# Modelos para datos continuos: MLG Gamma

Recuerde que $Y$ ~Gamma($\theta$,$\nu$), la fdp de $Y$ puede ser escrita como

$$f_Y(y) = \exp\left\{\frac{y\left(-\frac{1}{\mu}\right) - \log(\mu)}{\frac{1}{\nu}} + \nu\log\nu + (\nu-1)\log y - \log\Gamma(\nu)\right\},$$

donde $\mu = \nu/\theta$. El parámetro canónico es $-1/\mu$, entonces la función de enlace es

$$g(\mu) = -\frac{1}{\mu}.$$

El parámetro de dispersión es $\phi = 1/\nu$.

Asuma que se tiene v.a.s independientes $Y_1, \ldots, Y_n$ cada una con distribución Gamma. Si las v.a.s tienen un parámetro de dispersión común ($\phi$), luego también tienen parámetro de forma común ($\nu$), dado que $\nu = 1/\phi$. Si las v.a.s tienen medias distintas ($\mu_i$), entonces tienen distintos parámetros de escala ($\theta_i$), dado que $\theta_i = \nu/\mu_i$. En otras palabras, bajo estos supuestos, $Y_i \sim$Gamma($\theta_i$,$\nu$). Asumiendo que estas variables explicativas $\mathbf{x}_i$ pueden tomar en cuenta estas diferencias en las medias $\mu_i$'s, podemos usar el MLG para modelar estos datos:

$$g(\mu_i) = \sum_{j=1}^{p} x_{ij}\beta_j.$$

Note que el supuesto de que $Y_i \sim$Gamma($\theta_i$,$\nu$) implica que las $Y_i's$ tienen *coeficiente de variación*, $C_i$, constante:

$$\begin{aligned} C_i &= \frac{\text{s.d.}(Y_i)}{\mu_i} \\ &= \frac{\sqrt{\mu_i^2/\nu}}{\mu_i} = \frac{1}{\sqrt{\nu}}. \end{aligned}$$

Elección de la función de enlace

Las tres funciones de enlace comúnmente usadas para MLG gamma son

1. *inverse* : $g(\mu_i) = \frac{1}{\mu_i}$
2. *log* : $g(\mu_i) = \log \mu_i$
3. *identity* : $g(\mu_i) = \mu_i$

Estas funciones de enlace son implementadas en R.

**NOTA:** La función de enlace canónica ($g(\mu_i) = -\frac{1}{\mu_i}$) no está en esta lista! Sin embargo, la función de enlace canónica es equivalente a la *inverse* en el sentido de que si

$$g(\mu_i) = -\frac{1}{\mu_i} = \sum_{j=1}^{p} x_{ij}\beta_j$$

entonces

$$\begin{aligned}\frac{1}{\mu_i} &= \sum_{j=1}^{p} x_{ij}(-\beta_j) \\ &= \sum_{j=1}^{p} x_{ij}\beta_j^*,\end{aligned}$$

donde $\beta_j^* = -\beta_j$. La función de enlace *inverse* es más usada que la canónica.

**IMPORTANTE:** *NO* podemos realizar las pruebas usuales para evaluar la bondad de ajuste en base a la devianza cuando el parámetro de dispersión es desconocido. En particular, sabemos que

$$D \sim \chi^2_{n-p},$$

pero no podemos calcular $D$ si $\phi$ no es conocido. Por otro lado, para comparar dos modelos anindados, no podemos analizar el cambio de la devianza relativa a la distribución $\chi^2$. Sin embargo, si podemos realizar pruebas de Wald (multivariadas, si es necesario) para analizar formalmente la significancia de cada variable explicativa. Además podemos analizar el cambio de la devianza relativa a la distribución $F$.

Recuerde que, bajo la hipótesis nula el modelo propuesto se ajusta tan bien como el modelo saturado, la devianza $D$, tiene distribución asintótica

$$D \sim \chi^2_{n-p}, \tag{1}$$

donde $n$ es el número de observaciones y $p$ es el número de coeficientes de regresión.

Por otro lado, el cambio de devianza cuando sacamos q variables explicativas del modelo , $\Delta D$, tiene distribución asintótica

$$\Delta D \sim \chi^2_q \tag{2}$$

bajo la hipótesis nula de que el modelo reducido se ajusta tan bien como el modelo completo.

Recuerde que además $D = D^*/\phi$, donde $D^*$ es el "residual deviance", y $\phi$ es el parámetro de dispersión. De forma similar, $\Delta D = \Delta D^*/\phi$, donde $\Delta D^*$ es la "deviance" enla tabla de ANOVA.

Si el parámetro de dispersión es conocido (es decir, en elcaso de Poisson o binomial), podemos calcular $D$ y usar (1) para probar el ajuste global del modelo. Y, podemos calcular $\Delta D$ y usar (2) para realizar la prueba de significancia de $q$ variables explicativas via el comando en R `anova` con el argumento `test="Chisq"`.

En el caso donde $\phi$ no es conocido (es decir en los casos Gamma o normal), no podemos usar estos resultados para probar el ajuste global del modelo Sin embargo, *podemos* probar la significancia de las $q$ variables explicativas usando el comando `anova` con el argumento `test="F"`.

Para entender esto, recuerde que si $X_1 \sim \chi^2_q$, $X_2 \sim \chi^2_p$, y $X_1$ y $X_2$ son independientes, entonces

$$\frac{X_1/q}{X_2/p} \sim F_{q,p}.$$

Por lo tanto, asumiendo que el modelo completo se ajusta tan bien como el modelo saturado, asintóticamente,

$$\frac{\Delta D^*/q}{D^*/(n-p)} = \frac{(\Delta D^*/\phi)/q}{(D^*/\phi)/(n-p)} = \frac{\Delta D/q}{D/(n-p)} \sim F_{q,n-p} \qquad (3)$$

bajo la hipótesis nula de que el modelo reducido se ajusta tan bien como el modelo completo.

Este resultado es exacto en el caso de la normal. Note que la prueba F-test es también valida si $\phi$ es conocido.

**RESUMEN:** Si $\phi$ es conocido, la prueba $\chi^2$ y F-test son válidas, y son asintóticamente equivalentes. Sin embargo, la prueba $\chi^2$ es usualmente más precisa, y por ellos es preferible al F-test en este caso. Si $\phi$ no es conocido, solo el F-test es posible, ya que (a diferencia de la prueba $\chi^2$ ) no depende de $\phi$.

Podemos usar el comando `anova y F-test para realizar estas pruebas de hipótesis.`