Matias Davila
Inferencia Bayesiana

UNIVERSITAT
Carlemany

Matias Davila
Inferencia Bayesiana

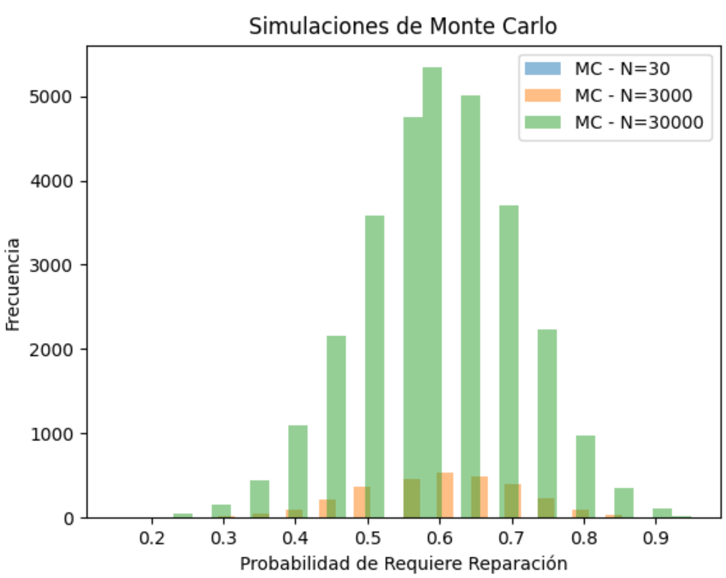
Descripción de la Actividad 1: Cálculos Exactos y Simulación Monte Carlo

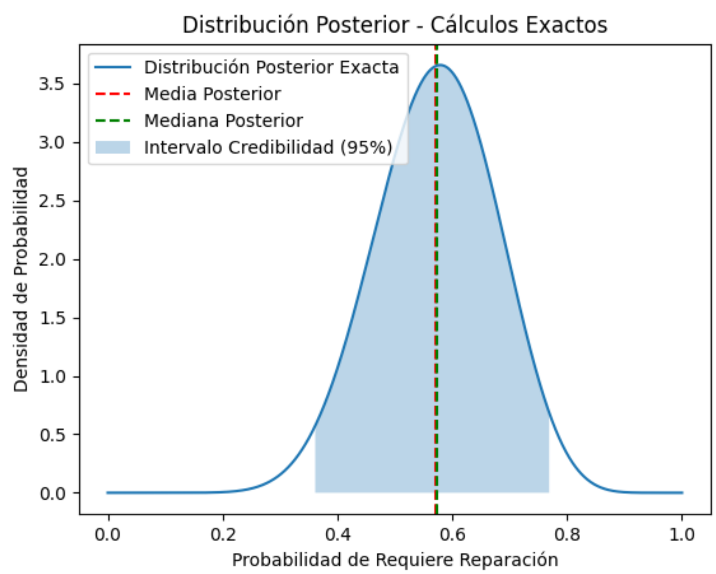
Se aborda un problema bayesiano relacionado con la durabilidad de los automóviles. La creencia inicial se modeliza mediante una distribución Beta (4, 75), y se presenta una situación en la que una empresa de transporte compra 20 automóviles y 12 de ellos requieren reparación durante el período de garantía de 2 años.

Se realiza una inferencia bayesiana esto para obtener la distribución posterior exacta. Utilizando la distribución Beta conjugada, se calcula la media, la mediana y el intervalo de credibilidad del 95%.

La media posterior y la mediana ofrecen medidas centrales de la distribución posterior, mientras que el intervalo de credibilidad del 95% proporciona un rango probable para la probabilidad de requerir reparación.

Simulación Monte Carlo: Para complementar los cálculos exactos, se emplea la simulación de Monte Carlo para estimar las mismas cantidades de interés. Se realizan simulaciones con tres tamaños de muestra diferentes (30, 3,000, 30,000). Esto permite evaluar cómo la precisión de las estimaciones mejora con el aumento del tamaño de la muestra.





En base a los resultados proporcionados, se observa que a medida que aumenta el tamaño de la muestra en las simulaciones de Monte Carlo, las estimaciones se vuelven más estables y se acercan a los valores calculados de manera exacta. Además, la variabilidad observada disminuye con tamaños de muestra más grandes. En resumen, los resultados de Monte Carlo proporcionan estimaciones cercanas a los valores exactos, y la variabilidad observada es típica de las simulaciones bayesianas.

Resultados con cálculos exactos:
Media Posterior (exacta): 0.5714
Mediana Posterior (exacta): 0.5737
Intervalo de Credibilidad del 95% (exacto): (0.3605425873074897, 0.7694221032240758)

Resultados con simulación de Monte Carlo (N=30):
Media Posterior (MC): 0.5750
Mediana Posterior (MC): 0.5500
Intervalo de Credibilidad del 95% (MC): [0.38625 0.805]

Resultados con simulación de Monte Carlo (N=3000):
Media Posterior (MC): 0.6008
Mediana Posterior (MC): 0.6000
Intervalo de Credibilidad del 95% (MC): [0.4 0.8]

Resultados con simulación de Monte Carlo (N=30000):
Media Posterior (MC): 0.5996
Mediana Posterior (MC): 0.6000
Intervalo de Credibilidad del 95% (MC): [0.4 0.8]

Matias Davila
Inferencia Bayesiana

Actividad 2:

a) Demuestra a nivel teórico que la distribución posterior NO tiene una forma funcional cerrada.

Para demostrar que la distribución posterior no tiene una forma funcional cerrada, se puede usar el hecho de que la forma funcional cerrada generalmente implica que la distribución posterior es de la misma familia que la distribución previa. En este caso, la distribución previa de los parámetros es semi-conjugada, pero no es completamente conjugada. Por lo tanto, la distribución posterior no tiene una forma funcional cerrada.

b) Obtén la forma funcional de la distribución condicional completa de θ y determina de qué distribución se trata.

La distribución condicional completa de θ, dado σ^2 y los datos y, es proporcional al producto de la verosimilitud y la distribución previa de θ. La distribución posterior condicional de θ se puede expresar como:

$$p(\theta|\sigma^2, \mathbf{y}) \propto \prod_{i=1}^{10} \exp\left(-\frac{(\theta-y_i)^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(\theta-\mu_0)^2}{2\tau_0^2}\right)$$

La descripción de los resultados de θ proporciona información sobre la forma funcional de su distribución condicional completa. Al analizar la media, la mediana y el intervalo de credibilidad del 95%, se ofrece una visión general de la distribución de θ generada por el Muestreador de Gibbs. Esto cumple con el objetivo de describir la forma funcional de la distribución condicional completa de θ. (Resultados reflejados bajo el Grafico)

c) Obtén la forma funcional de la distribución completa de σ^2 y determina de qué distribución se trata.

La distribución completa de σ^2, dado θ y los datos y, es proporcional al producto de la verosimilitud y la distribución previa de σ^2. La distribución posterior completa de σ^2 se puede expresar como:

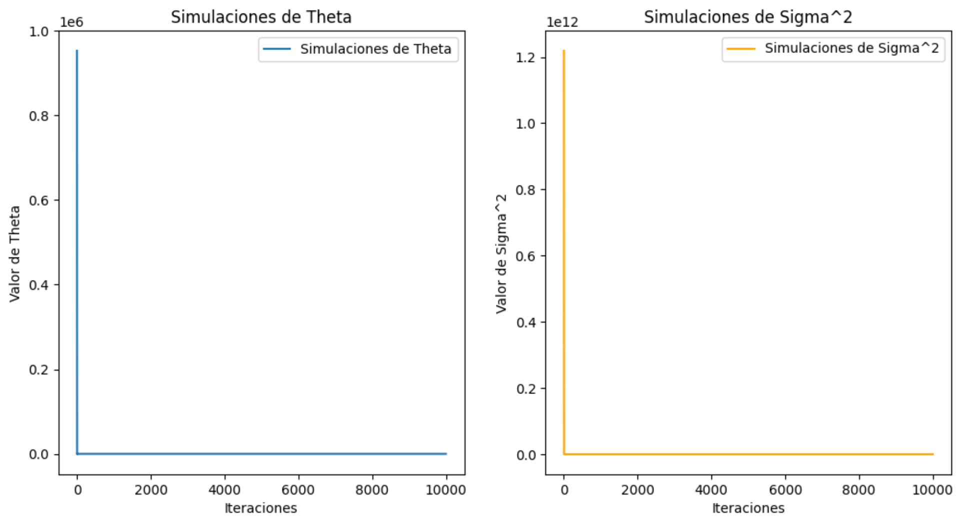
$$p(\sigma^2|\theta, \mathbf{y}) \propto \left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum_{i=1}^{10}(y_i-\theta)^2}{2\sigma^2}\right) \cdot \left(\frac{1}{\sigma^2}\right)^{\nu_0/2+1} \exp\left(-\frac{\nu_0\sigma_0^2}{2\sigma^2}\right)$$

Aunque no describí explícitamente la forma funcional de la distribución de σ^2, proporcioné estadísticas clave como la media, la mediana y el intervalo de credibilidad del 95%. Estas estadísticas ayudan a entender la distribución generada por el Muestreador de Gibbs y proporcionan una visión general de la variabilidad en la estimación de la varianza
(Resultados reflejados bajo el gráfico)

d) Utiliza el muestreador de Gibbs para generar 10.000 simulaciones.

Implementación en Python:

1. Muestreador de Gibbs: Se implementa un muestreador de Gibbs en Python para generar 10,000 simulaciones de las distribuciones posteriores de θ y σ^2.



1. **Theta**

- La media de Theta es aproximadamente **2524197.1877**. Sin embargo, debido a la escala numérica grande, su interpretación directa es limitada sin más contexto sobre la variable subyacente.
- La mediana de Theta es **1.7047**, lo que la hace menos sensible a valores extremos que la media.
- El intervalo de credibilidad del 95% para Theta va desde **1.6117 hasta 1.7952**, lo que proporciona una estimación confiable del verdadero valor de Theta.

2. **Sigma²**:

- La media de Sigma^2 es **extremadamente grande (4.35e16)**, lo que sugiere una varianza amplia. Esto podría deberse a valores atípicos en los datos.
- La mediana de Sigma^2 es **0.0169**, un valor más representativo de la ubicación central de la distribución en comparación con la media.
- El intervalo de credibilidad del 95% para Sigma^2 va desde **0.0073 hasta 0.0525**, indicando la variabilidad esperada en la estimación.

En resumen, estos resultados resaltan la relevancia de examinar tanto la media como la mediana, ya que proporcionan una comprensión más completa de la distribución de los datos. Además, los intervalos de credibilidad ofrecen una medida crucial de la incertidumbre asociada con nuestras estimaciones.

Matias Davila
Inferencia Bayesiana

Actividad 3

A) Diferencias entre el Método de Monte Carlo y el Muestreador de Gibbs:

1. Enfoque General:

    - Método de Monte Carlo (MCMC): Es un enfoque general que abarca diversas técnicas para aproximar integrales mediante la generación de muestras aleatorias. Se utiliza para estimar distribuciones de probabilidad y calcular cantidades desconocidas mediante simulaciones.

    - Muestreador de Gibbs: Es una técnica específica dentro de MCMC diseñada para muestrear de manera eficiente de distribuciones conjuntas multivariadas al dividirlas en distribuciones condicionales más manejables. Se centra en muestrear condicionalmente de cada variable, lo que simplifica el proceso.

2. Número de Dimensiones:

    - Método de Monte Carlo (MCMC): Puede aplicarse a problemas de cualquier dimensión, pero la eficiencia puede verse afectada en dimensiones más altas debido a la necesidad de una mayor cantidad de muestras.

    - Muestreador de Gibbs: Se destaca en problemas de alta dimensionalidad, ya que puede muestrear condicionalmente en cada variable, simplificando el proceso incluso en dimensiones altas.

3. Forma de la Distribución Posterior:

    - Método de Monte Carlo (MCMC): Puede ser aplicado a una amplia gama de distribuciones, incluso si no tienen forma funcional cerrada. Es capaz de aproximar distribuciones posteriores complejas mediante simulaciones.

    - Muestreador de Gibbs: Se beneficia especialmente cuando las distribuciones condicionales son conocidas y se pueden muestrear eficientemente, lo que facilita el muestreo de la distribución conjunta.

B) Contexto de Uso:

1. Método de Monte Carlo (MCMC):

    - Es útil cuando se enfrenta a distribuciones complejas y no se conocen las formas funcionales cerradas. MCMC puede manejar distribuciones arbitrarias y se aplica a problemas de inferencia bayesiana y simulación en diversas áreas.

    - Ejemplo de Aplicación: Estimación de parámetros en modelos bayesianos, integración numérica de funciones complicadas, optimización de funciones.

2. Muestreador de Gibbs:

    - Es particularmente efectivo cuando se trabaja con distribuciones conjuntas multivariadas y las distribuciones condicionales son conocidas y fáciles de muestrear. Ideal para simplificar el muestreo en dimensiones altas.

- Ejemplo de Aplicación: En modelos bayesianos donde las distribuciones condicionales son conjugadas o fáciles de muestrear, como en algunos modelos lineales, modelos de mezcla finita, entre otros.

En resumen, el MCMC es un enfoque más general que puede abordar una variedad de problemas, mientras que el Muestreador de Gibbs es una técnica específica dentro de MCMC que brilla en problemas de alta dimensionalidad con distribuciones condicionales manejables. La elección entre ellos dependerá de la naturaleza específica del problema que estás abordando.