Verosimilitud extendida

El proceso de extender la verosimilitud consiste en tratar al vector aleatorio X como una variable mezcla, donde la variable de mezcla u se considera como latente. Este método pretende encontrar los mejores estimadores para la variable de mezcla u, y a su vez que esta variable de la mejor probabilidad de haber observado a la variable X.

Por ejemplo, supongamos que X es una variable aleatoria Pareto(1,β), de la cual tenemos una realización X_0, y que nos interesa estimar el valor de β, luego sea $X|u$ una distribución exponencial con parámetro λ, y además supóngase que λ también es una variable exponencial con parámetro β. Entonces, X se puede expresar como una mezcla de distribuciones exponenciales como sigue:

$$.f(x) = \int_0^\infty f_{x|\lambda}(x) f(\lambda) d\lambda = \int_0^\infty \lambda \exp^{-\lambda x} \beta \exp^{-\beta\lambda} d\lambda = \frac{\beta}{(x+\beta)^2}$$

Ahora, si consideramos la distribución de $X|\lambda_i$, para cada λ_i tendríamos una densidad diferente. En la siguiente figura se ilustra la densidad de $X|\lambda_i$, para algunos valores de λ_i.

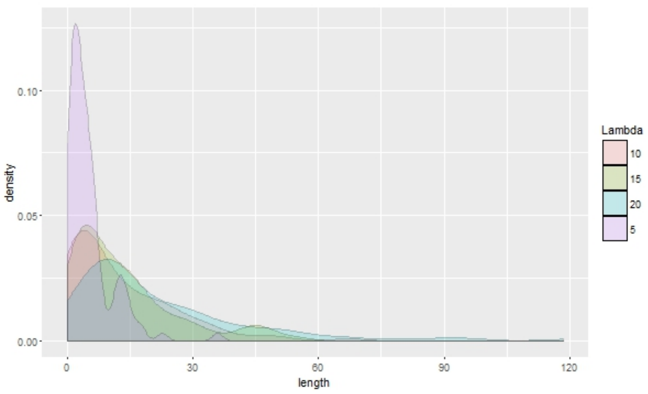

Figure 1:

De aquí, nos podríamos plantear la siguiente pregunta ¿Dada una observación X_0, cual sería la λ_i que le asigne mayor probabilidad u densidad a la ocurrencia de esta observación. Por ejemplo, en la gráfica anterior se puede observar que si $x_0 = 5$, la mejor lambda sería 5, mientras que si $x_0 = 30$, sería 15. En la misma línea, también nos podríamos preguntar por el mejor estimador de β donde la probabilidad de que ocurra la λ_i (que maximiza la la probabilidad de ocurrencia de x_0) sea máxima.

Entonces, si consideramos la función de verosimilitud para X_0, sería:

$$.Lik(\beta|X) = f_x(x) = \frac{\beta}{(x+\beta)^2} = \int\limits_0^\infty f_{x|\lambda}(x) f(\lambda) d\lambda$$

Y con el planteamiento anterior, en lugar de buscar el mejor estimador de β en Lik($\beta|X_0$), podríamos preguntarnos por la mejor λ_i inducida por β tal que β maximice la ocurrencia de λ_i, y a su vez, λ_i maximice la probabilidad de haber observado X_0, entonces con este enfoque, la función de verosimilitud para X_0 queda expresada de la siguiente madera:

$$.Lik(\lambda_i, \beta|X_0) = f_{x|\lambda_i}(x_0) f(\lambda i)$$

De mamera general, la verosimilitud extendida se define de la siguiente forma:

Sea X un vector aleatorio de dimención p con algún vector de parámetros θ, y sea u una variable aleatoria con algún vector de parámetros β, con soporte en R^+, tal que $f(x) = \int\limits_0^\infty f_{x|u}(x) f(u) du$, con $X_i|u_i \perp X_j|u_j$, y $u_i|\beta \perp u_j|\beta \; \forall i \neq j$, entonces:

$$.Lik(u_i, \theta, \beta|X) = \prod_{i=1}^{p} f_{x|u_i}(x) f(u_i)$$

donde u_i es una variable latente.

Trabajar con la verosimilitud extendida podría parecer más complicado al tener que estimar una variable latente para cada observación de X, pero en ocasiones resulta más sencillo optimizar la función $\prod_{i=1}^{p} f_{x|u_i}(x) f(u_i)$ que $f(X)$, y además, existen algoritmos computacionales, vía cadenas demarkov, que permiten encontrar los estimadores de los parámetros deseados.