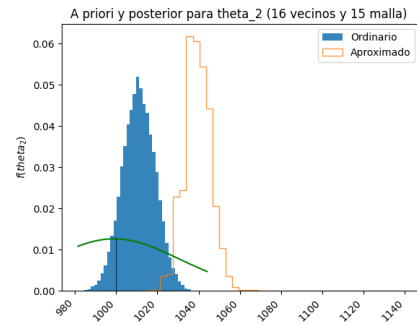


(d) Posterior con Forward aproximado en malla de 50x50

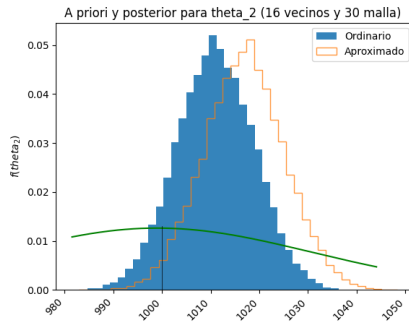
Figura 3.23: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.



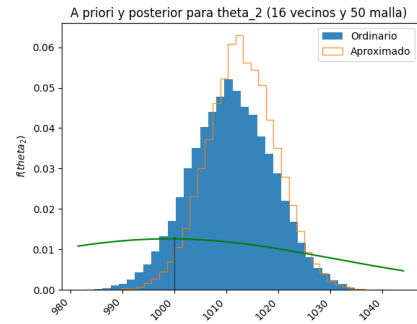
(a) Posterior con Forward aproximado en malla de 10x10



(b) Posterior con Forward aproximado en malla de 15x15



(c) Posterior con Forward aproximado en malla de 30x30



(d) Posterior con Forward aproximado en malla de 50x50

Figura 3.24: Distribución posterior para g con Forward ordinario y Forward aproximado con 5 vecinos cercanos.

Capítulo 4

Conclusiones y trabajo futuro

Referencias

Sears, F. W., Zemansky, M. W., Young, H. D., Vara, R. H., García, M. G., Gümes, E. R., ... Benites, F. G. (1986). *Física universitaria*. Fondo Educativo Interamericano Naucalpan de Juárez, México.

Apéndice A

Anexo

