Curso: Inferencia Bayesiana, Tarea # 2

Instructor: Imelda Trejo Lorenzo

Para entregar el 02 de Septiembre 2025, antes de clase.

Libros de clase:

1. Sea $Y_{1:n}$ una muestra aleatoria de una distribución exponencial $\text{Exp}(\theta)$, con parámetro $\theta > 0$.

   a) Escriba la función de verosimilitud y la log-verosimilitud de θ.

   b) Calcule el estimador de máxima verosimilitud (MLE) de θ.

2. En epidemiología y ciencias sociales es común trabajar con **datos de conteo**, por ejemplo, el número de casos de una enfermedad reportados en un día.

   Para estos casos, un modelo clásico es la distribución de Poisson, la cual supone que la varianza y la media de los conteos son iguales. Sin embargo, en muchos problemas reales los datos muestran **sobredispersión**, es decir, la varianza es mayor que la media. En tales casos, se usa la distribución **Binomial Negativa** (ver Ejercicio 3).

   Sea $Y_{1:n}$ una muestra aleatoria de una distribución de Poisson con parámetro $\theta > 0$, es decir,

   $$Y_i \sim \text{Poisson}(\theta), \quad i = 1, \ldots, n.$$

   Aquí, Y_i representa el número de casos observados en un intervalo de tiempo fijo (por ejemplo, número de infecciones diarias en un brote epidémico).

   a) Escriba la función de verosimilitud y la log-verosimilitud de θ.

   b) Calcule el estimador de máxima verosimilitud (MLE) de θ.

   c) Interprete el resultado obtenido en términos del promedio de los conteos observados.

   d) Genere un conjunto de datos de tamaño $n = 20$ a partir de una distribución Poisson con parámetro $\theta = 5$. Calcule el promedio muestral y compárelo con el valor teórico de la media. Haga una gráfica de barras con la distribución de frecuencias observada y, en la misma figura, superponga la distribución de Poisson teórica con $\theta = 5$. Comente si los datos simulados reflejan la forma de la distribución de Poisson. Repite el mismo ejercico, con $n = 200$ y luego con $n = 2000$. ¿Cuáles son tus conclusiones?

3. Sea Y una variables aleatoria con distribución Binomial Negativa y parámetros $r > 0$ y $\theta > 0$, y $Y_{1:n}$ una muestra aleatoria. Esto es:

$$Y_i \sim \text{NegBin}(r, \theta), \quad i = 1, \dots, n.$$

Para este ejecicio, la parametrización que usaremos para la probabilidad de la Binomial Negativa (con r conocido) es

$$f(y_i \mid \theta) = \binom{y_i + r - 1}{y_i} \left(\frac{\theta}{\theta + r}\right)^{y_i} \left(\frac{r}{\theta + r}\right)^{r}, \quad y_i = 0, 1, 2, \dots$$

La variable Y_i representa el número de fracasos (conteos) antes de observar r éxitos en una secuencia de ensayos de Bernoulli, bajo una probabilidad de éxito implícita que depende de θ, $p = \frac{r}{r+\theta}$ probabilidad de éxito.

Así, $Y_i = y_i$ es un conteo discreto $y_i \in \{0, 1, 2, \dots\}$ y se utiliza comúnmente para modelar datos de conteos con **sobredispersión.** Es decir, cuando la varianza es mayor que la media (por ejemplo, número de casos de enfermedad en un día, número de llamadas en un centro de atención, etc.).

Con esta parametrización los parametros

- $\theta > 0$: controla la *media* de los conteos, ya que

$$\mathbb{E}[Y_i] = \theta.$$

- $r > 0$: es un *parámetro de dispersión* que regula la varianza. En efecto,

$$\text{Var}(Y_i) = \theta + \frac{\theta^2}{r}.$$

a) Genera un conjunto de datos con ésta distribución de tamaño $n = 8$ con $r = 5$ y $\theta = 10$. Calcula la media y varianza de cada Y_i. Calcula el promedio muestral $\bar{Y}$

b) Escriba la función de verosimilitud y la log-verosimilitud de θ.

c) Calcule el estimador de máxima verosimilitud (MLE) de θ.

d) Interprete el resultado obtenido en términos del promedio de los conteos observados.