“Modelos conjugados discretos”, una aplicación Shiny para la enseñanza de estadística bayesiana: teoría.

¡Bienvenido! Este documento explica la teoría básica que sustenta los modelos conjugados que se trabajan en la aplicación basándose en las notas de clase de la profesora Ramírez,Isabel C.(2020).Estadística bayesiana. Además, proporciona instrucciones precisas para el manejo adecuado de la misma con el objetivo de garantizar una experiencia exitosa.

En un curso de estadística bayesiana los modelos conjugados son los primeros modelos en ser estudiados, en esta perspectiva se cuantifica la creencia sobre la ocurrencia de un evento de interés desde un concepto de probabilidad subjetiva. Para ello se establece una distribución de probabilidad para las observaciones de un evento p($\mathbf{y} \mid \theta$), denominada distribución de verosimilitud y otra distribución de probabilidad para sus parámetros correspondientes p(θ), conocidas como distribuciones a priori. Es decir, los parámetros de la distribución de verosimilitud se consideran variables aleatorias. Posteriormente, estas distribuciones se reemplazan en el teorema de Bayes con el propósito de actualizar la creencia a priori e inferir sobre parámetros y observaciones futuras a través de una distribución posterior que pertenece a una familia de distribuciones conocidas.

Considerando lo anterior, el objetivo de esta aplicación es evidenciar cómo los datos observados y la distribución a priori influyen en la distribución posterior.

Tabla de contenido.

1. Inferencia con enfoque Bayesiano

Para realizar un análisis de datos con enfoque Bayesiano se siguen estos pasos:

1. Se define distribución de verosimilitud p($\mathbf{y}$ | θ) y distribución a priori p(θ)

2. Se calcula la distribución conjunta para los datos y parámetros p(θ,y) mediante la definición de probabilidad condicional:

$$\mathrm{p}(\mathbf{y} \mid \theta) = \frac{\mathrm{p}(\theta, \mathbf{y})}{\mathrm{p}(\theta)}$$

$$\mathrm{p}(\theta, \mathbf{y}) = \mathrm{p}(\mathbf{y} \mid \theta)\mathrm{p}(\theta) \tag{1}$$

3. Se reemplaza (1) en el teorema de Bayes para encontrar la distribución posterior de θ condicionada sobre las observaciones:

$$\mathrm{p}(\theta \mid \mathbf{y}) = \frac{\mathrm{p}(\theta, \mathbf{y})}{\mathrm{p}(\mathbf{y})} = \frac{\mathrm{p}(\mathbf{y} \mid \theta)\mathrm{p}(\theta)}{\mathrm{p}(\mathbf{y})}$$

donde $\mathrm{p}(\mathbf{y}) = \sum_\theta \mathrm{p}(\theta)\mathrm{p}(\mathbf{y} \mid \theta)$ si θ es discreto y $\mathrm{p}(\mathbf{y}) = \int_\theta \mathrm{p}(\theta)\mathrm{p}(\mathbf{y} \mid \theta)$ si θ es continuo.

4. Se fija y de forma que:

$$\mathrm{p}(\theta \mid \mathbf{y}) \propto \mathrm{p}(\mathbf{y} \mid \theta)\mathrm{p}(\theta)$$

siendo el lado derecho de la ecuación la distribución posterior no normalizada.

2. Tipos de modelos.

En esta sección se describe la función de masa de probabilidad, según sea el caso, de las distribuciones de verosimilitud, a priori y posterior para cada modelo conjugado que analiza la aplicación. Además, de los valores y parámetros conocidos que deben ser ingresados para el cálculo de estas distribuciones.

2.1. Modelos conjugados discretos.

2.1.1. Modelo binomial.

Distribución de verosimilitud: sea un vector de observaciones $x_1, .., x_n$ donde cada una es el resultado de un ensayo de Benoulli (éxito/fracaso), Y representa el número de éxitos en estos n ensayos de Bernoulli, $y = \sum_i^n x_i$ y su distribución de probabilidad es:

$$\mathrm{p}(Y = y \mid \theta) = \binom{n}{y} \times \theta^y \times (1-\theta)^{n-y}$$

donde $y \sim \text{Binomial}(n, \theta)$, $y = \{0, 1, ..., n\}$ y $\binom{n}{y} = \frac{n!}{y!(n-y)!}$ es el coeficiente binomial

Distribución a priori: $\theta \sim \text{Beta}(\alpha, \beta)$

$$\mathrm{p}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \theta^{\alpha-1} \times (1-\theta)^{\beta-1}$$

donde $\theta \in [0, 1]$ porque representa una tasa, proporción o probabilidad. Además, α es el número de éxitos y β es el número de fracasos.

Distribución posterior: $\theta \mid Y = y \sim \text{Beta}(y + \alpha, n - y + \beta)$

$$\mathrm{p}(\theta \mid Y = y, n) = \frac{\Gamma(\alpha+n+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} \times \theta^{y+\alpha-1} \times (1-\theta)^{n-y+\beta-1}$$

donde $\theta \in (0, 1)$ y $\Gamma(y + \alpha + n - y + \beta) = \Gamma(\alpha + n + \beta)$

Valores y parámetros conocidos:

- **Número de observaciones**: Debe ser un valor mayor a 2 e indica la cantidad de observaciones de la distribución binomial que será utilizada para estimar la distribución de verosimilitud.

- **Número de ensayos Bernoulli(n):** Se requiere mínimo un ensayo (≥ 1) para que tenga sentido la distribución de verosimilitud.

- **Número de éxitos**: número de éxitos en **n** ensayos Bernoulli, el cual debe ser un valor ≥ 0 y es necesario para estimar la probabilidad de éxito de un único ensayo (θ).

Estos valores son necesarios para simular el conjunto de datos en el escenario de simulación.

- $\boldsymbol{\alpha}$: primer parámetro de forma de la distribución a priori. $\alpha > 0$

- $\boldsymbol{\beta}$: segundo parámetro de forma de la distribución a priori. $\beta > 0$

Los parámetros, α y β, definen la forma de la distribución a priori y toman valores mayores a 0.

2.1.2. Modelo Poisson.

Distribución de verosimilitud: la distribución de verosimilitud de un vector de observaciones $\mathbf{y} = (y_1, \ldots, y_n)$ es:

$$\mathrm{p}(\mathbf{y}|\theta) = \frac{e^{-n\theta} \times \theta^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_{i}!}$$

donde $y_i \sim \text{Poisson}(\theta)$.

Distribución a priori: la tasa de ocurrencia promedio de un evento está dada por $\theta \sim Gamma(\alpha, \beta)$:

$$\mathrm{p}(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \times \theta^{\alpha-1} \times e^{-\beta\theta}$$

donde $\theta \in (0, \infty)$ y α representa el número de ocurrencias en β intervalos de tiempo.

Distribución posterior: $\theta|\mathbf{y} \sim \text{Gamma}\left(\sum_{i=1}^{n} y_i + \alpha, n + \beta\right)$, esto es:

$$\mathrm{p}(\theta|\mathbf{y}) = \frac{\beta^{*(\alpha^*)}}{\Gamma(\alpha^*)} \times \theta^{\alpha^*-1} \times e^{-\beta^*\theta}$$

donde $\alpha^* = \sum_{i=1}^{n} y_i + \alpha$ y $\beta^* = n + \beta$

Valores y parámetros conocidos:

- **Número de observaciones (n)**: $n \geq 2$

- $\boldsymbol{\theta}$: parámetro principal de la distribución de verosimilitud. Corresponde a la tasa de ocurrencia promedio de la muestra del evento. $\theta \geq 0$

Ambos valores son necesarios para simular el conjunto de datos en el escenario de simulación.

- $\boldsymbol{\alpha}$: parámetro de forma de la distribución a priori. $\alpha > 0$

- $\boldsymbol{\beta}$: parámetro de escala inversa de la distribución a priori. $\beta > 0$