Modelos de la Familia Exponencial

Exponential Family Models

Roberto Herrera A.*, Adel Mendoza M.**, Daniel Mendoza C.***

RESUMEN

Muchas de las distribuciones utilizadas en las estadísticas hacen parte de la familia exponencial, dando a entender con ello, una ventaja considerable con respecto a otros modelos que en sí no pertenecen a esta familia, ventaja que se declara en forma significativa cuando se trata de calcular el estadístico $T(\tilde{x})$ de una muestra aleatoria $X_1, X_2, \ldots, X_n$. Entre los modelos que pertenecen a la familia exponencial tenemos la distribución de Poisson, Binomial, Normal, Gamma, Beta, entre otras, esto evidencia la importancia de la familia exponencial en la teoría estadística moderna.

Palabras clave: Estadística suficiente, Familia exponencial, Muestra, Parámetro, Función de probabilidad.

ABSTRACT

Many of the distributions utilized in the statistics do part of the exponential family, implying with it, a substantial advantage with regard to other models that itself do not belong to this family, advantage that is declared in significant form when is a matter of calculating the statistician $T(\tilde{x})$ of a random sample $X_1, X_2, \ldots, X_n$. Among the models that belong to the exponential family we have the distribution Poisson, Binomial, Normal, Gamma, Beta among others, this gives evidence of the importance of the exponential family in the modern statistical theory.

Key words: Sufficient statistic, Family exponential, sampling, Parameter, Probability function.

* *Magíster en Ciencias Estadísticas, Universidad Nacional de Colombia, Docente Asociado Universidad del Atlántico. Grupo de investigación 3i+D, Estadística Industrial y Gestión de la Calidad. robertoherrera@mail.uniatlantico.edu.co*

** *Magíster en Ingeniería Industrial, Universidad del Norte, Docente Asistente Universidad del Atlántico. Grupo de investigación 3i+D. adelmendoza@mail.uniatlantico.edu.co*

*** *Magíster en Ingeniería Industrial, Universidad de los Andes, Docente Asistente Universidad del Atlántico. Grupo de investigación 3i+D. danielmendoza@mail.uniatlantico.edu.co*

Fecha de recepción: Marzo 20 de 2012 • Fecha de aceptación: Abril 30 de 2012

1. INTRODUCCIÓN

Muchas de las distribuciones utilizadas en la estadística [1] hacen parte de una gran familia llamada ***familia exponencial***, implicando con ello, una ventaja sustancial con respecto a otros modelos que no pertenezcan a esta familia, ventaja que se manifiesta en forma significativa cuando se trata de calcular el estadístico $T(\tilde{x})$ de una muestra aleatoria $X_1, X_2, \ldots, X_n$.

Entre los modelos que pertenecen a la familia exponencial tenemos la distribución Poisson, Binomial, Normal, Gamma, Beta, entre otras [2], esto da evidencia de la importancia de la familia exponencial en la teoría estadística moderna [3].

2. DISTRIBUCIONES QUE PERTENECEN A LA FAMILIA EXPONENCIAL

Según [4], la familia de la distribución exponencial son modelos estadísticos de la forma $\rho = \{p_\theta : \theta \in \Theta\}$. Se denomina familia exponencial de un parámetro, si existen funciones de valor real $c(\theta)$, $d(\theta)$, $\theta \in \Theta$, funciones de valor real T y S definidos en R^n y un conjunto $G \subseteq R^n$ que no depende de θ, de tal forma que la función de densidad o de probabilidad $p(\tilde{x}, \theta)$ puede ser discreta de la siguiente forma:

$$p(\tilde{x}, \theta) = \exp\{c(\theta)T(\tilde{x}) + d(\theta) + s(\tilde{x})\} \times \mathrm{I}_G(\tilde{x}) \tag{1}$$

Donde $\mathrm{I}_G(\cdot)$ es una función indicadora de G.

$$\mathrm{I}_G(\tilde{x}) = \begin{cases} 1 & si \quad \tilde{x} \leq 0 \\ 0 & en \quad otro \quad caso \end{cases}$$

En la familia exponencial $T(\tilde{x})$ es un estadístico suficiente y completo. La suficiencia [5] de la estadística $T(\tilde{x})$ con recorrido $\mathrm{I}(= R_T)$ para $\tilde{\theta}$, se da si y solo si existe una función $g(t, \tilde{\theta})$, $t \in \mathrm{I}$, $\tilde{\theta} \in \Theta$ y una función $h(\tilde{x})$, $\tilde{x} \in R^n$, tales que normalmente se toman como,

$$p(\tilde{x}, \tilde{\theta}) = g(T(\tilde{x}), \tilde{\theta}) \times h(\tilde{x}) \tag{2}$$

para toda $\tilde{x} \in R^n$ y $\tilde{\theta} \in \Theta$.

3. PROPIEDADES DE LA FAMILIA EXPONENCIAL

3.1. Determinación de la suficiencia

En la familia exponencial la suficiencia es demostrada de la siguiente manera:

$$g(T(\tilde{x}),\theta) = \exp\{c(\theta)T(\tilde{x}) + d(\theta)\} \qquad (3)$$

$$h(\tilde{x}) = \exp\{s(\tilde{x})\} \times \mathrm{I}_G(\tilde{x})$$

3.2. Completez en la familia exponencial

Una estadística $T(\tilde{x})$ se llama completa si la única función $g(T(\tilde{x}),\tilde{\theta})$ tal que

$$E_\theta\left[g(T(\tilde{x}),\tilde{\theta})\right] = 0, \quad \theta \in \Theta$$

es la función $g \equiv 0$. En consecuencia en la familia exponencial de un parámetro,

$$E_\theta\left[g(T(\tilde{x}),\tilde{\theta})\right] = \int_0^\infty g(t)\exp\{c(\theta)T(\tilde{x}) + d(\theta)\} \qquad (4)$$

La única manera que se haga cero la ec(4) es cuando $g(t) = 0$, así $g \equiv 0$ y $T(\tilde{x})$ es completa.

En la familia Poisson [6], supóngase que se tiene un vector aleatorio adimensional con función de densidad $\vartheta(\lambda)$,

$$p(\tilde{x},\lambda) = \prod_{i=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} \qquad (5)$$

$$p(\tilde{x},\lambda) = \frac{\lambda^{\sum_{i=0}^{\infty} x_i} e^{-n\lambda}}{\prod_{i=0}^{\infty} x!}$$

Reescribiendo en forma exponencial la ec(5) tenemos:

$$p(\tilde{x},\lambda) = \sum_{i=0}^{\infty} x_i \log\lambda - n\lambda - \log\prod_{i=0}^{\infty} x!$$

$$= \sum_{i=0}^{\infty} x_i \log\lambda - n\lambda - \sum \log x_i! \qquad (6)$$

Los parámetros de esta distribución de X son los siguientes,

$$c(\lambda) = \log \lambda$$

$$d(\lambda) = n\lambda$$

$$T(\widetilde{x}) = \sum_{i=0}^{\infty} x_i, \tag{7}$$

$$s(\widetilde{x}) = \sum_{i=0}^{\infty} \log x_i$$

Supóngase un vector aleatorio adimensional de función de densidad Binomial $\beta(n, p)$: Sea X una de las variables del vector con función de probabilidad Binomial [7]

$$p(x, \lambda) = \binom{n}{x} p^x (1-p)^{n-x}$$

En consecuencia en el vector aleatorio la función de probabilidad es de la forma:

$$p(\widetilde{x}, n, p) = \prod_{i=0}^{\infty} \binom{n}{x} p^x (1-p)^{n-x} \tag{8}$$

Aplicando los criterios de la familia exponencial a la ec (8), el modelo es el siguiente:

$$p(\widetilde{x}, n, p) = \log \prod_{i=1}^{\infty} \binom{n}{x} + n^n \log(1-p) + \sum_{i=0}^{\infty} x_i \log \frac{p}{(1-p)} \tag{9}$$

$$= \sum_{i=0}^{\infty} x_i \log \frac{p}{(1-p)} + n^n \log(1-p) + \sum_{i=0}^{\infty} \log \binom{n}{x}$$

Los parámetros para este modelo [8] en su orden son,

$$c(p) = \log \frac{p}{(1-p)},$$

$$d(p) = n \log(1-p) \tag{10}$$

$$T(\widetilde{x}) = \sum_{i=0}^{\infty} x_i, \quad s(\widetilde{x}) = \sum_{i=0}^{\infty} \log \binom{n}{x}$$

Uno de los modelos más aplicados en las Ciencias Estadísticas es el modelo o función de densidad Normal, por lo que la transformación de este modelo a la familia exponencial es sumamente importante [9].

Supóngase un vector aleatorio adimensional de función de densidad $X \sim N\left(\mu, \sigma_0{}^2\right)$, $\sigma_0{}^2$ conocida

El resultado del vector aleatorio es de la forma:

$$p(\widetilde{x}, \mu) = \prod_{i=0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{\frac{1}{2}\left(\frac{x-\mu}{\sigma_0}\right)^2\right\}$$

En consecuencia la función de probabilidad es de la forma,

$$p(\widetilde{x}, \mu) = \frac{1}{\left(\sqrt{2\pi}\right)^n \sigma^n} \exp\left\{\frac{\mu \sum_{i=-\infty}^{\infty} x_i}{\sigma_0^n} - \frac{1}{2\sigma_0^n} \sum_{i=-\infty}^{\infty} x_i - \frac{\mu}{\sigma_0^n}\right\}$$

$$= \exp\left\{\frac{\mu \sum_{i=-\infty}^{\infty} x_i}{\sigma_0^n} - \frac{1}{2\sigma_0^n} \sum_{i=-\infty}^{\infty} x_i - \frac{\mu}{\sigma_0^n} + \log\left[\left(\sqrt{2\pi}\quad\right)^n \sigma_0^n\right]^{-1}\right\} \mathrm{I}_{R^n}(\widetilde{x}) \qquad (11)$$

Los parámetros para el modelo normal en un vector adimensional son:

$$c(\mu) = \frac{\mu}{\sigma_0^2},$$

$$d(\mu) = -\frac{\mu^2}{\sigma_0^2}$$

$$T(\widetilde{x}) = \sum_{i=-\infty}^{\infty} x_i, \qquad (12)$$

$$s(\widetilde{x}) = \frac{1}{2\sigma_0^2} \sum_{i=0}^{\infty} x_i + \log\left[\left(\sqrt{2\pi}\right)^2 \sigma_0^2\right]^{-1}$$

En la **Tabla 1**, se resumen los valores de los parámetros para algunos modelos que pertenecen a esta familia.

Tabla 1. Parámetros de algunos modelos de la familia exponencial

Familia de distribución	$c(\theta)$	$T(\tilde{x})$
POISSON	$\log \lambda$	$\sum_{i=o}^{\infty} x_i$
BINOMIAL	$\log \frac{p}{(1-p)},$	$\sum_{i=0}^{n} x_i$
NORMAL	$\frac{\mu}{\sigma_0^2}$	$\sum_{i=-\infty}^{\infty} x_i$
GAMMA	$-\lambda$	$p-1$
	$p-1$	$\log \sum_{i=0}^{\infty} x_i$
BETA	$s-1$	$\log\left(1-\sum_{i=0}^{\infty} x_i\right)$
	$r-1$	$\log \sum_{i=0}^{\infty} x_i$

Fuente: Elaboración del autor

4. REPARAMETRIZACIÓN DE LA FAMILIA EXPONENCIAL

Sea $\tilde{X}$ vector aleatorio adimensional con función de densidad o de probabilidad [10] en la familia exponencial

$$p(\tilde{x},\theta)$$

$$= \exp\{c(\theta)T(\tilde{x}) + d(\theta) + s(\tilde{x})\} \times \mathrm{I}_G(\tilde{x}) \tag{13}$$

Sea $\eta = c(\theta)$, $\theta \in \Theta$. Entonces se puede ver que,

$$p(\tilde{x},\eta) = \exp\{\eta T(\tilde{x}) + d_0(\eta) + s(\tilde{x})\} \times \mathrm{I}_G(\tilde{x}) \qquad (14)$$

Esta expresión se conoce como la reparametrización natural de $p\,(\tilde{x},\theta)$, donde,

$$d_0(\eta) = -\log \int \exp\{\eta T(\tilde{x}) + s(\tilde{x})\} d(\tilde{x}) \qquad (15)$$

Esta integral es reemplazada por sumatoria en caso de una función de probabilidad.

Si c es uno a uno, el cálculo de $d_o(\eta)$ no es necesario realizarlo, ya que $\eta = c(\theta)$, $c^{-1}(\eta) = \theta$, $d(\theta) = d(c^{-1}(\eta)) = d_0(\eta)$

Un ejemplo de esta reparametrización es la función de densidad de Rayleigh, que es aplicada para fenómenos estocásticos de "tiempo de falla" [11].

Supóngase $X_1,...,X_n$ una muestra aleatoria con función de densidad:

$$p(x,\theta) = \left(\frac{x}{\theta^2}\right)\exp\left(\frac{-x^2}{2\theta^2}\right),$$

$$x > 0, \theta > 0$$

Realizando los cálculos para la familia exponencial.

$$p(\tilde{x},\theta) = \prod_{i=1}^{n}\left(\frac{x}{\theta^2}\right)\exp\left(\frac{-x^2}{2\theta^2}\right)$$

$$\begin{aligned} p(\tilde{x},\theta) &= \frac{1}{\theta^{2n}}\prod_{i=1}^{n} x_i \exp\left\{-\frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2\right\} \\ &= -\log\theta^{2n} + \log\prod_{i=1}^{n} x_i - \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i \\ &= \exp\left[\frac{-1}{2\theta^2}\sum_{i=1}^{n} x_i^2 - n\log\theta^2 + \sum_{i=1}^{n}\log x_i\right] \end{aligned} \qquad (16)$$

Reparametrizando

$$\eta = c(\theta) = \frac{-1}{2\theta^2}, \quad \theta^2 = \frac{-1}{2\eta},$$

$d(\theta) = -\eta \log \theta^2$

Por lo que $d_0(\eta) = n \log(-2n)$

$E_\eta(T(\tilde{x})) = -d_0'(\eta)$

$$-\frac{\partial d_0}{\partial \eta} = \frac{\partial[n \log(-2\eta)]}{\partial \eta}$$

$$= -n\left[\frac{1}{-2\eta}(-2)\right] = -\frac{n}{\eta} = E(T(\tilde{x})) \qquad (17)$$

Al parametrizar la ec(17) se obtiene,

$$-\frac{n}{\eta} = \frac{-n}{-\frac{1}{2\theta^2}} = E_\theta(T(\tilde{x})) = 2n\theta^2$$

Para la varianza

$$\mathrm{var}_\eta(T(\tilde{x})) = -d_0''(\eta)$$

$$= \frac{-\partial}{\partial \eta}\left(-\frac{n}{\eta}\right) = \frac{n}{\eta^2}$$

$$\mathrm{var}_\eta(T(\tilde{x})) = \frac{n}{\eta^2}$$

Parametrizando se obtiene la siguiente expresión:

$$\mathrm{var}_\eta(T(\tilde{x})) = \frac{n}{\left(-\frac{1}{2\theta^2}\right)} = 4n\theta^2$$

$$\mathrm{var}_\eta(T(\tilde{x})) = 4n\theta^2 \qquad (18)$$

Con esta reparametrización se consigue de una manera sencilla la estimación de θ, veamos:

$$E_\theta(T(\tilde{x})) = T(X) = 2n\theta^2$$

Como $T(X)=\sum_{i=1}^{n} x_i^2$, entonces,

$$2n\theta^2 = \sum_{i=1}^{n} x_i^2$$

$$\theta = \sqrt{\frac{1}{2\eta}\sum_{i=1}^{n} x_i^2} \qquad (19)$$

5. CONCLUSIONES

Modificar una función de densidad o de probabilidad hacia la familia exponencial permite de una manera sencilla evaluar los estadísticos suficientes, y con ello determinar los estadísticos que permiten estimar los valores de los parámetros θ. Por ejemplo, en el modelo Normal [12], el estadístico es $\sum_{i=-\infty}^{\infty} x_i$, esta información permite calcular el estimador insesgado y completo en función de μ, de la forma $\bar{x} = \frac{1}{n}\sum_{i=-\infty}^{\infty} x_i$, conocido como promedio muestral.

El mismo análisis se debe realizar para la familia Binomial en donde el estadístico $\sum_{i=0}^{n} x_i$, calcula la suma de todos los éxitos del fenómeno estocástico y $\hat{p} = \frac{1}{n}\sum_{i=0}^{n} x_i$ es el estimador insesgado para n ensayos de la muestra aleatoria.

6. REFERENCIAS BIBLIOGRÁFICAS

[1] M. Levy, “A note on nonunique MLEs and sufficient statistics”, *Annals of Mathematical Statistics*, vol. 39, pp. 66-80, 1985.

[2] A. Mayorga, “Inferencia Estadística. Colección Textos”, *Universidad Nacional de Colombia*, vol. 1, pp. 88-100, 2004.

[3] W. Navidi, “A Graphical Illustration of the EM Algorithm”, *Annals of Mathematical Statistics*, 51(1), pp. 30-50, 1997.

[4] O. Barndorff N., *Information and Exponential Families in Statistical Theory*. New York: John Wiley, 1978.

[5] E. Dynkin, Necessary and sufficient statistics for families of distributions, *Uspekhi Mat. Nauk*, vol. 6, pp. 68-90, 1951.

[6] H. Mann, y A. Wald, “On stochastic limit and order relationships”, *Annals of Mathematical Statistics*, vol. 14, pp. 217-250, 1943.

[7] P. Bickel, y K. Doksum, *Mathematical Statistics.* San Francisco: Holden-Day, 1977.

[8] C. Rao, “Efficient Estimates and Optimum Inference Procedures in Large Samples”, *Journal of the Royal Statistical Society*, vol. 24, pp. 46-55, 1962.

[9] S. Shapiro, y R. Francia, “An Approximate Analysis of Variance Test for Normality”, *Journal of the American Statistical Association*, vol. 67, pp. 215-230, 1972.

[10] C. Mallows, “Some comments on Cp”. *Technometrics*, vol. 15, pp. 661-675, 1973.

[11] D. R. Cox, and D. Hinkley, *Theoretical Statistics*. London: Chapman and Hall, 1974.

[12] C.R. Rao, *Linear Statistical Inference and its Applications*. New York: Wiley, 1965.