

Taller 2
Estadística Bayesiana
Fecha de entrega: 8 de octubre de 2020

1. Suponga que se lleva a cabo un experimento tal que los datos se pueden modelar mediante una distribución Poisson(λ). Se asigna como distribución a priori para λ una Gamma con parámetros α y β . Dado que se tiene incertidumbre sobre los valores de estos hiperparámetros, se les asigna una distribución a priori Gamma a cada uno.
 - a) Describa el muestreador de Gibbs para muestrear de las distribuciones posteriores de α y β .
 - b) Simule 50 observaciones de una distribución Poisson con $\lambda = 5$. Seleccione distribuciones a priori no informativas para α y β . Implemente un algoritmo MCMC para obtener una estimación de la distribución posterior de λ , α y β . Calcule un intervalo de credibilidad del 90 % para λ y concluya. Verifique convergencia de las cadenas.
2. Se tienen 27 datos de un proceso químico, en donde se mide la cantidad de CO₂ (Y) en función de las siguientes variables predictoras:

Código Variable	Nombre variable
X1	Tiempo de Expansión.
X2	Porcentaje de Solvatación.
X3	Rendimiento de Aceite.
X4	Carbón Total.
X5	Solvente Total.
X6	Consumo de Hidrógeno.

- a) A partir de un análisis descriptivo, seleccione cuatro variables para ajustar un modelo de regresión lineal múltiple.
 - b) Con las variables seleccionadas en el numeral a), realice el ajuste del modelo de regresión lineal múltiple utilizando un enfoque bayesiano, suponga distribuciones a priori no informativas. Especifique claramente el modelo. Muestre el resultado de al menos 10000 simulaciones, con un quemado de 1000. Muestre la tabla de parámetros ajustados y escriba la ecuación ajustada. Determine si cada una de las variables predictoras tiene efecto sobre la variable de respuesta. Concluya a la luz de los datos.
 - c) Realice un análisis de convergencia de las cadenas MCMC.
 - d) Verifique los supuestos del modelo.
3. Se someten a estudio tres marcas de baterías. Se sospecha que el tiempo de vida (en semanas) de las tres marcas es diferente. Se prueban cinco baterías de cada marca con los siguientes resultados:

Semanas de vida		
Marca 1	Marca 2	Marca 3
100	76	108
96	80	100
92	75	96
96	84	98
92	82	100

- a) Realice un análisis descriptivo de los datos experimentales.
 - b) Especifique el modelo para este experimento utilizando un enfoque bayesiano. Seleccione distribuciones a priori no informativas.
 - c) Realice al menos 10000 simulaciones con un quemado de 1000. Estime e interprete los efectos de cada tipo de jabón sobre la media de pérdida de peso. Muestre la estimación puntual y los intervalos de credibilidad del 95 % para cada uno de los efectos. De acuerdo a los resultados, responda las siguientes preguntas.
 - i) ¿El tiempo promedio de vida de las tres marcas de baterías es diferente?
 - ii) ¿Existe diferencia entre el promedio de vida de las marcas 2 y 3?
 - iii) ¿Cuál marca presenta una mayor tiempo promedio de vida?
 - d) Realice un análisis de convergencia de las cadenas MCMC.
 - e) Verifique los supuestos del modelo.
4. Considere el ejercicio anterior. Evalúe si el modelo ajustado es consistente con los datos. Utilice al menos tres medidas de discrepancia.
5. (Corresponde al enunciado del problema 5.3 del libro Bayesian Modeling Using Winbugs) Considere una base de datos simulados que consiste en $p = 15$ covariables y $n = 50$ observaciones. Las covariables X_j (para $j = 1, \dots, 15$) fueron generadas de una distribución normal estándar, mientras que la variable de respuesta fue generada tal que $Y_i \sim N(X_{i4} + X_{i5}, 2.5^2)$ para $i = 1, \dots, 50$. Estime el modelo de regresión completo y proponga otros tres modelos justificando claramente su selección de variables. Utilice como medida de ajuste el DIC.