¿QUÉ SON LOS MODELOS LINEALES GENERALIZADOS (GLM)?

ALFREDO VASQUEZ S. (*)

RESUMEN.

Se define la clase de modelos llamados "Modelos Lineales Generalizados" compuestos de componentes sistemática y aleatoria. La componente aleatoria especifica que la distribución de las observaciones, pertenece a la familia exponencial. La componente sistemática es lineal o una transformación de ella es lineal. Los estimadores de los parámetros de la parte sistemática son obtenidos por la generalización del método iterativo usado en el probit análisis, dicha generalización se debe a Nelder. Se define la desvianza para generalizar el análisis de varianza.

(*) Departamento Matemáticas, Facultad de Ciencias, Universidad del Norte.

INTRODUCCION.

Los modelos lineales son combinaciones de componentes sistemáticas y aleatorias, por ejemplo, el modelo lineal clásico:

$$Y_i = \sum_j \alpha_j X_{ij} + \epsilon_i$$

está compuesto de

- i) $\sum_j \alpha_j X_{ij}$, que es una componente lineal sistemática con parámetros α_j que pueden ser conocidos o desconocidos pero estimables. Las variables X_{ij} son independientes de carácter cuantitativo, cualitativo o mixtas.
- ii) ϵ_i , es la componente aleatoria que especifica la distribución de las observaciones, así por ejemplo, si $\epsilon_i \sim N(0, \sigma^2)$ entonces $Y_i \sim N(\sum_j \alpha_j X_{ij}, \sigma^2)$.

El supuesto básico en el modelo lineal clásico es la distribución normal de la componente error, dándole también esta característica a la variable observada "Y"; pero existen aplicaciones distintas en que la linealidad está comprometida y el supuesto en cuestión no es realístico. En el probit análisis se observan conteos y es más adecuado suponer una distribución binomial (el parámetro p es una función de la tolerancia ψ , y esta última es lineal en la dosis o es una función de ella) de forma análoga, en las tablas de contingencia se observa frecuencias que se adecúan a una distribución multinomial y la parte sistemática es usualmente multiplicativa (lineal en logaritmo); por último, la clase que se basa en la distribución χ^2 o Gamma para estimar las componentes de la varianza cuyo modelo tiene como componente sistemática una estructura lineal.

Luego el problema consiste en encontrar un conjunto de modelos con componente sistemática lineal que nos de una respuesta adecuada

a cada uno de los requerimientos de distribución de las observaciones para cada situación específica. La respuesta fue presentada por Nedler y Wedderburn en el paper "Generalized Linear Models" (J.R. St. Soc.A, 1972) y sobre cuya base se estructurará esta monografía. A continuación daremos la idea en que se basa la generalización y la definición de GLM. Posteriormente se mostrará el modo de ajustar los modelos en base al criterio de máxima verosimilitud.

2.- GENERALIZACION.

Del estudio del modelo lineal clásico, vemos que las observaciones Z tienen distribución normal y ésta pertenece a la familia exponencial, como también la binomial, Poisson, Gamma, por lo tanto parece lógico derivar una generalización como sigue:

Sea Z una variable aleatoria con función de densidad

$$f(Z, \theta, \phi) = \exp [\alpha(\phi)\{Z\theta - g(\theta) + h(Z)\} + \beta(\phi, Z)]$$

donde:

- i) $\alpha(\phi) > 0$, y para cada valor fijo de ϕ tenemos que $f(Z, \theta, \phi)$ es una familia exponencial. El parámetro ϕ puede considerarse como un parámetro de perturbación, por ejemplo, la varianza en la distribución normal o p en la distribución Gamma.
- ii) $\theta = g(Y)$ es una transformación lineal de las medias, es decir, $g(Y) = \sum \alpha_i X_i$. Esta función g se llama función encadenadora y su función inversa existe.

Las tres funciones encadenadoras que cubren la mayoría de las aplicaciones son:

Identidad: $g = I$, es decir, $Y = \sum \alpha_i X_i$

Logaritmo: $\log Y = \sum \alpha_i X_i$ (para estructuras multiplicativas de θ con $\theta > 0$)

Logit : $\log \left(\frac{Y}{1-Y} \right) = \sum \alpha_i X_i$ (estructura binomial)

también señalaremos que estas funciones encadenadoras son transformaciones de la media teórica, más que transformaciones de los datos originales.

3.- DEFINICION.

Un Modelo Lineal Generalizado (GLM) está caracterizado por:

- i) Una variable dependiente Z (observación) con alguna distribución perteneciente a $f(Z, \theta, \phi)$.
- ii) Un conjunto de variables, independientes $X_1, X_2, \dots, X_n$, tales que $Y = \sum \beta_i X_i$ donde β_i pueden ser parámetros conocidos o desconocidos pero estimables, recordaremos que las variables X 'S pueden ser cuantitativas, cualitativas o bien mixtas.
- iii) Una función encadenadora $\theta = g(Y)$ que conecta el parámetro θ de la distribución de Z , con los Y 'S del modelo lineal.

El procedimiento para ajustar estos modelos a observaciones se basa en las ecuaciones de máxima verosimilitud y es una generalización de la estimación máximo verosimil del probit análisis. (Nelder, 1968, Biom. 23). Para este método se usa el procedimiento iterativo de Newton Raphson para obtener resultados para la primera y segunda derivada del logaritmo de la función máximo verosimil en términos de la media (μ), varianza de Z (V) y el factor escalar $\alpha(\phi)$.

Es bien conocido que:

$$E \left(\frac{\partial L}{\partial \theta} \right) = 0 \quad \text{y} \quad E \left(\frac{\partial^2 L}{\partial \theta^2} \right) = - E \left(\frac{\partial L}{\partial \theta} \right)^2$$

Sea $Z_1, Z_2, \dots, Z_n$ una muestra aleatoria, entonces

$$L = \alpha(\phi) [\{\theta \sum Z_i - ng(\theta) + \sum h(Z_i)\} + \sum \beta(\phi, Z_i)]$$

(familia exponencial)

$$\frac{\partial L}{\partial \theta} = \alpha(\phi) \{\sum Z_i - ng'(\theta)\}$$

$$\frac{\partial^2 L}{\partial \theta^2} = -\alpha(\phi) ng''(\theta)$$

$$E\left(\frac{\partial L}{\partial \theta}\right) = \alpha(\phi) (n\mu - ng'(\theta)) = 0 \quad \text{esto implica}$$

$$\mu = g'(\theta)$$

$$E\left(\frac{\partial L}{\partial \theta}\right)^2 = \alpha^2(\phi) n^2 E(\bar{Z} - g'(\theta))^2 = n\alpha^2(\phi) V(Z) =$$

$$= -E\left(\frac{\partial^2 L}{\partial \theta^2}\right) = \alpha(\phi) ng''(\theta)$$

$$\text{Luego: } \alpha^2(\phi) V(Z) = \alpha(\phi) g''(\theta)$$

En caso que $\alpha(\phi) = 1$ se tiene

$$g''(\theta) = V$$

$$\text{Además } g''(\theta) = \frac{dg'(\theta)}{d\theta} = \frac{d\mu}{d\theta} = V$$

Por ejemplo, para una familia exponencial de un parámetro ($\alpha(\phi) = 1$) se tiene:

$$f(Z, \theta) = \exp \{Z\theta - g(\theta) + h(Z)\}$$

$$\frac{\partial L}{\partial \theta} = Z - g'(\theta) = Z - \mu$$

$$-\frac{\partial^2 L}{\partial \theta^2} = g''(\theta) = V = \text{var}(Z)$$

4.- PROPOSICION.

La siguiente proposición provee de un método iterativo para obtener las soluciones de las ecuaciones máximo verosímiles para los parámetros de un G.L.M.

Proposición: La solución de las ecuaciones máximo verosímiles es equivalente a un proceso mínimo cuadrático iterativo con una función de peso

$$w = \left(\frac{d\mu}{dY} \right)^2 / V$$

y una variable dependiente modificada (la variable de trabajo probit en el probit análisis):

$$y = Y + (Z - \mu) / \left(\frac{d\mu}{dY} \right)$$

donde , Y y V están basados en los estimadores corrientes.

Dem.: $L = \ln f = \alpha(\phi) [Z\theta - g(\theta) + h(Z)] + \beta(\phi, Z)$

$$\frac{\partial L}{\partial \beta_i} = \frac{\partial L}{\partial \theta} \cdot \frac{\partial \theta}{\partial \mu} \cdot \frac{\partial \mu}{\partial Y} \cdot \frac{\partial Y}{\partial \beta_i} = \alpha(\phi) [Z - g'(\theta)] \frac{\partial \theta}{\partial \mu} \cdot \frac{\partial \mu}{\partial Y} x_i$$

pero $V = \frac{d\mu}{d\theta}$ $g'(\theta) = \mu$

$$\frac{\partial L}{\partial \beta_i} = \alpha(\phi) \frac{(Z - \mu)}{V} \cdot \frac{d\mu}{dY} x_i$$

$$\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} = \frac{\partial}{\partial \beta_i} \left(\frac{\partial L}{\partial \beta_j} \right) = \frac{\partial}{\partial \beta_i} \left(\frac{\partial L}{\partial Y} \frac{\partial Y}{\partial \beta_j} \right) = \frac{\partial^2 L}{\partial Y^2} x_i x_j$$

$$\frac{\partial^2 L}{\partial Y^2} = \frac{\partial}{\partial Y} \left(\frac{\partial L}{\partial \theta} \cdot \frac{\partial \theta}{\partial Y} \right) = \frac{\partial^2 L}{\partial \theta^2} \left(\frac{\partial \theta}{\partial Y} \right)^2 + \frac{\partial L}{\partial \theta} \left(\frac{\partial^2 \theta}{\partial Y^2} \right)$$

$$\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} = \left(\frac{\partial^2 L}{\partial \theta^2} \left(\frac{\partial \theta}{\partial \beta} \right)^2 + \frac{\partial L}{\partial \theta} \left(\frac{\partial^2 \theta}{\partial \beta^2} \right) \right) x_i x_j$$

pero $\frac{\partial \theta}{\partial Y} = \frac{d\theta}{d\mu} \cdot \frac{d\mu}{dY}$

$$\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} = (-\alpha(\phi) V \left(\frac{d\theta}{d\mu} \right)^2 \left(\frac{d\mu}{dY} \right)^2 + \alpha(\phi) (Z - \mu) \frac{d^2 \theta}{dY^2}) x_i x_j$$

$$E \left(\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = - \frac{\alpha(\phi)}{V} \left(\frac{d\mu}{dY} \right)^2 x_i x_j$$

Sea $W = \left(\frac{d\mu}{dY} \right)^2 \frac{1}{V}$, entonces:

$$\frac{\partial L}{\partial \beta_i} = \frac{\alpha(\phi) W x_i (Z - \mu)}{d\mu/dY}$$

y $E \left(- \frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = \alpha(\phi) W x_i x_j$

Usando el proceso de Newton-Raphson con la esperanza en las segundas derivadas en una muestra de tamaño n se tiene:

$$\frac{\partial L}{\partial \beta_i} = \alpha(\phi) \sum_{K=1}^n w_K x_{iK} (z_K - \mu_K) / \left(\frac{d\mu_K}{dY_K} \right)$$

$$E \left(- \frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = \alpha(\phi) \sum_{K=1}^n w_K x_{iK} x_{jK} = A_{ij}$$

o bien

$$\sum_j A_{ij} (\delta \beta^\circ)_j = \frac{\partial L}{\partial \beta_i} = \sum_K w_K x_{iK} (z_K - \mu_K) / \left(\frac{d\mu_K}{dY_K} \right)$$

más

$$\begin{aligned}
 \sum_j A_{ij} (\beta^\circ)_j &= \sum_j \left(\sum_K W_K X_{jK} (\beta^\circ)_j \right) = \\
 &= \sum_{K=1}^n W_K X_{iK} \left(\sum_j X_{jK} (\beta^\circ)_j \right) = \\
 &= \sum_{K=1}^n W_K X_{iK} Y_K
 \end{aligned}$$

Se obtiene:

$$\sum_j A_{ij} (\beta^\circ + \delta\beta)_j = \sum_j A_{ij} \beta_j^1 = \sum_{K=1}^n W_K X_{iK} \left(Y_K + \frac{Z_K - \mu_K}{\left(\frac{d\mu_K}{dY_K}\right)} \right)$$

o bien:

$$\sum_j A_{ij} \beta_j^1 = \sum_K W_K X_{iK} Y_K, \quad Y_K = Y_K + \frac{Z_K - \mu_K}{\left(\frac{d\mu_K}{dY_K}\right)}$$

donde, $\beta_j^1 = \beta^\circ + \delta\beta^\circ$ es el nuevo estimador de β_i .

La última ecuación puede ser escrita:

$$\sum_j \left(\sum_K W_K X_{iK} X_{jK} \right) \beta_j = \sum_K W_K X_{iK} Y_K$$

$$\sum_K W_K X_{iK} \left(\sum_j X_{jK} \beta_j \right) = \sum_K W_K X_{iK} Y_K$$

que es una regresión lineal ponderada.

La solución de las ecuaciones máximo verosímiles es la solu
ción mínimo cuadrática cuando

$$W = \left(\frac{d\mu}{dY} \right)^2 / V = \frac{1}{V(Y)}$$

y se calcula como primera aproximación haciendo $V(Y) = V / \left(\frac{d\mu}{dY} \right)^2$.

Enseguida procedemos a la estimación usando el modelo lineal que minimiza

$$\sum_K W_K (y_K^* - y_K)^2 \quad \text{con} \quad y_K = \sum_j \beta_j x_{jK}$$

se obtiene la solución correcta previo reemplazo de y_K^* por:

$$y_K = y_K + \frac{Z_K - \mu_K}{(d\mu_K/dy_K)}$$

Para realizar el ajuste se debe:

- 1.- Hacer $\mu = Z$ como primera aproximación y a partir de ésta calcular Y , entonces calcular W y hacer $y = Y$.
- 2.- Obtener la primera aproximación a los β 's por regresión.
- 3.- Para la distribución binomial modificar los valores extremos de Z , reemplazar $Z = \frac{1}{2}$ para el valor cero y $Z = n - \frac{1}{2}$ para el valor n .
- 4.- Usar modificaciones adecuadas en aquellos casos que conduzcan a valores infinitos para Y .

Ejemplo: Probit Análisis (Nelder 68 - Biom 23)

El problema consiste en que n sujetos son expuestos a un estímulo. Sea X la dosis a medir y r la respuesta.

Supuesto : r binomial con $E(r/n) = P$ y

$$P = \int_{-\alpha}^Y \rho(\mu) d\mu \quad \text{donde} \quad \rho(\mu) \quad \text{es una función de densidad de}$$

probabilidad y $Y = \beta_0 + \beta_1 X$ ($X = 0$ ó 1).

Luego:

$$P = g(Y) = g(\beta_0 X_0 + \beta_1 X_1)$$

usando la forma homogénea en X_0 y X_1 ($X_0 = 1$ y $X_1 = 0$ ó 1).

La función máxima verosimil

$$L = r \ln P + (n - r) \ln Q$$

$$\frac{\partial L}{\partial \beta_i} = \frac{p - P}{V(P)} g'(Y) X_i$$

y

$$E \left(\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = - \frac{[g'(Y)]^2}{V(P)} X_i X_j$$

Luego usando el proceso de Newton-Raphson que requiere valores iniciales $b_0^{(0)}$ y $b_1^{(0)}$, después los ajustes $\delta b_0^{(0)}$ y $\delta b_1^{(0)}$ se calculan de:

$$\sum a_{ij} \delta b_j^{(0)} = \frac{\partial L}{\partial \beta_i} \quad \text{y tomando el valor de la esperanza,}$$

$$a_{ij} = E \left(\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = \frac{(g'(Y))^2}{V(P)} X_i X_j = -W X_i X_j$$

entonces:

$$\sum_j \left(\sum_i W X_i X_j (b_j^{(0)} + (\delta b)_j^0) \right) = \sum_i W X_i \left(Y + \frac{p - P}{g'(Y)} \right)$$

$$\sum_i \left(\sum_j W X_i X_j b_j^{(1)} \right) = \sum_i W X_i \left(Y - \frac{p - P}{g'(Y)} \right)$$

que es el proceso iterativo.

5.- CASO ESPECIAL.

Cuando el parámetro de la distribución de Z y el modelo lineal Y coinciden, es decir,

$$\Theta = Y = \sum_j \beta_j X_j$$

la función máximo verosimil para una muestra de tamaño n es:

$$L = \alpha(\phi) \left\{ \sum_{K=1}^n (Z_K Y - g(Y) + h(Z_K)) \right\}$$

y

$$\frac{\partial L}{\partial \beta_i} = \alpha(\phi) \left\{ \sum_{K=1}^n Z_K X_{iK} - \sum_{K=1}^n g'(Y_K) \cdot X_{iK} \right\} = 0$$

pero $g'(Y_K) = \mu_K$

$$\text{Luego: } \sum Z_K X_{iK} - \sum \hat{\mu}_K X_{iK} = 0$$

$$\text{o bien: } \sum Z_K X_{iK} = \sum_{K=1}^n \hat{\mu}_K X_{iK}$$

Esta última ecuación dice que para una variable cualitativa independiente ($X_{iK} = 0$ ó $X_{iK} = 1$) los totales marginales ajustados con respecto a esa variable será igual a los observados para esa misma variable.

Por el teorema de la factorización vemos que $\sum Z_K X_{jK}$ es una estadística suficiente. Además el hecho que $Y = \Theta$ implica que:

$$\frac{d^2 \Theta}{dY^2} = 0$$

lo que nos lleva a:

$$\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} = E \left(\frac{\partial^2 L}{\partial \beta_i \partial \beta_j} \right) = - \alpha(\phi) \left(\frac{d\mu}{dy} \right)^2 \frac{x_i x_j}{v}$$

Cuando θ es también la media de la distribución, $\mu = \theta = y$, se tiene los modelos lineales usuales con errores normales, en este caso se cumple:

$$\begin{aligned} g'(\theta) &= \theta \\ g(\theta) &= \frac{\theta^2}{2} + \text{cte} \end{aligned}$$

la cual determina en forma única la distribución normal con varianza $1/\alpha(\phi)$ (Teor. de Patil y Schorrock, 1965).

6.- ANALISIS DE DESVIANZA.

Un modelo se dice que está jerarquizado si el ajuste de los coeficientes se realiza en el mismo orden en que están declarados, por lo tanto, dos modelos extremos existen:

- 1.- Modelo mínimo: Contiene el número mínimo de coeficientes que el modelo permite; se le llama modelo nulo y ajusta una gran media asignando toda la variación en los datos a la componente aleatoria del modelo.
- 2.- Modelo completo: Contiene el mayor número de términos (coeficientes) que el modelo permite y ajusta exactamente, luego, asigna toda la variación de los datos a la componente sistemática del modelo.

El proceso de ajuste del modelo ordenado consiste en ir paso a paso desde el modelo mínimo al modelo completo con el objeto de incrementar la bondad del ajuste a medida que el número de parámetros au-

menta en el modelo, esto es posible realizar al comparar modelos anidados. Comparamos los ajustes de los diferentes modelos con errores no-normales como se hace en el análisis de varianza con errores normales. El ajuste de los parámetros en cada etapa se hace maximizando la función máximo verosímil del modelo corriente y la medida cuantitativa de comparación será dada por la cantidad $-2 L$ máx que se llamará desvianza. La desvianza se mide desde el modelo completo, omitiéndose los términos que envuelven constantes, datos solamente, o el factor escalar.

La desvianza para las distribuciones más usuales de la familia exponencial son:

Distribución	Desvianza
Normal	$\sum (Z - \hat{\mu})^2 / \hat{\sigma}^2$
Poisson	$2 \{ \sum \ln (Z/\hat{\mu}) - \sum (Z - \hat{\mu}) \}$
Binomial	$2 [\sum Z \ln (Z/\hat{\mu}) + \sum (n - Z) \ln (\frac{n-Z}{n-\hat{\mu}})]$
Gamma	$2p \{ - \sum \ln (Z/\hat{\mu}) + \sum (Z - \hat{\mu}) / \hat{\mu} \}$

Los grados de libertad de un modelo es el número de parámetros linealmente independientes a estimar, o bien, el rango de la matriz de diseño X . Los grados de libertad, multiplicados por un factor escalar cuando fuere necesario, forman una escala para el conjunto de modelos secuenciales comparables con las desvianzas; cuando los grados de libertad del residuo multiplicados por el factor escalar es aproximadamente igual a la desvianza del modelo presente (ajustado), el proceso de continuar ajustando componentes sistemáticos termina. El factor escalar puede ser conocido o desconocido pero estimable directamente por ejemplo haciendo réplicas de observaciones, o indirectamente desde la desvianza después que un modelo adecuado ha sido ajustado. Para determinar si un modelo es adecuado se grafica las desvianzas sucesivas versus sus grados de libertad, aceptando como una medida del factor escalar la porción lineal desde el origen determinado por esos puntos de me

nos grados de libertad.

7. GENERALIZACION DEL ANALISIS DE VARIANZA.

Para distribuciones normales las primeras diferencias de desvianza son las sumas de cuadrados en el análisis de varianza para un ajuste secuencial como sigue:

Supongamos un modelo con componentes A, B y C, declaradas en ese mismo orden, entonces:

Modelo	Desvianza	Diferencias	Componente
Minimal	d_m		A
A		$d_m - d_A$	
	d_A		
B		$d_A - d_{AB}$	B dado que A está en el modelo.
	d_{AB}		
C		$d_{AB} - d_{ABC}$	C dado que A y B están en el modelo.
	d_{ABC}		
Completo		$d_{ABC} - d_0$	Residuo.
	d_0		

el análisis de varianza para el modelo secuencial se definió como el que tiene componentes dados por las primeras diferencias de las desvianzas con los grados de libertad respectivos. Estas componentes tienen distribuciones proporcional a una χ^2 , exacta para errores normales y aproximada para otros. (Este análisis fue sugerido por Good, 1967).

8.- OTROS TEST.

- Test Chi-cuadrado: La diferencia en ajuste entre dos modelos anidados se distribuye como χ^2 con los grados de libertad dados por la diferencia de los respectivos grados de libertad, de manera que podemos docimar el mejoramiento en el procedimiento de ajuste al incluir un término en el modelo. En particular, la comparación de cualquier modelo con el modelo máximo nos provee de una bondad del ajuste para ese modelo.
- Test-F: La desvianza para las distribuciones normal y Gamma es afectada por un parámetro perturbador y si este parámetro es desconocido los test chi-cuadrados no pueden usarse. Sin embargo, se puede usar el F-test en la misma forma como es usado en el análisis de varianza para comparar modelos anidados.

9.- ALGUNAS SIMPLIFICACIONES DE LA DESVIANZA.

En este apartado se desarrollará un teorema que simplifica la expresión de la desvianza para la distribución de Poisson y Gamma.

Teor.: Si a) $Y = \mu^\alpha$ ó b) $Y = \log \mu$ y un término constante es ajustado en el modelo (digamos $Y = \sum_{i=0}^n \beta_i X_i$, donde $X_0 = 1$), entonces

$$\sum \frac{(Z_K - \hat{\mu})\hat{\mu}}{\hat{V}} = 0$$

donde $\hat{\mu}$, $\hat{V}$ son los estimadores máximo-verosímil de μ y V .

Dem.: Sabemos que

$$\frac{\partial L}{\partial \beta_i} = \sum \frac{Z - \mu}{V} \frac{d\mu}{dY} X_i \quad (\text{se omite } \alpha(\phi)).$$

$$\text{Caso a): } Y = \mu^\alpha \implies \frac{d\mu}{dY} = \frac{\mu}{\alpha Y}$$

$$\text{luego } \frac{\partial L}{\partial \beta_i} = \sum \frac{Z - \mu}{V} \frac{\mu}{\alpha Y} x_i = 0$$

es decir:

$$\sum \frac{(Z - \hat{\mu}) \hat{\mu}}{\alpha \hat{V} \hat{Y}} \cdot x_i = 0 \quad x_i \neq 0$$

$$\therefore \sum \frac{(Z - \hat{\mu}) \hat{\mu}}{\hat{V} \hat{Y}} = 0$$

$$\text{pero } \sum \frac{(Z - \hat{\mu}) \hat{\mu}}{\hat{V}} = \sum \frac{Y (Z - \hat{\mu}) \hat{\mu}}{\hat{V} \hat{Y}} = 0$$

$$\text{Caso b): } Y = \log \mu \implies \frac{d\mu}{dY} = \mu$$

$$\text{luego } \frac{\partial L}{\partial \beta_i} = \sum \frac{(Z - \mu) \mu}{V} x_i = 0$$

$$\sum \frac{(Z - \hat{\mu}) \hat{\mu}}{\hat{V}} x_i = 0$$

$$\text{En el caso } \frac{\partial L}{\partial \beta_0} = 0 \implies \frac{(Z - \hat{\mu}) \hat{\mu}}{\hat{V}} = 0$$

Luego bajo las condiciones del Teorema las desvianzas toman los siguientes valores:

$$\text{Poisson} \quad : \quad 2 \sum Z \ln (Z/\mu)$$

$$\text{Gamma} \quad : \quad -2 \sum \ln (Z/\mu)$$

10.- COMENTARIOS.

Los modelos lineales generalizados dan una metodología que permite el análisis de problemas con observaciones que tienen una estructura distribucional perteneciente a la familia exponencial abarcando una gran cantidad de situaciones. Esta presentación no entra en detalles como por ejemplo aplicaciones del análisis de desviación para algunos diseños de experimentos, al respecto, queda claro que el análisis de varianza corriente es un caso particular del análisis de desviación cuando se respeta el esquema de cálculo de la desviación con distribución del error normal.

En su uso debe implementarse un programa computacional para la obtención de los estimadores máximo verosímil, ya que el proceso iterativo no es de fácil cálculo manual.

BIBLIOGRAFIA.

- [1] Bishop, Y.M.M. Full contingency tables, logits and split contingency tables. Biometrics 25, (1969) 383-399.
- [2] Kendall, M.G. and Stuart, A. (1967). The advance Theory of Statistics. 2nd ed. Vol. II. London: Griffin.
- [3] Nelder, J.A. Weighted regression, quantal response data and inverse polynomials. Biometrics, 24, 979-985.
- [4] Nelder, J.A. and Wedderburn, R.W.M. Generalized linear Models. J. R. Statist. Soc. A, 135, 370-384 (1972).
- [5] Patil G.P. and Richard Shorrock. On Certain Properties of the Exponential-type Families (1965). J.R. Statist. Soc. B, 27, 94-99.