

### Ejercicio 1

Suponga que  $X_i \sim U(0, \theta)$  independientes. ¿Cómo haría inferencia bayesiana (conjugada) sobre  $\theta$  para analizar la Hipótesis  $H_0 : \theta < 20$  vs  $H_1 : \theta \geq 20$ ? Ejemplo: Tenemos un equipo electrónico cuya confiabilidad se puede describir con una distribución uniforme. Los tiempos de falla en meses de 30 de estos aparatos son: 7.7, 1.2, 5.5, 5.5, 13.0, 12.0, 9.7, 11.0, 18.0, 19.0, 16.0, 9.3, 6.8, 3.7, 4.8, 6.1, 2.2, 3.6, 9.9, 8.0, 19.0, 11.0, 15.0, 8.1, 15.0, 5.1, 12.0, 7.3, 2.1, 14.0. ¿Qué se puede concluir entonces del contraste de hipótesis anterior?

### Solución:

Dado que se nos pide hacer un análisis conjugado para un modelo donde  $X_i \sim U(0, \theta)$  entonces es que se tiene que proponer una distribución a priori  $\theta \sim \text{Pareto}(\alpha_0, \beta_0)$ . Mostraremos que en efecto estamos ante un análisis conjugado. Primeramente, recordemos que la densidad de una distribución  $\text{Pareto}(\alpha, \beta)$  es

$$f(x) = \alpha \beta^\alpha x^{-(\alpha+1)} \mathbb{1}_{[\beta, \infty)}(x). \quad (1)$$

Notemos que el soporte de la distribución Pareto depende del parámetro  $\beta$ .

Por otro lado, la distribución posterior del parámetro  $\theta$  es

$$f(\theta|X^n) \propto f(\theta) \mathcal{L}(\theta), \quad (2)$$

con la distribución a priori  $\theta \sim \text{Pareto}(\alpha_0, \beta_0)$  se sigue que

$$f(\theta|X^n) \propto \alpha_0 \beta_0^{\alpha_0} \theta^{-(\alpha_0+1)} \mathbb{1}_{[\beta_0, \infty)}(\theta) \prod_{i=1}^n \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(X_i).$$

Observemos que la verosimilitud se anula si existe un  $i$  tal que  $X_i > \theta$ . Entonces, para tener verosimilitud no nula se requiere que para todo  $i$  se cumpla  $\theta \geq X_i$ . Dicho de otra forma, basta con que se cumpla que  $\theta \geq X_{(n)}$  donde  $X_{(n)} = \max_i \{X_i\}$  también conocida como la estadística de orden superior. Por tanto,

$$f(\theta|X^n) \propto \theta^{-(\alpha_0+n+1)} \mathbb{1}_{[\beta_0, \infty)}(\theta) \mathbb{1}_{[X_{(n)}, \infty)}(\theta).$$

Recordemos que el producto de dos funciones indicadores sobre dos conjuntos es la función indicadora de la intersección. Así

$$f(\theta|X^n) \propto \theta^{-(\alpha_0+n+1)} \mathbb{1}_{[\max\{\beta_0, X_{(n)}\}, \infty)}(\theta). \quad (3)$$

Contrastando (3) con (1) que es el kernel de una distribución Pareto, lo que muestra que en efecto son familias conjugadas.

Por tanto, la distribución posterior es también una distribución Pareto

$$\theta|X^n \sim \text{Pareto}(\alpha_p, \beta_p), \quad (4)$$

donde  $\alpha_p = \alpha_0 + n$  y  $\beta_p = \max\{\beta_0, X_{(n)}\}$ .

Previo a ver los datos, no tenemos conocimiento sobre  $\theta$ . Como se propondrá una distribución a priori Pareto, y el parámetro  $\beta$  controla el soporte de forma que  $\theta \in [\beta, \infty)$ . Conviene proponer una distribución cuyo soporte sean los reales no negativos, esto pasaría con la distribución Pareto solo si  $\beta$  es cercano a cero. Por tanto elegimos  $\beta = 1$ . Para el parámetro  $\alpha$ , tenemos distribuciones más *delgadas* para mayores

valores de  $\alpha$ . Luego, elegimos un valor pequeño, sin embargo, para  $\alpha \leq 1$  no tiene media y para  $\alpha \leq 2$  no tiene varianza. Por ello seleccionamos para la distribución a priori  $\alpha = 5$ .

Así, la distribución a priori es

$$\theta \sim \text{Pareto}(\alpha_0 = 5, \beta_0 = 1).$$

Como  $\alpha_p = \alpha_0 + n = 35$  y  $\beta_p = \max\{\beta_0, X_{(n)}\} = \max\{1, 19\} = 19$ , la distribución posterior es

$$\theta|X^n \sim \text{Pareto}(\alpha_p = 35, \beta_p = 19). \quad (5)$$

Notemos que el soporte de la distribución posterior decreció y además se tiene una distribución más *concentrada*.

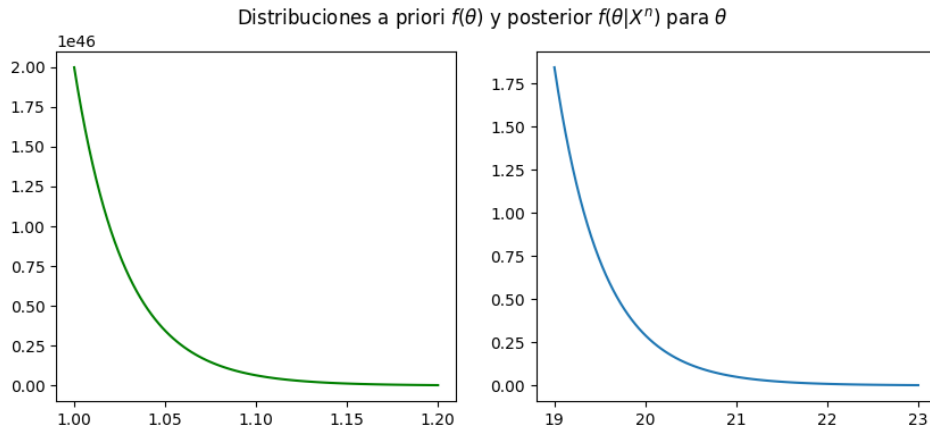


Figura 1: Distribución a priori y distribución posterior para  $\theta$ .

### Ejercicio 2

Si  $\mathbf{Y} \sim N_n(\mathbf{X}\theta, \Sigma)$  y  $\Sigma$ , la matriz de varianzas y covarianzas, conocida (definida positiva);  $\mathbf{X}$ , la matriz de diseño en la regresión, conocida. Encuentre la posterior  $\theta$  usando una  $N_p(\mu_0, \Sigma_0)$  como *a priori*. Encuentre la distribución predictiva de un dato futuro  $Y_{n+1}$ .

### Solución:

Recordemos que la densidad de una variable  $\mathbf{X} \sim N_n(\mu, \Sigma)$  (distribución normal n-variada) es

$$f(\mathbf{X}|\theta) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{X} - \mu)' \Sigma^{-1} (\mathbf{X} - \mu) \right\}, \quad (6)$$

con  $\mathbf{X}, \mu \in \mathbb{R}^n$ .

Empecemos por encontrar la distribución posterior para  $\theta$ . Tenemos un modelo normal  $n$ -variado y se propone una distribución a priori normal  $p$ -variada. Así la distribución posterior se sigue de

$$\pi(\theta|\mathbf{Y}) \propto f(\mathbf{Y}|\theta)\pi(\theta). \quad (7)$$

con  $\theta \in \mathbb{R}^p$ . Luego, el modelo de  $\mathbf{Y}$  tiene la siguiente verosimilitud

$$f(\mathbf{Y}|\theta) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{Y} - \mathbf{X}\theta)' \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\theta) \right\}. \quad (8)$$

Juntando la distribución a priori con (8) tenemos

$$\begin{aligned} \pi(\theta|\mathbf{Y}) &\propto (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{Y} - \mathbf{X}\theta)' \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\theta) \right\} \cdot \\ &\quad (2\pi)^{-p/2} |\Sigma_0| \exp \left\{ -\frac{1}{2} (\theta - \mu_0)' \Sigma_0^{-1} (\theta - \mu_0) \right\}. \end{aligned}$$

Prescindiendo de las constantes

$$\pi(\theta|\mathbf{Y}) \propto \exp \left\{ -\frac{1}{2} \left[ (\mathbf{Y} - X\theta)' \Sigma^{-1} (\mathbf{Y} - X\theta) + (\theta - \mu_0)' \Sigma_0^{-1} (\theta - \mu_0) \right] \right\}. \quad (9)$$

y observamos que es parecido al núcleo de una distribución normal  $p$ -variada. Para ello buscamos completar la forma cuadrática tal que tengamos un núcleo de la forma

$$\exp \left\{ -\frac{1}{2} (\theta - \mu_n)' \Sigma_n^{-1} (\theta - \mu_n) \right\}. \quad (10)$$

Consideremos

$$Q = (\mathbf{Y} - X\theta)' \Sigma^{-1} (\mathbf{Y} - X\theta) + (\theta - \mu_0)' \Sigma_0^{-1} (\theta - \mu_0).$$

Desarrollando

$$Q = \mathbf{Y}' \Sigma^{-1} \mathbf{Y} - \mathbf{Y}' \Sigma^{-1} X\theta - \theta' X' \Sigma^{-1} \mathbf{Y} + \theta' X' \Sigma^{-1} X\theta + \theta' \Sigma_0^{-1} \theta - \theta' \Sigma_0^{-1} \mu_0 - \mu_0' \Sigma_0^{-1} \theta + \mu_0' \Sigma_0^{-1} \mu_0.$$

Simplificando

$$Q = \mathbf{Y}' \Sigma^{-1} \mathbf{Y} - 2\mathbf{Y}' \Sigma^{-1} X\theta + \theta' X' \Sigma^{-1} X\theta + \theta' \Sigma_0^{-1} \theta - 2\theta' \Sigma_0^{-1} \mu_0 + \mu_0' \Sigma_0^{-1} \mu_0. \quad (11)$$

De igual forma para el argumento de (10)

$$\begin{aligned} (\theta - \mu_n)' \Sigma_n^{-1} (\theta - \mu_n) &= \theta' \Sigma_n^{-1} \theta - \theta' \Sigma_n^{-1} \mu_n - \mu_n' \Sigma_n^{-1} \theta + \mu_n' \Sigma_n^{-1} \mu_n, \\ &= \theta' \Sigma_n^{-1} \theta - 2\mu_n' \Sigma_n^{-1} \theta + \mu_n' \Sigma_n^{-1} \mu_n. \end{aligned} \quad (12)$$

Comparando los términos 3 y 4 de (11) con el primero de (12)

$$\theta' \Sigma_n^{-1} \theta = \theta' X' \Sigma^{-1} X\theta + \theta' \Sigma_0^{-1} \theta = \theta' \left( X' \Sigma^{-1} X + \Sigma_0^{-1} \right) \theta,$$

obtenemos la relación

$$\Sigma_n^{-1} = \Sigma_0^{-1} + X' \Sigma^{-1} X. \quad (13)$$

Luego, comparando los términos 2 y 5 de (11) con el segundo de (12)

$$\begin{aligned} -2\mu_n' \Sigma_n^{-1} \theta &= -2\mathbf{Y}' \Sigma^{-1} X\theta - 2\mu_0' \Sigma_0^{-1} \theta, \\ \mu_n' \Sigma_n^{-1} \theta &= \left( \mathbf{Y}' \Sigma^{-1} X + \mu_0' \Sigma_0^{-1} \right) \theta, \end{aligned}$$

transponiendo se sigue

$$\Sigma_n^{-1} \mu_n = X' \Sigma \mathbf{Y} + \Sigma_0^{-1} \mu_0,$$

lo que concluye que

$$\mu_n = \Sigma_n \left( X' \Sigma \mathbf{Y} + \Sigma_0^{-1} \mu_0 \right), \quad (14)$$

los términos restantes son constantes respecto a  $\theta$  por lo que no requiere análisis.

Así es claro que la distribución posterior es

$$\theta|\mathbf{Y} \sim N_p(\mu_n, \Sigma_n), \quad (15)$$

con  $\mu_n, \Sigma_n$  definidos en (14) y (13), respectivamente.

Por otro lado, para obtener la distribución predictiva para una nueva observación  $\mathbf{Y}_{n+1}$  dada las observaciones previas es necesario calcular

$$f(\mathbf{Y}_{n+1}|\mathbf{Y}) = \int f(\mathbf{Y}_{n+1}|\theta, \mathbf{Y})\pi(\theta|\mathbf{Y})d\theta,$$

y por independencia de las observaciones

$$f(\mathbf{Y}_{n+1}|\mathbf{Y}) = \int f(\mathbf{Y}_{n+1}|\theta)\pi(\theta|\mathbf{Y})d\theta. \quad (16)$$

Sustituyendo las respectivas densidades en (16)

$$\begin{aligned} f(\mathbf{Y}_{n+1}|\mathbf{Y}) &= \int (2\pi)^{-n/2}|\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{Y}_{n+1} - X\theta)' \Sigma^{-1} (\mathbf{Y}_{n+1} - X\theta) \right\} \cdot \\ &\quad \cdot (2\pi)^{-p/2}|\Sigma_n|^{-1/2} \exp \left\{ -\frac{1}{2} (\theta - \mu_n)' \Sigma_n^{-1} (\theta - \mu_n) \right\} d\theta. \end{aligned}$$

Sacando las constante de la integral

$$f(\mathbf{Y}_{n+1}|\mathbf{Y}) = (2\pi)^{-n/2}(2\pi)^{-p/2}|\Sigma|^{-1/2}|\Sigma_n|^{-1/2} \int \exp \left\{ -\frac{1}{2} Q(\theta) \right\} d\theta,$$

donde

$$\begin{aligned} Q(\theta) &= (\mathbf{Y}_{n+1} - X\theta)' \Sigma^{-1} (\mathbf{Y}_{n+1} - X\theta) + (\theta - \mu_n)' \Sigma_n^{-1} (\theta - \mu_n), \\ &= \mathbf{Y}_{n+1}' \Sigma^{-1} \mathbf{Y}_{n+1} - 2\mathbf{Y}_{n+1}' \Sigma^{-1} X\theta + \theta' X' \Sigma^{-1} X\theta + \\ &\quad + \theta' \Sigma_n^{-1} \theta - 2\mu_n' \Sigma_n^{-1} \theta + \mu_n' \Sigma_n^{-1} \mu_n. \end{aligned} \quad (17)$$

De manera análoga al calculo de la distribución posterior de  $\theta$ , ahora buscamos tener la forma cuadrática

$$\begin{aligned} (\theta - \mu_p)' \Sigma_p^{-1} (\theta - \mu_p) &= \\ &= \theta' \Sigma_p^{-1} \theta - 2\mu_p' \Sigma_p^{-1} \theta + \mu_p' \Sigma_p^{-1} \mu_p. \end{aligned} \quad (18)$$

Buscamos relaciones de  $\mu_p$  y  $\Sigma_p$  con los parámetros conocidos. Así, comparando el primer término de (18) con los términos 3 y 4 de (17) y además comparando los términos 2 y 5 de (17) con el segundo término de (18) obtenemos las relaciones

$$\Sigma_p^{-1} = X' \Sigma^{-1} X + \Sigma_n^{-1}, \quad (19)$$

$$\mu_p = \Sigma_p \left( X' \Sigma^{-1} \mathbf{Y}_{n+1} + \Sigma_n^{-1} \mu_n \right), \quad (20)$$

análogas a las relaciones (13) y (14) de la distribución posterior.

Por tanto, podemos escribir la distribución predictiva como

$$\begin{aligned} f(\mathbf{Y}_{n+1}|\mathbf{Y}) &= (2\pi)^{-n/2}(2\pi)^{-p/2}|\Sigma|^{-1/2}|\Sigma_n|^{-1/2} \cdot \\ &\quad \cdot \int \exp \left\{ -\frac{1}{2} \left[ (\theta - \mu_p)' \Sigma_p^{-1} (\theta - \mu_p) + \mathbf{Y}_{n+1}' \Sigma^{-1} \mathbf{Y}_{n+1} + \mu_n' \Sigma_n^{-1} \mu_n \right] \right\} d\theta. \end{aligned}$$

Reordenando

$$\begin{aligned} f(\mathbf{Y}_{n+1}|\mathbf{Y}) &= (2\pi)^{-n/2}|\Sigma|^{-1/2}|\Sigma_n|^{-1/2}|\Sigma_p|^{1/2} \exp \left\{ -\frac{1}{2} \left( \mathbf{Y}_{n+1}' \Sigma^{-1} \mathbf{Y}_{n+1} + \mu_n' \Sigma_n^{-1} \mu_n \right) \right\} \cdot \\ &\quad \cdot \int (2\pi)^{-p/2}|\Sigma_p|^{-p/2} \exp \left\{ -\frac{1}{2} (\theta - \mu_p)' \Sigma_p^{-1} (\theta - \mu_p) \right\} d\theta. \end{aligned}$$

Vemos que la integral corresponde a la integral de una distribución normal  $p$ -variada con media  $\mu_p$  y matriz de covarianzas  $\Sigma_p$ , por tanto integra uno. Finalmente

$$f(\mathbf{Y}_{n+1}|\mathbf{Y}) = (2\pi)^{-n/2}|\Sigma|^{-1/2}|\Sigma_n|^{-1/2}|\Sigma_p|^{1/2} \exp \left\{ -\frac{1}{2} \left( \mathbf{Y}_{n+1}' \Sigma^{-1} \mathbf{Y}_{n+1} + \mu_n' \Sigma_n^{-1} \mu_n \right) \right\}, \quad (21)$$

que es la función de densidad de la distribución predictiva.

### Ejercicio 3

Sea el modelo de crecimiento logístico  $\frac{dX}{dt} = \theta_1 X (\theta_2 - X)$  con  $X(0) = X_0$ . Suponga que tenemos observaciones  $y_i$  para  $X(t_i)$ ,  $t_1 < t_2 < \dots < t_n$ , con ruido gaussiano aditivo independiente, esto es

$$y_i = X(t_i) + \epsilon_i; \quad \epsilon_i \sim N(0, \sigma), \quad i = 1, 2, \dots, n.$$

Simule datos con  $X(0) = 100$ ,  $\theta_1 = 0.001$ ,  $\theta_2 = 1000$ ,  $\sigma = 30$ ,  $n = 26$  equiespaciados en  $t \in [0, 10]$ . Haga inferencia bayesiana para los parámetros  $\theta_1, \theta_2$ . Utilice a priores Gamma, centradas en los valores verdaderos. ¿Qué puede decir de la tasa de crecimiento que es  $\lambda = \theta_1 \theta_2$ ?

### Solución:

Para este ejercicio se tiene se implemento numéricamente la ecuación diferencial correspondiente al crecimiento logístico. Luego, simulamos 26 datos en un arreglo equiespaciado del tiempo con un ruido gaussiano tal como se pide en el enunciado. Podemos ver los datos graficados en la Fig 2

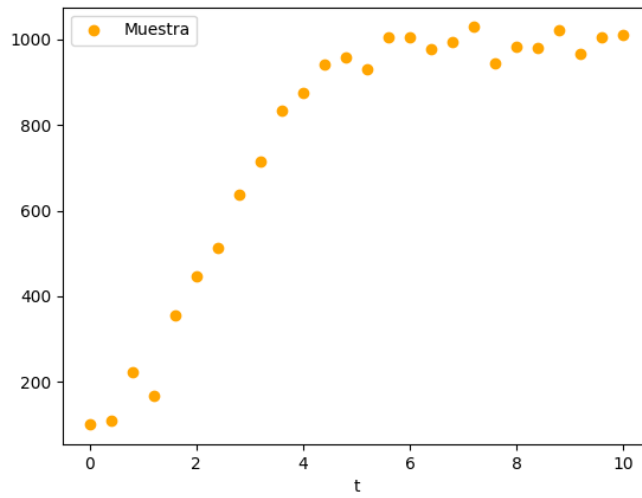


Figura 2: Simulación de datos de crecimiento logístico.

La distribución posterior es

$$\pi(\theta_1, \theta_2 | \mathbf{y}) \propto \mathcal{L}(\theta_1, \theta_2) \pi(\theta_1, \theta_2), \quad (22)$$

donde la verosimilitud se obtiene del modelo gaussiano. Así

$$\begin{aligned} \mathcal{L}(\theta_1, \theta_2) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - F_\theta(t_i))^2 \right\}, \\ &= \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - F_\theta(t_i))^2 \right\}. \end{aligned} \quad (23)$$

Elegimos las distribuciones a priori gammas independientes

$$\begin{aligned} \theta_1 &\sim \text{Gamma} \left( \alpha_0, \frac{0.001}{\alpha_0} \right), \\ \theta_2 &\sim \text{Gamma} \left( \beta_0, \frac{1000}{\beta_0} \right), \end{aligned}$$

con  $\alpha_0 = \beta_0 = 1000$ . Así

$$\pi(\theta_1, \theta_2) = \left( \frac{0.001}{\alpha_0} \right)^{\alpha_0} \frac{\theta_1^{\alpha_0-1} \exp \{ -0.001\alpha_0/\theta_1 \}}{\Gamma(\alpha_0)} \cdot \left( \frac{1000}{\beta_0} \right)^{\beta_0} \frac{\theta_2^{\beta_0-1} \exp \{ -1000\beta_0/\theta_2 \}}{\Gamma(\beta_0)}. \quad (24)$$

Las distribuciones a priori se dan en la Fig 3 con los parámetros anteriormente descritos.

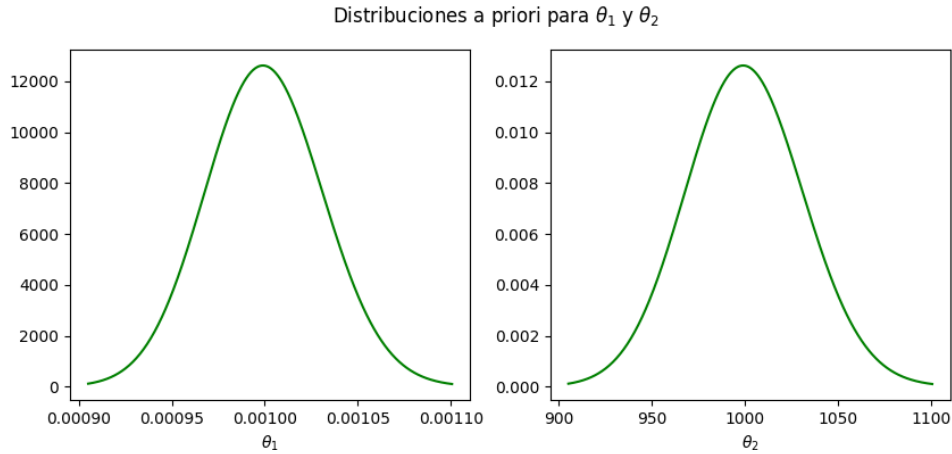


Figura 3: Distribuciones a priori para  $\theta_1$  y  $\theta_2$ .

Sustituyendo (23) y (24) en (22) tenemos la distribución objetivo. La cual por medio de métodos Monte Carlo podemos simular de dicha distribución. Con MCMC simulamos una cadena de tamaño  $T = 100000$  tomamos las marginales de la posterior para cada parámetro las cuales vemos en la Fig 4 y en la Fig 5.

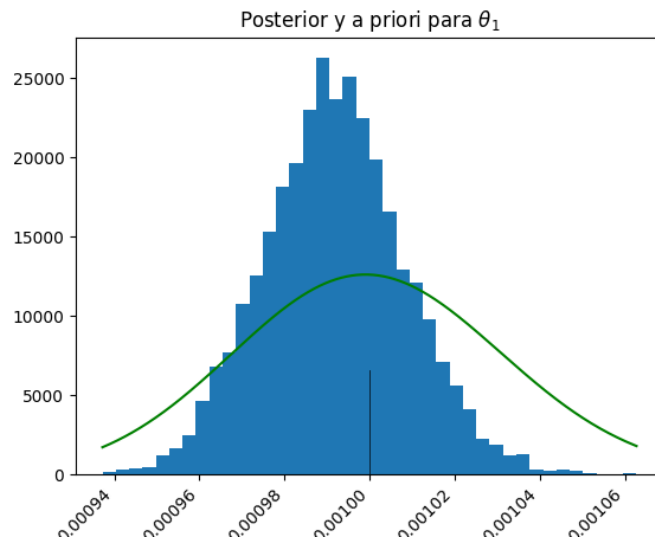


Figura 4: Distribución marginal posterior para  $\theta_1$  (azul) y distribución a priori (verde) con parámetro verdadero (negro).

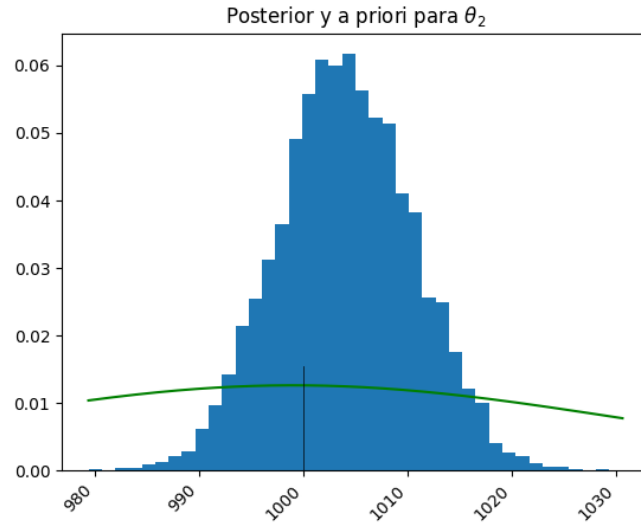


Figura 5: Distribución marginal posterior para  $\theta_2$  (azul) y distribución a priori (verde) con parámetro verdadero (negro).

De igual forma, para la tasa de crecimiento  $\lambda$  obtenemos el histograma posterior obtenido simplemente de tomar el producto de cada marginal posterior, graficado en la Fig 6.

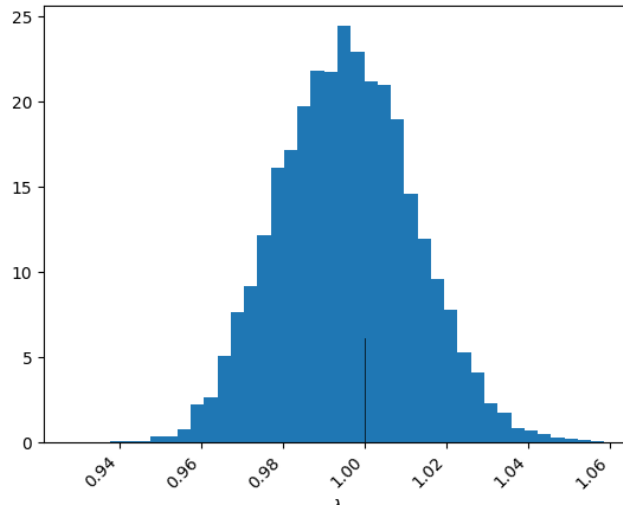


Figura 6: Posterior para parámetro de tasa de crecimiento  $\lambda$ .

Finalmente, graficamos la distribución predictiva de la cadena tomando submuestreos de la cadena fruto de MCMC y tomando la media en cada submuestreo para considerar la trayectoria generada por el forward map en la Fig 7.

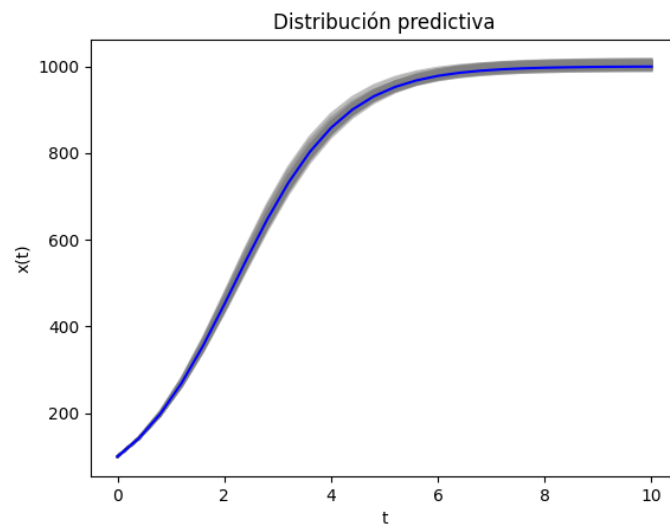


Figura 7: Distribución predictiva para crecimiento logístico (gris) y la trayectoria real simulada (azul).

Notamos que dicha distribución predictiva ajusta satisfactoriamente a la trayectoria simulada con los parámetros reales.