ACT-11302: Cálculo Actuarial III

–Notas de Clase–

Juan Carlos Martínez-Ovando

ITAM

Primavera 2018

Modelo binomial-beta

Especificación

Modelo binomial/Bernoulli

Definimos un conjunto de variables aleatorias $X_1, \ldots, X_n$, siendo n el tamaño de la muestra.
Suponemos que las X_is son intercambiables, con $X_i \sim \text{Ber}(x|\theta)$ con $\theta \in (0,1)$.
Así, tenemos que condicional en θ,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n|\theta) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{I}_{\{0,1\}}(x_i). \tag{1}$$

Distribución inicial uniforme

Siendo que $\theta \in (0,1)$, podemos suponer que cualquier intervalo de $(0,1)$ de la misma longitud thenga la misma probabilidad de pertenencia para θ, i.e.

$$\mathbb{P}(a < \theta < b) = \mathbb{P}(c < \theta < d), \tag{2}$$

para todo $a < b$ y $c < d$ en el intervalo unitario, tal que $(d-c) = (b-a)$.

Especificación

Modelo binomial/Bernoulli

Definimos un conjunto de variables aleatorias $X_1, \ldots, X_n$, siendo n el tamaño de la muestra.
Suponemos que las X_is son intercambiables, con $X_i \sim \text{Ber}(x|\theta)$ con $\theta \in (0,1)$.
Así, tenemos que condicional en θ,

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n|\theta) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{I}_{\{0,1\}}(x_i). \quad (1)$$

Distribución inicial uniforme

Siendo que $\theta \in (0,1)$, podemos suponer que cualquier intervalo de $(0,1)$ de la misma longitud thenga la misma probabilidad de pertenencia para θ, i.e.

$$\mathbb{P}(a < \theta < b) = \mathbb{P}(c < \theta < d), \quad (2)$$

para todo $a < b$ y $c < d$ en el intervalo unitario, tal que $(d-c) = (b-a)$.

Especificación

Distribución inicial uniforme

La condici"on anterior implica que

$$\pi(\theta) = \mathbb{P}(\theta = z) = 1\, \mathbb{I}_{(0,1)}(z). \tag{3}$$

Reflexionen sobre este resultado.

Distribución conjunta observables y parámetros

Considerando la distriución inicial uniforme, tenemos que la distribución conjunta esta dada por

$$\begin{aligned}
\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta) &= \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n | \theta) \times \pi(\theta) \\
&= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{I}_{\{0,1\}}(x_i)\, \mathbb{I}_{(0,1)}(\theta).
\end{aligned}$$

Especificación

Distribución inicial uniforme

La condici"on anterior implica que

$$\pi(\theta) = \mathbb{P}(\theta = z) = 1\,\mathbb{I}_{(0,1)}(z). \tag{3}$$

Reflexionen sobre este resultado.

Distribución conjunta observables y parámetros

Considerando la distriución inicial uniforme, tenemos que la distribución conjunta esta dada por

$$\begin{aligned}
\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta) &= \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n|\theta) \times \pi(\theta) \\
&= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{I}_{\{0,1\}}(x_i)\,\mathbb{I}_{(0,1)}(\theta).
\end{aligned}$$

Aprendizaje

Teorema de Bayes

Distribución posterior

Empleando el teorema de Bayes, podemos actualizar $\pi(\theta)$ a la luz de la información contenida en la muestra (datos) $X_1 = x_1, \ldots, X_n = x_n$,

$$\begin{aligned}\mathbb{P}(\theta|X_1 = x_1, \ldots, X_n = x_n) &= \frac{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta)}{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)} \\ &\propto \theta^{n_1} (1-\theta)^{n_0} \mathbb{I}_{(0,1)}(\theta),\end{aligned}$$

donde $n_1 = \#\{x_i = 1\}$ y $n_0 = \#\{x_i = 0\}$ (con $n = n_1 + n_0$).

Pregunta

- Es $\theta^{n_1} (1-\theta)^{n_0} \mathbb{I}_{(0,1)}(\theta)$ una función de probabilidad?
- Cómo calcular $\mathbb{P}(x_1, \ldots, x_n)$?

Teorema de Bayes

Distribución posterior

Empleando el teorema de Bayes, podemos actualizar $\pi(\theta)$ a la luz de la información contenida en la muestra (datos) $X_1 = x_1, \ldots, X_n = x_n$,

$$\begin{aligned}\mathbb{P}(\theta | X_1 = x_1, \ldots, X_n = x_n) &= \frac{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta)}{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)} \\ &\propto \theta^{n_1} (1-\theta)^{n_0} \, \mathbb{I}_{(0,1)}(\theta),\end{aligned}$$

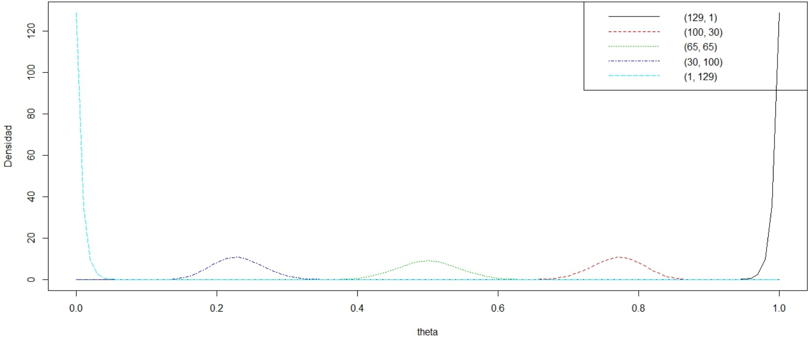
donde $n_1 = \#\{x_i = 1\}$ y $n_0 = \#\{x_i = 0\}$ (con $n = n_1 + n_0$).

Pregunta

- Es $\theta^{n_1} (1-\theta)^{n_0} \, \mathbb{I}_{(0,1)}(\theta)$ una función de probabilidad?
- Cómo calcular $\mathbb{P}(x_1, \ldots, x_n)$?

Teorema de Bayes

Figura: Distribución final de θ para diferentes valores de n_1 y n_0 con distribución inicial uniforme (a.k.a. **función de verosimilitud convecional**)



Teorema de Bayes

Constante de normalización

La constante de normalización de $\mathbb{P}(\theta|X_1 = x_1, \ldots, X_n = x_n)$ no depende de θ. Sin embargo, no es ajena a $\pi(\theta)$, pues se obtiene como el promedio de ésta, i.e.

$$\begin{aligned}
\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) &= \int \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta) \mathrm{d}\theta \\
&= \int_{(0,1)} \theta^{n_1} (1-\theta)^{n_0} \, \mathrm{d}\theta \\
&= \frac{\Gamma(n_1+1)\Gamma(n_0+1)}{\Gamma(n_1+n_0+2)}. \qquad (4)
\end{aligned}$$

Distribución final (revisada)

$$\theta|x_1, \ldots, x_n \sim \mathrm{Be}(\theta|n_1+1, n_0+1). \qquad (5)$$

Teorema de Bayes

Constante de normalización

La constante de normalización de $\mathbb{P}(\theta|X_1 = x_1, \ldots, X_n = x_n)$ no depende de θ. Sin embargo, no es ajena a $\pi(\theta)$, pues se obtiene como el promedio de ésta, i.e.

$$\begin{aligned} \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) &= \int \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n, \theta) \mathrm{d}\theta \\ &= \int_{(0,1)} \theta^{n_1} (1-\theta)^{n_0} \, \mathrm{d}\theta \\ &= \frac{\Gamma(n_1+1)\Gamma(n_0+1)}{\Gamma(n_1+n_0+2)}. \end{aligned} \tag{4}$$

Distribución final (revisada)

$$\theta|x_1, \ldots, x_n \sim \mathsf{Be}(\theta|n_1+1, n_0+1). \tag{5}$$

Conjugacidad

Distribución inicial conjugada

Para el modelo Bernououlli (y binomial también), la distribución inicial conjugada es

$$\pi(\theta) \quad = \quad \text{Be}(\theta | a_0, b_0)$$

donde $a_0, b_0 > 0$ son dos hiperparámetros (parámetros de la distribución inicial, usualmente fijados por nosotros previamente a haber observado los datos).

Distribución final conjugada

Siguiendo el razonamiento anterior, la distribución final conjugada para θ dado un conjunto de datos, $x_1, \ldots, x_n$ es

$$\pi(\theta | x_1, \ldots, x_n) \quad = \quad \text{Be}(\theta | a_0 + n_1, b_0 + n_0)$$

con n_1 y n_0 definidos como antes.

Conjugacidad

Distribución inicial conjugada

Para el modelo Bernouolli (y binomial también), la distribución inicial conjugada es

$$\pi(\theta) \quad = \quad \mathsf{Be}(\theta|a_0, b_0)$$

donde $a_0, b_0 > 0$ son dos hiperparámetros (parámetros de la distribución inicial, usualmente fijados por nosotros previamente a haber observado los datos).

Distribución final conjugada

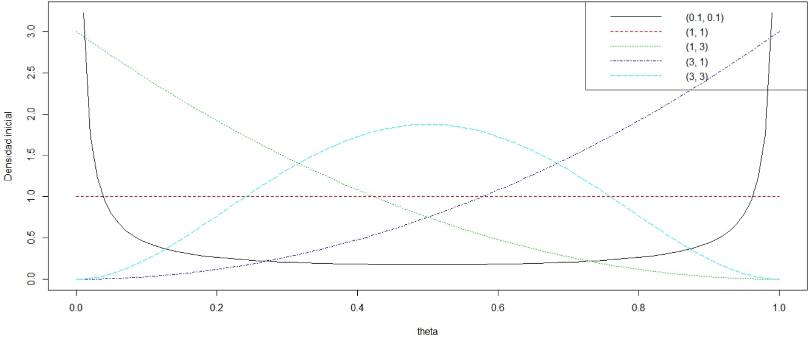
Siguiendo el razonamiento anterior, la distribución final conjugada para θ dado un conjunto de datos, $x_1, \ldots, x_n$ es

$$\pi(\theta|x_1, \ldots, x_n) \quad = \quad \mathsf{Be}(\theta|a_0 + n_1, b_0 + n_0)$$

con n_1 y n_0 definidos como antes.

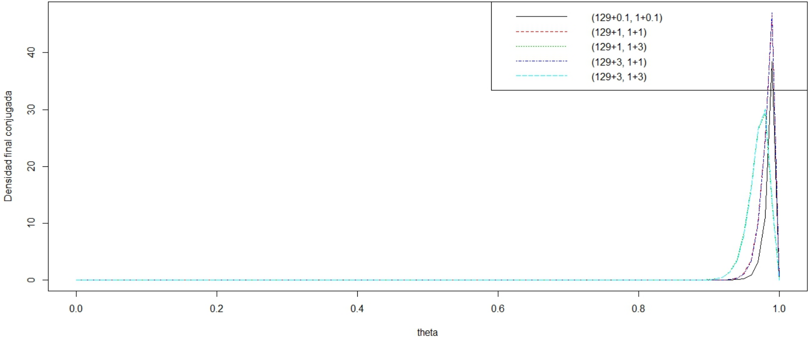
Conjugacidad

Figura: Distribución inicial conjugada para θ con diferentes hiperparámetros a_0 y b_0



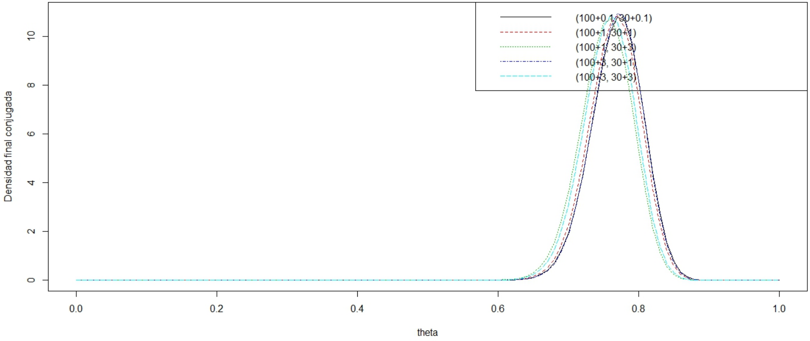
Conjugacidad

Figura: Distribución final conjugada para θ con diferentes hiperparámetros a_0 y b_0 y datos muestrales $n_1 = 129$ y $n_0 = 1$



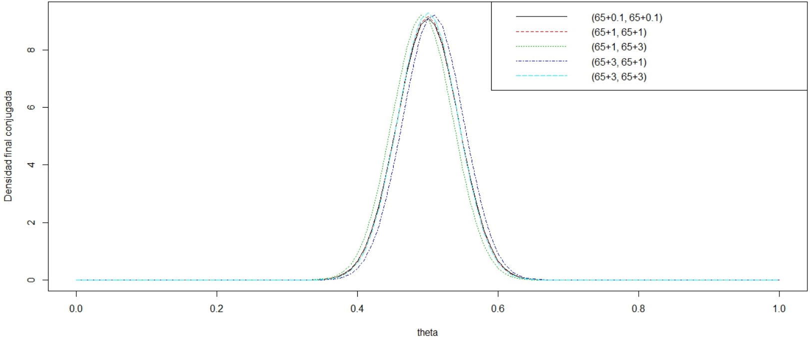
Conjugacidad

Figura: Distribución final conjugada para θ con diferentes hiperparámetros a_0 y b_0 y datos muestrales $n_1 = 100$ y $n_0 = 30$



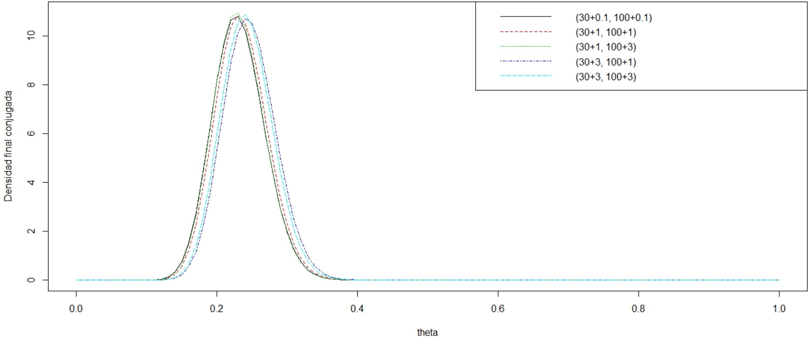
Conjugacidad

Figura: Distribución final conjugada para θ con diferentes hiperparámetros a_0 y b_0 y datos muestrales $n_1 = 65$ y $n_0 = 65$



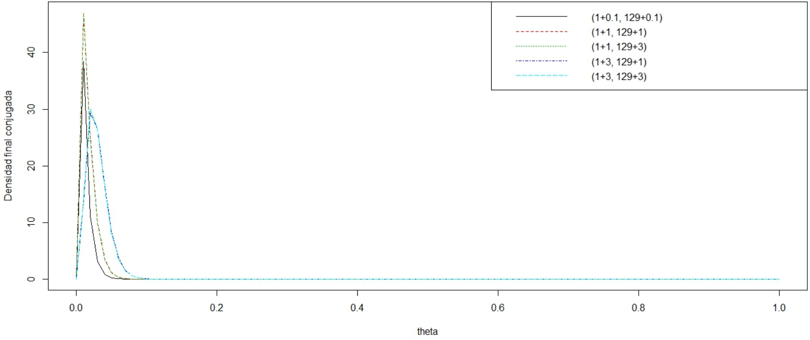
Conjugacidad

Figura: Distribución final conjugada para θ con diferentes hiperparámetros a_0 y b_0 y datos muestrales $n_1 = 30$ y $n_0 = 100$



Conjugacidad

Figura: Distribución final conjugada para θ con diferentes hiperparámetros a_0 y b_0 y datos muestrales $n_1 = 1$ y $n_0 = 129$



Gracias

juan.martinez.ovando@itam.mx