

Modelo Lineal Generalizado

y

Cálculo de la Provisión Técnica

Eva Boj del Val

Teresa Costa Cor

Departamento de Matemática Económica, Financiera y Actuarial
Facultad de Economía y Empresa
Universidad de Barcelona

Enero de 2014

Índice

1. Modelo lineal generalizado	3
1.1. Descripción del modelo	3
1.2. Estimación de parámetros.....	5
1.3. Desviación y residuos	6
1.4. Función de enlace “Identidad” y “Logarítmica”	8
1.5. Familia paramétrica de distribuciones	8
2. Cálculo de la provisión técnica	10
2.1. Datos	10
2.2. Caso particular: Modelo Poisson “sobredisperso”	10
2.3. Error de predicción. Fórmula “analítica”	11
2.4. El modelo lineal generalizado y la metodología bootstrap. Distribución predictiva....	12
Bibliografía	16

1. Modelo lineal generalizado

1.1. Descripción del modelo

Supongamos una variable aleatoria $\mathbf{Y}_{(n \times 1)}$, con (y_i) para $i=1,2,\dots,n$ observaciones independientes, que recogen la siniestralidad a explicar y que juegan el papel de variable respuesta en el modelo. Supongamos P predictores o factores potenciales de la estructura de riesgo $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_P$, vectores $(n \times 1)$: (f_{ij}) para $i=1,2,\dots,n$ y $j=1,2,\dots,P$. Recordemos el modelo clásico de regresión lineal por mínimos cuadrados ordinarios, en el que suponemos una distribución del error ε_i Normal centrada y con varianza constante, $\varepsilon_i \sim N(0, \sigma^2)$. La relación lineal de la respuesta con la estructura sistemática dada por los predictores es:

$$y_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij} + \varepsilon_i. \quad (1)$$

Disponemos de observaciones independientes $y_i \sim N(\mu_i, \sigma^2)$, con esperanza

$$E[y_i] = \mu_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij}, \quad (2)$$

y varianza constante

$$\text{Var}[y_i] = \sigma^2. \quad (3)$$

En el Modelo Lineal Generalizado (MLG) tenemos de nuevo (y_i) para $i=1,2,\dots,n$ observaciones independientes de la respuesta, y un predictor lineal determinista al que simbolizamos por η_i :

$$\eta_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij}. \quad (4)$$

Las dos extensiones respecto del modelo clásico son:

- 1) La distribución de \mathbf{Y} no tiene por qué ser la Normal. Consideramos en primer lugar que puede ser cualquier distribución derivada de la familia exponencial de McCullagh y Nelder [ver McCullagh y Nelder (1989) p.28]. Estas distribuciones se caracterizan por tener una función de densidad o de probabilidad en un punto de la forma:

$$f(y_i; \theta_i, \phi_i) = \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \right\}, \quad (5)$$

para funciones especificadas $a(\cdot)$, $b(\cdot)$ y $c(\cdot)$, donde θ_i se denomina parámetro canónico y ϕ_i parámetro de dispersión. Puede deducirse fácilmente a partir de la formulación más general que:

$$E[y_i] = \mu_i = b'(\theta_i), \quad (6)$$

$$Var[y_i] = b''(\theta_i) a(\phi_i). \quad (7)$$

Por tanto, la varianza es el producto de dos componentes:

- la primera, $b''(\theta_i)$, que depende únicamente de θ_i (y por tanto de la esperanza μ_i). A esta componente se le denomina *función de varianza* y se explicita su dependencia respecto de la esperanza: $b''(\theta_i) = V(\mu_i)$
- la segunda, $a(\phi_i)$, depende sólo del parámetro de dispersión ϕ_i y usualmente adopta la forma $a(\phi_i) = \frac{\phi}{w_i}$, con parámetro de dispersión constante para todas las observaciones, ϕ , y unos pesos especificados *a priori*, w_i , que varían de observación a observación.

Teniendo en cuenta lo anterior podemos describir:

$$Var[y_i] = \frac{\phi V(\mu_i)}{w_i}, \quad (8)$$

donde ϕ es el parámetro de dispersión, $V(\mu_i)$ es la función de varianza y w_i es el posible peso especificado *a priori* de la observación i . Notamos que suponiendo $\phi=1$, los recíprocos de los pesos pueden reinterpretarse como parámetros de escala no constantes:

$$\frac{1}{w_i} = \phi_i.$$

- 2) La respuesta está ligada con el predictor lineal a través de una función F :

$$E[y_i] = \mu_i = F(\eta_i) = F\left(\beta_0 + \sum_{j=1}^P \beta_j f_{ij}\right). \quad (9)$$

Para despejar el predictor lineal debemos hacer la función inversa de F , $g = F^{-1}$, a la que denominamos *función de enlace (link function)*, pues es la que nos enlaza la respuesta con el predictor lineal. A la función de enlace, g , le exigimos que sea monótona y diferenciable:

$$g(E[y_i]) = g(\mu_i) = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij}. \quad (10)$$

Existen, para algunas distribuciones de la familia exponencial, funciones de enlace “naturales”, también denominadas canónicas. Para estos enlaces canónicos se produce que el parámetro canónico coincide con el predictor lineal: $\theta(\mu_i) = \eta_i$.

Pero, en general, podemos optar por modelizar con cualquier otro enlace que no sea el canónico. Es usual utilizar uno derivado de la familia de enlaces paramétricos:

$$\eta_i = g(\mu_i) = \begin{cases} \mu_i^\lambda & \text{para } \lambda \neq 0 \\ \log(\mu_i) & \text{para } \lambda = 0. \end{cases} \quad (11)$$

En la Tabla 1 del Anexo se recogen las distribuciones más conocidas que forman parte de la familia exponencial definida en (5), junto con el enlace canónico asociado.

Resumiendo, en el MLG, tenemos dos extensiones respecto al modelo de regresión lineal clásico, una respecto a la distribución del error plasmada en $Var(\mathbf{Y})$, que podrá proceder de cualquier distribución de la familia exponencial y no tiene por qué ser la Normal; y otra respecto a la función de enlace, $g(\boldsymbol{\mu})$, que debe ser una función monótona diferenciable y no tiene por qué ser la Identidad.

Comentario: Al aplicar el MLG a unos datos, decidimos sobre la distribución del error y sobre la función de enlace a utilizar en el modelo. La elección del enlace canónico tiene la ventaja de simplificar la formulación, pero no tiene por qué implicar que sea el más adecuado para unos datos particulares. Si el objetivo es seleccionar un modelo, la simplicidad de la función de enlace no debe sustituir a la calidad del ajuste como criterio.

1.2. Estimación de parámetros

La estimación de los parámetros β_j del predictor lineal se realiza mediante la maximización del logaritmo de la función de verosimilitud:

$$l = \sum_{i=1}^n l_i = \sum_{i=1}^n \left\{ \frac{y_i \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \right\}. \quad (12)$$

Cabe notar que suponiendo una distribución del error Normal y la función Identidad como función de enlace, obtenemos como caso particular la solución por mínimos cuadrados ordinarios del modelo clásico.

La versión más general del MLG no exige que la distribución del error de \mathbf{Y} pertenezca a la familia de distribuciones exponenciales caracterizadas por (5). En dicha versión más general seguimos teniendo (y_i) para $i = 1, 2, \dots, n$ observaciones independientes de la respuesta, y para éstas se conocen sólo los dos primeros momentos:

$$E[y_i] = \mu_i, \quad (13)$$

$$Var[y_i] = \frac{\phi V(\mu_i)}{w_i}, \quad (14)$$

siendo $V(\mu_i)$ una función de varianza especificada, ϕ el parámetro de dispersión también especificado (en general positivo), y w_i los posibles pesos *a priori* de las observaciones.

En este contexto más amplio, aplicable a distribuciones que no pertenezcan a la familia exponencial, la estimación de los parámetros β_j se realiza maximizando los logaritmos de las funciones de cuasi-verosimilitud:

$$q(\mathbf{y}; \boldsymbol{\mu}) = \sum_{i=1}^n q_i = \sum_{i=1}^n w_i \int_{y_i}^{\mu_i} \frac{y_i - s}{\phi V(s)} ds. \quad (15)$$

Tal maximización nos lleva a resolver el sistema de ecuaciones lineales

$$\sum_{i=1}^n w_i \frac{y_i - \mu_i}{\phi V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0 \quad \forall j \quad (16)$$

mediante algún método numérico. Para este caso general de la familia exponencial, las cuasi-

verosimilitudes juegan el papel de verosimilitudes.

1.3. Desvianza y residuos

En el MLG la variabilidad no explicada por un modelo (fijada una función de enlace, una distribución del error y unos predictores) se plasma en la *desvianza escalada* $D^*(\mathbf{y}; \hat{\boldsymbol{\mu}})$. Si $L(\mathbf{y}; \hat{\boldsymbol{\mu}})$ denota la función de verosimilitud del modelo en estudio y $L(\mathbf{y}; \mathbf{y})$ la función de verosimilitud del modelo saturado¹³, entonces:

$$D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) = -2 \log \left(\frac{L(\mathbf{y}; \hat{\boldsymbol{\mu}})}{L(\mathbf{y}; \mathbf{y})} \right). \quad (17)$$

La desvianza escalada disminuye a mayor número de número de predictores incluidos en el modelo, hasta llegar a explicar la variabilidad total de los datos. En la Tabla 2 del Anexo, detallamos la expresión que toma en algunos de los casos particulares de la familia exponencial, en términos de desvianzas no escaladas:

$$D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = \phi D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}). \quad (18)$$

Podemos escribir las desvianzas en función de las cuasi-verosimilitudes descritas en (15):

$$D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = \sum_{i=1}^n d_i = \sum_{i=1}^n 2w_i \frac{y_i - s}{V(s)} d(s) = -2\phi q(\mathbf{y}; \hat{\boldsymbol{\mu}}). \quad (19)$$

Para el cálculo de las desvianzas asociadas a un modelo tan sólo necesitamos conocer los dos primeros momentos.

Se definen, entre otros, dos tipos de residuos:

- Residuos de Pearson:
$$r_i^P = \frac{y_i - \hat{\mu}_i}{\left(\frac{V(\hat{\mu}_i)}{w_i} \right)^{1/2}}, \quad (20)$$

- Residuos de desvianza:
$$r_i^D = \text{signo}(y_i - \hat{\mu}_i) \sqrt{d_i} \quad (21)$$

¹³ El modelo saturado es el que tiene tantos parámetros como individuos, y por tanto se cumple $\hat{\mu}_i = y_i$ para $i=1,2,\dots,n$.

(donde d_i es la i -ésima componente de (19)).

Para la estimación del parámetro de dispersión ϕ de un modelo con $P+1$ coeficientes podemos utilizar la fórmula:

$$\hat{\phi}^D = \frac{1}{n-P-1} \sum_{i=1}^n (r_i^D)^2 = \frac{1}{n-P-1} \sum_{i=1}^n d_i. \quad (22)$$

O alternativamente el estimador de momentos basado en los residuos generalizados de Pearson:

$$\hat{\phi}^P = \frac{1}{n-P-1} \sum_{i=1}^n (r_i^P)^2 = \frac{1}{n-P-1} \sum_{i=1}^n w_i \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)}. \quad (23)$$

Comentario: Usualmente se indica que el modelo lineal generalizado más apropiado para unos datos es aquél que nos ofrece una menor desviación. Como se intuye, tenemos diferentes maneras de reducirla: si variamos la función de enlace; si variamos la distribución del error; y/o si variamos los factores de riesgo incluidos en el predictor lineal.

1.4. Función de enlace “Identidad” y “Logarítmica”

En el MLG obtenemos un modelo aditivo si combinamos cualquier distribución del error con el enlace *identidad*, y obtenemos un modelo multiplicativo si utilizamos el enlace *logarítmico*. Según (10), que nos relaciona la respuesta con el predictor lineal, tenemos que para el enlace *identidad*:

$$\mu_i = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij}. \quad (24)$$

Y para el enlace *logarítmico*:

$$\log(\mu_i) = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j f_{ij},$$

$$\mu_i = \exp\left(\beta_0 + \sum_{j=1}^P \beta_j f_{ij}\right) = e^{\beta_0} \times e^{\beta_1 f_{i1}} \times e^{\beta_2 f_{i2}} \times \dots \times e^{\beta_p f_{ip}}. \quad (25)$$

Observamos que si partimos de la base β_0 en el modelo aditivo debemos sumar (o restar) las P cantidades $\sum_{j=1}^P \beta_j f_{ij}$. Y si partimos de la base e^{β_0} en el multiplicativo, debemos multiplicar por las P cantidades de incremento (o decremento) $\prod_{j=1}^P e^{\beta_j f_{ij}}$. En el caso particular de predictores binarios, partiendo del efecto global, sumaremos β_j o multiplicaremos por e^{β_j} , sólo cuando para el individuo i se dé la característica j , para $j=1, \dots, P$, ya que en tal caso se cumplirá $f_{ij} = 1$.

1.5. Familia paramétrica de distribuciones

Al igual que con la función de enlace, donde podemos trabajar con la familia paramétrica (11), también podemos utilizar para la distribución del error la de familia paramétrica

$$V(\mu_i) = \mu_i^\zeta. \quad (26)$$

Observamos que con esta familia de distribuciones del error obtenemos algunos casos particulares de la familia exponencial, por ejemplo:

- ⇒ Si $\zeta = 1$, la estructura de error es Poisson
- ⇒ Si $\zeta = 2$, la estructura de error es Gamma
- ⇒ Si $\zeta = 3$, la estructura de error es Inversa Gaussiana

Referencia: Se puede encontrar una descripción detallada del modelo lineal generalizado en McCullagh and Nelder (1989), y una descripción de su aplicación en tarificación *a priori* en la referencia Boj *et al.* (2004), de la cual se ha extraído el resumen de este primer apartado.

2. Cálculo de la provisión técnica

2.1. Datos

Denotamos por c_{ij} a las cuantías de los siniestros ocurridos en el año de origen i y pagados en el año de desarrollo j , con $i = 0, 1, \dots, k$ y $j = 0, 1, \dots, k$. El triángulo *run-off* que recoge dichos datos es el siguiente:

		Años de desarrollo, $j = 0, 1, \dots, k$					
		0	1	2	...	$k-1$	k
Años de origen, $i = 0, 1, \dots, k$	0	$c_{0,0}$	$c_{0,1}$	$c_{0,2}$...	$c_{0,(k-1)}$	$c_{0,k}$
	1	$c_{1,0}$	$c_{1,1}$	$c_{1,2}$...	$c_{1,(k-1)}$	
		
	$k-2$	$c_{k-2,0}$	$c_{k-2,1}$	$c_{k-2,2}$			
	$k-1$	$c_{k-1,0}$	$c_{k-1,1}$				
	k	$c_{k,0}$					

2.2. Caso particular: Modelo Poisson “sobredisperso”

Si utilizamos el MLG para modelar dichos datos en el cálculo de la provisión técnica, y aplicamos el caso en que la distribución del error es Poisson “sobredispersa” junto con la función de enlace canónica logarítmica, obtenemos como caso particular el modelo Chain-Ladder clásico determinista, es decir, la estimación de los pagos futuros coincide en ambos modelos.

El modelo Poisson “sobredisperso” aplicado a las cuantías del triángulo supone los siguientes valores para la distribución:

$$\mu_{ij} = E[c_{ij}] \quad V(\mu_{ij}) = \mu_{ij} \quad \phi > 1 \quad w_{ij} = 1.$$

Con lo que:

$$\mu_{ij} = E[c_{ij}] \quad \text{Var}[c_{ij}] = \phi\mu_{ij}.$$

Y combinado con la función de enlace logarítmica:

$$\log \mu_{ij} = \eta_{ij}.$$

Concretamente, el predictor lineal es de la forma $\eta_{ij} = c_0 + \alpha_i + \beta_j$ siendo α_i el factor correspondiente a los años de origen $i = 1, \dots, k$, β_j el factor correspondiente a los años de desarrollo $j = 1, \dots, k$, y c_0 el término que se correspondería al año de origen y desarrollo 0. Cabe notar que en la construcción de las estimaciones el término c_0 siempre debe ser añadido en la predicción, junto con el coeficiente correspondiente al año de origen y el correspondiente al año desarrollo. A partir de aquí podemos realizar las estimaciones de las cuantías del resto del rectángulo mediante la expresión:

$$\hat{c}_{ij} = \exp(c_0 + \alpha_i + \beta_j),$$

y con ellas calcular los pagos futuros por año de origen, \hat{R}_i , sumando las cuantías estimadas por filas para $i = 1, \dots, k$, y los pagos futuros totales, \hat{R} , como la suma de los individuales.

2.3. Error de predicción. Fórmula “analítica”

Podemos calcular el error cometido en la predicción de los pagos futuros utilizando la fórmula analítica del modelo que tiene en cuenta la distribución supuesta en los datos, mediante las expresiones que a continuación detallamos (nos referimos a England y Verrall (1999), (2002) y (2006) para un mayor detalle). O bien se pueden calcular haciendo uso de metodología *bootstrap*, como explicaremos en el siguiente apartado.

Nuestro objetivo es calcular el error de predicción, el cuál vendrá recogido por el error cuadrático medio (mean square error, MSE, en inglés) de predicción, definido y calculado aproximadamente como:

$$MSE(\hat{c}_{ij}) = E[(c_{ij} - \hat{c}_{ij})^2] \cong \text{Var}[c_{ij}] + \text{Var}[\hat{c}_{ij}]. \quad (27)$$

Observamos que el error de predicción (27) es la suma de dos componentes: la primera se

refiere a la variabilidad de los datos (varianza del proceso) y la segunda a la variabilidad de la estimación (varianza de la estimación). La primera la estimaremos con la fórmula de la varianza de la distribución y la segunda utilizando la aproximación

$$Var[\hat{c}_{ij}] \cong \left| \frac{\partial \mu_{ij}}{\partial \eta_{ij}} \right|^2 Var[\eta_{ij}],$$

de lo que se obtiene:

- Error cuadrático medio para cada estimación de los pagos futuros:

$$MSE(\hat{c}_{ij}) \cong \phi \hat{\mu}_{ij} + \hat{\mu}_{ij}^2 Var[\hat{\eta}_{ij}]. \quad (28)$$

- Error cuadrático medio para la estimación de los pagos futuros por año de origen:

$$MSE(\hat{R}_i) \cong \sum_{j=k-i+1}^k \phi \hat{\mu}_{ij} + \sum_{j=k-i+1}^k \hat{\mu}_{ij}^2 Var[\hat{\eta}_{ij}] + 2 \sum_{\substack{j_1=k-i+1 \\ j_2=k-i+1 \\ j_2 > j_1}}^k \hat{\mu}_{ij_1} \hat{\mu}_{ij_2} Cov[\hat{\eta}_{ij_1}, \hat{\eta}_{ij_2}] \quad i=1, \dots, k. \quad (29)$$

- Error cuadrático medio para la estimación de los pagos futuros totales:

$$MSE(\hat{R}) \cong \sum_{i=1}^k \sum_{j=k-i+1}^k \phi \hat{\mu}_{ij} + \sum_{i=1}^k \sum_{j=k-i+1}^k \hat{\mu}_{ij}^2 Var[\hat{\eta}_{ij}] + 2 \sum_{i=1}^k \sum_{\substack{j_1=k-i+1 \\ j_2=k-i+1 \\ j_2 > j_1}}^k \hat{\mu}_{ij_1} \hat{\mu}_{ij_2} Cov[\hat{\eta}_{ij_1}, \hat{\eta}_{ij_2}]. \quad (30)$$

2.4. El modelo lineal generalizado y la metodología *bootstrap*. Distribución predictiva.

La estimación de los errores estándar de predicción (raíz cuadrada del MSE (28), (29) ó (30)) la podemos realizar alternativamente haciendo uso de la metodología de remuestreo *bootstrap*. Con esta metodología obtenemos además la distribución predictiva de los pagos futuros individuales, por año de origen y totales. De estas distribuciones, si es de interés, podemos calcular estadísticos como la media, la desviación estándar, cuantiles, asimetría, curtosis, realizar histogramas, etc. Nos referimos al capítulo 10 de Kaas *et al.* (2008) dónde encontramos con detalle el código en R del cálculo del error estándar de predicción tanto con fórmula analítica como con uso de *bootstrap* y, también, del análisis estadístico de la distribución predictiva de los pagos futuros obtenida mediante *bootstrap*.

La idea o procedimiento general de la metodología *bootstrap* (en este texto como caso particular del denominado *bootstrapping residuals*) es la siguiente:

- Se aplica el modelo Chain-Ladder clásico al triángulo de los datos originales de cuantías acumuladas, se desacumulan y se calculan los residuos de Pearson (20),

$$r_{ij}^P = \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}}}.$$

- Dichos residuos se remuestran B veces y con ellos se construyen B nuevas muestras de triángulos de cuantías no acumuladas teniendo en cuenta la fórmula (20) de la que se despeja

$$c_{ij}^* = r_{ij}^{P*} (\hat{c}_{ij})^{1/2} + \hat{c}_{ij}.$$

- Se estima el modelo lineal generalizado Poisson “sobredisperso” (o Chain-Ladder clásico) a cada una de las B muestras y con él los pagos futuros individuales, por año de origen y totales.
- El resultado es que disponemos de B valores de los pagos futuros por año de origen y totales, lo que nos proporciona la distribución predictiva de dichos pagos. Al tener la distribución predictiva podemos (si es de interés) calcular estadísticos informativos como son por ejemplo la media, la desviación estándar o cuantiles de las distribuciones.
- Respecto al error de predicción (prediction error, PE, en inglés) estimado mediante metodología *bootstrap*, éste se corresponde con la raíz cuadrada de la suma de la varianza de los pagos futuros según la distribución supuesta (en este caso Poisson) más la varianza estimada del error mediante *bootstrap*, de forma que:

$$PE_{boot}(c_{ij}) \cong \sqrt{\hat{\phi}^P \hat{c}_{ij} + \frac{n}{n-p} (SE_{boot}(c_{ij}))^2}, \quad (31)$$

$$PE_{boot}(R_i) \cong \sqrt{\hat{\phi}^P \hat{R}_i + \frac{n}{n-p} (SE_{boot}(R_i))^2}, \quad i=1, \dots, k, \quad (32)$$

$$PE_{boot}(R) = \sqrt{\hat{\phi}^P \hat{R} + \frac{n}{n-p} (SE_{boot}(R))^2}, \quad (33)$$

donde el parámetro de dispersión se estima con (23) y SE_{boot} significa error estándar

(standard error, SE, en inglés) de la distribución *bootstrap* considerada (pagos futuros individuales, por año de origen o totales). Observamos que para la estimación del error estándar se utiliza una estimación corregida por sesgo, siendo $n - p$ el número de grados de libertad, con $p = 2k + 1$ el número de parámetros del modelo.

Comentarios: en el proceso de simulación de residuos y de estimación de los B modelos involucrados en el proceso *bootstrap* debemos tener en cuenta la distribución del error supuesta (si es Poisson como es el caso que hemos detallado o por ejemplo si es Gamma), de este modo aplicaremos la fórmula correspondiente en la función de varianza.

Alternativamente, en el proceso *bootstrap* se puede trabajar con los residuos de Pearson “escalados”, los cuales incluyen en el denominador también el parámetro de dispersión que multiplica a la varianza (14) de la v.a.:

$$\frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\phi \hat{c}_{ij}}}.$$

O bien se puede incluir la corrección por sesgo al igual que cuando estimamos el parámetro de escala:

$$\sqrt{\frac{n}{n-p}} \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}}}.$$

En tal caso no sería necesario incluir la corrección por sesgo en el cálculo del error estándar estimado con *bootstrap*.

Finalmente, cabe notar que a mayor B mejor ajuste del método de remuestreo. Si se utiliza el caso Poisson “sobredisperso”, a mayor B , la estimación de los pagos futuros medios debería parecerse cada vez más a la del modelo Chain-Ladder.

ANEXO. Tablas de propiedades para casos particulares del MLG

	Normal o Gaussiana $N(\mu, \sigma^2)$	Binomial $B(m, \pi) / m$	Poisson $P(\mu)$	Gamma $G(\mu, \nu)$	Gaussiana inversa $GI(\mu, \sigma^2)$
Rango de y	$(-\infty, +\infty)$	0,1,...,m/m	0,1,2,..., ∞	$(-\infty, +\infty)$	$(-\infty, +\infty)$
Peso w	1	1	1	1	1
Parámetro de dispersión: ϕ	σ^2	$1/m$	1	ν^{-1}	σ^2
$b(\theta)$	$\frac{\theta^2}{2}$	$\log(1 + e^\theta)$	$\exp(\theta)$	$-\log(-\theta)$	$-(-2\theta)^{1/2}$
$c(y; \theta)$	$-\frac{1}{2} \left(\frac{y^2}{\phi} + \log(2\pi\phi) \right)$	$\log \binom{m}{my}$	$-\log y!$	$\nu \log(\nu y) - \log(y) - \log \Gamma(\nu)$	$-\frac{1}{2} \left(\log(2\pi\phi y^3) + \frac{1}{\phi y} \right)$
$\mu(\theta) = E(Y; \theta)$	θ	$\frac{e^\theta}{(1 + e^\theta)}$	$\exp(\theta)$	$\frac{-1}{\theta}$	$(-2\theta)^{-1/2}$
Función de enlace canónica: $\theta(\mu)$	Identidad: μ	Logit: $\log(\mu/(1-\mu))$	Logarítmico: $\log(\mu)$	Recíproco: $1/\mu$	$1/\mu^2$
Función de varianza: $V(\mu)$	1	$\mu(1-\mu)$	μ	μ^2	μ^3

Tabla 1. Principales características.

Distribución:	Desvianzas D :
Normal o Gaussiana $N(\mu, \sigma^2)$	$\sum_{i=1}^n (y_i - \hat{\mu}_i)^2$
Binomial $B(m, \pi) / m$	$2 \sum_{i=1}^n \{ y_i \log(y_i / \hat{\mu}_i) + (m - y_i) \log[(m - y_i) / (m - \hat{\mu}_i)] \}$
Poisson $P(\mu)$	$2 \sum_{i=1}^n \{ y_i \log(y_i / \hat{\mu}_i) - (y_i - \hat{\mu}_i) \}$
Gamma $G(\mu, \nu)$	$2 \sum_{i=1}^n \{ -\log(y_i / \hat{\mu}_i) + (y_i - \hat{\mu}_i) / \hat{\mu}_i \}$
Gaussiana inversa $GI(\mu, \sigma^2)$	$\sum_{i=1}^n (y_i - \hat{\mu}_i)^2 / (\hat{\mu}_i^2 y_i)$

Tabla 2. Desvianzas.

Referencias

Boj, E., Claramunt, M.M. and Fortiana, J. (2004). *Análisis multivariante aplicado a la selección de factores de riesgo en la tarificación*. Cuadernos de la Fundación MAPFRE, número 88. Fundación MAPFRE Estudios, Madrid.

http://www.mapfre.com/documentacion/publico/i18n/catalogo_imagenes/grupo.cmd?path=1050569

McCullagh, P. and Nelder, J. (1989). *Generalized linear models (2nd edition)*. Chapman and Hall, London.

England, P.D. and Verrall, R.J. (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics* 25, 281–293.

England, P.D. and Verrall, R.J. (2002). Stochastic claims reserving in general insurance (with discussion). *British Actuarial Journal* 8, 443–544.

England, P.D. and Verrall, R.J. (2006). Predictive Distributions of Outstanding *Liabilities in General Insurance*. *Annals of Actuarial Science* 1 (II), 221–270.

Kaas, R., Goovaerts, M., Dhaene, J. and Denuit, M. (2008). *Modern actuarial risk theory: using R. Second edition*. Springer, Heidelberg.