

Conjugacidad con la familia exponencial

Caso Uniparámetro

Una característica de la familia exponencial, es que existe un estadístico suficiente de dimensión fija.

Para la familia exponencial:

$$p(\theta) \propto \exp\{\alpha\phi(\theta) + \beta b(\theta)\}$$

$$\Rightarrow p(\theta|x) \propto \exp\{[\alpha + u(x)]\phi(\theta) + [\beta + 1]b(\theta)\}$$

Sea $k(\alpha, \beta)$ la constante asociada a la definición de $p(\theta)$. Entonces $k(\alpha + u(x), \beta + 1)$ es la constante asociada a $p(\theta, x)$.

De la ecuación $p(\theta|x) = \frac{p(x|\theta).p(\theta)}{p(x)}$, se obtiene:

$$p(x) = \frac{p(x|\theta).p(\theta)}{p(\theta|x)}$$

Sustituyendo las densidades anteriores se obtiene:

$$p(x) = \frac{a(x)\exp\{u(x)\phi(\theta) + b(\theta)\}k(\alpha, \beta)\exp\{\alpha\phi(\theta) + \beta b(\theta)\}}{\exp\{[\alpha + u(x)]\phi(\theta) + [\beta + 1]b(\theta)\}k(\alpha + u(x), \beta + 1)}$$

$$\Rightarrow p(x) = \frac{a(x)k(\alpha, \beta)}{k(\alpha + u(x), \beta + 1)}$$

(Distribución marginal predictiva: $p(x) = \int p(x|\theta).p(\theta)d\theta$).

Ejemplo:

Modelo binomial

$$\begin{aligned} p(x) &= \frac{\binom{n}{x} \theta^x (1-\theta)^{n-x} B^{-1}(\alpha, \beta) \theta^{\alpha-1} (1-\theta)^{\beta-1}}{B^{-1}(\alpha+x, \beta+n-x) \theta^{\alpha+x-1} (1-\theta)^{\beta+n-x-1}} \\ &= \binom{n}{x} \frac{B(\alpha+x, \beta+n-x)}{B(\alpha, \beta)} \end{aligned}$$

Esta es la distribución beta-binomial.

Para una muestra de tamaño n de la familia exponencial obtenemos la densidad conjunta:

$$p(\mathbf{x}|\theta) = \left[\prod_{i=1}^n a(x_i) \right] \exp\left\{ \left[\sum_{i=1}^n u(x_i) \right] \phi(\theta) + nb(\theta) \right\}$$

Si tomamos la distribución a priori conjugada:

$$p(\theta) = k(\alpha, \beta) \exp[\alpha \phi(\theta) + \beta b(\theta)]$$

$$\Rightarrow p(\theta|\mathbf{x}) = k(\alpha + \sum_{i=1}^n u(x_i)) (\beta + n) \times \exp\left\{ \left[\alpha + \sum_{i=1}^n u(x_i) \right] \phi(\theta) + [\beta + n] b(\theta) \right\}$$

Distribución marginal predictiva

$$p(\mathbf{x}) = \frac{\prod_{i=1}^n [a(x_i)] k(\alpha, \beta)}{k(\alpha + \sum u(x_i), \beta + n)}$$

Principales Familias Conjugadas

- Distribución Binomial (ó Bernoulli): La familia Beta es conjugada para la binomial
- Distribución normal con varianza conocida: Para una muestra de tamaño n :

$$l(\theta; \mathbf{x}) \propto \exp\left\{-\frac{n}{2\sigma^2}(\bar{x} - \theta)^2\right\}$$

Si $p(\theta) \sim N(\mu, \tau^2)$, por teorema anterior:

$$\theta|\mathbf{x} \sim N(\mu_1, \tau_1^2)$$

donde:

$$\begin{aligned}\mu_1 &= \frac{n\sigma^{-2}\bar{x} + \tau^{-2}\mu}{n\sigma^{-2} + \tau^{-2}} \\ \tau_1^{-2} &= n\sigma^{-2} + \tau^{-2}\end{aligned}$$

- Distribución Poisson:

Sea $\mathbf{X} = (X_1, \dots, X_n)$ una muestra aleatoria de una Poisson(θ).

Entonces:

$$p(\mathbf{x}|\theta) = \prod_{i=1}^n p(x_i|\theta) = \prod_{i=1}^n \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

$$l(\theta|\mathbf{x}) \propto e^{-n\theta}\theta^{\sum x_i}$$

El kernel de la verosimilitud tiene la forma de una Gamma: $\theta^\alpha e^{-b\theta}$.

Si $\theta \sim G(\alpha, \beta)$

$$p(\theta|\mathbf{x}) \propto \theta^{\alpha + \sum x_i - 1} \exp\{-(\beta + n)\theta\}$$

$$\Rightarrow \text{Gamma}(\alpha + \sum x_i, \beta + n)$$

- Distribución exponencial:

Sea $\mathbf{X} = (X_1, \dots, X_n)$ una muestra aleatoria de una distribución exponencial $Exp(\theta)$.

$$p(\mathbf{x}|\theta) = \theta^n \exp\left\{-\theta \sum_{i=1}^n x_i\right\}$$

El kernel pertenece a la familia Gamma. Si

$\theta \sim Gamma(\alpha, \beta)$ ($p(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta}$)

$$\begin{aligned} p(\theta|\mathbf{x}) &\propto \theta^n \exp\left\{-\theta \sum_{i=1}^n x_i\right\} \theta^{\alpha-1} \exp(-\beta\theta) \\ &= \theta^{n+\alpha-1} \exp\left\{-\left(\sum x_i + \beta\right)\theta\right\} \end{aligned}$$

$$\Rightarrow Gamma(\alpha + n, \beta + \sum x_i)$$

- Distribución Multinomial:

Sea $\mathbf{X} = (X_1, \dots, X_p)$ = número de casos observados y

$\theta = (\theta_1, \dots, \theta_p)$ las probabilidades asociadas a cada categoría en una muestra de tamaño n . Se tienen las restricciones: $\sum_{i=1}^p X_i = n$ y $\sum_{i=1}^p \theta_i$.

$$l(\theta; \mathbf{x}) \propto \prod \theta_i^{x_i}$$

Este es el kernel correspondiente a la familia Dirichlet.

La familia Dirichlet con parámetros a_1, a_2, \dots, a_p es una familia natural conjugada para la multinomial:

$$p(\theta|\mathbf{x}) \propto \left[\prod_{i=1}^p \theta_i^{x_i}\right] \left[\prod_{i=1}^p \theta_i^{a_i-1}\right] \sim Dir(a_1 + x_1, \dots, a_p + x_p)$$

Constante de Proporcionalidad:

$$\frac{\Gamma(a + n)}{\prod_{i=1}^n \Gamma(a_i + x_i)}$$

Esta es la generalización de la binomial con distribución Beta a priori.