00321

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

75

FACULTAD DE CIENCIAS

INTRODUCCION A LOS MODELOS LINEALES GENERALIZADOS, APLICACION A LAS ENCUESTAS ENADID 1997 Y ENSA 2000

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A :

PEREZ RICO RAYUNDO

DIRECTORA DE TESIS: M. en A. P. MARIA DEL PILAR ALONSO REYES

PAGINACION DISCONTINUA

DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:
INTRODUCCION A LOS MODELOS LINEALES GENERALIZADOS, APLICACION A LAS ENCUESTAS ENADID 1997 y ENSA 2000

realizado por PEREZ RICO RAYMUNDO

con número de cuenta 09527499-1, quién cubrió los créditos de la carrera de ACTUARIO

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario M. en A.P. MARIA DEL PILAR ALONSO REYES

Propietario M. en C. JOSE ANTONIO FLORES DIAZ

Propietario ACT. JAIME VAZQUEZ ALAMILLA

Suplente M. en C. MARIA DE LOURDES GUERRERO ZARCO

Suplente ACT. MARYPAOLA JANETT MAYA LOPEZ

Consejo Departamental de Matemáticas

M. en C. JOSE ANTONIO FLORES DIAZ

FACULTAD DE CIENCIAS
DE
MATEMATICAS

AGRADECIMIENTOS

A Dios por haberme colocado en este camino y permitirme concluir esta etapa de mi vida.

A mis padres, Juan Pérez B. y Catalina Rico V. por haberme apoyado de manera incondicional en todo momento y brindarme su amor y cariño. Sé que sin su apoyo nunca hubiera logrado esta meta.

A mis hermanos; Juan Manuel, Verónica y Adriana, por el cariño que me brindan.

A mi directora de tesis, M. en A. P. Ma. del Pilar Alonso Reyes, por brindarme sus valiosas aportaciones y dedicarme su tiempo durante el proceso de este trabajo.

A M. en C. José Antonio Flores, Act. Jaime Vázquez Alamilla, M. en C. Lourdes Guerrero Zarco y Act. Marypaola Maya , por sus interesantes cometarios en el desarrollo de este trabajo.

A mis amigos; Pedro, Joaquín, Guillermo, Iván, Oliver, Víctor, Armando, Gabriela, Mariana y Liliana. Por su valiosa amistad y compañía otorgada.

Amelia, por su comprensión y apoyo.

A la UNAM y en particular a la Facultad de Ciencias por darme la oportunidad de aprender en esta gran Institución.

A mis Padres

Con todo mi cariño

ÍNDICE

INTRODUCCIÓN

El presente trabajo proporciona una introducción a los Modelos Lineales Generalizados, los cuales son empleados en gran cantidad de proyectos en el ámbito laboral y de investigación debido a su gran capacidad para describir el comportamiento de diferentes conjuntos de datos. Asimismo, como ejemplo, se realizó un ejercicio aplicando estos modelos a una situación real. El ejercicio consistió en calcular la probabilidad del uso de algún método anticonceptivo en mujeres en edad fértil y que se encontraban con alguna relación conyugal.

Para la ejecución del proyecto, se utilizó un Modelo Lineal Generalizado (MLG), en el cual se consideró una variable de respuesta con distribución Binomial y una función liga Probit. Los instrumentos de medición utilizados fueron la Encuesta Nacional de la Dinámica Demográfica 1997 (ENADID 1997) y la Encuesta Nacional de Salud 2000 (ENSA 2000).

El desarrollo de este trabajo se presenta en cuatro capítulos, los cuales se describen brevemente a continuación.

En el primer capítulo se presenta una introducción a los modelos lineales generalizados, en donde se definen cuales son los componentes que los conforman. Asimismo, se realiza una clasificación respecto a los más utilizados, mostrándose en forma explícita cada unos de los componentes que definen un MLG.

El capitulo dos presenta los procedimientos de estimación de los modelos lineales generalizados, incluyendo el análisis de dos de los métodos más utilizados por paquetes estadísticos, como son el Método de Newton Raphson y el Método Iterativo Ponderado de Mínimos Cuadrados. Adicionalmente, se muestran las pruebas de bondad de ajuste más utilizadas.

En el tercer capítulo se realiza la estimación de algunos modelos seleccionados a través del método iterativo ponderado de mínimos cuadrados, desarrollándose de manera explícita la estimación del modelo que será utilizado en la aplicación para el capítulo final.

El capítulo cuatro tiene como objetivo calcular la probabilidad de uso de algún método anticonceptivo en mujeres en edad fértil y que se encontraban con alguna relación conyugal. Para la medición de dicho efecto se utilizó la ENADID 1997. Como primer paso se seleccionó un conjunto de variables, las cuales se analizaron a través de tablas de contingencia. Posteriormente, las variables que mostraron alguna relación con la variable de respuesta fueron consideradas para formar parte del conjunto de variables explicativas en el modelo lineal generalizado; dicho modelo como se mencionó anteriormente esta definido con una distribución Binomial y una función liga Probit. Finalmente, se ejecutó el modelo verificando la bondad de ajuste y realizando inferencias respecto a las variables más importantes.

El documento concluye con una discusión sobre diversos aspectos del trabajo realizado.

CAPÍTULO 1

Introducción a los Modelos Lineales Generalizados

Definición

Un modelo lineal generalizado es definido en términos de un conjunto de variables aleatorias independientes e idénticamente distribuidas Y_1 , Y_2 , ..., Y_n , donde cada una de éstas tiene una distribución que pertenece a la familia exponencial. Por lo tanto, Y_i para i = 1, 2, ..., n, tiene las siguientes propiedades:

1. La función de densidad de cada Y_i es de la forma canónica[1] y depende de un único parámetro θ_i donde θ_i no debe ser el mismo para todas las Y_i, es decir,

$$f(y;\theta)=\exp\{y_i \times b_i(\theta_i) + c_i(\theta_i) + d_i(y_i)\} \qquad (1.1)$$

donde $b_i(\theta_i)$ es llamado el parámetro natural.

2. La distribución de todas las Y_i 's es la misma, es decir, por ejemplo todas normales o todas binomiales.

Por lo tanto la función de densidad de probabilidad conjunta de Y_1 , Y_2 , ..., Y_n está dada por:

$$f(Y_1 , Y_2 , ..., Y_n ; \theta_1,\theta_2,...,\theta_n) = \exp\left\{\sum_{i=1}^{n} y_i b(\theta_i) + \sum_{i=1}^{n} c(\theta_i) + \sum_{i=1}^{n} d(y_i)\right\} \qquad (1.2)$$

[1] La forma canónica se refiere a considerar la función c(y) = y en la expresión general de la familia exponencial, es decir, f(y;θ)= a(θ) b(y) exp{c(y)Q(θ)}

Para la especificación del modelo, los parámetros θ_i no son de usual interés dado que puede haber uno por cada observación. Por lo tanto para un modelo lineal generalizado se considera a un conjunto menor de parámetros $\beta_1, \beta_2, ..., \beta_p$ ($p<n$), tal que una combinación lineal de β_i sean igual a alguna función del valor esperado μ_i de Y_i , es decir existe una función $g(\mu_i) = X_i^t \beta$.

En donde g(.) es una función diferenciable y monótona, ésta es llamada la función liga.

En términos formales, la definición **un modelo lineal generalizado está formado por los tres componentes siguientes**[2]:

1. Un componente aleatorio, el cual identifica la distribución de probabilidad de la variable respuesta[3] Y, en donde dicha distribución de probabilidad debe de pertenecer a la familia exponencial. El cual está constituido por un vector de variables aleatorias independientes e idénticamente distribuidas con media μ.

2. Un componente lineal sistemático relacionado al predictor lineal, sea

$\eta = X\beta$ el cual es un vector de $n \times 1$, dado que

$$\underline{X}_{nxp} = \begin{bmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_n \end{bmatrix}$$

y $X'_i = [X_{i1}, X_{i2}, \ldots, X_{ip}]$ es un vector de variables explicativass para la observación i, para i = 1, ..., p

y $\underline{\beta}^t$ es un vector de $p \times 1$ parámetros $\beta = (\beta_1, \beta_2, ..., \beta_p)^t$, los cuales son desconocidos y tienen que ser estimados a partir de la matriz de datos $\underline{X}$.

[2] Según Mccullagh y Nelder 1989
[3] También llamada variable dependiente

Por lo tanto cada elemento de η está dado por

$$\eta_i = \sum_{j=1}^{p} \beta_j X_{ij} = \underline{X}\beta \tag{1.3}$$

3. Una función liga g(.), la cual es una función diferenciable, uno a uno y monótona, tal función relaciona al predictor lineal y a los valores estimados. Dado que la función es uno a uno tiene función inversa, ésta relaciona la esperanza de Y_i con el predictor lineal tal como

$$E[Y_i] = \mu_i = g^{-1}(\eta_i) \tag{1.4}$$

Por lo que aplicando g(.) a (1.4) dado que tiene inversa, se tiene

$$g(\mu_i) = \eta_i = \sum_{j=1}^{p} \beta_j X_{ij} \qquad \text{la cual es (1.3)}$$

Por lo tanto se tiene que los modelos lineales generalizados especifican una relación entre la esperanza de la variable aleatoria Y y una función que es combinación lineal de los predictores. Está generalización da lugar a la especificación de un modelo que admite resultados continuos o discretos.

Ejemplo 1

Dobson[4] propone considerar algún lenguaje, el cual es descendiente de otro lenguaje como por ejemplo, Greco moderno es descendiente del Greco antiguo, o un lenguaje Romance es descendiente del latín.

Se propone entonces un modelo simple para determinar si hay un cambio de vocabulario, tal que si los lenguajes están separados por el tiempo t, entonces la

[4] Dobson, Annette J., *An Introduction to Generalized Linear Models*, pág. 32. Chapman and Hall, 1990

probabilidad de que ellos tengan palabras afines para un particular significado es e^{-0t} donde 0 es un parámetro. Se afirma que 0 es aproximadamente el mismo para muchos significados usados comúnmente. Para probar esto se consideró una lista de n significados usados regularmente, suponiendo que un lingüista determina para cada significado, si las correspondientes en los dos lenguajes son cognados o no son cognados. Para tal situación se puede desarrollar un modelo lineal generalizado de la siguiente manera.

Se definieron las variables aleatorias $Y_1, Y_2, \ldots, Y_n$ como sigue:

$$Y_i = \begin{cases} 1 & \text{Si los lenguajes tienen palabras afines para el significado } i \\ 0 & \text{Si las palabras no son cognados o afines} \end{cases}$$

Entonces $P(Y_i = 1) = e^{-\theta}$ es la probabilidad de que los lenguajes tengan palabras afines para el significado i, y $P(Y_i = 0) = 1 - e^{-\theta}$ es la probabilidad de que los lenguajes no tengan palabras afines, es decir, se tuvo una distribución Bernoulli(π), donde se tiene que su esperanza es π

$$E(Y_i) = \pi = e^{-\theta} \tag{1.5}$$

Entonces se tiene el primer elemento de un modelo lineal generalizado, el cual es el componente aleatorio, dado por $Y \sim$ Bernoulli(π), en donde además pertenece a la familia exponencial, ya que

$$\begin{aligned} P(Y = y) &= \pi^{y}(1-\pi)^{1-y} \\ &= \exp\{\, y\ln(\pi) + (1-y)\ln(1-\pi)\,\} \\ &= \exp\{\, y\ln(\pi) + \ln(1-\pi) - y\ln(1-\pi)\,\} \\ &= \exp\left\{\ln(1-\pi) + y\ln\left(\frac{\pi}{1-\pi}\right)\right\} \qquad \text{para } y = 0, 1 \end{aligned}$$

Renombrando

$$b(\pi) = \ln\left(\frac{\pi}{1-\pi}\right); \qquad c(\pi) = \ln(1-\pi) \quad y \quad d(y) = 0$$

Sustituyendo en la ecuación anterior se tiene

$$P(Y = y) = \exp\{yb(\pi) + c(\pi) + d(y)\}$$

Con lo anterior se ha mostrado que dicha distribución pertenece a la familia exponencial.

El segundo y tercer componente del modelo es el sistemático y la función liga respectivamente, en donde se tiene que la función liga y el predictor lineal tienen relación intrínseca, ya que

$$g(\mu) = \eta = X\beta \tag{1.6}$$

además se tiene que g(.) está determinada por la esperanza de Y, es decir, $E(Y) = \mu$ y por (1.5) se tiene entonces que

$$\begin{aligned} g(\pi) &= \ln(\pi) \\ &= \ln(e^{-\theta t}) \\ &= -\theta t \end{aligned}$$

Si $X_i = [-t]$ para toda i y $\beta = [\theta]$

Por último se tiene que la función diferenciable y monótona esta definida por $g(\pi) = \ln(\pi)$, la cual tiene como función inversa a la función exponencial, por lo que aplicando esta a (1.6) se obtiene

$$\begin{aligned} g^{-1}(\eta) &= E(Y) \\ &= e^{-\theta t} \\ &= \mu \end{aligned} \tag{1.7}$$

Ejemplo 2

Dobson propone un ejemplo similar[5] para el caso de Australia. El siguiente ejemplo se realizará con datos de México.

En una población grande, la probabilidad de seleccionar aleatoriamente a un individuo que posee una enfermedad crónica en un tiempo particular es pequeña. Si se asume que la incidencia[6] de la enfermedad entre diferentes individuos son eventos independientes, entonces el número de casos nuevos Y en un periodo de tiempo fijo, puede ser modelado por una distribución Poisson, donde su función de densidad de probabilidad está dada por:

$$f(y;\lambda) = \frac{\lambda^y e^{-y}}{y!}, \quad y = 0,1,\ldots; \tag{1.8}$$

y λ es el número promedio de enfermos para tal enfermedad por periodo de tiempo.

La tendencia en la incidencia de una cierta enfermedad puede ser modelada tomando variables independientes Y_1 , Y_2 , ..., Y_n que representan el número de casos ocurridos en intervalos de tiempos sucesivos enumerados por i=1, 2, ..., n, y sea $E(Yi) = \lambda$, obviamente ésta variará al tiempo t.

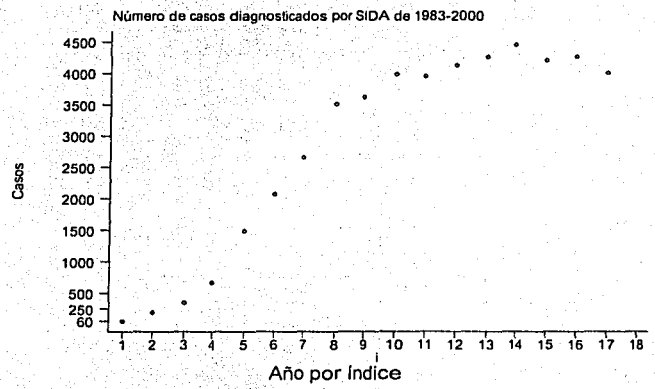
Se tiene por ejemplo, el número de casos diagnosticados por SIDA (Síndrome de Inmuno Deficiencia Adquirida) en México[7] por periodos anuales de 1983 a 2000, los datos se muestran en la siguiente tabla:

[5] Dobson propone modelar la tendencia en la mortalidad, teniendo como datos al número muertos por SIDA por trimestre en los años de 1983 a 1986.
[6] La incidencia es definida como el número de casos por una cierta enfermedad en el último año.
[7] Datos oficiales de CONASIDA, Boletín de Información Epidemiológica año 1993-2000.

Índice (i)	Año	Diagnosticados en el año
1	1983	60
2	1984	198
3	1985	349
4	1986	673
5	1987	1,485
6	1988	2,069
7	1989	2,661
8	1990	3,517
9	1991	3,625
10	1992	3,988
11	1993	3,950
12	1994	4,129
13	1995	4,268
14	1996	4,467
15	1997	4,212
16	1998	4,275
17	1999	4,014

Casos diagnosticados por SIDA por año para el periodo de 1983 a 2000

Graficando los datos, se tiene



Se observa en la gráfica que el número de casos se incrementa con el tiempo. Para estos datos un posible modelo es la distribución Poisson con

$$\lambda_i = i^{\theta}$$

Donde θ es el parámetro a ser estimado.

Estos datos pueden ser descritos por un modelo lineal generalizado con función liga :

$$g(\mu) = g(\lambda_i) = \theta \log(i)$$

en donde $\theta = \beta$ y $\log(i) = x_i$, en la ecuación (1.3) dado por $g(\mu) = \eta = X\beta$

El componente aleatorio está dado por una distribución Poisson definido como en (1.8), el cual pertenece a la familia exponencial.

Entonces se tienen los tres elementos para describir este comportamiento como un modelo lineal generalizado.

Tipos de Ligas en un modelo lineal generalizado

Se han observado dos ejemplos de modelos lineales generalizados, en donde se especifica a η como un predictor lineal generado por $X_1,\ldots,X_p$ que son las variables explicativas que se asumen en el modelo. Además se tiene que la relación entre η y las variables está dada como en (1.3) por $g(\mu) = \eta = X\beta$.

Por lo tanto se tiene que la función que relacione a η y μ, cualquiera que sea, debe ser especificada; así que una vez hecho esto, está liga distingue a un miembro de la familia de los modelos lineales generalizados.

Así la función liga es muy importante en los modelos lineales generalizados, por lo que la elección de ésta o del modelo estadístico a utilizar depende de la distribución de los datos y el marco teórico del estudio, el cual le permite al investigador entender la naturaleza de los datos. Específicamente, **la distribución del componente aleatorio de Y, es decir, la parte que no puede ser explicada sistemáticamente por *p* variables, determina la función liga y el tipo de modelo lineal generalizado.**

Considerando los diferentes tipos de datos que toma el componente aleatorio Y se realizará una clasificación, en donde se presentan los modelos con su liga canónica y los más usados frecuentemente. Tal clasificación es la siguiente:

1.1 Modelos de Respuesta Binomial
1.2 Modelos de Respuesta Continua
1.3 Modelos de Respuesta de Conteo

1.1 Modelos de Respuesta Binomial

En estos modelos se considera a la distribución Binomial en el componente aleatorio, así también se considera a la distribución Bernoulli, la cual es un caso particular de la distribución Binomial cuando n=1.

1.1.1 Distribución Bernoulli

Se tiene a la distribución Bernoulli la cual es para variables aleatorias binarias, es decir sólo pueden tomar dos valores, 0 y 1 por ejemplo. En donde las probabilidades que pueden tomar dichos valores están especificadas por

$$\begin{aligned} P(Y = 1) &= \pi \qquad y \\ P(Y = 0) &= 1 - \pi \end{aligned} \tag{1.9}$$

Donde $E(Y) = \pi$ y la función de densidad es:

$$\begin{aligned} f(y;\pi) &= \pi^y (1-\pi)^{1-y} \\ &= (1-\pi)\left(\frac{\pi}{1-\pi}\right)^y \\ &= \exp\left(\ln(1-\pi)\right) \times \exp\left\{y \ln\left(\frac{\pi}{1-\pi}\right)\right\} \\ &= \exp\left\{y \ln\left(\frac{\pi}{1-\pi}\right) + \ln(1-\pi)\right\} \end{aligned}$$

Renombrando

$$b(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) \; ; \; c(\pi) = \ln(1-\pi) \; y \; d(y) = 0$$

Por lo tanto $f(y;\pi)$ pertenece a la familia exponencial.

El predictor lineal dado por $g(\mu) = \eta = X\beta$

Renombrando al parámetro natural, se tiene

$$\begin{aligned}\theta &= \ln\left(\frac{\pi}{1-\pi}\right)\\ &= \eta\\ &= X\beta\\ &= g(\mu)\end{aligned}$$

Por lo tanto la función liga está dada por

$$g(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) \tag{1.10}$$

Y la función inversa se obtiene primero aplicando la función exponencial

$$f(x) = \exp(x) \tag{1.11}$$

a la ecuación (1.10) y utilizando la ecuación (1.6) se tiene

$$\begin{aligned}\exp(\eta) &= \exp(X\beta)\\ &= \frac{\pi}{1-\pi}\end{aligned}$$

Despejando π se tiene

$$\begin{aligned}\pi &= (1-\pi)\exp(X\beta)\\ \pi &= \exp(X\beta) - \pi\exp(X\beta)\\ \pi\exp(X\beta)+\pi &= \exp(X\beta)\\ \pi(1+\exp(X\beta)) &= \exp(X\beta)\end{aligned}$$

Por lo tanto la funcion inversa de g esta determinada por

$$\pi = \frac{\exp(X\beta)}{1+\exp(X\beta)} \tag{1.12}$$

Por lo que cumple que

$$\begin{aligned} E(y) &= \mu \\ &= \pi \\ &= g^{-1}(\eta) \end{aligned}$$

y la función liga como en (1.10) dada por

$$g(\pi) = \ln\left(\frac{\pi}{1-\pi}\right)$$

que es una función diferenciable y por lo tanto se tiene el tercer componente de los modelos lineales generalizados.

1.1.2 **Modelos de Regresión Binomial**

Son los más usados en análisis que tienen respuestas discretas, es decir correspondiente al número de éxitos en k eventos o proporcionales.

Para el modelo Binomial los datos de respuesta proporcional incluyen dos variables identificadas por el número de éxitos Y_i de cada población de K_i eventos, ambas variables están indexadas por i, ya que los datos proporcionales no requieren que el número eventos sea el mismo para cada observación.

Por lo tanto una variable aleatoria de respuesta Y_i con distribución Binomial toma $k+1$ valores y tiene su función de densidad dada por

$$f(y;k;\pi) = \frac{k!}{y!(k-y)!}\pi^{y}(1-\pi)^{k-y} \qquad (1.13)$$

Donde π es la probabilidad de éxito, y es el número de éxitos y k es el total de la población en cuestión, está función de densidad pertenece a la familia exponencial ya que

$$
\begin{aligned}
f(y;k;\pi) &= \frac{k!}{y!(k-y)!}\pi^{y}(1-\pi)^{k-y} \\
&= \frac{k!}{y!(k-y)!}(1-\pi)^{k}\left(\frac{\pi}{1-\pi}\right)^{y} \\
&= \exp\left(\ln(1-\pi)\right)\exp\left(y\ln\left(\frac{\pi}{1-\pi}\right)\right)\exp\left(\ln\left(\frac{k!}{y!(k-y)!}\right)\right) \\
&= \exp\left[y\ln\left(\frac{\pi}{1-\pi}\right) + k\ln(1-\pi) + \ln\left(\frac{k!}{y!(k-y)!}\right)\right]
\end{aligned}
$$

Renombrando se tiene

$$b(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) \; ; \; c(\pi) = k\ln(1-\pi) \; y \; d(y) = \ln\left(\frac{k!}{y!(k-y)!}\right)$$

por lo tanto $f(y;k;\pi)$ pertenece a la familia exponencial ya que puede expresarse de la forma

$$f(y;k,\pi) = \exp\left\{y \times b(\pi) + c(\pi) + d(y)\right\}$$

Renombrando al parámetro natural se tiene

$$
\begin{aligned}
\theta &= \ln\left(\frac{\pi}{1-\pi}\right) \\
&= \eta \\
&= X\beta \\
&= g(\mu)
\end{aligned}
$$

Por definición, se tiene

$$
\begin{aligned}
E(Y) &= \mu \\
&= k\pi
\end{aligned}
\tag{1.14}
$$

Por lo tanto $\pi = \frac{\mu}{k}$ y la función liga en términos de π queda como

$$g(\mu) = \ln\left(\frac{\pi}{1-\pi}\right) \tag{1.15}$$

Es decir, se aplica (1.14) a (1.15) quedando de la siguiente manera

$$\begin{aligned} g(\mu) &= \ln\left(\frac{\frac{\mu}{k}}{1 - \frac{\mu}{k}}\right) \\ &= \ln\left(\frac{\frac{\mu}{k}}{\frac{k-\mu}{k}}\right) \\ &= \ln\left(\frac{\mu}{k-\mu}\right) \end{aligned} \tag{1.16}$$

Para obtener la función inversa se aplica primero la función exponencial

$$f(x) = \exp(x) \tag{1.17}$$

Aplicando (1.17) a (1.16) y utilizando la ecuación (1.3) se tiene

$$\begin{aligned} \exp(\eta) &= \exp(X\beta) \\ &= \frac{\mu}{k - \mu} \end{aligned} \tag{1.18}$$

Despejando a $\mu = E(Y)$ se tiene

$$\begin{aligned} \frac{\mu}{k - \mu} &= \exp(X\beta) \\ \mu &= (k - \mu)\exp(X\beta) \\ \mu &= k\exp(X\beta) - \mu\exp(X\beta) \\ \mu\exp(X\beta) + \mu &= k\exp(X\beta) \\ \mu(1 + \exp(X\beta)) &= k\exp(X\beta) \\ \mu &= \frac{k\exp(X\beta)}{1 + \exp(X\beta)} \end{aligned} \tag{1.19}$$

Y dado que

$$g^{-1}(\eta) = \mu$$

Entonces se tienen al segundo y tercer componente de los modelos lineales generalizados,(1.3) y (1.4) respectivamente.

1.1.3 El modelo Probit

El modelo Probit fue usado por primera vez en ensayos biológicos. Típicamente la probabilidad de muerte fue medida contra el logaritmo de la dosis de alguna toxina, en donde la muerte depende de la tolerancia que el sujeto tiene al agente de la toxina, por lo que los sujetos con una menor tolerancia son más probables a morir.

La tolerancia es asumida como una variable aleatoria que se distribuye normalmente. Agresti[8] propuso el siguiente ejemplo. En el que se supone que se tiene x que es la dosis de un químico tóxico.

Se define entonces la siguiente variable aleatoria:

$$Y = \begin{cases} 1 & \text{Si el sujeto al que se le aplica la dosis muere} \\ 0 & \text{Si no muere el sujeto al aplicarle la dosis} \end{cases}$$

Posteriormente se seleccionó un sujeto de forma aleatoria. Sea Y = 1 si el sujeto muere, se supone además que τ es la tolerancia del sujeto a la toxina, es decir, el sujeto muere si $\mathcal{T} \leq x$. Por lo que el sujeto puede sobrevivir a la dosis x si ésta es menor que la tolerancia $\mathcal{T}$ y muere si la dosis x es mayor o igual a la tolerancia $\mathcal{T}$. Dado que el nivel de tolerancia varía de un sujeto a otro, entonces T es una variable aleatoria definida como sigue:

$$G(t) = P(\mathcal{T} \leq t) \qquad (1.20)$$

[8] Agresti A., *Categorical Data Analysis*, pág. 103, John Wiley & Sons,2000.

Entonces para una dosis fija, la probabilidad de que un sujeto muera es:

$$\begin{aligned} P(Y = 1) &= \pi(x) \\ &= P(T \leq x) \\ &= G(x) \end{aligned} \tag{1.21}$$

Por lo tanto Y tiene una distribución Bernoulli como en (1.9) ya que solo toma dos valores, morir codificado como 1 ó sobrevive codificado como 0.

Entonces en (1.21) el sujeto muere cuando la dosis es mayor que la tolerancia con una probabilidad $\pi(x)$.

En muchos experimentos toxicológicos la distribución del logaritmo de la dosis es aproximadamente normal con media μ y varianza σ^2 . Por lo tanto si G(.) tiene una distribución normal, entonces

$$\begin{aligned} \pi(x) &= G(x) \\ &= P(T \leq x) \\ &= \phi\left[\frac{(x - \mu)}{\sigma}\right] \end{aligned} \tag{1.22}$$

Donde ϕ tiene distribución normal estándar. Renombrando a

$\frac{1}{\sigma} = \beta$; $-\frac{\mu}{\sigma} = \alpha$ se tiene que

$$\phi\left(\frac{(x - \mu)}{\sigma}\right) = \phi(\alpha + \beta x) \tag{1.23}$$

aplicando ϕ^{-1} a (1.23) que es la función inversa, se tiene

$$g^{-1}(\pi(x)) = \alpha + \beta x$$

La cual es una combinación lineal de parámetros desconocidos y variables explicativas.

Por lo tanto se tienen los tres componentes, se tiene el componente aleatorio, ya que Y que se distribuye Bernoulli ($\pi(x)$) para una x dada; la función liga y el componente sistemático los cuales están dados por

$$g(\mu) = \phi^{-1}(\bullet) \text{ y } \quad g^{-1}(\pi(x)) = \alpha + \beta x$$

Así la liga está dada por ϕ^{-1}, la cual es la función inversa de una distribución normal estándar.

Donde:

$E(Y) = \mu$

$= \pi(x)$ para x dada porque Y~Bernoulli ($\pi(x)$).

Y el predictor lineal dado por

$$\begin{aligned} \eta &= g^{-1}(\mu) \\ &= g^{-1}(\pi(x)) \\ &= \phi(\pi(x)) \\ &= \alpha + \beta x \end{aligned}$$

Por lo que se obtienen los componentes (1.3) y (1.4) de un modelo lineal generalizado.

1.1.4 El modelo C-Log-Log

La liga logit y probit son simétricas en el sentido que

$$liga(\pi) = -liga(1-\pi)$$

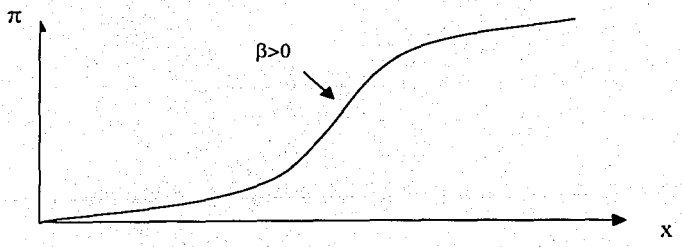
ya que

$$\begin{aligned} \text{logit}(\pi) &= \log\left[\frac{\pi}{1-\pi}\right] \\ &= -\log\left[\frac{1-\pi}{\pi}\right] \\ &= -\text{logit}[1-\pi] \end{aligned}$$

Esto implica que la curva de respuesta para $\pi(x)$ tiene una aparente simetría cerca de 0.5. Es decir $\pi(x)$ se aproxima a cero a la misma velocidad que se aproxima a 1.

El logit o probit son inadecuados cuando $\pi(x)$ se incrementa desde cero bastante lento y se aproxima a 1 rápidamente.

La función siguiente cumple con el comportamiento descrito anteriormente;
$\pi = 1 - \exp[-\exp(\alpha + \beta x)]$, ésta es asimétrica y su gráfica es :



La liga para este modelo y el predictor lineal están dados por :

$$\begin{aligned} \eta &= g(\mu) \\ &= \log\left(-\log\left(1-\frac{\mu}{k}\right)\right) \\ &= \alpha + \beta x \end{aligned} \qquad (1.24)$$

en donde

$$\begin{aligned} \mu_i &= k_i \pi_i(x) \\ &= g^{-1}(\eta_i) \\ &= k_i\left[1 - \exp[-\exp(\alpha + \beta x)]\right] \end{aligned}$$

donde $\pi_i(x) = \frac{\mu_i}{k_i}$

La ecuación (1.24) es llamada la liga complementaria log-log y dado que el componente aleatorio Y_i está dado por la distribución Binomial (k_i,π_i), se tienen entonces los tres componentes de un modelo lineal generalizado.

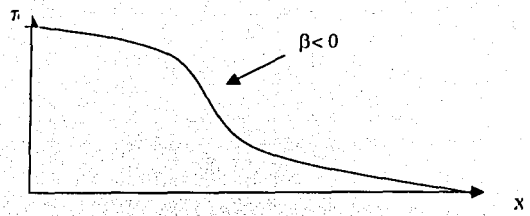
1.1.5 El modelo Log-Log

Otro comportamiento de la probabilidad se da cuando $\pi(x)$ parte de uno lentamente pero se aproxima a cero rápidamente, al incrementarse x y cuando la curva es monótona decreciente, se tiene que $b > 0$ y cuando es monótona creciente se tiene que $b<0$, para este caso se tiene al modelo log-log.

La función descrita anteriormente está dada por

$$\pi(x) = \exp[-\exp(\alpha + \beta x)] \qquad (1.25)$$

cuya gráfica es:



La variable aleatoria está dada por una distribución Binomial(k, π), donde $\pi(x)$ sigue el comportamiento descrito anteriormente. La función liga está dada por

$$\begin{aligned} \eta &= g(\mu) \\ &= \log(-\log(\pi(x))) \\ &= \alpha + \beta x \end{aligned} \qquad (1.26)$$

Lo anterior para el caso univariado, para el caso múltiple es de manera similar considerando el vector $X\beta$.

Aplicando g^{-1} a (1.26) se tiene entonces (1.25), con lo que se tienen los tres elementos de un modelo lineal generalizado.

Los modelos log-log y c-log-log son usados cuando se tiene un número pequeño o un número grande de éxitos con respecto a la población total, es decir, cuando la función $\pi(x)$ es asimétrica.

1.2 Modelos de Respuesta Continua

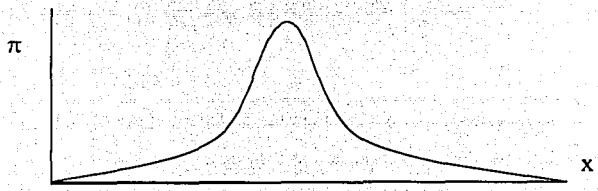
1.2.1 Distribución Gaussiana

La distribución Gaussiana es mejor conocida como la distribución normal. Los modelos de regresión basados sobre la distribución normal son comúnmente referidos como modelos de mínimos cuadrados. Dicha distribución es miembro de la familia exponencial y por lo tanto pertenece a los modelos lineales generalizados.

La ecuación considera un término de error dado por:

$$Y = X\mathrm{B} + \varepsilon$$

La cual se distribuye normal y tiene la siguiente forma:



Y su función de densidad de probabilidad está dada por:

$$f(y;\mu;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(y-\mu)^2\right\}$$

La cual pertenece a la familia exponencial ya que

$$f(y;\mu;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{-1}{2\sigma^2}(y-\mu)^2\right\}$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{-1}{2\sigma^2}(y^2 - 2y\mu + \mu^2)\right\} \qquad (1.27)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{-y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2})\right\}$$
$$= \exp\left\{-\frac{1}{2}\ln(2\pi\sigma^2)\right\}\exp\left\{\frac{-y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2})\right\}$$
$$= \exp\left\{\frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right\}$$

Renombrando, para σ^2 conocida

$$b(\mu) = \frac{\mu}{\sigma^2}$$
$$c(\mu) = -\frac{\mu^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2)$$
$$d(y) = -\frac{y^2}{2\sigma^2}$$

Por lo tanto $f(y;\mu;\sigma^2)$ pertenece a la familia exponencial para σ^2 conocida, o de forma equivalente la función $f(y;\theta)$ pertenece a la familia exponencial si se puede expresar de la siguiente manera[9]:

$$f(y;\theta) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y;\phi)\right\} \qquad (1.28)$$

[9] Hardin James, Hilbe Joseph. *Generalized Linear Models and Extensions*, STATA PRESS, 2001, pag. 54.

Donde se tiene que el parámetro natural está dado por θ y ϕ es la dispersión requerida para producir los errores estándares, donde a(ϕ) es también llamado factor de expansión. Para la distribución Poisson, Binomial y Binomial Negativa $a(\phi) = 1$.

Por lo que rescribiendo la ecuación (1.27) se tiene que

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right\} \\
&= \exp\left\{-\frac{1}{2}\ln(2\pi\sigma^2)\right\} \exp\left\{\frac{-y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right\} \\
&= \exp\left\{\frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right\} \\
&= \exp\left\{\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right\}
\end{aligned}
$$

Para esta distribución la función liga está dada por la identidad, es decir,

$$
\begin{aligned}
g(\mu) &= \mu \\
&= \eta \\
&= X\beta
\end{aligned}
$$

dicha función liga es uno a uno diferenciable, teniéndose así los tres elementos que conforman un modelo lineal generalizado.

1.2.2 **Modelo con Liga Log-Normal**

Este modelo está basado en la distribución Gaussiana, éste emplea como liga al logaritmo en vez de la identidad. La liga logarítmica generalmente es usada para datos de respuesta que sólo pueden tomar valores positivos sobre la escala continua o valores mayores a cero.

La función de densidad de probabilidad está dada por:

$$f(y;\mu;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(y-\log(\mu))^2\right\} \qquad (1.29)$$

Desarrollando para verificar que pertenece a la familia exponencial y determinar su función liga dado el parámetro natural.

$$f(y;\mu;\sigma^2) = \exp\left\{\frac{y\log(\mu)}{\sigma^2} - \frac{(\log(\mu))^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right\}$$

$$= \exp\left\{\frac{y\log(\mu) - \dfrac{(\log(\mu))^2}{2}}{\sigma^2} - \frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}\right\}$$

Por lo que rescribiendo

$$b(\theta) = \frac{(\log(\mu))^2}{2}$$

$$c(y;\phi) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{y^2}{2\sigma^2}$$

En donde se observa que la función liga está dada por

$$\begin{aligned} g(\mu) &= \log(\mu) \\ &= \eta \\ &= X\beta \end{aligned}$$

Por lo que la función inversa está determinada por

$$\begin{aligned} \mu &= g^{-1}(\eta) \\ &= \exp(X\beta) \end{aligned}$$

Cumpliéndose (1.3) y (1.4) para un modelo lineal generalizado, en donde el componente aleatorio está dado por la distribución Normal($\log(\mu)$, σ^2), para σ^2 conocida.

1.2.3 Distribución Gamma con Liga Canónica

El modelo Gamma es usado para situaciones cuando la variable respuesta solo puede tomar valores mayores o iguales a cero. Este modelo es usado cuando las respuestas positivas tienen un coeficiente de variación.

Es importante señalar que dentro del marco tradicional de los modelos lineales generalizados delimitan el modelo a un único parámetro, que es la media o μ. Por lo que el valor del coeficiente de desviación en ocasiones se determina por el usuario.
La función de densidad está dada por[10]

$$f(y;\mu;\phi) = \frac{y}{\Gamma(1/\phi)}\left(\frac{y}{\mu\phi}\right)^{1/\phi} exp\left(\frac{-y}{\mu\phi}\right) \qquad (1.30)$$

Tal distribución pertenece a la familia exponencial, ya que se puede expresar de la forma (1.28), es decir;

$$f(y;\theta) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y;\phi)\right\}$$

Por lo que transformando la ecuación se tiene

$$\begin{aligned} f(y;\mu;\phi) &= \frac{y}{\Gamma(1/\phi)}\left(\frac{y}{\mu\phi}\right)^{1/\phi} \exp\left(\frac{-y}{\mu\phi}\right) \\ &= \exp\left\{\ln(y) - \ln(\Gamma(1/\phi))\right\} \exp\left\{\frac{1}{\phi}(\ln(y) - \ln\mu - \ln\phi)\right\} \exp\left\{\frac{-y}{\mu\phi}\right\} \\ &= \exp\left\{\ln(y) - \ln(\Gamma(1/\phi)) + \frac{\ln(y) - \ln\mu - \ln\phi}{\phi} - \frac{y}{\mu\phi}\right\} \end{aligned}$$

[10] Hardin James, op. cit., pág. 64

$$= \exp\left\{\frac{-y}{\mu\phi} - \frac{\ln\mu}{\phi} + \ln(y) + \frac{\ln(y)}{\phi} - \ln(\Gamma(1/\phi)) - \frac{\ln\phi}{\phi}\right\}$$

$$= \exp\left\{\frac{-y/\mu - (-\ln\mu)}{-\phi} + \left(1+\frac{1}{\phi}\right)\ln(y) - \frac{\ln\phi}{\phi} - \ln(\Gamma(1/\phi))\right\}$$

$$= \exp\left\{\frac{-y/\mu - (-\ln\mu)}{-\phi} + \left(\frac{\phi+1}{\phi}\right)\ln(y) - \frac{\ln\phi}{\phi} - \ln(\Gamma(1/\phi))\right\}$$

Por lo tanto la función pertenece a la familia exponencial.

De acuerdo a la forma (1.28) se tiene que

$$\theta = 1/\mu$$

$$b(\theta) = -\ln(\theta)$$

$$c(y;\phi) = -\frac{\ln\phi}{\phi} - \ln(\Gamma(1/\phi))$$

Por su parte la función liga queda identificada por

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

$$= 1/\mu$$

Determinándose que la función inversa esta dada por

$$g^{-1}(\eta) = \mu$$

$$= \frac{1}{X\beta}$$

Y dado que se supone la existencia de una combinación lineal de parámetros y variables explicativas se tiene

$$g(\mu) = \eta = X\beta$$

Teniéndose así los tres componentes de un modelo lineal generalizado para un coeficiente de variación dado.

1.3 Modelos de Respuesta de Conteo

Este tipo de modelos son utilizados cuando la variable respuesta y las variables explicativas son categóricas, es decir están medidas sobre escalas nominales u ordinales. En donde cada escala puede tener más de dos valores. Las observaciones consisten de conteos o frecuencias en las celdas de una tabla de contingencia formada por una clasificación cruzada de varias variables.

Como ejemplo se tienen los datos de un estudio de pacientes con un cáncer en la piel llamado melanoma maligno. Para una muestra de 400 pacientes con cáncer, se les clasificó por la ubicación en la que se tenía el tumor y el tipo de tumor. Quedando la tabla de la siguiente manera:

	Ubicación del tumor			
Tipo de tumor	Cabeza y cuello	Tronco	Extremidades	Total
Peca melatónica de Hutchinson	22	2	10	34
Melanoma extendido superficial	16	58	115	185
Nodular	19	33	73	125
Indeterminado	11	17	28	56
Total	68	106	226	400

Tabla 1.

En este ejemplo hay dos variables de respuesta, ubicación y tipo del tumor. Las frecuencias en las celdas son consideradas como variables aleatorias, las cuales tienen como restricción que el número de sujetos debe ser fijo.

La pregunta de interés, es saber si existe alguna relación entre las dos variables de respuesta.

Las distribuciones de probabilidad más comunes y usadas que se consideran para este tipo de modelos son la distribución Poisson y la Multinomial

Supóngase que se tiene una tabla bidimensional con J categorías para la variable A y K categorías para la variable B, considérese la siguiente forma:

	B_1	B_2	...	B_k	Total
A_1	Y_{11}	Y_{12}	...	Y_{1k}	$Y_{1.}$
A_2	Y_{21}	Y_{22}	...	Y_{2k}	$Y_{2.}$
.			.		
.			.		
.			.		
A_j	Y_{j1}	Y_{j2}	...	Y_{jk}	$Y_{j.}$
Total	$Y_{.1}$	$Y_{.2}$	...	$Y_{.k}$	$n = Y_{..}$

Tabla 2.

En donde cada Y_{jk} denota la frecuencia para la celda (j,k), $Y_{j.}$ y $Y_{.k}$ denotan las sumas parciales por renglón y por columna respectivamente y n representa el total general.
En general para una tabla de $J \times K \times ... \times L$ se escriben las frecuencias como $Y_{jk...l}$.
Ahora se consideran a los modelos lineales generalizados para tablas de contingencia de dos variables para cada tipo de distribución.
Ahora surge la siguiente pregunta ¿Cada una de las distribuciones pertenecen a la familia exponencial?
La respuesta es afirmativa como se muestra a continuación:

a) Poisson

Dada la función de densidad

$$\begin{aligned} f(y;\mu) &= \frac{e^{-\mu}\mu^{y}}{y!} \\ &= e^{-\mu}e^{y\ln\mu}e^{-\ln(y!)} \\ &= \exp\{y\ln\mu - \mu + -\ln(y!)\} \end{aligned} \tag{1.31}$$

Renombrando se tiene

$$\begin{aligned} b(\mu) &= \ln(\mu) \\ c(\mu) &= -\mu \\ d(y) &= -\ln(y!) \end{aligned}$$

Donde $b(\mu)$ es el parámetro natural.

Por lo tanto la distribución Poisson pertenece a la familia exponencial.

b) Multinomial

La distribución Multinomial se utiliza cuando una secuencia de experimentos independientes e idénticos son ejecutados. Donde se supone que en cada experimento puede resultar cualquiera de los r posibles resultados.

La función de densidad de probabilidad está dada por la siguiente función de densidad;

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_r = y_r) = \frac{y!}{y_1!\,y_2!\ldots y_r!}\pi_1^{y_1}\pi_2^{y_2}\ldots\pi_r^{y_r}$$

$$Donde \sum_{i=1}^{r} y_i = y. \tag{1.32}$$

Tal distribución pertenece a la familia exponencial ya que

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_r = y_r) = \frac{y!}{y_1!\,y_2!\ldots y_r!}\pi_1^{y_1}\pi_2^{y_2}\ldots\pi_r^{y_r}$$

$$= \exp\left\{\ln\left[\frac{y!}{y_1!\,y_2!\ldots y_r!}\right]\right\}\exp\{y_1\ln(\pi_1)\}\ldots\exp\{y_r\ln(\pi_r)\}$$

$$= \exp\left\{\sum_{i=1}^{r} y_i\ln(\pi_i) + \ln\left[\frac{y!}{y_1!\,y_2!\ldots y_r!}\right]\right\} \qquad (1.33)$$

Renombrando se tiene

$$b(\mu) = \ln(\pi_i)$$
$$c(\mu) = 0$$
$$d(y) = \ln\left[\frac{y!}{y_1!\,y_2!\ldots y_r!}\right]$$

Mostrándose así que la distribución Multinomial pertenece a la familia exponencial.

Cada unas de estas distribuciones pertenecen a la familia exponencial, y por lo que tienen asociada una liga canónica con la cual definen un elemento en la familia de los modelos lineales generalizados.

Las distribuciones mencionadas son utilizadas cuando se tiene una tabla de contingencia y se quiere ajustar un modelo el cual sea capaz de predecir los valores de las celdas, estos modelos son llamados Modelos Log-lineales.

1.3.1 1 **Distribución Poisson**

Para la distribución Poisson con frecuencias en las celdas $Y_1, Y_2, \ldots, Y_n$ y parámetros $\lambda_1, \lambda_2, \ldots, \lambda_m$, las frecuencias esperadas están dadas por $E(Y_i)=\lambda_i$
Es decir se tiene una fila de variables de la siguiente manera:

Y_1	Y_2	...	Y_n

En la ecuación (1.31) se encuentra que el parámetro natural está dado por $g(\mu_i) = \ln(\mu_i)$ y asumiendo la existencia de un predictor lineal

$$\begin{aligned} \eta_i &= g(\mu_i) \\ &= \ln(\mu_i) \\ &= X\beta \end{aligned}$$

Se tienen entonces los tres componentes de un modelo lineal generalizado, la distribución que pertenece a la familia exponencial, el predictor lineal y la función liga que se relacionan para determinar la esperanza de la distribución.

Por último la función inversa de g determina la esperanza de la variable aleatoria Y, presentando la siguiente forma:

$$\begin{aligned} \mu_i &= g^{-1}(\eta_i) \\ &= \exp(X\beta) \end{aligned} \qquad (1.34)$$

1.3.2 **Distribución Multinomial**

Supóngase que se tiene una muestra de una distribución Multinomial de tamaño n dado por $N = JK$ celdas en una tabla de contingencia. En donde las probabilidades π_{ij} son de una distribución conjunta para dos variables de respuesta.

Se tiene que la distribución multinomial pertenece a los modelos lineales generalizados, ya que presenta una función liga determinada por

$$\begin{aligned} \eta_{ij} &= g(\mu_{ij}) \\ &= \ln(\mu_{ij}) \\ &= X\beta \end{aligned} \tag{1.35}$$

Así como su esperanza definida por

$$\begin{aligned} \mu_{ij} &= E(Y_{ij}) \\ &= n\pi_{ij} \end{aligned} \tag{1.36}$$

Por lo tanto la función liga está definida por (1.35), teniéndose entonces que las frecuencias esperadas están determinadas por

$$\begin{aligned} \mu_{ij} &= g^{-1}(\eta_{ij}) \\ &= \exp(X\beta) \end{aligned}$$

Conformándose así los tres componentes que conforman a un modelo lineal generalizado.

Se observa que los valores estimados para este caso son las frecuencias que conforman a una tabla de contingencia. Tales modelos parten del supuesto de independencia para calcular sus frecuencias estimadas, es decir, supóngase que se tienen dos variables de respuesta, las cuales conforman una tabla de contingencia con N celdas, definidas como la variable que toma x_{ij} valores $i = 1,2,...,I$, $j = 1,2,...,J$. El supuesto es considerar a las dos variables de respuesta estadísticamente independientes, es decir;

$$\pi_{ij} = \pi_{i+}\pi_{+j} \qquad i = 1, 2, ..., I; \quad j = 1, 2, ..., J \tag{1.37}$$

Por lo tanto como se determinó en (1.36) la función que relaciona las frecuencias esperadas es

$$E(Y_{ij}) = n\pi_{ij} \tag{1.38}$$

Entonces si se sustituye (1.37) en la ecuación (1.38) se tiene

$$E(Y_{ij}) = n\pi_{i+}\pi_{+j} \tag{1.39}$$

Ahora si se sustituye la función liga para encontrar un predictor lineal se tiene

$$\ln(\mu) = \ln\left[E(Y_{ij})\right] = \ln(n) + \ln\pi_{i+} + \ln\pi_{+j} \tag{1.40}$$

En donde se observa que el logaritmo de las frecuencias esperadas en la celda (i,j) es una función aditiva de un efecto en el i-ésimo renglón y en la j-ésima columna. Denotando a la variable renglón por X y la variable columna por Y, se re-expresa la ecuación (1.40) como

$$\ln(\mu_{ij}) = \mu + \lambda_i^X + \lambda_j^Y \tag{1.41}$$

donde

$$\lambda_i^X = \ln(\pi_{i+}) - \frac{\left(\sum_h \ln(\pi_{h+})\right)}{I}$$

$$\lambda_j^Y = \ln(\pi_{+j}) - \frac{\left(\sum_h \ln(\pi_{+h})\right)}{J}$$

$$\mu = \ln(n) + \frac{\left(\sum_h \ln(\pi_{h+})\right)}{I} + \frac{\left(\sum_h \ln(\pi_{+h})\right)}{J}$$

Los parámetros $\{\lambda_i^X\}$ y $\{\lambda_j^Y\}$ satisfacen que

$$\sum \lambda_i^x = \sum \lambda_j^y = 0$$

El modelo (1.41) es llamado **modelo log-lineal bajo independencia** para una tabla de contingencia bidimensional.

Si se supone que existe dependencia entre las variables, entonces el modelo log-lineal cambia a la siguiente estructura.

En donde para toda $m_{ij} > 0$, se define a

$$\eta_{ij} = \ln(m_{ij})$$

entonces

$$\eta_{i.} = \frac{\sum_j n_{ij}}{J} \quad , \quad \eta_{.j} = \frac{\sum_i n_{ij}}{I}$$

$$\mu = \eta_{..} = \frac{\sum_i \sum_j n_{ij}}{IJ}$$

Que denota la media de $\ln(m_{ij})$

$$\lambda_i^x = \eta_{i.} - \eta_{..} \quad , \quad \lambda_j^y = \eta_{.j} - \eta_{..}$$
$$\lambda_{ij}^{xy} = \eta_{ij} - \eta_{i.} - \eta_{.j} + \eta_{..} \qquad (1.42)$$

Despejando a η_{ij} de la ecuación (1.42) se tiene

$$\ln(m_{ij}) = \mu + \lambda_i^x + \lambda_j^y + \lambda_{ij}^{xy} \qquad (1.43)$$

Este modelo describe perfectamente cualquier conjunto de frecuencias esperadas positivas. Éste es llamado **el modelo saturado**, y es el modelo más general para tablas de contingencia de dos variables.

Los parámetros $\{\lambda_i^X\}$ y $\{\lambda_j^Y\}$ en (1.43) son las desviaciones alrededor de la media, y por lo tanto satisfacen $\sum_i \lambda_i^X = \sum_j \lambda_j^Y = 0$, por lo que se tiene entonces que hay *I-1* parámetros de renglón linealmente independiente y *J-1* parámetros columna linealmente independiente.

Además $\{\lambda_{ij}^{XY}\}$ satisface que $\sum_i \lambda_{ij}^{XY} = \sum_j \lambda_{ij}^{XY} = 0$ (1.44)

Dado $\{\lambda_{ij}^{XY}\}$ en las *(I-1)(J-1)* celdas en las primeras *(I-1)* renglones y *(J-1)* columnas, estas restricciones determinan los parámetros para las celdas en la última columna o último renglón. Por lo tanto se tienen que *(I-1)(J-1)* de esos términos son linealmente independientes.

El término adicional $\{\lambda_{ij}^{XY}\}$ en el modelo (1.43) con respecto a (1.41) es debido a la asociación de los parámetros que reflejan las desviaciones de independencia de X y Y.

Por lo tanto el número de parámetros linealmente independientes es igual a $I+(I-1)+(J-1) = I+J+I$ para el modelo bajo independencia y $I+(I-1)+(J-1)+(I-1)(J-1) = IJ$ para el modelo saturado.

Para tablas de mayor dimensión, el número de parámetros en el modelo saturado log-lineal es igual a el número de celdas en la tabla.

CAPÍTULO 2

Estimación de los Modelos Lineales Generalizados

La estimación de los modelos lineales generalizados se realiza en dos etapas, primero bajo el concepto de máxima verosimilitud y posteriormente se utilizan métodos iterativos, tales son el método de Newton - Raphson y el método Iterativo Ponderado de Mínimos Cuadrados (IRLS : Iterative Re-weight Least Squares).

2.1 Método de Newton-Raphson. Versión Multidimensional

El método de Newton-Raphson es un método numérico que encuentra el mínimo o máximo de una función de varias variables.
Sea la función

$$f : U \longrightarrow R^n$$

y las variables $\underline{X} = (X_1, X_2, \ldots, X_n)$. Adicionalmente se asume que la función es de clase C^2, es decir, las derivadas parciales de orden 2 existen y son continuas.
Para un punto dado sea $\underline{x}^o \in U$, es decir, $\underline{x}^o = (x_1^o, x_2^o, \ldots, x_n^o)$ se puede aproximar a f con la expansión de Taylor para $\underline{x}^o$, por lo que se tiene la siguiente aproximación;

$$f(\underline{x}) = f(\underline{x}^o) + \sum_{i=1}^{n} \frac{\partial f(\underline{x}^o)}{\partial x_i}(x_i - x_i^o) + \\ + \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\underline{x}^o)}{\partial x_j \partial x_i}(x_i - x_i^o)(x_j - x_j^o) + \varepsilon(\underline{x} - \underline{x}^o)\|\underline{x} - \underline{x}^o\|^2$$

Donde $\varepsilon(\underline{x} - \underline{x}^o) \to 0$

Como $\underline{x} \to \underline{x}^o$. Se puede utilizar el siguiente polinomio

$$f_2(\underline{x}) = f(\underline{x}^o) + \sum_{i=1}^{n} \frac{\partial f(\underline{x}^o)}{\partial x_i}(x_i - x_i^o) + \\ + \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\underline{x}^o)}{\partial x_j \partial x_i}(x_i - x_i^o)(x_j - x_j^o)$$

El objetivo es encontrar el mínimo o máximo $\underline{x}^*$ de f, donde se sabe que

$$\frac{\partial f(\underline{x}^*)}{\partial x_i} = 0$$

Para toda $i \in \{1,2,\ldots,n\}$. Dado que f_2 es una aproximación a f se tiene que

$$\frac{\partial f_2(\underline{x}^*)}{\partial x_k} = 0$$

Para toda $k \in \{1,2,\ldots,n\}$. Al calcular las derivadas del polinomio f_2 se obtuvo

$$\frac{\partial f_2(\underline{x})}{\partial x_k} = \frac{\partial f(\underline{x}^o)}{\partial x_k} + \sum_{i=1}^{n} \frac{\partial^2 f(\underline{x}^o)}{\partial x_k \partial x_i}(x_i - x_i^o) + \\ + \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^3 f(\underline{x}^o)}{\partial x_k \partial x_j \partial x_i}(x_i - x_i^o)(x_j - x_j^o)$$

Dado que se asume que f es de clase C^2, sólo se aproximará con los primeros dos términos de la ecuación anterior, ya que se desconoce si las derivadas de orden 3 existan y sean continuas. Por lo que la aproximación queda de la siguiente forma:

$$\frac{\partial f_2(\underline{x})}{\partial x_k} = \frac{\partial f(\underline{x}^o)}{\partial x_k} + \sum_{i=1}^{n} \frac{\partial^2 f(\underline{x}^o)}{\partial x_k \partial x_i}(x_i - x_i^o) \qquad (2.1)$$

Evaluando $\underline{x}^*$ en f_2 para obtener las condiciones

$$\frac{\partial f(\underline{x}^o)}{\partial x_k}+\sum_{i=1}^{n}\frac{\partial^2 f(\underline{x}^o)}{\partial x_k \partial x_i}\left(x_i^* - x_i^o\right)=0 \qquad k=1,2,\ldots,n$$

Dado que las ecuaciones se resuelven de manera simultánea, se re-escribe la ecuación anterior en forma matricial, quedando como;

$$\underline{x}^* = \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} \quad \text{y} \quad \underline{x}^o = \begin{bmatrix} x_1^o \\ x_2^o \\ \vdots \\ x_n^o \end{bmatrix}$$

$$p(\underline{x}^o) = \begin{bmatrix} \dfrac{\partial f(\underline{x}^o)}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\underline{x}^o)}{\partial x_n} \end{bmatrix} \quad \text{y la matriz } H(\underline{x}^o) = \begin{bmatrix} \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\underline{x}^o)}{\partial x_n \partial x_n} \end{bmatrix}$$

Por lo que la ecuación en forma matricial queda como

$$p(\underline{x}^o) + H(\underline{x}^o)(\underline{x}^* - \underline{x}^o) = 0$$

La ecuación anterior tiene solución para $\underline{x}^*$ si y sólo si $\det H(\underline{x}^o) \neq 0$, es decir,

$$\underline{x}^* = \underline{x}^o - H^{-1}(\underline{x}^o)p(\underline{x}^o)$$

Pero ésta es sólo una aproximación, ya que se está generando por f_2, así que el mínimo o máximo está dado de manera iterativa dado un valor x^o

$$\underline{x}^1 = \underline{x}^o - H^{-1}(\underline{x}^o)p(\underline{x}^o) \qquad (2.2)$$

Repitiendo el proceso se puede tener una mejor aproximación utilizando ahora x^1

$$\underline{x}^2 = \underline{x}^1 - H^{-1}(\underline{x}^1)p(\underline{x}^1)$$

Realizando de la misma manera el proceso anterior hasta encontrar el punto de convergencia se puede encontrar una buena aproximación a x^*, es decir,

$$x^{m+1} = x^m - H^{-1}(x^m)p(x^m) \tag{2.3}$$

Por lo tanto se tiene un método iterativo, con un valor inicial de x^0, y finaliza cuando la diferencia entre los valores es pequeña, es decir

$$x^{m+1} = x^m + \delta \qquad \text{para } \delta \text{ cercano a cero}$$

Donde δ es un valor que determina el usuario o el software empleado.

2.2 Estimación de los MLG

Como primer paso para estimar los MLG se aplica el método de máxima verosimilitud y posteriormente se recurre a utilizar métodos de aproximación numérica.

Por lo que se toma la función de máxima verosimilitud (1.2) y posteriormente se aplica el logaritmo, quedando así la ecuación:

$$l(\theta; y) = \left\{ \sum y_i b(\theta_i) + \sum c(\theta_i) + \sum d(y_i) \right\} \tag{2.4}$$

Se tiene además por lo tratado en el Apéndice A que

$$E(Y_i) = \mu_i = \frac{-c'(\theta_i)}{b'(\theta_i)} \tag{2.5}$$

y $$Var(Y_i) = \frac{b''(\theta_i)c'(\theta_i) - b'(\theta_i)c''(\theta_i))}{[b'(\theta_i)]^3} \tag{2.6}$$

Y suponiendo la existencia de un componente sistemático dado por

$$g(\mu_i) = \sum_{j=1}^{p} \beta_j X_{ij} = \eta_i \tag{2.7}$$

Donde $g(.)$ es una función monótona y diferenciable

Una propiedad de las distribuciones de la familia exponencial es que satisfacen las condiciones de regularidad, las cuales aseguran que el máximo global de la función log-verosímil $l(0;y)$ está dado de manera única por la solución de las ecuaciones $\frac{\partial l}{\partial \beta} = 0$. Entonces el interés consiste en derivar $l(0;y)$ con respecto a β_j, obteniéndose la función puntaje (score) dada por

$$U_j = \frac{dl(\theta;y)}{d\beta_j} = \sum_{i=1}^{n} \frac{\partial l_i}{\partial \beta_j} \tag{2.8}$$

donde

$$l_i = y_i b(\theta_i) + c(\theta_i) + d(y_i)$$

Para obtener U_j se aplica la regla de la cadena, es decir;

$$\frac{\partial l_i}{\partial \beta_j} = \frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j} \tag{2.9}$$

Para encontrar tal expresión se calcula por partes.

Tomando la ecuación log-verosímil y derivando con respecto a θ se obtiene

$$\frac{\partial l_i}{\partial \theta} = y_i b'(\theta_i) + c'(\theta_i) \tag{2.10}$$

Despejando $c'(\theta)$ de (2.5) y sustituyendo en (2.10) se tiene

$$\begin{aligned} \frac{\partial l_i}{\partial \theta} &= y_i b'(\theta_i) + c'(\theta_i) \\ &= b'(\theta_i)(y_i - \mu_i) \end{aligned} \tag{2.11}$$

Ahora tomando la ecuación (2.5)

$$\mu_i = \frac{-c'(\theta_i)}{b'(\theta_i)}$$

y calculando su derivada

$$\frac{\partial \mu_i}{\partial \theta_i} = \frac{-c''(\theta_i)b'(\theta_i) + c'(\theta_i)b''(\theta_i)}{[b'(\theta_i)]^2}$$

multiplicando por $\frac{b'(\theta_i)}{b'(\theta_i)}$ la ecuación anterior queda como

$$\frac{\partial \mu_i}{\partial \theta_i} = \frac{-c''(\theta_i)b'(\theta_i) + c'(\theta_i)b''(\theta_i)}{[b'(\theta_i)]^2}$$

$$= \frac{b'(\theta_i)}{1}\left(\frac{-c''(\theta_i)b'(\theta_i) + c'(\theta_i)b''(\theta_i)}{[b'(\theta_i)]^3}\right)$$

y sustituyendo (2.6) se obtiene que

$$\frac{\partial \mu_i}{\partial \theta_i} = b'(\theta_i)Var(Y_i) \tag{2.12}$$

Calculando el reciproco de la derivada anterior se obtiene

$$\frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{b'(\theta_i)Var(Y_i)}$$

Por último derivando (2.7) se tiene

$$\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial \beta_j}$$

$$= x_{ij}\frac{\partial \mu_i}{\partial \eta_i}$$

Por lo tanto la ecuación queda como

$$\frac{\partial l_i}{\partial \beta_j} = \frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}$$

$$= \cancel{b'(\theta_i)}(y_i - \mu_i)\frac{1}{\cancel{b'(\theta_i)}Var(Y_i)}x_{ij}\frac{\partial \mu_i}{\partial \eta_i}$$

$$= \frac{(y_i - \mu_i)x_{ij}}{Var(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) \tag{2.13}$$

Los elementos de la matriz de información están definidos por

$$\begin{aligned}\wp_{jk} &= E[U_j U_k] \\ &= E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] \quad \text{Por el apéndice B}\end{aligned}$$

Por (2.13), para cada Y_i la contribución a $\wp_{jk}$ está dada por

$$\begin{aligned}E\left[\frac{\partial l}{\partial \beta_j}\frac{\partial l}{\partial \beta_k}\right] &= E\left[\frac{(y_i-\mu_i)^2 x_{ij}x_{ik}}{(\mathrm{Var}(Y_i))^2}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right] \\ &= \left[\frac{x_{ij}x_{ik}}{(\mathrm{Var}(Y_i))^2}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right] E\left((y_i-\mu_i)^2\right)\end{aligned} \tag{2.14}$$

pero por definición

$$Var[Y] = E[(y-\mu)^2]$$

por lo que la ecuación anterior (2.14) queda como

$$\begin{aligned}E\left[\frac{\partial l}{\partial \beta_j}\frac{\partial l}{\partial \beta_k}\right] &= \left[\frac{x_{ij}x_{ik}}{(\mathrm{Var}(Y_i))^2}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right] E\left((y_i-\mu_i)^2\right) \\ &= \left[\frac{x_{ij}x_{ik}}{(\mathrm{Var}(Y_i))}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right]\end{aligned} \tag{2.15}$$

Por lo tanto la contribución total está dada por (ver Apéndice D)

$$\wp_{jk} = \sum_{i=1}^{n}\left[\frac{x_{ij}x_{ik}}{(\mathrm{Var}(Y_i))}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right] \tag{2.16}$$

Renombrando la ecuación (2.13) se tiene la función puntaje definida por

$$\frac{\partial l_i}{\partial \beta_j} = U_j = \frac{(y_i-\mu_i)x_{ij}}{\mathrm{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) \tag{2.17}$$

donde x_{ij} es el j-ésimo elemento de x_i'. En general las ecuaciones $U_j = 0$ (j=1,...,p) son no lineales y deben ser resueltas por iteraciones numéricas.

Por lo que se utiliza el método de Newton-Raphson descrito en la seccion 2.1 de este capítulo, por lo que tomando la ecuación (2.2) se tiene

$$x^1 = x^o - H^{-1}(x^o)p(x^o)$$

Renombrando los elementos; $x^o = \underline{\beta}^o$ es el vector inicial con el cual se calcula la primera aproximación, $p(x^o) = U(\underline{\beta}^o)$ es el vector que contiene las primeras derivadas de la función $l(\theta;y)$ evaluadas en $\underline{\beta}^o$, y por último $H^{-1}(x^o) = H^{-1}(\underline{\beta}^o)$ es la matriz inversa de segundas derivadas de la función $l(\theta;y)$ evaluada en $\underline{\beta}^o$.

Dado que el proceso es iterativo, se toma la m-ésima aproximación, es decir, la ecuación **(2.3)**

$$x^m = x^{m-1} - H^{-1}(x^{m-1})p(x^{m-1})$$

De igual manera, renombrando se tiene que $x^m = \beta^m$ es la m-ésima aproximación del vector de parámetros que maximizan la función log máximo verosímil $l(\theta;y)$.

$p(\underline{x}^{m-1}) = U(\beta^{m-1})$ es el vector que contiene las primeras derivadas de la función $l(\theta;y)$ evaluadas en $\underline{\beta}^{m-1}$, y $H^{-1}(x^m) = H^{-1}(\underline{\beta}^{m-1})$ es la matriz inversa de segundas derivadas de la función $l(\theta;y)$ evaluada en $\underline{\beta}^{m-1}$ (m-1)-ésima aproximación.

Por lo tanto la ecuación iterativa queda de la siguiente forma:

$$\underline{\beta}^m = \underline{\beta}^{m-1} - H^{-1}(\underline{\beta}^{m-1})U(\underline{\beta}^{m-1}) \qquad (2.18)$$

En donde se debe de iniciar con un valor $\underline{\beta}^o$ el proceso iterativo hasta encontrar una convergencia de los parámetros según el método de Newton-Raphson. Por lo que se observa que el cálculo de los parámetros puede ser un proceso tedioso en caso de iniciar con un vector inicial "malo", ya que el cálculo tardará en converger.

Y por el contrario, será eficiente en caso de proporcionar un vector inicial muy cercano a los parámetros reales.

2.3 Método Iterativo Ponderado de Mínimos Cuadrados.

El método Iterativo Ponderado de Mínimos Cuadrados es relacionado con el método de Puntajes de Fisher(Scoring's Fisher), el cual consiste en tomar la ecuación (2.18) y reemplazar la matriz de las segundas derivadas por la esperanza de la matriz de los valores esperados, es decir,

$H(\underline{\beta}^{m-1})$ con $E\left(H(\underline{\beta}^{m-1})\right) = \wp(\underline{\beta}^{m-1})$

En el apéndice B se muestra que

$$\begin{aligned}\wp_{jk} &= E\left[U_j U_k\right] \\ &= E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] \\ &= -E\left[\frac{\partial^2 l}{\partial \theta_j \partial \theta_k}\right]\end{aligned}$$

Por lo que en la ecuación (2.18) se hace la sustitución quedando de la siguiente forma

$$\underline{\beta}^{m} = \underline{\beta}^{m-1} + \left[\wp(\underline{\beta}^{m-1})\right]^{-1} U(\underline{\beta}^{m-1})$$

En donde $\wp(\underline{\beta}^{m-1})$ denota la matriz de información evaluada en $\underline{\beta}^{m-1}$.

Multiplicando $\wp(\underline{\beta}^{m-1})$ por la izquierda a ambos lados de la ecuación se obtiene

$$\begin{aligned}\wp(\underline{\beta}^{m-1})\underline{\beta}^{m} &= \wp(\underline{\beta}^{m-1})\underline{\beta}^{m-1} + \wp(\underline{\beta}^{m-1})\left[\wp(\underline{\beta}^{m-1})\right]^{-1} U(\underline{\beta}^{m-1}) \\ &= \underline{\wp(\underline{\beta}^{m-1})\beta^{m-1}} + U(\underline{\beta}^{m-1})\end{aligned} \qquad (2.19)$$

En donde el elemento *(j,k)* de la matriz $\wp(\underline{\beta}^{m-1})$ está dado por (2.15), es decir;

$$\wp_{jk} = \left[\sum_{i=1}^{n} \frac{x_{ij}x_{ik}}{\mathrm{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right] \tag{2.20}$$

Por lo que se puede expresar a $\wp$ como sigue

$$\wp = X^t W X$$

Donde W es una matriz diagonal con elementos;

$$w_{ii} = \frac{1}{\mathrm{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \tag{2.21}$$

La expresión del lado derecho de la ecuación (2.19) es el vector con elementos

$$\sum_{k=1}\sum_{i=1}^{n} \frac{x_{ij}x_{ik}}{\mathrm{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \beta_k^{m-1} + \sum_{i=1}^{n} \frac{(y_i-\mu_i)x_{ij}}{\mathrm{Var}(Y_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)$$

Evaluado en la (m-1)-ésima aproximación de β y rescribiéndola utilizando (2.20) y (2.21) queda como

$$X^T W z$$

Donde los elementos de z son

$$z_i = \sum_k x_{ik}\beta_k^{(m-1)} + (y_i - \mu_i)\left(\frac{\partial \eta_i}{\partial \mu_i}\right)$$

en tal ecuación μ_i y η_i están evaluadas en $\underline{\beta}^{m-1}$

Entonces el método de puntajes puede ser escrito como la ecuación iterativa siguiente:

$$X^T W X \underline{\beta}^{(m)} = X^T W z \tag{2.22}$$

De manera equivalente, se encuentra la inversa de X^tWX y se multiplica por ambos lados se tiene

$$\underline{\beta}^{(m)} = (X^T W X)^{-1} X^T W z$$

En el método Iterativo Ponderado de Mínimos Cuadrados(IRLS) no es necesario indicar valores iniciales como en el método de Newton – Raphson, el cual necesita valores iniciales de los parámetros, con este método basta únicamente con dar un valor aproximado para los valores ajustados de $\hat{\mu}_i$, lo cual es más sencillo de implementar. Una propuesta es la siguiente:

$\hat{\mu}_i = \frac{(y_i + \overline{y})}{2}$	Para un modelo con Distribución No Binomial
$\hat{\mu}_i = \frac{k_i(y_i + .5)}{(k_i + 1)}$	Para un modelo con Distribución Binomial

Tal propuesta es hecha por Hardin[11].

Asimismo se utilizará la Devianza definida como $D = 2[l(b_{max}; y) - l(b; y)]$. Más adelante se mostrará una definición formal.

En otras palabras, cuando se está estimando un modelo particular, el objetivo es encontrar los valores de los parámetros que minimicen la Devianza, entonces el método de IRLS se detiene cuando la diferencia entre las Devianzas de sucesivas iteraciones es muy pequeña, es decir, menor que la tolerancia definida por el usuario o en su defecto el paquete estadístico utilizado. Los valores de los parámetros que minimizan la devianza son los mismos que maximizan la función log-verosímil[12]. Más adelante se dará un argumento de la utilización de la Devianza como medida de bondad de ajuste, lo anterior será discutido en la sección **(2.6)**.

[11] Hardin James, Hilbe Joseph, Generalized Linear Models and extensions, STATA Press pág. 23,2001
[12] Hardin James, Hilbe Joseph, Generalized Linear Models and extensions, STATA Press. 2001.

En términos prácticos, el siguiente algoritmo es un propuesta que se utiliza en el método Iterativo Ponderado de Mínimos Cuadrados (IRLS) para estimar los parámetros del modelo en cuestión.

Algoritmo para estimar Modelos Lineales Generalizados por el método Iterativo Ponderado de Mínimos Cuadrados:

1.Inicializar el valor de μ_i

2.Calcular el valor de la función liga, es decir, $\eta = g(\mu)$

3.Inicializar el valor de Devianza_Temporal_1, i.e., Dev_ant=0

4.Inicializar el valor de Devianza_Temporal_2, i.e., Dev_nva=1

5.Inicializar el valor de la Diferencia entre Devianzas, i.e. $\Delta D = 1$

6.Iniciar un proceso que se repita hasta que la condición no se cumpla, i.e.,

$While(|\Delta D| > tolerancia)$ {

calcular W

calcular z

realizar el proceso de $\underline{\beta}^{(m)} = (X^T W X)^{-1} X^T W z$

Calcular $\eta = X\beta$

Calcular $\mu = g^{-1}(\eta)$

Haz $Dev\_ant = Dev\_nva$

Calcula la devianza nueva $Dev\_nva = 2[l(b_{max}; y) - l(b; y)]$

Calcula la diferencia entre Devianzas $\Delta D = Dev\_nva - Dev\_ant$

} ***se repite el proceso hasta que la condicion no se cumpla*

Calcula la matriz de Varianzas y covarianzas de los estimadores $V = \left[E\left[-\frac{\partial^2 l}{\partial\theta_j \partial\theta_k}\right]\right]^{-1}$

7. Fin de Proceso Hasta que no se cumpla la condición.

Este Algoritmo es aplicado por STATA para estimar modelos lineales generalizados.

2.4 Inferencia

Una vez obtenido el estimador máximo verosímil en la sección anterior, ahora se analizará si el estimador es insesgado y posteriormente se calculará un intervalo de confianza.

2.4.1 Insesgamiento y distribución de los estimadores.

Supóngase que la función log-verosímil $l(\theta;y)$ tiene un único máximo b y que este estimador está cerca del verdadero valor del parámetro β. Utilizando el polinomio de Taylor de primer orden para el vector de puntajes $U(\beta)$ alrededor del punto $\beta = b$, se obtiene

$$U(\beta) \cong U(b) + H(b)(\beta - b)$$

Lo anterior utilizando la ecuación (2.1) y renombrando a $U(b)$ por el vector de las primeras derivadas.

Donde $H(b)$ denota la matriz de las segundas derivadas de la función log-verosímil evaluada en $\beta = b$, se tiene además por el apéndice B que asintóticamente la esperanza de H es igual a la matriz de información, es decir;

$$\mathcal{O}_{jk} = E\left[U_j U_k\right] = E\left[\frac{\partial l}{\partial \beta_j}\frac{\partial l}{\partial \beta_k}\right] = E(-H)$$

Por lo tanto para muestras grandes se sustituye la matriz H por la matriz de información $\mathcal{O}$, quedando como

$$U(\beta) \cong U(b) + \wp(\beta - b)$$

Pero $U(b) = 0$, ya que b es el punto en el cual se maximiza la función log – verosímil $l(\theta; y)$ y por lo tanto al evaluar en las primeras derivadas éstas son cero. Por lo que aproximadamente se tiene que

$$U(\beta) \cong \wp(\beta - b)$$

Suponiendo que $\wp$ es no singular se calcula la inversa obteniendo

$$\wp^{-1}U(\beta) \cong (\beta - b)$$
$$(\beta - b) \cong \wp^{-1}U(\beta)$$

Calculando la esperanza a la ecuación anterior y considerando a $\wp$ como constante, queda como

$$E(\beta - b) \cong \wp^{-1}E(U(\beta))$$

Pero por el apéndice B se tiene que $E(U)=0$,

Por lo que se tiene que b es un estimador insesgado de β, al menos asintóticamente, es decir;

$$E(b) = E(\beta) = \beta$$

Por otra parte la matriz de varianzas y covarianzas de b está dada por

$$\begin{aligned} E\left[(\beta - b)(\beta - b)^T\right] &= E\left[\wp^{-1}U(\wp^{-1}U)^T\right] \\ &= E\left[\wp^{-1}UU^T\left(\wp^{-1}\right)^T\right] \\ &= \wp^{-1}E(UU^T)\wp^{-1} \end{aligned}$$

ya que $\left(\wp^{-1}\right)^T = \wp^{-1}$ por ser simétrica

Además por el apéndice B se tiene que

$$E(UU^T) = \wp$$

Sustituyendo se tiene

$$E\left[(\beta - b)(\beta - b)^T\right] = (\wp^{-1}\wp)\wp^{-1}$$
$$= \wp^{-1}$$

Por lo que la inversa de matriz de información es la matriz de varianzas y covarianzas de los estimadores.

Se sabe además para muestras grandes y aplicando el teorema del limite central que

$$(b - \beta) \sim N(0, \wp^{-1}) \tag{2.23}$$

O de manera equivalente

$$(b - \beta)^T \wp (b - \beta) \sim \chi^2_p \tag{2.24}$$

ambas distribuciones de manera asintótica.

La estadística $(b - \beta)^T \wp (b - \beta)$ es algunas veces llamada la estadística de Wald y se utiliza para hacer inferencias acerca de β.

2.4.2 Intervalos de confianza

Obtenida la ecuación (2.23) se deduce que para muestras grandes y de manera asintótica que

$$b \sim N(\beta, \wp^{-1}) \tag{2.25}$$

Por lo que esta distribución se utiliza para:

- Asegurar la confiabilidad de los estimadores b_j por medio de las magnitudes de sus errores estándares

$$s.e.(b_j) = \sqrt{v_{jj}}$$

donde v_{jj} es el j-ésimo término de la matriz $\wp^{-1}$

- Calcular intervalos de confianza para los parámetros de manera individual, por ejemplo, para un intervalo de confianza al 95% para b_j, se tiene;

$$b_j \pm 1.96\sqrt{v_{jj}}$$

- Examinar las correlaciones entre los estimadores usando

$$corr(b_j, b_k) = \frac{v_{jj}}{\sqrt{v_{jj}}\sqrt{v_{kk}}}$$

Lo anterior será utilizado en secciones posteriores.

2.5 Bondad de Ajuste

Una vez obtenidos los estimadores se desea analizar si el modelo en realidad ajusta al conjunto de datos. Esto puede hacerse comparando la verosimilitud de un modelo particular con la verosimilitud del modelo saturado, el cual es definido como sigue:

1. El modelo saturado es un modelo lineal generalizado usando la misma distribución del modelo en cuestión.
2. El modelo saturado tiene la misma función liga que el modelo en cuestión.
3. El número de parámetros en el modelo saturado es igual al número total de observaciones.

El punto 3. dice que el modelo saturado es aquel modelo que se ajusta de manera perfecta al conjunto de datos.

Las funciones de probabilidad para el modelo saturado y el modelo de interés deben de ser evaluadas en su respectivos máximos estimados para obtener los valores de $L(b_{max}; y)$ y $L(b; y)$ respectivamente. Si el modelo de interés describe a los datos bien entonces $L(b; y)$ debe de ser aproximadamente igual a $L(b_{max}; y)$.

Si el modelo es pobre en ajuste entonces $L(b; y)$ debe ser mucho más pequeño que $L(b_{max}; y)$. Lo anterior sugiere el uso de la siguiente estadística.

$$\lambda = \frac{L(b_{max}; y)}{L(b; y)}$$

Como una medida de bondad de ajuste. O equivalentemente tomando logaritmos se tiene la diferencia entre funciones log-verosímiles, dado por

$$\log \lambda = l(b_{max}; y) - l(b; y)$$

Para valores grandes de log λ, la estadística sugiere que el modelo es pobre respecto a la descripción de los datos. Para determinar una región critica y realizar una prueba se necesita saber cual es la distribución muestral.

2.6 Distribución muestral de la estadística log-verosímil

Supóngase que un cierto modelo de interés tiene p parámetros denotado por el vector parametral β. Una aproximación en serie de Taylor para $l(\beta; y)$ puede ser obtenida expandiendo alrededor del estimador máximo verosímil $\boldsymbol{b}$.

$$l(\beta; y) \cong l(b; y) + (\beta - b)U(b) + \frac{1}{2}(\beta - b)^T H(b)(\beta - b) \qquad (2.26)$$

Donde U(b) es el vector de puntajes, es decir, $\frac{\partial l}{\partial \beta_j}$ evaluado en $\boldsymbol{b}$ y H($\boldsymbol{b}$) es la matriz de segundas derivadas

$$\frac{\partial^2 l}{\partial \beta_j \partial \beta_i}$$

evaluado en **b**. Por definición de **b**, se tiene que $U(b) = 0$. Para muestras grandes $-H(b)$ puede ser aproximado por la matriz de información $\mathcal{I} = E(-H)$ ya que $H(b)$ es igual a su valor esperado $E(H(b))$ en forma asintótica. Utilizando la ecuación (2.26) se puede obtener el siguiente arreglo

$$l(b; y) - l(\beta; y) \cong \frac{1}{2}(b - \beta)^T \mathcal{I}(b - \beta)$$

Pero por la ecuación (2.24) se tiene que $(b-\beta)^T \wp (b-\beta) \sim X^2_p$, por lo que despejando de la ecuación anterior se tiene que

$$2[l(b;y)-l(\beta;y)] \sim X^2_p \tag{2.27}$$

Esta estadística será utilizada para asegurar el ajuste de un modelo y para comparar modelos alternativos.

2.6.1 La Devianza como medida de Bondad de Ajuste

Nelder y Wedderburn[13] definen a la estadística del Cociente de log-verosimilitud como la "Devianza", definida como

$$D = 2\log\lambda = 2[l(b_{max};y)-l(b;y)] \tag{2.28}$$

Tal ecuación puede ser arreglada como sigue

$$\begin{aligned} D = 2\{&[l(b_{max};y)-l(\beta_{max};y)] \\ &-[l(b;y)-l(\beta;y)] \\ &+[l(\beta_{max};y)-l(\beta;y)]\} \end{aligned} \tag{2.29}$$

El primer elemento de corchetes del lado derecho tiene una distribución X^2_N, esto por la ecuación (2.27) y dado que es el modelo saturado tiene N parámetros. Por el mismo argumento, el segundo elemento del lado derecho tiene una distribución X^2_p. Por último, si el modelo con p parámetros describe a los datos tan bien como el modelo saturado entonces el tercer elemento es una constante positiva la cual debe ser cercana a cero. Es decir, si las variables aleatorias definidas por los primeros dos términos son independientes y el tercer término es cercano a cero, entonces

$$D \sim X^2_{N-p}$$

[13] Nelder y Wedderburn, *Generalized Linear Models*, Chapman and Hall, 1972

si el modelo es bueno en ajuste. Es decir, si el valor de la devianza del modelo es menor que el valor predecido por una Ji-Cuadrada al 100(1-α)%, entonces el modelo tiene buen ajuste.

Si el modelo es pobre en ajuste el tercer término de la ecuación (2.29) debe de ser grande y por lo tanto D será más grande que el valor predecido de una distribución Ji-Cuadrada con N-p grados de libertad, ya que $D \sim X^2_{N-p}$.

Lo anterior es la justificación también de porque se utiliza la diferencia entre devianzas de sucesivas iteraciones como finalización del proceso iterativo. Ya que cuando la diferencia entre éstas es menor que una cierta tolerancia, entonces el modelo con p parámetros ajusta de buena manera al modelo saturado con lo que se obtiene el estimador máximo verosímil.

Por lo que se tiene que la distribución muestral de la estadística del Cociente de log-verosimilitud puede ser usado para investigar el ajuste de un modelo si se estima la Devianza D de los datos y se compara con el valor apropiado de una distribución Ji-Cuadrada.

En términos prácticos, $\hat{\mu}_i$ denotará el estimador máximo verosímil de μ_i bajo el modelo de interés y $\tilde{\mu}_i = y_i$ denotará el estimador máximo verosímil bajo el modelo saturado. Dichos valores serán sustituidos en la ecuación (2.28).

2.6.2 Devianza Binomial

Para el caso de la distribución Binomial, la estadística de la Devianza será utilizada sólo cuando existan datos agrupados, es decir $k_i > 1$, por lo que la estadística de la Devianza converge a una distribución Ji-Cuadrada con k-p grados de libertad cuando

$k_i \to \infty$, para toda i, donde p es el número de parámetros incluida la constante. Por lo que para grupos razonablemente grandes, la Devianza proporciona un buen ajuste por el modelo.

Cuando se tiene que $k_i = 1$ para toda i, es decir con datos individuales la distribución de la Devianza no converge a una distribución Ji-Cuadrada, u otra conocida. Por lo que no puede ser usada como una medida de bondad de ajuste.[14]

En términos generales, la diferencia de las devianzas entre modelos tiene una distribución asintótica Ji-Cuadrada si el número de grupos $k \to \infty$ o el tamaño de cada grupo $k_i \to \infty$.

2.7 Bondad de Ajuste: Matriz de la Confusión

Esta medida es fácilmente calculada para cuando se tienen datos de una distribución Binomial, es decir en donde se involucra una probabilidad.
La matriz de confusión simplemente clasifica el número de casos bien predecidos y los predecidos erróneamente.
Lo anterior se observa en la siguiente tabla:

Observado	Predecido	Resultado de predicción
y=1	y=1	Correcto
y=1	y=0	Incorrecto
y=0	y=1	Incorrecto
y=0	y=0	Correcto

[14] http://data.princeton.edu/wws509, Princeton University.

Para calcular el valor predecido se utiliza la probabilidad generada por el modelo lineal generalizado, ya sea utilizando la liga logit o probit, ya que ambas nos proporcionan una probabilidad, tomando la siguiente regla de decisión

$$\hat{y}_{id} = \begin{cases} 1 \text{ si } \pi_i(\bar{x}) \geq 0.5 \\ 0 \text{ si } \pi_i(\bar{x}) < 0.5 \end{cases} \qquad (2.30)$$

Donde en el caso de una Distribución Binomial $\hat{y}$ es considerado el número de éxitos en el grupo con k_i. Con este valor, se determina si una Y de cierto grupo tiene probabilidad de obtener un éxito dada la regla anterior.
En términos de una matriz queda de la siguiente manera:

Clasificado	Observado	
	0	1
0	α_1	δ_1
1	δ_2	α_2

En esta se determinará el número de valores predichos correctamente y posteriormente se tomará el cociente entre el total de valores predichos. En términos matemáticos se tiene, la siguiente medida:

(Porcentaje de Clasificación Correcta de la Matriz de Confusión)

$$PMC = \frac{(\alpha_1 + \alpha_2)}{(\alpha_1 + \alpha_2 + \delta_1 + \delta_2)}$$

En donde PMC determina porcentaje de clasificaciones realizadas correctamente.

Por ejemplo. Supóngase la siguiente tabla:

MC =

Clasificado	Observado	
	0	1
0	25	3
1	15	45

El PMC es igual a ***PMC = (25+45)/(25+45+15+3) = 70/88 = 87.5 %*** de respuestas clasificadas correctamente.

Ésta medida será utilizada posteriormente en el capitulo 4.

2.8 Pruebas de Hipótesis

Una vez obtenidas las distribuciones de los parámetros y del cociente de log-verosimilitud, se pueden realizar pruebas de hipótesis de los parámetros, asimismo para comparar modelos con distinto número de parámetros. Es decir, las hipótesis acerca de los parámetros pueden ser probadas usando las distribuciones asintóticas definidas en las ecuaciones (2.23) y (2.24), dadas por $(b-\beta) \sim N(0,\mathcal{O}^{-1})$ y $(b-\beta)^T \mathcal{O}(b-\beta) \sim X^2_p$ respectivamente.

Como caso general, suponga que se desea comparar dos modelos, uno con p parámetros y el otro con q.

Es decir, considérese la hipótesis nula

$$H_0 = \beta = \beta_o = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_q \end{bmatrix}$$

y una hipótesis más general, dada por

$$H_1 = \beta = \beta_1 = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad \text{donde } q < p < N$$

Se puede probar H_0 contra H_1 utilizando como estadístico de prueba a la diferencia del cociente de log-verosimilitudes, dadas por

$$\begin{aligned} \Delta D = D_0 - D_1 &= 2[l(b_{max};y) - l(b_0;y)] - 2[l(b_{max};y) - l(b_1;y)] \\ &= 2[l(b_1;y) - l(b_0;y)] \end{aligned}$$

Si ambos modelos describen los datos con un buen ajuste entonces $D_0 \sim X^2_{N-q}$ y $D_1 \sim X^2_{N-p}$, implicando que $\Delta D \sim X^2_{p-q}$ (suponiendo que existen condiciones de independencia). Si el valor de ΔD es consistente con una distribución X^2_{p-q}

entonces se seleccionaría el modelo correspondiente a H_0 ya que es más simple, es decir tiene menos parámetros y ajusta bien a los datos. Ahora si el valor de ΔD está en la región critica (es decir, es mayor que el punto (1-α)100% de una distribución Ji-Cuadrada), entonces se debe de rechazar la hipótesis nula y aceptar H_1, ya que proporciona un mejor ajuste de los datos, aun cuando el número de parámetros es mayor.

CAPÍTULO 3

Estimación de algunos modelos seleccionados

El objetivo de este capítulo es definir en forma explicita las ecuaciones que se utilizan para poder llevar a cabo la estimación de algunos modelos lineales generalizados. Esta selección ha sido realizada con base a que dichos modelos serán utilizados posteriormente.

Como se mencionó se mostrará el proceso de estimación explicado en el capítulo anterior aplicado a algunos modelos lineales generalizados, la estimación se realizará utilizando el método iterativo ponderado de mínimos cuadrados (Iterative Reweigthed Least Squares).

Las distribuciones seleccionadas son:

a) La distribución Binomial con su liga canónica Logit
b) La distribución Binomial con la liga Probit
c) La distribución Poisson con su liga canónica.

Asimismo se calculará la devianza de cada modelo, con el fin de utilizarlo posteriormente como una medida de bondad de ajuste entre el modelo saturado y el modelo con p-parámetros.

3.1 Modelo Logit

Supóngase que se tiene una variable aleatoria de respuesta con distribución Binomial, es decir, $Y \sim Bin(k_i, \pi_i)$,

Con función de densidad

$$f(y_i; k_i; \pi_i) = \frac{k_i!}{y_i!(k_i - y_i)!} \pi_i^{y_i} (1 - \pi_i)^{k_i - y_i} \tag{3.1}$$

El método de estimación requiere calcular las siguientes ecuaciones

$$z_i = \sum_k x_{ik} \beta_k^{(m-1)} + (y_i - \hat{\mu}_i)\left(\frac{\partial \eta_i}{\partial \mu_i}\right) = \left[\hat{\eta}_i + (y_i - \hat{\mu}_i)\left(\frac{\partial \eta_i}{\partial \mu_i}\right)\right]_{\beta_i^{(m-1)}} \tag{3.2}$$

y

$$w_{ii} = \frac{1}{(Var(Y_i))}\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \tag{3.3}$$

Para después evaluarlas en la ecuación

$$X^T W X \underline{\beta}^{(m)} = X^T W z$$

equivalentemente

$$\underline{\beta}^{(m)} = (X^T W X)^{-1} X^T W z \tag{3.4}$$

Se tiene que la relación entre el componente lineal sistemático y la función liga están dados por

$$\eta_i = \sum_{j=1}^{p} \beta_j X_{ij} \qquad \text{y} \qquad \eta_i = g(\mu_i) = \ln\left(\frac{\mu_i}{k_i - \mu_i}\right)$$

Calculando entonces las derivadas requeridas, se obtiene

$$\begin{aligned}\frac{\partial \eta_i}{\partial \mu_i} &= \frac{\partial}{\partial \mu_i}\left(\ln(\mu_i) - \ln(k_i - \mu_i)\right) \\ &= \frac{1}{\mu_i} - \frac{-1}{k_i - \mu_i} = \frac{1}{\mu_i} + \frac{1}{k_i - \mu_i} \\ &= \frac{k_i - \cancel{\mu_i} + \cancel{\mu_i}}{(\mu_i)(k_i - \mu_i)} = \frac{k_i}{(\mu_i)(k_i - \mu_i)}\end{aligned}$$

Por lo que la ecuación iterativa (3.2) queda como

$$z_i = \left[\hat{\eta}_i + \frac{(y_i - \hat{\mu}_i)k_i}{(\hat{\mu}_i)(k_i - \hat{\mu}_i)}\right]_{\beta_k^{(m-1)}}$$

Por otra parte, para calcular la derivada $\frac{\partial \mu_i}{\partial \eta_i}$ se utilizará derivación implícita.

Es decir, se toma la función liga

$$\begin{aligned}\eta_i &= g(\mu_i)\\ &= \ln\left(\frac{\mu_i}{k_i - \mu_i}\right)\\ &= \ln(\mu_i) - \ln(k_i - \mu_i)\end{aligned}$$

Tomando la ecuación anterior y derivando de manera implícita considerando a η_i como una función de μ_i

$$\begin{aligned}\frac{\partial \mu_i}{\partial \eta_i} &= \frac{\partial}{\partial \eta_i}\left(\ln(\mu_i) - \ln(k_i - \mu_i)\right)\\ &= \frac{\partial \ln(\mu_i)}{\partial \eta_i} - \frac{\partial \ln(k_i - \mu_i)}{\partial \eta_i}\\ &= \frac{\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{\mu_i} - \frac{\left(-\frac{\partial \mu_i}{\partial \eta_i}\right)}{k_i - \mu_i} = \frac{\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{\mu_i} + \frac{\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{k_i - \mu_i}\\ &= \frac{(k_i - \mu_i)\left(\frac{\partial \mu_i}{\partial \eta_i}\right) + (\mu_i)\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{(\mu_i)(k_i - \mu_i)}\\ &= \frac{k_i\left(\frac{\partial \mu_i}{\partial \eta_i}\right) - \cancel{\mu_i\left(\frac{\partial \mu_i}{\partial \eta_i}\right)} + \cancel{(\mu_i)\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}}{(\mu_i)(k_i - \mu_i)}\\ &= \frac{k_i\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{(\mu_i)(k_i - \mu_i)}\end{aligned}$$

En el lado izquierdo de la ecuación se tiene que $\frac{\partial \mu_i}{\partial \eta_i} = 1$, por lo que la ecuación queda como

$$1 = \frac{k_i\left(\frac{\partial \mu_i}{\partial \eta_i}\right)}{(\mu_i)(k_i - \mu_i)}$$

Despejando $\frac{\partial \mu_i}{\partial \eta_i}$

$$\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = \frac{(\mu_i)(k_i - \mu_i)}{k_i}$$

Entonces la ecuación w_{ii} (3.3), queda de la siguiente forma

$$w_{ii} = \frac{1}{(Var(Y_i))}\left[\frac{(\mu_i)(k_i - \mu_i)}{k_i}\right]^2 \qquad (3.5)$$

Pero se tiene que

$Y_i \sim \text{Bin}(k_i, \pi_i)$, entonces $E(Y_i) = k_i\pi_i = \mu_i$

$$\begin{aligned} Var(Y_i) &= k_i\pi_i(1-\pi_i) \\ &= \mu_i\left(1 - \frac{\mu_i}{k_i}\right) \\ &= \mu_i\left(\frac{k_i - \mu_i}{k_i}\right) \end{aligned}$$

Sustituyendo la $Var(Y_i)$ en la ecuación (3.5)

$$w_{ii} = \frac{1}{\left(\mu_i(1-\frac{\mu_i}{k_i})\right)}\left(\frac{(\mu_i)(k_i-\mu_i)}{k_i}\right)^2$$

$$= \frac{k_i}{\left(\mu_i(k_i-\mu_i)\right)}\left(\frac{(\mu_i)^2(k_i-\mu_i)^2}{k_i^2}\right)$$

$$= \left(\frac{(\mu_i)(k_i-\mu_i)}{k_i}\right)$$

Una vez obtenidas las ecuaciones iterativas, éstas se sustituyen en la ecuación (3.4), quedando de la siguiente manera:

$$\underline{\beta}^{(m)} = (X^T W X)^{-1} X^T W z$$

Partiendo del hecho que se tiene una matriz de información dada por los valores observados de p variables explicativas, así como una muestra de tamaño n.
Las dimensiones de las matrices son: X es de tamaño $n \times p$, W es de tamaño $n \times n$ y Z es de $p \times 1$;

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \qquad W = \begin{bmatrix} \frac{\mu_1(k_1-\mu_1)}{k_1} & 0 & \cdots & 0 \\ 0 & \frac{\mu_2(k_2-\mu_2)}{k_2} & 0 \cdots & 0 \\ \vdots & \vdots & & \\ 0 & \cdots & & \frac{\mu_n(k_n-\mu_n)}{k_n} \end{bmatrix}$$

$$X^t = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix} \qquad Z = \begin{bmatrix} \hat{\eta}_1 + \dfrac{(y_1 - \hat{\mu}_1)k_1}{\mu_i(k_1 - \hat{\mu}_1)} \\ \vdots \\ \hat{\eta}_n + \dfrac{(y_n - \hat{\mu}_n)k_n}{\mu_n(k_n - \hat{\mu}_n)} \end{bmatrix}$$

Por lo que los elementos de la matriz $(X^T WX)$ están dados por

$$(X^T WX)_{ij} = \sum_{l=1}^{n} x_{li} \frac{\mu_l(k_l - \mu_l)}{k_l} x_{lj} \qquad i, j = 1, 2, \ldots, p$$

Asimismo los elementos de la matriz $(X^T WZ)$, quedan definidos por

$$(X^T WZ)_i\big|_{\beta^{m-1}} = \sum_{l=1}^{n} x_{il} \frac{\hat{\mu}_l(k_l - \hat{\mu}_l)}{k_l} \hat{\eta}_l + \sum_{l=1}^{n} x_{il}(y_l - \hat{\mu}_l) \qquad i = 1, 2, \ldots, p$$

Una vez calculadas éstas matrices se sustituyen en la ecuación iterativa dada por (3.4) y comienza el proceso de estimar los p parámetros utilizando el algoritmo descrito en la sección **(2.3)**.

Ahora se debe utilizar una medida de bondad de ajuste, en el capítulo 2 se plantea utilizar la devianza. La estimación de la devianza queda definida por

$$D = 2\log\lambda = 2\left[l(b_{max}; y) - l(b; y)\right]$$

la cual se contrasta con el valor de una distribución Ji – Cuadrada con (N-p) grados de libertad, ya que $D \sim X^2_{N-p}$.

Suponiendo que se tiene una muestra de tamaño n y p variables explicativas, $l(b; y)$ queda como

$$l(\theta;y)=\sum_{i=1}^{n} l_i(\theta;y_i) \qquad \text{donde } y=[Y_1,\ldots,Y_n]^T.$$
$$\text{y} \qquad \theta=(\theta_1,\theta_2,\ldots,\theta_p)$$

Tomando la función de densidad conjunta de la distribución Binomial definida en (3.1) y aplicando el logaritmo natural, se obtiene la función log-verosimil, quedando de la siguiente manera

$$\sum_{i=1}^{n} l(y_i;k_i;\pi_i)=\sum_{i=1}^{n}\ln(f(y_i;k_i;\pi_i))=\sum_{i=1}^{n}\ln\left(\frac{k_i!}{y_i!(k_i-y_i)!}\right)+y_i\ln(\pi_i)+(k_i-y_i)\ln(1-\pi_i)$$

Renombrando en términos de la media μ_i, es decir $\mu_i=k_i\pi_i \quad \Rightarrow \quad \pi_i=\frac{\mu_i}{k_i}$

$$\sum_{i=1}^{n} l(y_i;\mu_i)=\sum_{i=1}^{n}\ln\left(\frac{k_i!}{y_i!(k_i-y_i)!}\right)+y_i\ln\left(\frac{\mu_i}{k_i}\right)+(k_i-y_i)\ln\left(\frac{k_i-\mu_i}{k_i}\right)$$

Lo anterior es para el modelo de interés, es decir, se considera $\hat{\mu}_i$ como el estimador máximo verosímil de μ_i y bajo el modelo saturado se considerará a $\tilde{\mu}_i=y_i$ quedando como

$$\sum_{i=1}^{n} l(y_i;\mu_i)=\sum_{i=1}^{n}\ln\left(\frac{k_i!}{y_i!(k_i-y_i)!}\right)+y_i\ln\left(\frac{y_i}{k_i}\right)+(k_i-y_i)\ln\left(\frac{k_i-y_i}{k_i}\right)$$

Dichos valores serán sustituidos en la función de la Devianza, es decir;

$$
\begin{aligned}
D &= 2\left|l(b_{max};y) - l(b;y)\right| \\
&= 2\left[\left(\sum_{i=1}^{n} \ln\left(\frac{k_i!}{y_i!(k_i-y_i)!}\right) + y_i \ln\left(\frac{y_i}{k_i}\right) + (k_i - y_i)\ln\left(\frac{k_i - y_i}{k_i}\right)\right)\right. \\
&\qquad \left. - \left(\sum_{i=1}^{n} \ln\left(\frac{k_i!}{y_i!(k_i-y_i)!}\right) + y_i \ln\left(\frac{\mu_i}{k_i}\right) + (k_i - y_i)\ln\left(\frac{k_i - \mu_i}{k_i}\right)\right)\right] \\
&= 2\sum_{i=1}^{n}\left[\left(y_i \ln\left(\frac{y_i}{k_i}\right) + (k_i - y_i)\ln\left(\frac{k_i - y_i}{k_i}\right)\right) - \left(y_i \ln\left(\frac{\mu_i}{k_i}\right) + (k_i - y_i)\ln\left(\frac{k_i - \mu_i}{k_i}\right)\right)\right] \\
&= 2\sum_{i=1}^{n}\left[\left(y_i\left(\ln\left(\frac{y_i}{k_i}\right) - \ln\left(\frac{\mu_i}{k_i}\right)\right) + (k_i - y_i)\left(\ln\left(\frac{k_i - y_i}{k_i}\right) - \ln\left(\frac{k_i - \mu_i}{k_i}\right)\right)\right)\right]
\end{aligned}
$$

En donde por las propiedades de los logaritmos se cancela $\ln(k_i!)$, quedando como sigue

$$
D = 2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{y_i}{\mu_i}\right) + (k_i - y_i)\ln\left(\frac{k_i - y_i}{k_i - \mu_i}\right)\right]
$$

Por último aplicando el algoritmo descrito en el capítulo anterior, se obtienen las estimaciones de los parámetros requeridos.

3.2 Modelo Probit

El modelo Probit tiene como variable dependiente a la distribución Binomial igual que el logit $Y \sim Bin(k_i, \pi_i)$. De igual manera es necesario calcular las ecuaciones iterativas (3.2) y (3.3).

Por definición se tiene que $\mu_i = k_i * \pi_i(x)$, teniéndose que el valor k_i es conocido. Asimismo el predictor lineal esta dado por :

$$\eta_i = \sum_{j=1}^{p} \beta_j X_{ij}$$

Por último la función liga y su inversa quedan definidas por:

$$\text{Si } g(\mu_i) = \phi^{-1}(\frac{\mu_i}{k_i}) = \eta_i$$

$\Rightarrow$ la inversa es

$$\mu_i = k_i * \phi(\eta_i) = g^{-1}(\eta_i)$$

O de forma equivalente

$$\mu_i = k_i * \int_0^{\eta_i} \phi(u)du \qquad (3.6)$$

donde $\phi(u)$ es la función de densidad normal estándar dada por

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

Por lo que para calcular $\frac{\partial \eta_i}{\partial \mu_i}$ se obtendrá $\frac{\partial \mu_i}{\partial \eta_i}$ y posteriormente se calculará su recíproco, es decir, $\left(\frac{\partial \mu_i}{\partial \eta_i}\right)^{-1} = \frac{\partial \eta_i}{\partial \mu_i}$

Derivando la ecuación (3.6) y aplicando el Teorema Fundamental del Cálculo (TFC), se obtiene lo siguiente

$$\frac{\partial \mu_i}{\partial \eta_i} = k_i * \phi(\eta_i)$$

Calculando el recíproco queda como

$$\frac{\partial \eta_i}{\partial \mu_i} = \frac{1}{k_i * \phi(\eta_i)}$$

Dado lo anterior, la ecuación iterativa (3.2) queda como

$$z_i = \left[\hat{\eta}_i + (y_i - \hat{\mu}_i)\left(\frac{1}{k_i * \phi(\eta_i)}\right)\right]\Bigg|_{\beta_i^{(m-1)}}$$

Asimismo la ecuación (3.3) queda dada por

$$w_{ii} = \frac{1}{(Var(Y_i))}(k_i * \phi(\eta_i))^2$$

De igual manera, la varianza está dada por

$$\mu_i = k_i * \pi_i(x) \qquad \text{y} \qquad Var(Y_i) = k_i\pi_i(1-\pi_i) = \mu_i(1-\frac{\mu_i}{k_i})$$

Por lo que sustituyendo, se obtiene

$$w_{ii} = \frac{1}{\left(\mu_i(1-\frac{\mu_i}{k_i})\right)}(k_i * \phi(\eta_i))^2$$

Dadas las ecuaciones iterativas, el siguiente paso es incorporarlas a la ecuación (3.4) e iniciar el proceso iterativo determinado por:

$$\underline{\beta}^{(m)} = (X^T W X)^{-1} X^T W z\Big|_{\beta_k^{(m-1)}}$$

Dado que ambos modelos solo difieren en la liga, pero mantienen la misma distribución, la Devianza queda definida de igual forma que en el modelo con liga logit. Es decir lo único que cambia es la media, ya que ésta es definida por la liga probit como se mencionó anteriormente.

3.3 Modelo Poisson

Para el modelo lineal generalizado con distribución Poisson se utiliza su liga canónica descrita en el capitulo 2, la cual queda definida por

$$\begin{aligned} \eta_i &= g(\mu_i) \\ &= \ln(\mu_i) \end{aligned}$$

Al calcular la derivada $\frac{\partial \mu_i}{\partial \eta_i}$ se obtendrá al igual que en los modelos anteriores, por derivación implícita considerando la función anterior,

$$\frac{\partial \eta_i}{\partial \eta_i} = \frac{\frac{\partial \mu_i}{\partial \eta_i}}{\mu_i}$$

donde $\frac{\partial \eta_i}{\partial \eta_i} = 1$, por lo que despejando $\frac{\partial \mu_i}{\partial \eta_i}$ se obtiene

$$\frac{\partial \mu_i}{\partial \eta_i} = \mu_i$$

Asimismo la derivada $\frac{\partial \eta_i}{\partial \mu_i}$ queda como

$$\frac{\partial \eta_i}{\partial \mu_i} = \frac{1}{\mu_i}$$

Sustituyendo las derivadas en las ecuaciones iterativas, éstas quedan como

$$z_i = \sum_k x_{ik}\beta_k^{(m-1)} + (y_i - \hat{\mu}_i)\left(\frac{1}{\mu_i}\right) = \left[\hat{\eta}_i + (y_i - \hat{\mu}_i)\left(\frac{1}{\mu_i}\right)\right]_{\beta_k^{(m-1)}}$$

Así como $w_{ii} = \frac{1}{(Var(Y_i))}\left(\mu_i\right)^2$

La distribución Poisson tiene como media y varianza las siguientes:

$$Y_i \sim \text{Poisson}(\mu_i)$$
$$E(Y_i) = \mu_i$$
$$Var(Y_i) = \mu_i$$

Por lo que la ecuación iterativa w_{ii} queda de la siguiente forma

$$w_{ii} = \frac{1}{\left(\mu_i\right)}\left(\mu_i\right)^2$$

Tales ecuaciones se sustituyen en la ecuación iterativa general dada por (3.4).

De igual manera se calcula la devianza como medida de bondad de ajuste, ésta queda definida por $2\{l(y;y) - l(y;\mu\}$

Por lo que sustituyendo en la ecuación log-verosímil estimador

$$\begin{aligned} D &= 2\{l(y;y) - l(y;\mu\} \\ &= 2\left[\left(\sum_{i=1}^{n} y_i \ln(y_i) + \cancel{ln(y_i!)} - y_i\right) - \left(\sum_{i=1}^{n} y_i \ln(\mu_i) + \cancel{ln(y_i!)} - \mu_i\right)\right] \\ &= 2\left[\left(\sum_{i=1}^{n} y_i \ln(y_i) - y_i - y_i \ln(\mu_i) + \mu_i\right)\right] \\ &= 2\left[\left(\sum_{i=1}^{n} y_i \ln\left(\frac{y_i}{\mu_i}\right) - (y_i - \mu_i)\right)\right] \end{aligned}$$

CAPÍTULO 4

Aplicación de los Modelos Lineales Generalizados

Análisis de Planificación Familiar en México
Utilización de métodos anticonceptivos

4.1 Introducción

En este capítulo se realizará un análisis estadístico respecto al conocimiento y la utilización de los servicios de planificación familiar, asimismo se determinarán las principales variables que intervienen en la decisión de utilizar métodos anticonceptivos en las mujeres casadas o unidas en edad reproductiva fértil.
Posteriormente se empleará un modelo lineal generalizado para calcular la probabilidad de que una mujer en edad fértil y además con alguna relación conyugal utilice algún método anticonceptivo.

4.2 Instrumento de medición

Como instrumento para realizar el análisis se utilizará la Encuesta Nacional de la Dinámica Demográfica 1997 (ENADID 1997). Tal encuesta tiene sus antecedentes en La Encuesta Nacional de la Dinámica Demográfica 1992 que tuvo como objetivo captar información tal como fecundidad, la migración y la mortalidad y una amplia gama de factores condicionantes.

La segunda Encuesta Nacional de la Dinámica Demográfica de 1997 trata de obtener información que permita obtener tendencias con los temas captados en la ENADID

1992, agregando además temas como la salud materno-infantil, las preferencias reproductivas, la historia de uniones y una profundización en la anticoncepción.

La ENADID 1997 se definió como una encuesta en hogares que cubriría a todos sus integrantes y de ellos a las mujeres en edad fértil; su ejecución estuvo a cargo del Instituto Nacional de Estadística Geografía e Informatica (INEGI). El método de recolección de la información sería una entrevista directa al jefe del hogar o a su cónyuge (o en su ausencia, una persona de 15 años o más residente en la vivienda) y a las mujeres entre 15 y 54 años de edad. El periodo del levantamiento quedó comprendido entre septiembre y diciembre de 1997.

La Encuesta Nacional de la Dinámica Demográfica 1997 está conformada de la siguiente manera:

I Características de la vivienda. Tal base tiene preguntas como; la disponibilidad de agua entubada, servicio sanitario, drenaje, energía eléctrica y el material predominante en pisos, esto con el fin de conocer las características de la vivienda.

II Datos generales. Esta base contiene información de todos los miembros del hogar. Obteniéndose características de tipo:

- *Generales.* Como son sexo, edad, estado civil.
- *Económicas.* Esto con el objetivo de tener información sobre la población económicamente activa. Por ello se investigó la condición de actividad de la población en edad activa, la ocupación principal, la situación en el trabajo y los ingresos monetarios derivados del trabajo y de otras fuentes (pensión, jubilación, etcétera).

- *Educativas.* Con el objeto de conocer la situación educativa del país y su relación con la dinámica demográfica se investigó el nivel de alfabetismo de la población, su asistencia escolar y nivel de escolaridad.
- *Servicios de salud.* Su propósito fue contabilizar a la población que tiene derecho a seguridad social, es decir acceso a los servicios de salud y la población que no cuenta con el mismo.
- *Migración.* Con el propósito de dar cuenta de las corrientes migratorias ocurridas en el interior del país, entre entidades y municipios, así como del flujo migratorio a otros países, se indagó sobre el lugar de nacimiento de la población, su residencia en junio de 1992, el tiempo y lugar de residencia anterior y actual.

III Mortalidad. Se abordó la mortalidad general, la materna y de la niñez, así como en los menores de un año de edad, con el objetivo de presentar las diferentes tasas de mortalidad calculadas tomando la referencia de enero 1992 a la fecha de la entrevista. Se investigó la fecha de la defunción, su certificación y registro, el sexo y la edad de la persona fallecida. Además para las mujeres entre 15 y 54 años de edad al momento de morir se investigó si se encontraban embarazadas, los meses de embarazo y la causa de la muerte.

IV Salud Materno Infantil del último y penúltimo hijo. Se indagó información sobre el tiempo de revisión prenatal, el lugar y personal donde fue revisada la madre antes, durante y posterior al alumbramiento, y se preguntó de lo anterior si existieron complicaciones, el tipo de parto y la edad gestacional del producto. Del recién nacido, el peso al nacimiento, la condición de lactancia, número de revisiones en su primer año de vida, el lugar donde se efectuó, las características de la revisión e información sobre su cobertura de vacunación.

V Características de la mujer. En esta base se capta información relacionada a:

- *Fecundidad.* Con el propósito de calcular indicadores que permitan analizar el comportamiento de esta variable, se elaboró la historia de embarazos de las mujeres de 15 a 54 años de edad y su condición de habla indígena. Con el recuento de embarazos se estableció el número de hijos nacidos vivos, sobrevivientes, fallecidos, abortos, mortinatos e intervalos inter genésicos.
- *Preferencias reproductivas.* De la población femenina en edad fértil se investigó sobre sus ideales en cuanto al número de hijos, el espaciamiento de los mismos, sus preferencias en cuanto al sexo de sus hijos y su motivación en cuanto a regular la fecundidad, con el propósito de identificar grupos prioritarios para ser atendidos por los programas de planificación familiar.
- *Anticoncepción.*- El interés fue actualizar la información disponible sobre el conocimiento, acceso y uso de métodos anticonceptivos por parte de las mujeres de 15 a 54 años y/o en sus parejas. Se indagó por la historia anticonceptiva de los últimos 5 años (1992-1997) con el fin de asociarla con los niveles y tendencias de la fecundidad, y para ello se buscó información sobre la condición de uso actual de algún método anticonceptivo, el tipo, lugar de obtención y la razón de uso del método actual o la razón de no uso.
- *Estado conyugal y número de uniones.*- Con el primer tema el interés fue conocer la situación conyugal de la población de 12 años y más y con el segundo se buscó conocer los patrones de nupcialidad de las mujeres en edad fértil y la exposición del riesgo de concebir, variables directamente asociadas con el nivel y la estructura de la fecundidad.

El diseño de la encuesta permite generar información a nivel nacional, por entidad federativa y para los siguientes cuatro tamaños de localidad:

- Menos de 2,500 habitantes.
- De 2,500 a 14,999 habitantes.
- De 15,000 a 99,999 habitantes.
- 100,000 y más habitantes.

Con esta clasificación se puede construir la agrupación para población rural o urbana, definida como rural, si la localidad posee menos de 2,500 habitantes y urbano, si la localidad posee más de 2,500 habitantes[15].
Como puede observarse las bases están separadas, por lo que el primer paso para desarrollar el análisis es unir las bases de interés. Tales bases son Características de la Mujer, Características de la Vivienda y Datos Generales.

Las dimensiones de estas bases son:
Características de la Mujer: 88,022 registros
Características de la Vivienda: 73,412 registros
Datos Generales: 325,558 registros.

Cada una de estas bases tiene una llave única, la cual identifica a cada uno de los registros, lo anterior permite unir las bases descritas anteriormente, tal llave es proporcionada por la encuesta y viene especificada en la descripción de archivos.

Por lo que la base maestra tendrá información de cada una de las mujeres de 12 a 54 años, con sus respectivas características de vivienda y datos generales. Obteniéndose una base final con un total de 88,022 registros.

[15] Clasificación según el Índice de Marginación, Consejo Nacional de Población.

Por otra parte, observando las dimensiones de las bases anteriores se puede pensar que existen incongruencias ya que hay más mujeres que viviendas, pero la razón es que en una vivienda pueden vivir más de dos mujeres. Por su parte, la base de datos generales contiene la información de todos los miembros del hogar, por lo que cada uno de los registros es vinculado con la base de características de la mujer de manera única.

Para realizar el análisis, se propone exclusivamente tener a las mujeres casadas o unidas en edad fértil, es decir, mujeres casadas por el civil, iglesia, ambas ó en unión libre con un periodo de edad de 15 a 49 años.

La razón de solo considerar a las mujeres casadas o unidas es porque son las usuarias potenciales de servicios de planificación familiar, y por que si bien es cierto que las mujeres solteras ejercen su sexualidad, resulta más difícil establecer el análisis dado el tabú del tema.

El proceso realizado para obtener la base final para posteriormente iniciar el análisis es el siguiente; tomando la base de 88,022 registros se seleccionan sólo a las mujeres de 15 a 49 años, que se conoce como el periodo fértil de la mujer. Por lo que el conjunto se reduce a 83,216 registros, de estás ahora únicamente se consideran a las mujeres que se encontraban casadas o unidas al momento de la encuesta, reduciéndose el conjunto a 50,010 registros.

Una vez obtenidas las mujeres casadas o unidas en edad fértil, se procedió a depurar dicha base, eliminando inconsistencias de registros duplicados, los cuales son comunes en este tipo de encuestas. Asimismo se procedió a eliminar los registros vacíos que no contenían información alguna.

Una vez realizada la depuración, la base contiene un total de 49,628 mujeres casadas o unidas en edad fértil, que contienen información relevante y con representatividad a nivel nacional, estatal y por tamaño de localidad (específicamente zona rural y urbana).

A través de un análisis estadístico se pretende medir el conocimiento y utilización de servicios de planificación familiar en México, a nivel nacional, estatal y zona rural y urbana, utilizando la base obtenida de la ENADID 97, que será el instrumento de medición.

Los pasos a seguir son:

i. Analizar tablas de contingencia de ciertas variables con respecto a la condición de utilización de métodos anticonceptivos, esto con el fin de encontrar algunas tendencias y relaciones con la variable de interés.
ii. Definir un modelo lineal generalizado que permita realizar inferencias y poder comparar medidas a nivel estatal.
iii. Conclusiones de aplicar el modelo lineal generalizado al instrumento de medición.

4.3 Análisis de tablas de contingencia

El objetivo de analizar las tablas de contingencia es encontrar alguna relación estadística existente entre la variable dependiente o de respuesta que es "Usa" o "No Usa" algún método anticonceptivo con las variables explicativas que se suponen tienen cierto efecto en la variable de respuesta. Todo éste análisis se realiza en primera instancia a nivel nacional y posteriormente se analizarán en el modelo a nivel estatal.

4.3.1 Variables socio-económicas

4.3.1.1 Ingreso mensual familiar. Está variable fue generada sumando los ingresos mensuales de cada miembro del hogar en la base de datos generales y posteriormente se le asigno a cada mujer del hogar correspondiente, este ingreso mensual percibido contempla además ingresos por pensiones, alquileres, becas u otro tipo de ayuda.

Para tabular está variable, primero se generaron cuartiles para poder tener una medida concreta del ingreso mensual familiar.
El corte de cada uno de los cuartiles es el siguiente:

Variable	Obs	Percentil	Centile	[Intervalo al	95% Conf.]
Ing_mh	49628	0	-	-	0*
(Ingreso		25	1,000.00	1,000.00	1,000.00
Mensual		50	1,988.00	1,923.46	2,000.00
Familiar)		75	3,800.00	3,773.99	3,852.00
		100	306,200.00	306,200.00	306200*

* Indican el mínimo y máximo de la muestra

Tabla 4 - 1

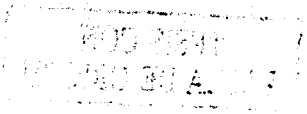
Lo que nos indica que alrededor de un 25% de la muestra percibe un ingreso mensual familiar de $1,000.00, el 50% percibe un ingreso mensual familiar de $1,988.00, y así de manera similar:

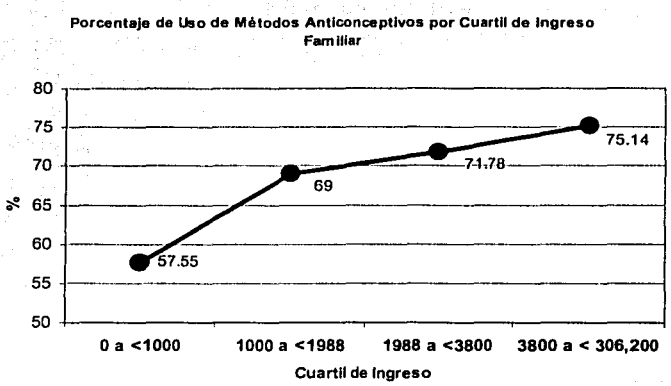
Lo importante ahora es determinar si existe alguna relación de esta variable de respuesta. La siguiente tabla presenta la relación existente entre estas variables:

Cuartiles de Ingreso familiar	Usa		Total
	No	Si	
Q1: con rango de 0 a <1000	1,571,552	2,130,301	3,701,853
% Renglón	42.45	57.55	100
% Columna	33.48	20.94	24.9
Q2: con rango de 1000 a <1988	1,099,273	2,446,338	3,545,611
% Renglón	31	69	100
% Columna	23.42	24.05	23.85
Q3: con rango de 1988 a <3800	1,083,912	2,757,467	3,841,379
% Renglón	28.22	71.78	100
% Columna	23.09	27.11	25.84
Q4: con rango de 3800 a < 306,200	938,868	2,838,081	3,776,949
% Renglón	24.86	75.14	100
% Columna	20	27.9	25.41
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 2

En donde se observa una tendencia positiva entre la variable "uso de algún método anticonceptivo" con el nivel de ingreso familiar. Para observar tal tendencia se puede observar la siguiente gráfica





Gráfica 4 - 1

En donde se observa una tendencia de mayor porcentaje en el uso de métodos anticonceptivos en mujeres casadas o unidas a medida que aumenta el cuartil de ingreso.

Por lo que la variable ingreso mensual familiar podría indicar que puede ser considerada para influir en la decisión de la mujer o de su pareja en utilizar algún método de anticoncepción y posteriormente utilizarla en el modelo.

4.3.1.2 Número de servicios que posee la vivienda. Está variable se refiere al número de servicios que tiene la vivienda, tales variables son dicotómicas (tiene o no el servicio) con las que se generará una sola, la cual suma el número de servicios que tiene la vivienda, tales variables se encuentran en la base de características de la vivienda y posteriormente se une con cada una de las mujeres que se encuentra en la vivienda. Las variables que se consideran son;

Tipo de piso:	0	Si tiene piso de tierra.
	1	Si tiene piso de cemento, madera, mosaico u otro recubrimiento.
Disposición de Servicio Sanitario:	0	Si es fosa (séptica), hoyo negro o no disponen.
	1	Si es excusado, retrete o letrina.
Disposición de Conexión de Agua:	0	Si no tiene servicio de conexión de agua
	1	Si tiene servicio de conexión de agua
Disposición de Electricidad	0	Si no tiene servicio de electricidad
	1	Si tiene servicio de electricidad

Tabla 4 - 3

Con esta variable de tipo económico también se puede inferir acerca de la decisión de usar o no métodos anticonceptivos. La siguiente tabla presenta los resultados:

Número de servicios que tiene la vivienda	Usa		Total
	No	Si	
0 Servicios	405,832	387,453	793,285
% Renglón	51.16	**48.84**	100
% Columna	8.65	3.81	5.34
1 Servicio	645,416	691,931	1,337,347
% Renglón	48.26	**51.74**	100
% Columna	13.75	6.8	9
2 Servicios	286,405	404,641	691,046
% Renglón	41.45	**58.55**	100
% Columna	6.1	3.98	4.65
3 Servicios	1,529,625	2,933,334	4,462,959
% Renglón	34.27	**65.73**	100
% Columna	32.59	28.84	30.02
4 Servicios	1,826,327	5,754,828	7,581,155
% Renglón	24.09	**75.91**	100
% Columna	38.91	56.57	51
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	**68.43**	100
% Columna	100	100	100

Tabla 4 - 4

En la tabla anterior se muestra que a medida que aumenta el número de servicios disminuye la tasa de no uso de métodos anticonceptivos o que es lo mismo aumenta la tasa de uso. Es decir, las mujeres que en su vivienda tienen 0 servicios poco más de *51 %* no utiliza ningún método, en contraste con las mujeres que en su vivienda tienen 4 servicios la tasa de no uso es alrededor de sólo *24 %*.

Lo que permite suponer una relación positiva entre número de servicios y el uso. En el caso anterior se tiene una diferencia de uso de métodos de alrededor de 27% (75.91 - 48.84) entre la categoría de 0 servicios y 4 servicios, la cual es considerable en términos de utilización de servicios de planificación familiar.

Es de observarse también que las dos tablas anteriores, en la categoría más alta de las variables explicativas se encuentra una tasa cerca de 75 % de uso de anticoncepción, con lo que se están validando ambas variables en términos de confiabilidad en ámbito socio-económico con la variable de respuesta.

Con lo expresado anteriormente se puede decir de manera preliminar que las variables socio-económicas de la familia influyen en las decisiones de utilización de métodos anticonceptivos.

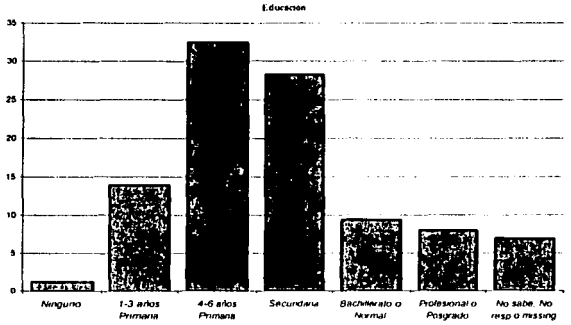
4.3.2 Variables en Relación a Características de la Mujer

4.3.2.1 Educación de la mujer: Esta variable es construida a través del nivel máximo de estudios al que haya asistido la mujer, excepto en el caso de nivel primaria, para la cual se consideran 2 categorías. En esta variable no se considera si una mujer ha terminado o no el nivel de estudios reportado. Por ejemplo, si una mujer estudió 2 años de secundaria, en el estudio únicamente se reporta que estudió nivel secundaria, de igual forma para los demás niveles.

Por lo que la variable cuenta con las siguientes categorías y su porcentaje correspondiente, quedando como sigue:

Educación de la mujer	Frecuencia	Porcentaje	Distribución Acumulada
Ninguno	178,113	1.2	1.2
1-3 años Primaria	2,066,708	13.9	15.1
4-6 años Primaria	4,822,033	32.44	47.54
Secundaria	4,200,794	28.26	75.8
Bachillerato o Normal	1,393,350	9.37	85.17
Profesional o Postgrado	1,183,151	7.96	93.13
No sabe, No resp o missing	1,021,643	6.87	100
Total	14,865,792	100	

Tabla 4 - 5 Los datos anteriores se pueden ver en la siguiente gráfica.

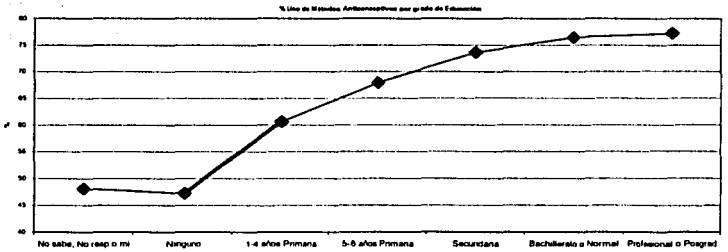


En dicha gráfica se muestra una distribución normal de los datos de educación, la cual está un poco sesgada a la izquierda. Ahora tabulando la educación de la mujer con respecto al uso de métodos anticonceptivos, se tiene lo siguiente:

Educación de la mujer	Usa		Total
	No	Si	
Ninguno	93,964	84,149	178,113
% Renglón	52.76	47.24	100
% Columna	2.00	0.83	1.20
1-3 años Primaria	814,344	1,252,364	2,066,708
% Renglón	39.40	60.60	100
% Columna	17.35	12.31	13.90
4-6 años Primaria	1,549,425	3,272,608	4,822,033
% Renglón	32.13	67.87	100
% Columna	33.01	32.17	32.44
Secundaria	1,108,667	3,092,127	4,200,794
% Renglón	26.39	73.61	100
% Columna	23.62	30.40	28.26
Bachillerato o Normal	327,568	1,065,782	1,393,350
% Renglón	23.51	76.49	100
% Columna	6.98	10.48	9.37
Profesional o Postgrado	269,457	913,694	1,183,151
% Renglón	22.77	77.23	100
% Columna	5.74	8.98	7.96
No sabe, No responde o Missing	530,180	491,463	1,021,643
% Renglón	51.89	48.11	100
% Columna	11.30	4.83	6.87
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 6

Asimismo se puede observar su cambio en el siguiente gráfico.



La variable de educación se considera una de las variables con la que se pueda realizar inferencias fuertes, ya que en forma teórica una mujer con un cierto nivel de educación tiene mayor probabilidad de utilizar algún método anticonceptivo en relación con alguna que no tiene educación. Este argumento es dado porque parte de la información acerca del uso de métodos anticonceptivos se difunde a través de programas otorgados en instituciones escolares y por medio de programas de salud a través de pláticas en instituciones de salud.

En la tabla se observa que a medida de un mayor nivel de educación se presenta una tasa mayor de uso de anticonceptivos, lo cual es de esperarse si se acepta el argumento que se dio anteriormente.

Es de notarse el cambio más considerable en las dos primeras categorías con relación al uso de métodos anticonceptivos, en donde van de un nivel nulo de educación a al menos 4 años de educación primaria, con un aumento de proporción de poco más de 13 %, dicho porcentaje es el mayor que se da entre categorías continuas, lo anterior también se muestra en la gráfica anterior.

4.3.2.2 Edad de la mujer: Con esta variable se puede realizar una inferencia acerca de la utilización de métodos anticonceptivos por parte de la mujer o su pareja con respecto a la edad que tiene la mujer.

Para tabular tal variable se va a generar una variable construida por quinquenios comenzando en 15-19, 20-24, y así sucesivamente hasta 45-49.

Se ha de recordar que el estudio sólo se está realizando para mujeres casadas o unidas, por lo que se espera que en el primer grupo haya menor proporción de mujeres. La tabulación queda de la siguiente forma:

Grupo de Edad	Usa		Total
	No	Si	
15-19	410,855	334,991	745,846
% Renglón	55.09	44.91	100
% Columna	8.75	3.29	5.02
20-24	910,661	1,313,756	2,224,417
% Renglón	40.94	59.06	100
% Columna	19.4	12.92	14.96
25-29	927,761	1,957,001	2,884,762
% Renglón	32.16	67.84	100
% Columna	19.77	19.24	19.41
30-34	697,834	2,126,516	2,824,350
% Renglón	24.71	75.29	100
% Columna	14.87	20.91	19
35-39	614,504	1,965,363	2,579,867
% Renglón	23.82	76.18	100
% Columna	13.09	19.32	17.35
40-44	523,577	1,503,839	2,027,416
% Renglón	25.82	74.18	100
% Columna	11.16	14.78	13.64
45-49	608,413	970,721	1,579,134
% Renglón	38.53	61.47	100
% Columna	12.96	9.54	10.62
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 7

La tabla anterior permite observar el cambio en la decisión de una mujer de tener hijos y por ende como utilizar métodos anticonceptivos para evitar embarazos no planeados en la pareja. Tal tendencia en el uso de anticonceptivos va en aumento

conforme aumenta la edad hasta llegar al grupo de edad de 40-44, y posteriormente en el siguiente grupo presenta un descenso, el cual es bastante entendible, ya que a esta edad algunas mujeres ya han dejado de ser fértiles, es decir, han entrado a la etapa de menopausia.

La variable de edad en primera instancia parece ser una variable importante para poder predecir que una mujer utilice o no algún método anticonceptivo.

4.3.2.3 Tipo de relación conyugal: Dado que sólo se están considerando a las mujeres casadas o unidas en el análisis, la intención es revisar si el tipo de relación que posee la mujer influye en la utilización de servicios de planificación familiar. Tal variable contiene las siguientes categorías; Unión Libre, Casada por el Civil, Casada por la Iglesia y Casada por Civil e Iglesia. La siguiente tabulación presenta la relación existente de esta variable con la utilización.

Grupo de Edad	Usa		Total
	No	SI	
Union Libre	1,095,239	1,801,799	2,897,038
% Renglón	37.81	62.19	100.00
% Columna	23.33	17.71	19.49
Civil	1,047,454	2,573,769	3,621,223
% Renglón	28.93	71.07	100.00
% Columna	22.32	25.30	24.36
Iglesia	251,676	270,294	521,970
% Renglón	48.22	51.78	100.00
% Columna	5.36	2.66	3.51
Civil e Iglesia	2,299,236	5,526,325	7,825,561
% Renglón	29.38	70.62	100.00
% Columna	48.99	54.33	52.64
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100.00
% Columna	100.00	100.00	100.00

Tabla 4 - 8

En este caso no se espera una tendencia positiva o negativa con respecto al uso, ya que en este caso se tienen valores de tipo categórico, es decir, es indistinto tener cualquier tipo de relación conyugal. Lo que se busca es encontrar diferencias en el uso.

En donde se puede ver que el mínimo uso existe en las mujeres que sólo están casadas por la iglesia, por lo que se puede asumir que el bajo uso esta marcado por cuestiones religiosas. Por otra parte el máximo es alcanzado por aquellas mujeres que están unidas por el civil, ya sea en forma única o conjunta con la iglesia.

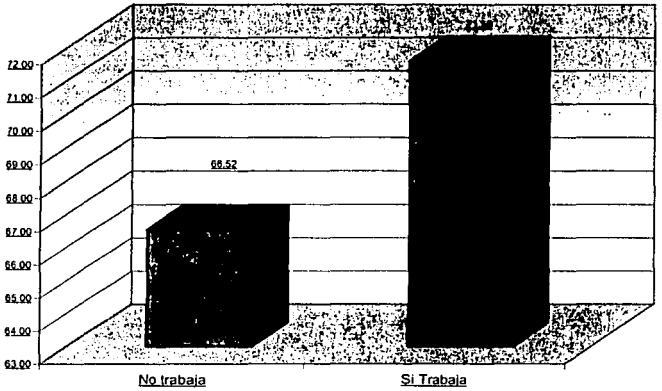
4.3.2.4 Condición de trabajo: Se refiere únicamente a si la mujer trabaja o no y la posible relación que pueda existir con respecto a su decisión de utilizar servicios de planificación familiar. Algunos estudios sociales sospechan que el hecho que una mujer que trabaje afecta de manera positiva en el uso de anticonceptivos, ya que retrasa el aumento en el número de hijos, debido al posible crecimiento laboral que la mujer pueda desarrollar.

Condición de Trabajo	Usa		Total
	No	Si	
No trabaja	3,100,521	6,160,714	9,261,235
% Renglón	33.48	66.52	100
% Columna	66.06	60.56	62.30
Si Trabaja	1,593,084	4,011,473	5,604,557
% Renglón	28.42	71.58	100
% Columna	33.94	39.44	37.70
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 9

La gráfica siguiente podría mostrar un efecto visual de el porcentaje de uso dada su condición de trabajo.

Porcentaje de Uso de Métodos Anticonceptivos por Condición de Trabajo



En donde no se observa una relación marcada como en las tablas anteriores, ya que el incremento es de 5 puntos porcentuales en la utilización de servicios de planificación con respecto a la condición de trabajo de la mujer. En ambas condiciones la utilización de anticonceptivos es alta.

4.3.2.5 Número de Hijos: Esta variable se considera importante, ya que el número de hijos que tiene una pareja puede influir de manera significativa y de manera positiva en la utilización de anticonceptivos. Ya que a medida que aumenta el tamaño de la familia también se deben de medir cuestiones de ingreso y estabilidad económica de la familia, ya que no es lo mismo proveer de alimento, servicios médicos y escolares, entre otros gastos a una familia de 3 miembros que a una familia de 6 miembros. Es por esta situación que se espera un incremento de utilización con respecto al número de hijos.

Número de Hijos	Usa		Total
	No	Si	
0	799,086	246,128	1,045,214
% Renglón	76.45	23.55	100
% Columna	17.02	2.42	7.03
1	1,031,539	1,534,200	2,565,739
% Renglón	40.20	59.80	100
% Columna	21.98	15.08	17.26
2	862,368	2,639,748	3,502,116
% Renglón	24.62	75.38	100
% Columna	18.37	25.95	23.56
3	574,465	2,370,658	2,945,123
% Renglón	19.51	80.49	100
% Columna	12.24	23.31	19.81
4	380,416	1,442,826	1,823,242
% Renglón	20.86	79.14	100
% Columna	8.10	14.18	12.26
más de 5	1,045,731	1,938,627	2,984,358
% Renglón	35.04	64.96	100
% Columna	22.28	19.06	20.08
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 10

Como se esperaba, se presentó un aumento en el uso de anticonceptivos según el número de hijos que tiene la mujer. En donde se observa el máximo de utilización de anticonceptivos se presenta en la categoría de 3 y 4 hijos, y descendiendo de forma extraña en la categoría de más de 5 hijos, aunque ésta podría ser explicada si las mujeres que reportan este dato son mayores, esto corroboraría de igual manera la tendencia aplicada en la tabulación de grupo de edad contra uso, en donde también se presenta un descenso en la última categoría.

Es decir que las mujeres mayores que no utilizaron anticonceptivos por consecuencia tienen un mayor número de hijos, pero estas mujeres ya tienen una edad mayor.

4.3.2.6 Religión: Las cuestiones religiosas se podrían considerar en la decisión de utilizar anticonceptivos, por lo que se analiza la relación del tipo de religión con el uso, presentando lo siguiente:

Religión	Usa		
	No	Si	Total
Católica	4,150,979	9,077,284	13,228,263
% Renglón	31.38	68.62	100
% Columna	88.44	89.24	88.98
Otra	542,626	1,094,903	1,637,529
% Renglón	33.14	66.86	100
% Columna	11.56	10.76	11.02
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 11

En las categorías Católica u Otra no hay diferencia significativa en el uso de anticonceptivos, ya que están entre 68.62% y 66.86%. De forma anticipada se puede decir que esta variable no presenta alguna determinación en la utilización de anticonceptivos. Además observando el segundo porcentaje, se aprecia que cerca del 90% posee la religión católica, con lo que no se tiene variabilidad.

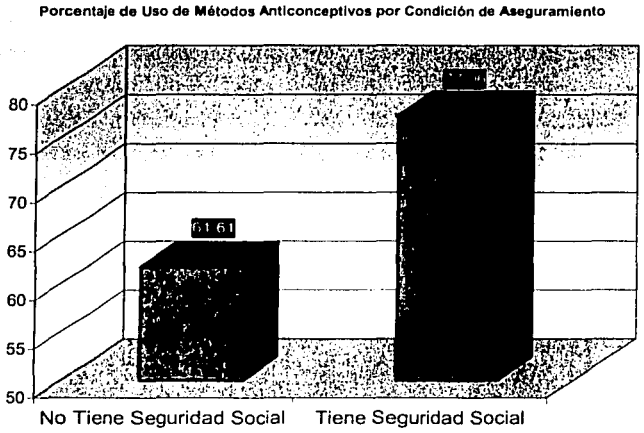
4.3.2.7 Condición de Seguridad Social: En este caso se refiere a si la mujer posee algún tipo de servicio médico como por ejemplo IMSS, ISSSTE, PEMEX, etc. entre otras. La variable es dicotómica en donde toma el valor de 1 si posee algún servicio de seguridad social o privado y 0 si no posee ningún plan de aseguramiento.
Con esta variable se pretende medir la accesibilidad a los métodos anticonceptivos, ya que se puede encontrar alguna correlación del uso con la forma de conseguir. Es decir, se pretende medir la influencia que tiene el hecho que una mujer cuente con servicio médico con una mayor probabilidad de tener acceso y uso a algún método anticonceptivo.
La tabla presenta el siguiente resultado:

Condición de Seguridad Social	Usa		Total
	No	Si	
No Tiene Seguridad Social	3,179,692	5,102,009	8,281,701
% Renglón	38.39	61.61	100
% Columna	67.75	50.16	55.71
Tiene Seguridad Social	1,506,610	5,060,951	6,567,561
% Renglón	22.94	77.06	100
% Columna	32.1	49.75	44.18
No responde	7,303	9,227	16,530
% Renglón	44.18	55.82	100
% Columna	0.16	0.09	0.11
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 12

En la tabla anterior se observa que existe una pequeña tasa de no respuesta de 0.11%, lo cual no es significativo y no afecta la distribución de la tabulación. Por otra parte existe un incremento de un 16% de uso de anticonceptivos para aquellas mujeres que tienen algún sistema de servicio médico, lo anterior era de esperarse ya que se supone

que tienen una mayor información y accesibilidad a los servicios de planificación familiar. Los resultados anteriormente mencionados se observan de manera más clara en la siguiente gráfica.



Esta gráfica permite tener en cuenta el hecho que una mujer tenga algún plan de aseguramiento, lo cual se ve reflejado en el uso de algún método anticonceptivo.
Con lo anterior, se determina que la variable de aseguramiento debe ser incluida en el modelo, dado que al parecer tiene un efecto significativo.

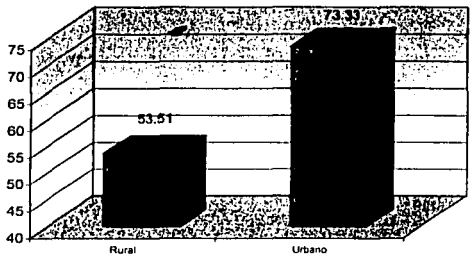
4.3.3 Variables en Relación al Entorno de Residencia

4.3.3.1 Condición de Residencia: Es de saberse que en áreas rurales existe un menor número de servicios tales como centros médicos, farmacias, escuelas, entre otro tipo de servicios, los cuales se esperan que influyan de forma indirecta en la decisión de una mujer a utilizar servicios de planificación familiar. En este caso la condición de residencia es una variable que afecta en forma externa, ya que puede ser considerada como una barrera natural que determina el hecho de que una mujer utilice servicio de planificación. La variable presenta la siguiente tabulación.

Condición de Residencia	Usa		Total
	No	Si	
Rural	1,709,748	1,967,808	3,677,556
% Renglón	46.49	53.51	100
% Columna	36.43	19.34	24.74
Urbano	2,983,857	8,204,379	11,188,236
% Renglón	26.67	73.33	100
% Columna	63.57	80.66	75.26
Total	4,693,605	10,172,187	14,865,792
% Renglón	31.57	68.43	100
% Columna	100	100	100

Tabla 4 - 13

Porcentaje de Uso de Métodos Anticonceptivos por Lugar de Residencia



En la gráfica anterior se corrobora lo mencionado, es decir, existe un menor uso de servicios de planificación familiar en las zonas rurales. Se observa un incremento de cerca del 20 % en el uso de anticonceptivos en las zonas urbanas con respecto a las zonas rurales.

En términos comparativos, ésta diferencia es muy significativa, ya que dice que el lugar de residencia es determinante en el uso de algún método. Con lo que se puede afirmar la importancia de la variable en un análisis posterior.

4.4 Especificación del modelo lineal generalizado.

En esta sección se definirá el modelo lineal generalizado que se utilizará para poder predecir la probabilidad de que una mujer que posee ciertas características haga uso de los servicios de planificación familiar. Una vez definido el modelo lineal generalizado se realizará la estimación a nivel estatal, esto dado que se tiene representatividad a este nivel. Y por ultimo se realizarán una serie de comparaciones entre varios estados.

Por lo que el objetivo es analizar un modelo lineal generalizado que presente la probabilidad de que una mujer use o no algún método anticonceptivo.

Con lo anterior se estaría analizando en forma completa el tema de uso de servicios de planificación familiar en México.

4.4.1 Modelo lineal generalizado para calcular la probabilidad de que una mujer utilice servicios de planificación familiar.

En esta sección se utilizará un modelo lineal generalizado con distribución Binomial y liga Probit, esto con el objetivo de calcular la probabilidad de que una mujer utilice algún servicio de planificación familiar dado un conjunto de variables explicativas.
Como primer método se propone utilizar a todas las variables explicativas que se mencionaron en el análisis de tablas de contingencia expresadas en la sección anterior esto con el fin de identificar las variables que no tienen significancia estadística.
El modelo que se propone es el siguiente:

$$\eta=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta_7X_7+\beta_8X_8+\beta_9X_9+\beta_{10}X_{10} \quad (4.1)$$

Donde
X_1 = Ingreso mensual familiar

X_2 = Número de servicios

X_3 = Educación de la mujer

X_4 = Edad de la mujer

X_5 = Tipo de relación conyugal

X_6 = Condición de trabajo.

X_7 = Número de hijos

X_8 = Religión de la mujer

X_9 = Lugar de Residencia

X_{10} = Condición de seguridad social

La liga está definida como en el capítulo dos, dada por

$g(\pi(x)) = \phi^{-1}(\bullet)$

y su función inversa por $g^{-1}(\pi(x)) = \eta$ recordando que $\mu = n * \pi(x)$

Con lo que se tienen los tres componentes de un modelo lineal generalizado.

El análisis en todas sus etapas, se desarrolla a través del paquete estadístico STATA 7.0, el cual tiene implementado los modelos lineales generalizados para las siguientes distribuciones:

- Normal
- Normal Inversa
- Bernoulli
- Binomial
- Poisson
- Binomial Negativa
- Gamma.

Y las ligas siguientes;

- Identidad

- Log
- Logit
- Probit
- log-log complementaria
- Binomial Negativa
- Log-log
- Log-complement

En el primer modelo realizado se consideraron observaciones individuales, es decir se tienen a las 49,628 mujeres en edad fértil que se encontraban casadas o unidas al momento de la encuesta.

Al ejecutar el modelo lineal generalizado se encontraron un total de 82 valores extraños o missings, por lo que estos se eliminaron de la muestra, quedando así un total de 49,546, que en términos expandidos corresponde a un total de 14,849,262 mujeres.

Por lo que la distribución utilizada es Binomial con parámetros $Bin(1, \pi(x))$.

Para este modelo no es posible realizar una medida de bondad de ajuste, ya que para datos individuales como se mencionó en el capítulo dos, la distribución no converge a alguna distribución conocida, por lo que únicamente se puede probar la significancia de cada uno de los parámetros.

El primer análisis se realizó a nivel nacional, por lo que no se utilizó el factor de expansión, esto además con el fin de identificar las variables que no tienen peso estadístico, es decir, al momento de hacer su prueba de hipótesis individual resulta no ser significativo.

El modelo fue introducido en STATA 7.0 con el siguiente comando

"glm usa ing_mh num_serv educ1 edad edo_cvl trab n_hijos rlg aseg ru, family(binomial) link(probit) irls"

Este modelo es el descrito por la ecuación (4.1), a dicho modelo se le indica el comando **glm,** posteriormente la variable de respuesta y después las variables explicativas, asimismo la distribución que sigue la variable independiente, la función liga y el tipo de algoritmo que se desea para realizar la estimación.

Una vez ejecutado el comando anterior, éste presenta los siguientes resultados;

Sección 1

```
Iteration 1   :   deviance  59479.7361
Iteration 2   :   deviance  59432.8191
Iteration 3   :   deviance  59432.8008
Iteration 4   :   deviance  59432.8008
Iteration 5   :   deviance  59432.8008
```

Sección 2

```
Generalized linear Models                    No. of obs          =      49546
Optimization     :  MQL Fisher scoring       Residual df         =      49535
                    (IRLS EIM)               Scale param         =          1
Deviance         =  59432.80079              (1/df) Deviance     =   1.199814
Pearson          =  49644.59614              (1/df) Pearson      =   1.002212

Variance function:              V(u) = u*(1-u)          [Bernoulli]
Link function    :              g(u) = invnorm(u)       [Probit]
Standard errors  :  EIM
```

Sección 3

usa	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
ing_mh	9.99E-07	1.37E-06	0.73	0.465	0.0000	0.0000
num_serv	0.1063734	0.0058392	18.22	0.000	0.0949	0.1178
educ1	-0.030537	0.0026548	-11.50	0.000	-0.0357	-0.0253
edad	0.0060775	0.0008603	7.06	0.000	0.0044	0.0078
edo_cvl	0.0117499	0.0027204	4.32	0.000	0.0064	0.0171
trab	0.0751513	0.0126938	5.92	0.000	0.0503	0.1000
n_hijos	0.0068953	0.0009465	7.28	0.000	0.0050	0.0088
rlg	0.0194662	0.0078042	2.49	0.013	0.0042	0.0348
aseg	0.2586334	0.0133344	19.40	0.000	0.2325	0.2848
ru	0.2603803	0.0144752	17.99	0.000	0.2320	0.2888
_cons	-0.401734	0.031336	-12.82	0.000	-0.4632	-0.3403

En donde en la sección 1, se presentan las iteraciones realizadas para poder terminar el proceso (tal algoritmo fue presentado en el capítulo dos), en la sección 2 se obtienen una descripción del tipo de método iterativo que desarrollo el paquete

estadístico, en este caso IRLS, además presenta el número de observaciones, el número de grados de libertad, la devianza del modelo con p-parámetros, la función liga, el tipo de distribución y la función que tiene la varianza.

En la sección 3, se describen las estimaciones de los parámetros, su desviación estándar y posteriormente, se calcula la estadística de Wald para cada estimador, en la cual se calculó la probabilidad para rechazar la hipótesis nula a un nivel de significancia 0.05, tal hipótesis está dada por $\beta_i = 0 \quad i = 1,...,n$. Y por último se presenta un intervalo de confianza para cada una de las estimaciones de los parámetros.

En la sección 3, se observa a un nivel de significancia 0.05, que la única variable que no pesa en el modelo es el ingreso familiar. Por lo que tal variable se eliminará, ya que además se supone que la encuesta no está diseñada para captar cuestiones de ingreso con una precisión confiable, ya que para captar ingreso se encuentra la Encuesta Nacional de Ingreso Gasto de los Hogares (ENIGH), la cual capta en forma más precisa está situación. Asimismo el número de servicios que se considera que tiene relación directa con el ingreso.

Por otra parte en el análisis de las tablas de contingencia, se encontró que la variable de religión no tenía un efecto importante en la decisión de una mujer para utilizar servicios de planificación, por lo que de igual manera será removida del modelo. Esto es debido a que cerca del 89% de las mujeres tienen religión católica.

Con el fin de utilizar una estadística de bondad de ajuste, se procedió a realizar grupos con base a las variables restantes del modelo y agregando una variable que se considera importante, tal es el *número de métodos anticonceptivos* que conoce una mujer en edad fértil casada o unida. Tal variable fue categorizada de acuerdo a la

siguiente manera; 0 métodos, 1-3 métodos, 4-7 métodos y 8-10 métodos. El supuesto para poder ejecutar esta situación es, que dentro de cada grupo de tamaño k_i, cada mujer tiene la misma probabilidad de utilizar algún método anticonceptivo.

Teniendo así una variable aleatoria definida por una distribución

$$Y_i \sim Bin(k_i, \pi_i(\bar{x}))$$

la cual toma valores en el siguiente rango $0 < Y_i \leq k_i$ y $0 \leq \pi_i(\bar{x}) \leq 1$

En donde Y_i, es el número de mujeres que utilizan algún método anticonceptivo dentro de cada grupo. Por su parte $\pi_i(\bar{x})$ es la probabilidad de que una mujer utilice algún método anticonceptivo dentro del grupo.

Obteniéndose entonces las frecuencias de uso de algún método anticonceptivo por las categorías de las variables mencionadas anteriormente, éstas son; **educl num_m4 gedad edo_cvl trab n_hijos aseg ru.**

Una vez calculadas las frecuencias por grupo, éstas quedaron de acuerdo a la siguiente tabla:

educl	Num_met	gedad	edo_cvl	trab	n_hijos	aseg	ru	Usa
Ninguno	0	15-19	Unión Libre	Si	0	Si	Urbano	Y1
Ninguno	0	15-19	Unión Libre	Si	0	Si	Rural	Y2
Ninguno	0	15-19	Unión Libre	Si	0	No	Urbano	Y3
Ninguno	0	15-19	Unión Libre	Si	0	No	Rural	Y4
Ninguno	0	15-19	Unión Libre	Si	1	Si	Urbano	Y5
Ninguno	8-10	45-49	Iglesia	Si	5	Si	Rural	YK
Profesional	0	15-19	Unión Libre	Si	0	Si	Urbano	Yj
Profesional	8-10	45-49	Casada	no	5 y más	No	Rural	Ym

Tabla 4 - 14

En donde dicha tabla contiene un total de 8,434 grupos dada la selección de variables realizada.

Respecto al factor de expansión se tiene que éste viene dado en forma individual por lo que en cada grupo se sumaron los diferentes factores de expansión para así obtener la representatividad de 14,849,262 mujeres en edad fértil y con alguna relación conyugal.

Como ya se comentó, la razón principal por la que se agrupó es para realizar una prueba de bondad de ajuste entre el modelo con p-parámetros y el modelo saturado, tal prueba como se mencionó en el capítulo dos, se realiza con la estadística de la Devianza y es comparando con el valor de una distribución Ji-Cuadrada.

Después de realizar diferentes pruebas, los elementos del modelo lineal generalizado final para el análisis nacional quedan de la siguiente manera:

El predictor lineal definido por

$$\eta=\beta_0+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta_7X_7+\beta_9X_9+\beta_{10}X_{10}+\beta_{11}X_{11} \qquad (4.2)$$

X_3 = Educación de la mujer

X_4 = Edad de la mujer

X_5 = Tipo de relación conyugal

X_6 = Condición de trabajo.

X_7 = Número de hijos

X_8 = Religión de la mujer

X_{10} = Condición de seguridad social

X_{11} = Número de métodos

La distribución de la variable de respuesta está definida por $Bin(k_i, \pi_i(\bar{x}))$, en donde se sabe que $\mu_i = k_i * \pi_i(x)$, teniéndose que el valor k_i es conocido. Y por último la función liga queda definida por:

$$\eta = g(\mu_i) = \phi^{-1}\left(\frac{\mu_i}{k_i}\right)$$

Asimismo la función inversa está definida por

$$\text{Si } g(\mu_i) = \phi^{-1}\left(\frac{\mu_i}{k_i}\right) = \eta_i$$
$$\Rightarrow \quad \text{la inversa es}$$
$$\mu_i = k_i * \phi(\eta_i) = g^{-1}(\eta_i)$$

Una vez definidos los elementos que se necesitan para un modelo lineal generalizado, se procedió a calcular las estimaciones de los parámetros, utilizando STATA 7.0, con la siguiente instrucción:

glm sum_usa educ1 num_m4 gedad edo_cvl trab n_hijos aseg ru [w=factor], family(binomial ki) link(probit) irls

El cual presenta los siguientes resultados.

```
(sum of wgt is  2.5059e+06)

Iteration 1   : deviance = -6710.3786
Iteration 2   : deviance = -6705.5600
Iteration 3   : deviance = -6706.2019
Iteration 4   : deviance = -6706.1610
Iteration 5   : deviance = -6706.1603
Iteration 6   : deviance = -6706.1602
Iteration 7   : deviance = -6706.1602
Iteration 8   : deviance = -6706.1602

Generalized
linear              models                  No. of obs      =          8434
Optimization     :  MQL Fisher scoring      Residual df     =          8425
                    (IRLS EIM)              Scale param     =             1
```

Deviance	=	6706.160218	(1/df) Deviance	=	0.7959834
Pearson	=	13894.46428	(1/df) Pearson	=	1.649195

Variance function: V(u) = u*(1-u/ni_n) [Binomial]
Link function : g(u) = invnorm(u/ni_n) [Probit]
Standard errors : EIM

BIC = 6624.79998

sum_usa_n	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
educ1	0.0223495	0.002648	-8.44	0	-0.0275395	-0.0171595
num_m4	0.4645009	0.0098386	47.21	0	0.4452176	0.4837843
Gedad	0.0217508	0.0042091	5.17	0	0.0135012	0.0300004
edo_cvl	0.0071906	0.0027348	2.63	0.009	0.0018304	0.0125507
Trab	0.0489808	0.0127051	3.86	0	0.0240792	0.0738823
n_hijos	0.0109519	0.0009561	11.45	0	0.0090779	0.0128259
Aseg	0.2035138	0.0133534	15.24	0	0.1773417	0.2296859
Ru	0.2235257	0.0150592	14.84	0	0.1940103	0.2530411
_cons	-1.089352	0.0300708	-36.23	0	-1.14829	-1.030415

En las estimaciones anteriores se incluyó el factor de expansión promedio de cada grupo de mujeres, el cual da la representatividad a nivel nacional y estatal. Esto es debido a que se considera una distribución binomial, en donde cada mujer de un mismo grupo tiene la misma probabilidad, considerando únicamente una mujer de cada grupo entonces se toma el promedio de los factores de expansión, esto con el fin de obtener representatividad nacional y estatal. Además se dispone del factor de expansión de las 14,849,262 de mujeres, esto con el fin de calcular probabilidades promedio por diferentes características.

El primer realizado fue ejecutar el modelo sin factor de expansión, esto para tratar de asegurar significancia estadística con estas variables.

Así suponiendo que una vez que dé validez en términos muestrales también daría el mismo ajuste a nivel poblacional. Los datos de la corrida sin factor de expansión se adjuntan en el apéndice C.

En ambos casos se obtienen resultados similares, en términos de las pruebas de bondad de ajuste y pruebas de hipótesis de cada uno de los parámetros. La tabla de resultados anterior, presenta el número de grupos que se utilizaron con el factor de expansión promedio de cada grupo de mujeres con características similares, obteniéndose un total de 8434 grupos. Un total de 8425 grados de libertad, adicionalmente se indica el tipo de distribución y liga que utilizó.

Lo anterior se hizo dado que se tiene una distribución binomial, la cual determina que cada mujer de un cierto grupo tiene la misma probabilidad.

La estadística de la Devianza arroja un valor 6706.16, que es comparado con el valor de una distribución Ji-Cuadrada con 8425 grados de libertad al 0.95 de probabilidad, dicho valor es 8639.6452, obviamente el valor de la Devianza es mucho menor que este valor, no rechazándose así la hipótesis que el modelo con p – parámetros ajusta bien con respecto al modelo saturado.

De igual manera se prueba que cada una de las estimaciones no son cero, ya que su *p-value* es menor en todos los casos que a nivel de significancia de 0.05.
El siguiente paso consiste en calcular las probabilidades de que una mujer con ciertas características haga uso de algún método anticonceptivo de planificación familiar.
El resumen de la probabilidad a nivel nacional de usar algún método anticonceptivo para mujeres en edad fértil que se encontraban casadas o unidas al momento de la encuesta fue de:

Variable	Observaciones Total de Grupos Muestrales	Mujeres con factor de expansión	Media	Desv. Estándar	Mínimo	Máximo
pr_usa	8,434	14,849,262	0.680369	0.1454782	0.0995875	0.887388

En donde se puede observar que la media de probabilidad, de que una mujer en edad fértil y que se encuentra casada o unida para todas las características de uso de métodos anticonceptivos está alrededor de 0.68.

En lo que respecta a la elección del modelo, se realizaron diferentes modelos con un número menor de variables, esto con el fin de encontrar un modelo que explicara la relación de uso de algún método anticonceptivo con un número menor de variables explicativas. En los modelos se eliminaron en primera instancia 3 variables, estas fueron; tipo de relación conyugal, condición de trabajo y lugar de residencia. Eliminando estas tres variables el modelo no es significativo en términos de la bondad de ajuste, ya que se obtiene una Devianza muy grande. De igual manera se eliminaron sólo dos variables, estas fueron tipo de relación conyugal y condición de trabajo. Y de nuevo se tiene una respuesta negativa en términos de una prueba de bondad de ajuste.

Asimismo se procedió a eliminar solo tipo de relación conyugal, pero el resultado fue el mismo. Por lo que el modelo propuesto anteriormente contiene el número necesario de variables explicativas para poder explicar la relación con respecto a la utilización de algún método anticonceptivo.

Ahora como otra medida de bondad de ajuste se utilizará la matriz de la confusión, presentada en el capítulo dos.

La cual consiste en predecir con valor de uno dada una probabilidad mayor de 0.50. La variable fue generada y arrojo los siguientes resultados:

	Variable Usa		
Valor Predicho	no	si	Total
P(Dep=1) <=0.5	1,469	694	2,163
P(Dep=1) > 0.5	2,155	4,116	6,271
Total	3,624	4,810	8,434
Correctas	1,469	4,116	5,585
% Correctas	40.54%	85.57%	66.22%
% Incorrectas	59.46%	14.43%	33.78%

En donde se toman los grupos muestrales existentes, se considera que en cada grupo todas las mujeres tienen la misma probabilidad de utilizar algún método anticonceptivo. Por lo que se pueden hacer inferencias de un grupo como la probabilidad de una mujer que comparte características de dicho grupo.

La medida de la matriz de la confusión, presenta un porcentaje de 66.22, dicho valor indica que se está prediciendo un 66.22 % en forma correcta, según la regla de predicción. Este valor está combinado por 41% de respuestas correctas predichas como no usuarias y un 85.57% de respuestas correctas de personas predichas por usuarias de métodos anticonceptivos.

Con las medidas de bondad de ajuste presentadas, como son; la Devianza, probar la hipótesis del valor nulo de cada uno de los estimadores y la matriz de la confusión, se puede asumir que el modelo está describiendo en buena forma a los datos.

El siguiente punto consiste en observar el cambio de probabilidad de utilizar algún método dado un cierto conjunto de variables, en donde dichas variables tienen un valor fijo, dicho valor estuvo dado por la media.

4.4.2 Análisis de los resultados del Modelo Lineal Generalizado. *Análisis Nacional*

Como objetivo de esta sección se analizará el efecto que tiene una mujer con ciertas características en la probabilidad de utilizar métodos anticonceptivos. Dado que se tienen grupos, se supone que cada mujer de cada grupo tiene la misma probabilidad de utilizar algún método anticonceptivo. Es decir, las observaciones entre grupos se consideran independientes.

Por lo tanto para analizar la manera en que cambia la probabilidad de uso de algún método anticonceptivo según las características de mujer se observarán diferentes escenarios, esto con el fin de observar el efecto que tiene cada una de las variables con respecto a la probabilidad de uso.

Como primer caso se analiza el efecto que tienen las variables de educación y número de métodos que conoce la mujer en la probabilidad de uso. En este caso se mantienen fijas las variables de grupo de edad, tipo de relación conyugal (estado civil), condición de trabajo, lugar de residencia, condición de aseguramiento y número de hijos. Por lo que se plantean los escenarios en donde las variables varían en forma conjunta, es decir, si aumenta el nivel en una variable también aumentará en las demás.

Como ejemplo del párrafo anterior se plantea lo siguiente, supóngase como primer caso que el grado de educación es nulo, entonces el número de métodos que se conocen es nulo también, asimismo no tienen hijos, ahora como segundo caso si aumenta el grado de educación a 1-3 años de educación primaria, entonces aumentará también el número de métodos a 4-7, métodos, y así sucesivamente. Los valores fijos que toman las variables son, grupo de edad de 25-29 años, la relación conyugal es por el civil, no trabajan, no tienen seguridad social y viven en área rural. Dadas las características anteriores se tiene la Tabla 4 – 15, que se presenta en la siguiente página.

Grupo de Edad	Edo_civil	Trabaja	Aseguramiento	Número de hijos	Lugar de residencia	Educación	Número de métodos	Eta=XB	Pr_(usa | X)
25-29	Civil	NO	NO	2	Rural	Ninguna	0	-0.9662428	0.166961343
25-29	Civil	NO	NO	2	Rural	1-3 años Primaria	1-3	-0.5240914	0.300107466
25-29	Civil	NO	NO	2	Rural	4-6 años Primaria	4-7	-0.08194	0.467347146
25-29	Civil	NO	NO	2	Rural	Secundaria	8-10	0.3602114	0.640655416

Tabla 4 - 15

Grupo de Edad	Edo_civil	Trabaja	Aseguramiento	Número de hijos	Lugar de residencia	Educación	Número de métodos	Eta=XB	Pr_(usa | X)
25-29	Civil	Si	Si	2	Urbano	Ninguna	0	-0.4902225	0.31198823
25-29	Civil	Si	Si	2	Urbano	1-3 años Primaria	1-3	-0.0480711	0.480829723
25-29	Civil	Si	Si	2	Urbano	4-6 años Primaria	4-7	0.3940803	0.653239075
25-29	Civil	Si	Si	2	Urbano	Secundaria	8-10	0.8362317	0.798487771

Tabla 4 - 16

En la tabla 4-15 se tiene el primer conjunto de variables fijas, en la cual se observa un cambio en la probabilidad de utilizar algún método anticonceptivo a medida que mejoran las categorías de las variables no fijas.

Al observar la tabla anterior se cumplen las hipótesis planteadas en las tablas de contingencia, es decir, la educación tiene un efecto positivo con respecto al uso, asimismo combinándose con un mayor conocimiento de métodos existentes para la prevención de métodos (ésta variable se incluye ya que se supone que a través de las instituciones de educación se propaga información acerca del número métodos que conoce una mujer).

En la tabla 4-15 se tiene en el primer renglón el peor escenario, no siendo en algún sentido peyorativo, ya que se tiene a una mujer que no cuenta con educación, ni seguridad social, no tiene trabajo y se ubica en zona rural. Éstas características se consideran el peor escenario, ya que en la zona rural existe menor número de servicios, asimismo no tiene algún grado de educación, no tiene conocimiento de algún método anticonceptivo.

Las variables restantes presentan únicamente un cierto grupo en particular. Para el caso mencionado anteriormente, se tiene que una mujer que presente tales características tiene una probabilidad de 0.1669, la cual es baja. Ahora si se toma a una mujer que presente las mismas características fijas, pero tiene de 4-6 años de educación y tiene conocimiento de 4-7 métodos de planificación familiar, ésta es de 0.4673, la cual es menor que la media que es 0.6804. Ahora comparando estos dos grupos de mujeres que presentan las mismas características fijas pero que varían en grado de educación y número de métodos, se tiene un aumento en la probabilidad de la segunda mujer de casi 2 veces más que la primera, es decir, 0.4673 contra 0.1669.

Con lo anterior parece indicar que dichas variables tienen un efecto muy fuerte en la probabilidad.

Ahora se corroborará tal efecto, utilizando las mismas variables fijas pero cambiando las categorías de vivir en zona urbana, tiene seguridad social y trabaja, tales datos se presentan en la tabla 4-16.

Dicha tabla en el primer renglón presenta a una mujer que pertenece al grupo de edad de 25-29 años, esta casada por el civil, trabaja, tiene seguridad social, tiene 2 hijos, vive en zona urbana, pero no tiene algún grado de educación ni tiene conocimiento de algún método anticonceptivo, tal mujer presenta una probabilidad de 0.3120, la cual sigue siendo baja, ya que esta por debajo de 0.5.

Igual que en el caso anterior, se toma una mujer que presenta las mismas características fijas pero tiene de 4-6 años de primaria y tiene conocimiento de 4-7 métodos, obteniéndose una probabilidad de 0.6532. En este caso la probabilidad aumenta casi el doble con respecto a la mujer anterior.

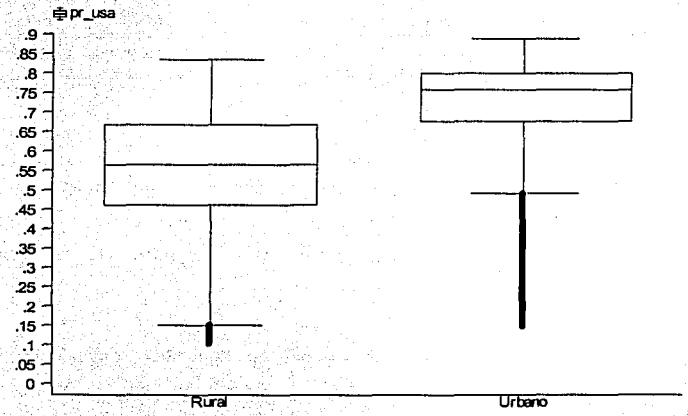
En esta tabla el máximo de la probabilidad se obtiene cuando una mujer mantiene las mismas características fijas, pero la mujer tiene el nivel de secundaria y tiene conocimiento de 8-10 métodos anticonceptivos, tal probabilidad es 0.7984. La cual es bastante alta.

Con lo anterior se ha presentado la importancia que tienen las variables de educación y conocimiento de número de métodos en la probabilidad de uso. Es decir, se ha probado en dos diferentes escenarios, el primero considera a una mujer que vive en zona rural, que no tiene seguridad social y no trabaja. Como segundo escenario, se presenta una mujer que vive en zona urbana, trabaja y tiene seguridad social. Ambas

mujeres comparten el encontrarse en el mismo grupo de edad de 25-29 años, están casadas sólo por el civil y tienen ambas dos hijos.

Considerado lo anterior, se observa el cambio en la probabilidad al avanzar de una categoría a otra en las variables fijas. Asimismo se corroboran las hipótesis que se plantearon en las tablas de contingencia. Como siguiente punto se analizará a nivel nacional algunas variables consideradas importantes que presentan algún cambio en la probabilidad de uso, como son el lugar de residencia, condición de aseguramiento y educación.

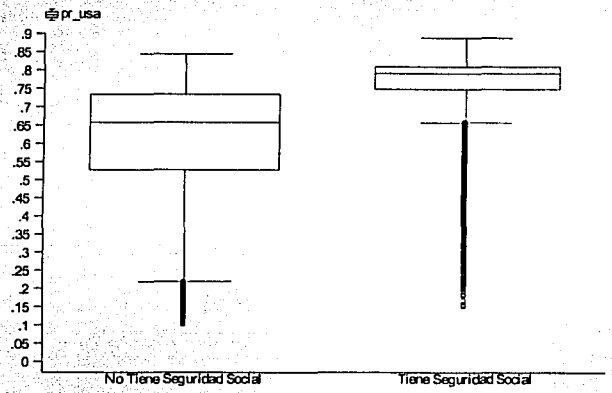
Haciendo un análisis comparativo de la probabilidad de uso por lugar de residencia, la presente gráfica muestra como varía la probabilidad en los estratos rural y urbano a nivel nacional.



La gráfica anterior presenta resultados interesantes, entre ellos se tiene que en el medio rural existe una menor media de probabilidad comparada con el medio urbano, asimismo se tiene una menor dispersión en el medio urbano, lo cual se

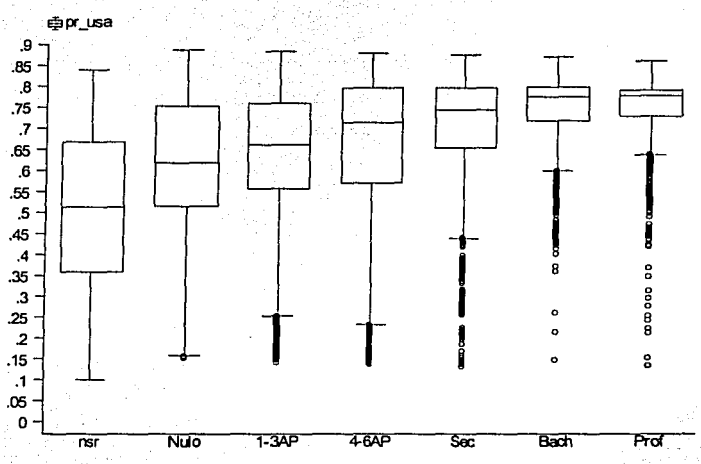
observa en la caja derecha, la cual presenta menor dispersión. Las medias de probabilidades son para la zona de residencia rural de 0.5433 y de 0.7253 para la zona urbana.

Otro resultado interesante que se asume que influye en la probabilidad de que una mujer utilice o no servicios de planificación familiar, es la condición de aseguramiento de la mujer. En la siguiente gráfica de caja se observa comparaciones de importancia.



En ésta gráfica es muy clara la diferencia existente en la probabilidad de uso de métodos anticonceptivos por condición de aseguramiento, es decir, el hecho que una mujer posea algún servicio médico influye en la decisión de utilizar algún método de anticoncepción. Las instituciones públicas son fuente importante de difusión de métodos anticonceptivos y de distribución gratuita de éstos. Por lo que es justificado la marcada diferencia existente el grupo de mujeres que tiene algún sistema de seguridad social. Las medias de la probabilidad para las mujeres no aseguradas es de 0.6157 contra 0.7619 para el grupo de mujeres aseguradas.

Por último, otra variable que juega un papel importante en la decisión de utilizar o no algún método anticonceptivo es la variable de educación. La gráfica muestra particularidades.



En esta gráfica, al observar las medianas en cada unas de las cajas se presenta una tendencia ascendente a medida que el nivel de educación va aumentando. Además de presentar una menor dispersión en las ultimas categorías de educación. Lo cual dice que la probabilidad de uso es alta y cercana a la media. Estas probabilidades son :

Educación	Media	Mediana
NSR	0.499	0.514
Nulo	0.613	0.618
1-3Años Prim	0.638	0.661
4-6 Años Prim	0.670	0.713
Secundaria	0.716	0.743
Bachillerato	0.746	0.775
Profesional	0.756	0.777

4.4.3 Análisis de los resultados del Modelo Lineal Generalizado. *Análisis Estatal*

El objetivo de ésta sección es desarrollar el mismo modelo lineal generalizado con liga probit y distribución binomial en cada estado de la República Mexicana. Al igual que en el caso anterior se analizará la bondad de ajuste del modelo.

La construcción de las variables se realiza de la misma forma que el modelo a nivel nacional. Es decir se agrupa en base a las mismas variables como son; **educ1 num_m4 gedad edo_cvl trab n_hijos aseg ru.**

A esta agrupación de datos se le aplica el mismo modelo presentando los siguientes resultados.

4.4.3.1 Análisis de la Devianza

El análisis de bondad de ajuste de los estados resultaron ser significativos en todos los estados con la prueba de la Devianza. A continuación se presenta la tabla que refleja la estadística de bondad de ajuste. Es decir, se analiza el valor de una Ji-Cuadrada con n-p grados de libertad, donde n es el número de grupos y p el número de parámetros ambos por estado al nivel de 95% de confianza. Este valor se compara con la Devianza proporcionada por el modelo.

Clave	Nombre Entidad	Devianza	Valor de una Ji-Cuadrada al 95%		Diferencia ((Ji_(n-p)_95%) - Devianza)
			n-p	Valor	
1	Aguascalientes	530.9	895	965.7	434.76
2	Baja California Norte	541.0	973	1,046.7	505.68
3	Baja California Sur	437.6	908	979.2	541.59
4	Campeche	758.9	1,149	1,229.0	470.06
5	Coahuila	396.0	1,098	1,176.2	780.17
6	Colima	507.8	943	1,015.6	507.73
7	Chiapas	1,259.3	1,319	1,404.6	145.29
8	Chihuahua	544.2	1,059	1,135.8	591.62
9	Distrito Federal	253.9	872	941.8	687.93
10	Durango	571.4	980	1,053.9	482.50
11	Guanajuato	796.7	1,105	1,183.4	386.77
12	Guerrero	1,083.0	1,210	1,292.0	209.00
13	Hidalgo	1,060.8	1,225	1,307.5	246.77
14	Jalisco	533.7	944	1,016.6	482.90
15	Edo. de México	492.0	1,286	1,370.5	878.49
16	Michoacán	784.8	1,055	1,131.7	346.89
17	Morelos	599.7	1,020	1,095.4	495.72
18	Nayarit	666.0	1,102	1,180.3	514.34
19	Nuevo León	386.5	971	1,044.6	658.13
20	Oaxaca	1,054.7	1,127	1,206.2	151.54
21	Puebla	1,037.4	1,221	1,303.4	265.97
22	Querétaro	690.7	1,006	1,080.9	390.24
23	Quintana Roo	742.9	1,095	1,173.1	430.18
24	San Luis Potosí	809.0	1,107	1,185.5	376.56
25	Sinaloa	557.2	1,133	1,212.4	655.24
26	Sonora	524.8	1,017	1,092.3	567.52
27	Tabasco	873.0	1,261	1,344.7	471.71
28	Tamaulipas	677.6	1,059	1,135.8	458.19
29	Tlaxcala	719.5	1,141	1,220.7	501.15
30	Veracruz	962.5	1,291	1,375.7	413.16
31	Yucatán	688.4	1,064	1,141.0	452.59
32	Zacatecas	390.8	720	783.5	392.72

Tabla 4 - 17

La tabla anterior presenta el comparativo entre el modelo saturado y el modelo con p parámetros. A través de la estadística de la Devianza. La cual se describe a continuación.

Como se observa el análisis se realiza a nivel estatal, indicado por la primera y segunda columna. La tercer columna presenta el valor de la devianza generada por el modelo con p parámetros. En la cuarta y quinta columna se tienen el número de grados de libertad y el valor de la Ji- Cuadrada al 95% con n-p grados de libertad respectivamente.

Para determinar la bondad de ajuste del modelo se calcula la quinta columna, la cual corresponde a la diferencia entre el valor la Ji-Cuadrada y Devianza. Según la regla determinada en el capítulo dos. La cual resume que; si el valor de la Devianza generada por el modelo es mayor que el valor predicho de una Ji- Cuadrada al $(1-\alpha)100\%$ se rechaza la hipótesis que el modelo ajusta de buena manera a los datos. En caso contrario, si el valor de la Devianza es menor que el valor de la Ji-Cuadrada, entonces el modelo describe en forma adecuada a los datos, no rechazándose así la hipótesis.

Por lo que para determinar dicha medida se calculó la diferencia mencionada anteriormente, tal diferencia será considerada de la siguiente manera; si el valor de la diferencia es negativo entonces se rechaza la hipótesis que el modelo estatal ajusta de buena manera a los datos. En caso contrario, el modelo ajusta de buena manera al conjunto de datos a nivel estatal.

Al aplicar la regla anterior se concluye que el modelo ajusta en buena forma a todos los estados sin excepción alguna. Por lo que se pueden realizar inferencias por entidad federativa.

4.4.3.2 Análisis de la Matriz de la Confusión

Ahora por otra parte, se determinará el porcentaje de valores predichos correctamente por el modelo. Lo anterior se realizará con la matriz de la confusión, medida de bondad de ajuste explicada en el capítulo dos.

La siguiente tabla presenta el porcentaje de valores predecidos correctamente;

Clave	Nombre Entidad	Porcentaje de valores predecidos correctamente
1	Aguascalientes	63.27%
2	Baja California Norte	74.95%
3	Baja California Sur	76.77%
4	Campeche	71.42%
5	Coahuila	72.72%
6	Colima	73.53%
7	Chiapas	64.38%
8	Chihuahua	73.41%
9	Distrito Federal	76.39%
10	Durango	68.76%
11	Guanajuato	65.17%
12	Guerrero	66.94%
13	Hidalgo	68.23%
14	Jalisco	66.84%
15	Edo. de México	72.82%
16	Michoacán	66.54%
17	Morelos	72.01%
18	Nayarit	76.33%
19	Nuevo León	70.41%
20	Oaxaca	67.78%
21	Puebla	67.80%
22	Querétaro	68.18%
23	Quintana Roo	70.92%
24	San Luis Potosí	67.65%
25	Sinaloa	75.48%
26	Sonora	74.56%
27	Tabasco	71.89%
28	Tamaulipas	72.38%
29	Tlaxcala	72.35%
30	Veracruz	73.23%
31	Yucatán	70.55%
32	Zacatecas	68.18%

Tabla 4 - 18

La tabla anterior presenta en su última columna el porcentaje de valores que se están prediciendo correctamente, decir se considera como 1, si su probabilidad de uso es mayor o igual a 0.5 y 0 si su probabilidad es menor que 0.5. De esta manera se obtiene el porcentaje de valores predecidos correctamente.

La manera en que se construyó es la siguiente

Aguascalientes
Valores muestrales

Valor predecido según la regla de prediccion de la matriz de confusión	Valor observado de Usa algún método anticonceptivo		
	No	Si	Total
0	115	77	192
1	255	457	712
Total	370	534	904
Correctas	115	457	572
% Correctas	31.08%	85.58%	63.27%
% Incorrectas	68.92%	14.42%	36.73%

Tabla 4 - 19

La tabla anterior muestra el caso de Aguascalientes, esta solo se presenta para mostrar la manera en que se construye el porcentaje de valores predichos correctamente. Los estados restantes se construyen de la misma manera.

En el reporte nacional se obtuvo una media de porcentaje de 66.22, es decir que esta prediciendo correctamente cerca del 66.22 %. En este caso se tiene que el mínimo se presenta en el estado de Aguascalientes con 63.27 y seguido de Chiapas con un valor de 64.38%.

Por otra parte los estados que tienen un mayor porcentaje de valores predichos correctamente son: Nayarit con 76.33, El Distrito Federal con un valor de 76.39 y Baja California Sur con 76.77.

En términos generales la mayoría de los estados están en un porcentaje mayor que el reportado por el nivel nacional. Además se puede decir que el modelo explica en un porcentaje mayor al 66% los datos a nivel estatal.

Concluyéndose por lo tanto que ambas medidas presentan resultados satisfactorios.

4.5 Análisis de los Resultados. Estados seleccionados.

Una vez verificada la bondad de ajuste del modelo lineal generalizado con liga Probit y distribución Binomial, el siguiente punto consiste en analizar los resultados arrojados por el modelo.

La tabla que se presenta a continuación muestra el promedio, la desviación estándar, el mínimo y máximo de la probabilidad de uso de algún método anticonceptivo por estado. Posteriormente se seleccionarán de dicha tabla algunos estados que presenten ciertas particularidades.

Clave	Nombre Entidad	Probabilidad Promedio Estatal	Desviación Estándar	Mínimo	Máximo
1	Aguascalientes	0.632	0.109	0.175	0.778
2	Baja California	0.743	0.130	0.172	0.966
3	Baja California Sur	0.770	0.110	0.149	0.972
4	Campeche	0.717	0.126	0.139	0.924
5	Coahuila	0.756	0.094	0.277	0.941
6	Colima	0.736	0.107	0.196	0.936
7	Chiapas	0.536	0.204	0.066	0.851
8	Chihuahua	0.743	0.077	0.323	0.841
9	Distrito Federal	0.780	0.069	0.285	0.912
10	Durango	0.678	0.122	0.145	0.864
11	Guanajuato	0.548	0.169	0.067	0.843
12	Guerrero	0.476	0.227	0.015	0.819
13	Hidalgo	0.634	0.169	0.105	0.891
14	Jalisco	0.665	0.122	0.123	0.832
15	Edo. de México	0.751	0.149	0.072	0.931
16	Michoacán	0.613	0.175	0.058	0.894
17	Morelos	0.736	0.119	0.202	0.926
18	Nayarit	0.738	0.122	0.192	0.918
19	Nuevo León	0.731	0.099	0.175	0.939
20	Oaxaca	0.549	0.211	0.058	0.903
21	Puebla	0.593	0.206	0.050	0.898
22	Querétaro	0.608	0.185	0.040	0.838
23	Quintana Roo	0.699	0.158	0.075	0.922
24	San Luis Potosí	0.606	0.163	0.114	0.867
25	Sinaloa	0.769	0.109	0.170	0.953
26	Sonora	0.762	0.073	0.355	0.889
27	Tabasco	0.689	0.126	0.176	0.900
28	Tamaulipas	0.726	0.086	0.333	0.926
29	Tlaxcala	0.721	0.110	0.228	0.866
30	Veracruz	0.699	0.144	0.133	0.917
31	Yucatán	0.702	0.153	0.076	0.902
32	Zacatecas	0.674	0.134	0.048	0.931

Observando los datos anteriores se tiene que los valores mínimos, pertenecen a los estados de Guerrero y Chiapas. Asimismo, el máximo se presenta en el Distrito Federal y Sinaloa. La probabilidad de uso en términos generales se encuentra poco arriba de 0.60

En este caso el interés es seleccionar 6 estados que representen diferentes contrastes, es decir, que presenten diferencias marcadas en algún sentido. Esto con el fin de calcular la probabilidad dado un cierto conjunto de variables explicativas en cada uno los estados seleccionados.

Estos estados fueron seleccionados en base al índice de marginación. Índice que es presentado por el Consejo Nacional de Población (CONAPO). Es este sentido se tomará el índice calculado para el año 2000 y se tomaron los tres primeros estados y en base al ranking de dicho índice se tomaran los tres estados que son considerados como grandes ciudades.

Los estados seleccionados fueron los siguientes:

1) Chiapas
2) Oaxaca
3) Guerrero
4) Jalisco
5) Nuevo León
6) Distrito Federal

Ahora se analizan por separado cada uno de los estados, con el fin de encontrar diferencias internas. Las principales variables empleadas son educación, lugar de residencia y número de hijos.

4.5.1 *Chiapas 07*

En la ***tabla 4-20*** presentada en la siguiente página, se presenta de manera conjunta diferentes resultados, la cual permite hacer inferencias dentro del estado.

Esta tabla presenta 3 grandes grupos en la primera columna; zona rural, zona urbana y total. Dentro de cada una de estas se tiene en el primer renglón la probabilidad promedio de uso de métodos anticonceptivos, en el segundo renglón se tiene el número promedio de hijos que tiene una mujer de cierto grupo, el tercer renglón presenta el número de observaciones que se tiene y por último el cuarto renglón que proporciona el porcentaje que representa en términos de cantidad de observaciones. Ahora todas las características expresadas anteriormente están divididas por grado de educación.

Observando la tabla se encuentra que en los totales por lugar de residencia una probabilidad de uso de anticonceptivos viviendo en zona rural de 0.45 y contrastado contra la zona urbana una probabilidad de 0.628, lo cual nos dice que existen diferencias dentro del estado. Asimismo, comparando estas dos características en cada uno de los niveles de educación se tiene una relación de mayor uso en casi todos los niveles, exceptuando el nivel de profesional o postgrado el cual es de 0.763 en zona rural contra 0.737 de zona urbana, pero sin embargo si se considera el porcentaje que representa la población rural el cual es muy pequeño en comparación del porcentaje de población urbana en el mismo nivel de educación, es decir, un 0.25% contra 5.85% respectivamente. La menor probabilidad de uso se encuentra en el grupo de mujeres que no respondieron o no saben que grado de educación tienen y que su lugar de residencia es una zona rural, tal probabilidad es de 0.345. Aquí se supone que la mujer no quiere responder por no tener algún grado de educación u otra causa.

Estado : 07 Chiapas		Educación de la mujer						
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	4-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Postgrado	Total
Rural								
Probabilidad Promedio de Uso	0.345	0.419	0.492	0.473	0.574	0.612	0.763	0.45
Número Promedio de Hijos	12	14	12	5	3	2	10	9
Observaciones	81,860	7,313	78,392	103,708	20,648	2,030	750	294,701
Porcentaje del Grupo Rural	27.78%	2.48%	26.60%	35.19%	7.01%	0.69%	0.25%	100.00%
								51.81%
Urbano								
Probabilidad Promedio de Uso	0.52	0.611	0.58	0.625	0.666	0.71	0.737	0.628
Número Promedio de Hijos	11	16	10	5	2	2	2	6
Observaciones	38,504	6,458	49,399	64,283	67,284	32,170	16,044	274,142
Porcentaje del Grupo Urbano	14.05%	2.36%	18.02%	23.45%	24.54%	11.73%	5.85%	100.00%
								48.19%
Total								
Probabilidad Promedio de Uso	0.401	0.509	0.526	0.531	0.644	0.705	0.738	0.536
Número Promedio de Hijos	12	15	11	5	3	2	2	7
Observaciones	120,364	13,771	127,791	167,991	87,932	34,200	16,794	568,843
Porcentaje del Grupo Total	21.16%	2.42%	22.47%	29.53%	15.46%	6.01%	2.95%	100.00%

Tabla 4 - 20

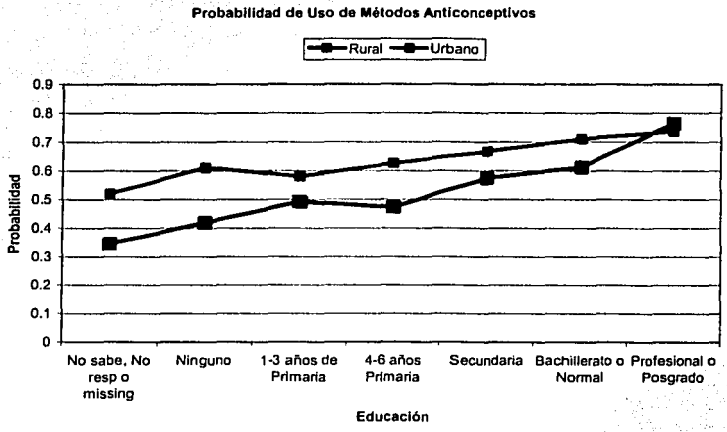
Siguiendo con el análisis de probabilidades de uso de métodos anticonceptivos, se observa que la probabilidad para las mujeres que reportaron no tener ningún grado de educación es de 0.419 en zona rural contra 0.611 para zona urbana, asimismo el número promedio de hijos en estas categorías de educación son la más altas con 14 y 16 respectivamente. Al observar este último dato se corrobora que se tenga una probabilidad de uso tan baja, ya que al no usarse algún método anticonceptivo esto implica que el número promedio de hijos sea alto.

Tomando un punto medio de educación la categoría de 4-6 años de primaria, para zona rural se tiene una probabilidad de uso de 0.473 contra una de 0.625 para zona urbana. Asimismo aquí se observa un menor número promedio de hijos con 5 en ambos lugares de residencia.

Observando únicamente la zona de residencia urbana, se aprecia un aumento de la probabilidad de uso de métodos anticonceptivos, de igual manera se observa una relación inversa con respecto al número promedio de hijos por categoría, lo cual era de esperarse, ya que al haber una mayor probabilidad de uso, esto implica que el número de hijos se esta controlando y por lo tanto va disminuyendo de 16 hasta 2 para las mujeres.

El máximo de probabilidad de uso lo tiene la categoría de profesional o postgrado con 0.705, lo cual es esperarse ya que se supone que en las instituciones educativas son una fuente potencial de recibir información de planificación familiar. Además de tener una cultura diferente respecto a la procreación de hijos.

En total se determinó una probabilidad final de 0.536, la cual es baja, ya que apenas esta por encima de 0.5, pero está por debajo de la media nacional. Además se observó un número promedio de 7 hijos por mujer, dicho número es alto.



En esta gráfica se tiene una tendencia muy marcada respecto a la probabilidad de utilizar algún método anticonceptivo y el nivel de educación que presenta una cierta mujer.

Asimismo como se comentó en la página anterior la probabilidad de uso es mayor en la zona urbana que en la zona rural. Tal aseveración se puede observar claramente en la gráfica anterior, con excepción del grupo de profesional o postgrado.

4.5.2 *Guerrero 12*

Los resultados del estado de Guerrero se presentan en la **tabla *4-21*** que se muestra en la página 133. Observando el total se tiene que la probabilidad de que una mujer en el estado de Guerrero utilice algún método anticonceptivo es **0.476**, la cual es muy baja, de hecho posee el mínimo respecto a los demás estados.

Considerando los totales en la variable promedio de hijos contra la educación, se observa que a medida que aumenta el grado de educación se reduce el número promedio de hijos corriendo de 14 a 3 hijos[16], de igual manera la probabilidad de utilizar algún método anticonceptivo aumenta al incrementarse el grado de educación, ya que ésta tiene su recorrido de 0.454 hasta 0.644.

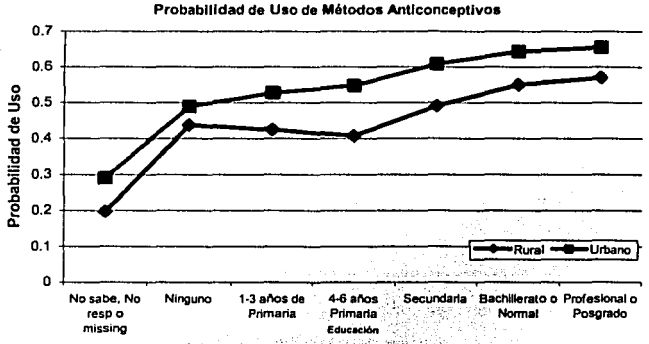
Ahora analizando los datos por lugar de residencia se aprecia que la probabilidad total en la zona rural es de 0.374 contrastado contra 0.551 de la zona urbana, aunque ésta sigue siendo pequeña. En este caso existe una gran diferencia ya que la probabilidad de uso entre zona rural y zona urbana son muy diferentes. Lo cual indica que para el estado el lugar de residencia es una variable que presenta diferencias muy marcadas.

El máximo en la probabilidad de uso como era de esperarse se presenta en las mujeres que tienen grado de profesional o postgrado, tal probabilidad es de 0.644. Y el mínimo se presenta al igual que en el caso de Chiapas en aquellas mujeres que no saben o no responden acerca del nivel de educación, tal probabilidad es de 0.233, en este caso se puede suponer que la mujer no responde porque no entiende la pregunta o le apena el responder la falta de educación.
Por último, la mayor diferencia por lugar de residencia, se presenta en el nivel de 1-3 años de educación, ya que para la zona rural esta es de 0.424 contra 0.527 para zona

[16] No se considera la categoría de no sabe o no responde en la variable de educación.

urbana, lo cual indica que existen diferencias en el desempeño de la educación, que se reflejan en el conocimiento de métodos anticonceptivos y por lo tanto en el uso de métodos de planificación familiar.

El estado de Guerrero es uno de los estados que presentan un mayor índice de marginación[17]. Con lo que se corroboran los resultados obtenidos en el presente análisis. En términos generales Guerrero presenta problemas respecto al uso de métodos anticonceptivos, ya que la probabilidad de uso que presenta es demasiado baja, la cual influye de manera inversa con el número de hijos, con lo que se infiere que se deben de tomar medidas para reducir esta falta de uso de métodos anticonceptivos y por ende ayudar a regular el número de nacimientos en el estado.
Los resultados expuestos anteriormente, se pueden observar en la siguiente gráfica, la cual presenta una tendencia positiva con respecto al nivel de educación. Asimismo también se observa que la probabilidad de uso es mayor en todos los niveles de educación para la zona urbana con respecto a la zona rural.



[17], Índices de Marginación Estatal 2000. Información CONAPO

Estado : 12 Guerrero		Educación de la mujer						
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	4-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Postgrado	Total
Rural								
Probabilidad Promedio de Uso	0.197	0.437	0.424	0.407	0.49	0.55	0.57	0.374
Número Promedio de Hijos	13	14	13	7	3	2	2	9
Observaciones	50,441	6,152	37,169	57,206	23,245	8,001	3,987	186,201
Porcentaje del Grupo Rural	27.09%	3.30%	19.96%	30.72%	12.48%	4.30%	2.14%	100.00%
								42.79%
Urbano								
Probabilidad Promedio de Uso	0.291	0.489	0.527	0.547	0.608	0.643	0.655	0.551
Número Promedio de Hijos	12	13	12	6	3	3	3	6
Observaciones	30,862	3,017	29,794	64,219	61,346	33,603	26,124	248,965
Porcentaje del Grupo Urbano	12.40%	1.21%	11.97%	25.79%	24.64%	13.50%	10.49%	100.00%
								57.21%
Total								
Probabilidad Promedio de Uso	0.233	0.454	0.469	0.481	0.575	0.625	0.644	0.476
Número Promedio de Hijos	13	14	13	7	3	3	3	8
Observaciones	81,303	9,169	66,963	121,425	84,591	41,604	30,111	435,166
Porcentaje del Grupo Total	18.68%	2.11%	15.39%	27.90%	19.44%	9.56%	6.92%	100.00%

Tabla 4 - 21

4.5.3 *Oaxaca 20*

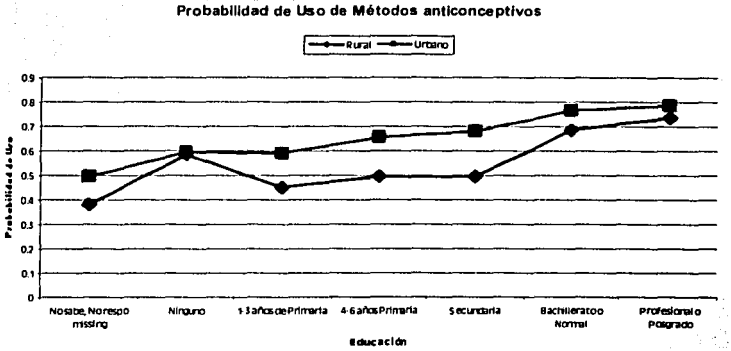
Se continúa con el tercero de los estados que presentan un mayor índice de marginación, éste al igual que los anteriores presenta problemas respecto al control de la fecundidad.

Observando la probabilidad de uso de métodos anticonceptivos del estado, ésta es de 0.549, la cual al igual que en el caso anterior es baja ya que está por debajo de 0.6803 que es la media nacional. Observando diferencias entre estratos, en este estado se tiene una diferencia de casi 0.20 en la probabilidad de uso, ya que para la zona rural es de 0.467 contra una de 0.652 para la zona urbana. Tales resultados también se pueden observar en la gráfica presentada en la página siguiente, la cual permite ver que la probabilidad por estrato en todos los niveles es mayor en la zona urbana que en la zona rural.

De igual manera se observa una relación inversa respecto al número de hijos y la probabilidad de uso, con excepción del grupo de ningún grado de educación, aunque este grupo sólo representa menos del 1% de la población, por lo que se puede omitir. Este aumento en la probabilidad va de 0.493 hasta 0.781 para los grupos de 1-3 años de educación y profesional respectivamente. Asimismo el número de hijos va de 11 hasta 1, para los mismos grupos del caso anterior. Lo cual como se comentó anteriormente es de esperarse esta asociación, ya que a mayor nivel de educación se supone un mayor conocimiento de métodos y una cultura diferente respecto al concepto sobre la procreación de hijos.

Al igual que en los estados anteriores se tiene una tasa alta de personas que no saben el grado de educación o no responden por ciertas razones. El supuesto que se hace aquí es que tales mujeres no tienen algún grado de educación por lo que no saben que responder o simplemente por pena.

Observando el grupo de lugar de residencia rural se presenta que existen desigualdades y diferencias dentro del mismo estrato, ya que el mínimo de probabilidad es de 0.45 para el grupo que tiene de 1-3 años de educación primaria hasta un 0.734 para las mujeres que tienen un nivel profesional. Con lo anterior se puede afirmar que la educación dentro del estrato es importante ya que la probabilidad de uso es bastante alta con respecto a los demás estados. Lo anterior también se refleja en el número promedio de hijos, ya que para el grupo de profesional o postgrado se tiene 1 hijo en promedio. En contraste con el grupo de las mujeres que tienen de 1-3 años de educación primaria, el número promedio de hijos es de 13, esta cifra es alarmante ya que con esto se prueba que no existe un control en la fecundidad.



El efecto de la caída en la zona rural en el nivel nulo a 1-3 años de educación se puede explicar debido a la poca representatividad de mujeres que se encuentran en dicho nivel, ya que tal grupo sólo representa menos de 1%. Este puede ser considerado un error de muestreo en la encuesta.

Estado : 20 Oaxaca		Educación de la mujer						
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	4-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Postgrado	Total
Rural								
Probabilidad Promedio de Uso	0.382	0.585	0.45	0.495	0.493	0.685	0.734	0.467
Número Promedio de Hijos	14	17	13	7	4	3	1	9
Observaciones	48,099	1,669	67,701	119,170	22,153	3,264	1,122	263,178
Porcentaje del Grupo Rural	18.28%	0.63%	25.72%	45.28%	8.42%	1.24%	0.43%	100.00%
								55.87%
Urbano								
Probabilidad Promedio de Uso	0.494	0.593	0.591	0.656	0.681	0.763	0.785	0.652
Número Promedio de Hijos	10	15	9	6	3	2	2	5
Observaciones	26,874	1,486	29,485	67,398	42,994	27,018	12,613	207,868
Porcentaje del Grupo Urbano	12.93%	0.71%	14.18%	32.42%	20.68%	13.00%	6.07%	100.00%
								44.13%
Total								
Probabilidad Promedio de Uso	0.422	0.589	0.493	0.553	0.617	0.755	0.781	0.549
Número Promedio de Hijos	13	16	11	6	3	2	2	8
Observaciones	74,973	3,155	97,186	186,568	65,147	30,282	13,735	471,046
Porcentaje del Grupo Total	15.92%	0.67%	20.63%	39.61%	13.83%	6.43%	2.92%	100.00%

Tabla 4 - 22

4.5.4 *Jalisco 14*

Con este estado inicia el grupo de 3 estados que son considerados los más avanzados en cuestiones de producción económica y de mayor densidad poblacional. La tabla 4-23 presenta los resultados respecto a la probabilidad de uso de métodos anticonceptivos por varias características.

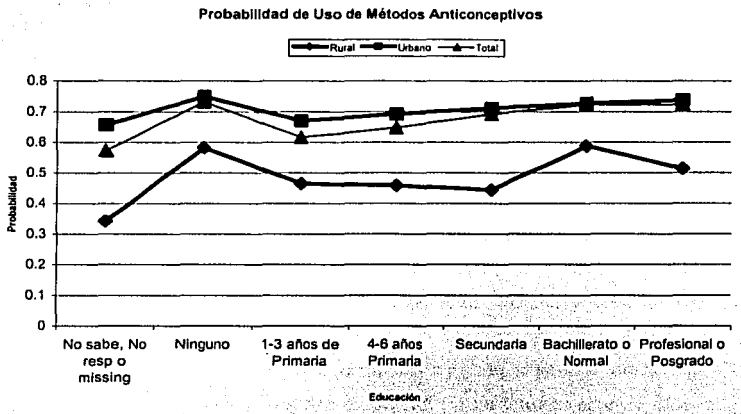
La probabilidad de uso a nivel estatal es de 0.665, en donde a nivel estatal se puede observar el cambio de uso respecto a los estados expuestos anteriormente. Se puede notar que la proporción que no responde o no saben el grado de educación que tiene es menor que los estados anteriores, lo cual habla que las mujeres respondieron en forma más concisa, este porcentaje sólo es de 3.49%.

Aquí como se comentó anteriormente, es considerada una ciudad, lo cual se refleja en el porcentaje de población de mujeres que viven en zona urbana, el cual es de 85.33%. Las probabilidades de uso son 0.454 para el grupo de mujeres en zona rural y de 0.702 para el grupo de mujeres de zona urbana. Aquí se tiene una diferencia de cerca 0.25 en la probabilidad. Lo cual puede comentar que en las zonas rurales al igual que en los estados anteriores la probabilidad de que una mujer utilice algún método anticonceptivo es menor que en zonas urbanas.

El mínimo se encuentra en las mujeres que reportaron tener de 1-3 años de educación, esto es sin considerar la categoría de ningún grado y los que no responden o no saben.
El máximo se encuentra como en los casos anteriores en el grupo de profesional o postgrado para la zona urbana.

Estado : 14 Jalisco		Educación de la mujer						
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	4-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Postgrado	Total
Rural								
Probabilidad Promedio de Uso	0.342	0.582	0.464	0.458	0.443	0.587	0.513	0.454
Número Promedio de Hijos	13	19	13	8	4	8	4	9
Observaciones	8,547	907	34,505	65,476	18,644	1,556	4,670	134,305
Porcentaje del Grupo Rural	6.36%	0.68%	25.69%	48.75%	13.88%	1.16%	3.48%	100.00%
								14.67%
Urbano								
Probabilidad Promedio de Uso	0.657	0.75	0.67	0.692	0.711	0.727	0.738	0.702
Número Promedio de Hijos	12	13	10	7	4	3	3	6
Observaciones	23,402	7,614	98,063	274,883	237,059	69,470	70,725	781,216
Porcentaje del Grupo Urbano	3.00%	0.97%	12.55%	35.19%	30.34%	8.89%	9.05%	100.00%
								85.33%
Total								
Probabilidad Promedio de Uso	0.573	0.732	0.616	0.647	0.692	0.724	0.724	0.665
Número Promedio de Hijos	12	14	11	7	4	3	3	6
Observaciones	31,949	8,521	132,568	340,359	255,703	71,026	75,395	915,521
Porcentaje del Grupo Total	3.49%	0.93%	14.48%	37.18%	27.93%	7.76%	8.24%	100.00%

Tabla 4 - 23



Esta gráfica permite observar la ligera tendencia positiva respecto a la probabilidad de uso de métodos anticonceptivos con el nivel de educación. Esto es observando únicamente la línea que representa el total. Ya que observando las líneas a nivel de estratos se presenta una caída entre el grado de ninguno y de 1-3 años de educación. Tal caída se explica al igual que en los estados anteriores, por el poco porcentaje que representa el grupo, el cual no permite realizar hipótesis si dicha caída es válida. Por lo que se puede omitir tal categoría y considerar las restantes para emitir alguna aseveración.

Asimismo en la gráfica se observa que la línea del total y la zona urbana están muy cercanas, esto es debido al porcentaje de grupo de mujeres que se tienen por tipo de estrato, ya que cerca de un 85% del total es de zona urbana, ésta es la razón por la cual las líneas están casi juntas.

4.5.5 *Nuevo León 19*

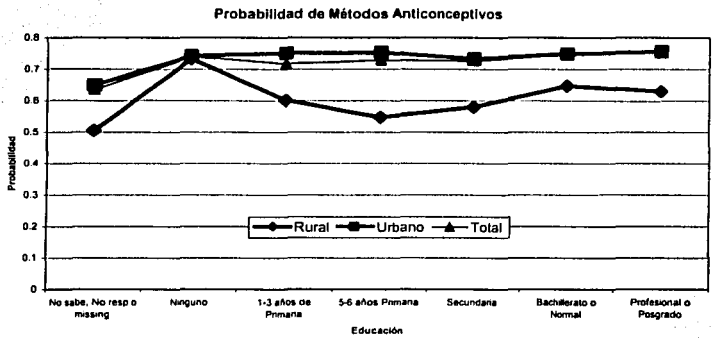
Para el estado de Nuevo León los resultados se presentan en la página 142 en la tabla 4-24, los cuales permiten observar que la probabilidad de uso de métodos anticonceptivos en el estado es alta, dicho valor es 0.731, el cual es mayor que en los estados anteriores.

Observando ahora la comparación entre los estratos rural y urbano, se tiene una probabilidad de uso de 0.571 para la zona rural contra 0.743 para la zona urbana. Lo anterior presenta una diferencia de un poco más de 0.15.

En este estado también se presenta una menor tasa de no respuesta o que no saben, ya que sólo se tiene en total un 1.12%, por lo que aquí además de hacer supuestos se pueden afirmar hipótesis acerca de la probabilidad de uso con la variable educación.

Al observar las probabilidades de uso para la zona urbana y el total, se tiene que éstas son cercanas, esto se podría explicar debido a que en Nuevo León se tiene que el 93% del grupo de mujeres viven en zona urbana. Por lo que se supone que para el total este estrato pesa en gran medida.

En esta gráfica también se puede observar un efecto raro para el grupo de las mujeres que viven en zona rural, las cuales tienen mayor probabilidad de utilizar algún método anticonceptivo en la categoría de educación de 1-3 años de educación contra el grupo que tiene de 4-6 años de educación. Por su parte el mínimo se encuentra en la zona rural para la población que no sabe que grado de educación posee, este valor es de 0.505, aunque de este valor sólo se pueden hacer pocas inferencias dado que solo representa poco menos del 2% de la zona rural.



Aquí se presenta un caso algo extraño en la zona de residencia rural, ya que según los datos, las mujeres que pertenecen al grado de educación ninguna tiene mayor probabilidad de utilizar algún método anticonceptivo, pero deteniéndose a observar el porcentaje que representa del total, se tiene que éste es muy pequeño es de menos de 0.5%, por lo cual se podría evitar el hacer conclusiones acerca de este dato, dado su poco nivel de representatividad.

En contraste con los grupos de 1-3 años de educación, 4-6 años de educación primaria y secundaria, los cuales representan para la zona rural cerca de 92% del total de la zona. En este sentido las probabilidades de uso se consideran confiables, teniéndose éstas cercanas de la media para la zona rural, ya que está alrededor de 0.571.

Estado : 19 Nuevo León	Educación de la mujer							
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	4-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Postgrado	Total
Rural								
Probabilidad Promedio de Uso	0.505	0.733	0.601	0.547	0.58	0.647	0.629	0.571
Número Promedio de Hijos	9	21	12	6	2	6	2	6
Observaciones	758	144	9,460	21,778	10,115	792	1,620	44,667
Porcentaje del Grupo Rural	1.70%	0.322%	21.18%	48.76%	22.65%	1.77%	3.63%	100.00%
								7.11%
Urbano								
Probabilidad Promedio de Uso	0.65	0.742	0.751	0.753	0.733	0.749	0.756	0.743
Número Promedio de Hijos	10	10	10	7	2	3	3	4
Observaciones	6,272	1,085	33,163	161,283	245,209	66,581	69,722	583,315
Porcentaje del Grupo Urbano	1.08%	0.19%	5.69%	27.65%	42.04%	11.41%	11.95%	100.00%
								92.89%
Total								
Probabilidad Promedio de Uso	0.635	0.741	0.717	0.728	0.727	0.748	0.754	0.731
Número Promedio de Hijos	10	11	10	7	3	3	3	4
Observaciones	7,030	1,229	42,623	183,061	255,324	67,373	71,342	627,982
Porcentaje del Grupo Total	1.12%	0.20%	6.79%	29.15%	40.66%	10.73%	11.36%	100.00%

Tabla 4 - 24

4.5.6 *Distrito Federal 09*

El Distrito Federal, es el estado que presenta las mejores condiciones respecto a instituciones educativas, fuentes de trabajo y un considerable número de centros médicos. Por lo que hace suponer que la probabilidad de uso de métodos anticonceptivos sea alta.

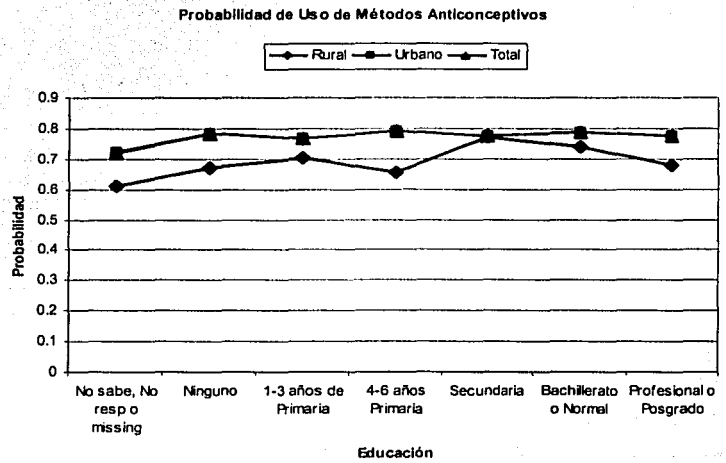
Los datos se presentan en la tabla *4-25*. En esta tabla se nota que el Distrito Federal está compuesto en su totalidad por zonas urbanas, excepto en ciertas pequeñas zonas, tales como las colindancias del Distrito Federal con los estado de México y Morelos. Dado este porcentaje de zona urbana se espera que la probabilidad en el estado sea muy cercana al total. Tales porcentajes de grupos de mujeres que viven en zona rural y urbana, son de 99.62 para zona urbana y 0.37 para zona rural.

Dado el poco nivel de representatividad del grupo de mujeres que viven en zonas rurales, se evitará hacer inferencias respecto a dicha zona, para evitar confusiones en las conclusiones finales, por lo que únicamente se analizará la zona urbana.

Deteniéndose a observar la gráfica y la tabla se tiene que el valor máximo por categorías de educación se presenta en aquellos que tienen de 4-6 años de primaria, pero en este estado no existe diferencia significativa respecto a la variación en la probabilidad de uso respecto al grado de educación. Ya que omitiendo el grado de no sabe o no responde el mínimo se presenta en aquellas mujeres que tienen de 1-3 años de primaria con un valor de 0.767.

Ahora las probabilidades en términos generales son altas por tipo de educación, lo que se puede comentar es la tendencia negativa en el número promedio de hijos, con respecto a la educación.

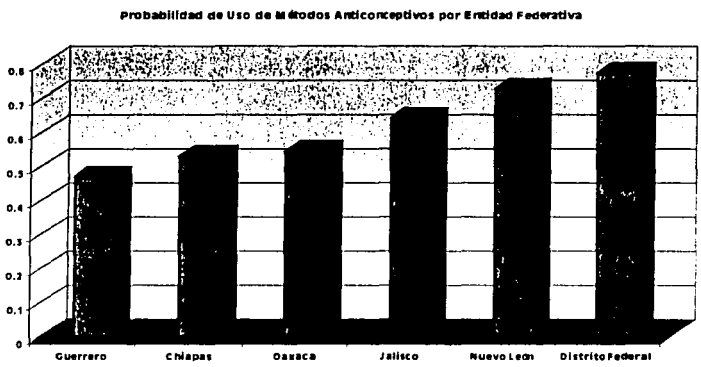
Como se esperaba en el Distrito Federal la probabilidad de uso de métodos anticonceptivos es alta en todos los niveles, esto debido a que es el estado que tiene el mayor nivel promedio en grado de educación, no tiene zonas rurales y existe una mayor fuente de empleo. Es decir, el entorno favorece a una mayor utilización de algún método anticonceptivo bajo el modelo.



Estado : 09 Distrito Federal	Educación de la mujer							
Lugar de Residencia	No sabe, No resp o missing	Ninguno	1-3 años de Primaria	5-6 años Primaria	Secundaria	Bachillerato o Normal	Profesional o Posgrado	Total
Rural								
Probabilidad Promedio de Uso	0.609		0.671	0.704	0.655	0.77	0.739	0.681
Número Promedio de Hijos	9		7	7	2	2	1	5
Observaciones	640		832	852	1,672	668	296	4,960
Porcentaje del Grupo Rural	12.90%	0.00%	16.77%	17.18%	33.71%	13.47%	5.97%	100.00%
								0.38%
Urbano								
Probabilidad Promedio de Uso	0.723	0.785	0.769	0.792	0.776	0.786	0.777	0.78
Número Promedio de Hijos	8	5	8	6	3	2	2	4
Observaciones	28,164	5,873	53,880	302,868	515,676	199,299	208,097	1,313,857
Porcentaje del Grupo Urbano	2.14%	0.45%	4.10%	23.05%	39.25%	15.17%	15.84%	100.00%
								99.62%
Total								
Probabilidad Promedio de Uso	0.72	0.785	0.767	0.792	0.776	0.786	0.777	0.78
Número Promedio de Hijos	8	5	8	6	3	2	2	4
Observaciones	28,804	5,873	54,712	303,720	517,348	199,967	208,393	1,318,817
Porcentaje del Grupo Total	2.18%	0.45%	4.15%	23.03%	39.23%	15.16%	15.80%	100.00%

Tabla 4 - 25

En la siguiente gráfica se presentan las entidades seleccionadas con su respectiva probabilidad promedio.



Como se mencionó anteriormente el máximo de la probabilidad de uso lo tiene el Distrito Federal y el mínimo el estado de Guerrero, teniendo una diferencia entre estos de un poco más de 0.30. Tal diferencia es muy grande, la cual implica que existen marcadas diferencias entre estos estados y posiblemente entre los estados restantes.

Se determinó que en cada uno de los estados seleccionados con alto grado de marginación presentan las mismas diferencias en la probabilidad promedio, es decir, en el lugar de residencia rural se tiene una menor probabilidad comparado con el lugar de residencia urbana.

4.6 Análisis de Tendencia a Nivel Nacional

En esta sección el objetivo es calcular la probabilidad de uso de métodos anticonceptivos para el año 2000, dado que la encuesta que se utilizó fue levantada en al año 1997. Una vez que se calcule la probabilidad para el año 2000, ésta se compara con la probabilidad de uso para el año 1997, teniéndose así un seguimiento de tal efecto. Tal probabilidad únicamente será calculada a nivel nacional.

4.6.1 Instrumento de medición. ENSA 2000.

La encuesta que se utiliza para el año 2000, es la Encuesta Nacional de Salud, 2000, tal encuesta fue levantada por la Secretaría de Salud a través del Instituto Nacional de Salud Pública.

Tal encuesta en términos generales está compuesta de 5 bases de datos; Niños, Adolescentes, Adultos, Hogar y Utilizadores. En las tres primeras bases se pregunta principalmente por cuestiones de salud para cada uno de los grupos de edad, como por ejemplo, para la base de adultos, se pregunta a la persona si tiene diabetes, y si está tomando tratamiento. Para calcular prevalencias de enfermedades esta encuesta es muy buena. En la base de hogar, se preguntan características de todos los miembros del hogar, como por ejemplo sexo, edad, educación, tipo de seguridad social, estado civil, religión, condición de habla de lengua indígena, entre otras.

Para desarrollar el análisis se utilizan casi todas las variables empleadas en el modelo anterior. En ésta encuesta se tienen las variables de educación de la mujer, tipo de relación conyugal (unión libre o casada), lugar de residencia (rural o urbano), condición de trabajo (si o no), grupo de edad al que pertenece la mujer. La variable de

edad presenta un intervalo menor que en la encuesta anterior, ya que el grupo de edad es de mujeres de 20-49 años de edad, la cual será utilizada como aproximación.

La variable número de hijos no se tiene en forma explícita en la encuesta, pero sin embargo se puede construir la variable de número de miembros en el hogar, la cual está muy correlacionada en hogares nucleares. Por lo que tal variable será utilizada como una aproximación de número de hijos.

Lamentablemente la variable número de métodos que conoce la mujer, no se dispone, por lo que el modelo se ejecutará con las variables restantes. El modelo que se realizará es el mismo que se realizó para correr el análisis nacional en la ENADID 97. La variable dependiente en esta encuesta esta construida exactamente igual que el caso anterior.

4.6.2 Especificación del Modelo Lineal Generalizado para ENSA 2000.

En términos formales los elementos del modelo lineal generalizado para esta encuesta quedan definidos por:
El predictor lineal establecido por

$$\eta = \beta_0 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_9 X_9 + \beta_{10} X_{10} \qquad (4.3)$$

X_3 = Educación de la mujer

X_4 = Edad

X_5 = Tipo de relación conyugal

X_6 = Condición de trabajo.

X_7 = Número de miembros en el hogar

X_9 = Lugar de Residencia

X_{10} = Condición de seguridad social

La distribución al igual que antes es definida por $Bin(k_i, \pi_i(\bar{x}))$, en donde se sabe que $\mu_i = k_i * \pi_i(x)$, para k_i conocido. Y la función liga que está definida por:

$$\eta = g(\mu_i) = \phi^{-1}(\frac{\mu_i}{k_i})$$

Por su parte la función inversa está definida por

$$\text{Si } g(\mu_i) = \phi^{-1}(\frac{\mu_i}{k_i}) = \eta_i$$

$$\Rightarrow \text{ la inversa es}$$

$$\mu_i = k_i * \phi(\eta_i) = g^{-1}(\eta_i)$$

De igual manera se procedió a agrupar respecto a las características de las mismas variables explicativas. Para tener un cierto número de éxitos (en este caso mujeres que utilizan algún método anticonceptivo) con su respectivo número de mujeres en dicho grupo.

Otra vez para ejecutar el modelo lineal generalizado se utilizará el paquete estadístico STATA 7.0, el comando empleado fue;

glm sum_usa educ1 gedad edo_cvl trab n_hijos ru aseg [aw=factor1], family(binomial ki) link(probit) irls

Los resultados que presenta se muestran a continuación:

```
Iteration 1 : deviance = -2919.9208
Iteration 2 : deviance = -2925.2153
Iteration 3 : deviance = -2924.9898
Iteration 4 : deviance = -2925.0122
Iteration 5 : deviance = -2925.0103
Iteration 6 : deviance = -2925.0104
Iteration 7 : deviance = -2925.0104
Iteration 8 : deviance = -2925.0104
Iteration 9 : deviance = -2925.0104

Generalized linear models                   No. of obs      =       3111
Optimization     :  MQL Fisher scoring      Residual df     =       3103
                    (IRLS EIM)              Scale param     =          1
```

Deviance	=	2925.010	(1/df) Deviance	=	0.9426395
Pearson	=	4357.722	(1/df) Pearson	=	1.404358

Variance function: V(u) = u*(1-u/ni_n) [Binomial]
Link function : g(u) = invnorm(u/ni_n) [Probit]
Standard errors : EIM

BIC = 2860.668837

sum_usa_n	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
educ1r	-0.0093956	0.0042624	-2.2	0.0280	-0.0177497	-0.0010415
gedad	0.0531516	0.0069462	7.65	0.0000	0.0395372	0.066766
edo_cvl	0.0224808	0.0063297	3.55	0.0000	0.0100747	0.0348868
trab	0.0320953	0.0303338	1.06	0.2900	-0.0273579	0.0915485
num_ind1	0.0416825	0.005879	7.09	0.0000	0.0301598	0.0532051
aseg	0.1731043	0.023206	7.46	0.0000	0.1276215	0.2185872
ru	-0.1680601	0.0223989	-7.5	0.0000	-0.2119611	-0.124159
_cons	0.1230227	0.0482243	2.55	0.0110	0.0285048	0.2175405

4.6.3 Bondad de Ajuste. ENSA 2000.

Al observar la significancia del modelo en términos de la Devianza, se tiene que el modelo resulta ser significativo en conjunto, ya que el valor que proporciona la devianza es de 2925.010432 contra un valor de una Ji- Cuadrada al 0.95 con 3103 grados de libertad determinado por 3233.7053. Al comparar tales valores, se tiene que la devianza es menor que el cuantil al 0.95% de confianza de una Ji-Cuadrada, por la regla de decisión dada en el capítulo dos, el modelo ajusta bien al conjunto de datos.

Observando la significancia de cada uno de los parámetros, resulta que la variable condición de trabajo no tiene significancia estadística, más sin embargo, esta se dejará ya que se considera importante en términos de decisión en la utilización de métodos anticonceptivos.

4.6.4 Análisis de los Resultados Nivel Nacional. ENSA 2000.

Por otra parte observando la siguiente tabla, ésta nos presenta una probabilidad promedio de

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
pr_usar	3111	15,163,183	.7217175	0.0628659	0.525343	0.9209764

El dato de probabilidad promedio es una media ponderada que nos arrojo un total de 15,163,183 mujeres casadas o unidas y en edad fértil al momento de la encuesta, el promedio de la probabilidad de uso de métodos anticonceptivos entre grupos de mujeres con ciertas características fue de 0.7217.

Comparando esta tabla con la tabla del análisis para el año de 1997 presentada en la página 110, se tiene un incremento en la probabilidad promedio de utilización de métodos anticonceptivos. Tal incremento va de 0.6803 hasta 0.7217175, el cual es de 0.0414, el cual es significativo en términos de un periodo de 3 años.

El cambio en la probabilidad se podría explicar a través de un cambio en la distribución del grado de educación, en la distribución de aseguramiento, un incremento en la condición de trabajo de la mujer, entre otras. Lo anterior es únicamente una especulación, ya que no se desarrolló el análisis para no desviar el objetivo de dicho análisis. Una forma de ejecutar dicho análisis es realizar un análisis de determinantes.

CONCLUSIONES

Los modelos lineales generalizados son empleados en gran proporción de estudios realizados en el ámbito de investigación y laboral. Estos son capaces de modelar una gran cantidad de situaciones de la vida cotidiana, esto es debido a la gran diversidad de distribuciones que se pueden emplear y funciones ligas asociadas a éstas.

En el presente trabajo se muestra como aplicación la ejecución de un modelo lineal generalizado utilizando la distribución binomial y la función liga probit, tal modelo fue ejecutado para medir la probabilidad de utilizar algún método anticonceptivo en una mujer en edad fértil y que además se encontraba casada o unida al momento de la encuesta.

Calcular la probabilidad de utilizar algún método anticonceptivo en una mujer en edad fértil y que tiene alguna relación conyugal tiene una gran importancia, ya que con esto se puede tener un parámetro de conocimiento y utilización respecto a métodos anticonceptivos, los cuales influyen directamente en las tasas de fecundidad y nacimientos evitados para un cierto periodo. Además de controlar el número de nacimientos en un hogar, la planificación familiar tiene como repercusión el dar un mejor nivel de vida a los hijos que desea tener, además al obtener un espaciamiento entre los hijos, éstos obtienen un mayor cuidado y atención. Por lo que la planificación familiar tiene grandes repercusiones a nivel global como a nivel hogar.

La probabilidad promedio de que una cierta mujer utilice algún método anticonceptivo a nivel nacional fue de 0.6803 para el año de 1997 y de 0.7217 para el año 2000. Para el año 1997, se realizó un análisis detallado, tales resultados se

presentan a continuación. Por el contrario para el año 2000, sólo se calcula la probabilidad de utilizar algún método anticonceptivo a nivel nacional.

Para el año 1997, el análisis de los resultados muestra una relación importante entre la educación y el hecho de utilizar algún método anticonceptivo, ya que a medida que aumenta el grado de educación aumenta también la probabilidad de que una mujer utilice algún método anticonceptivo, lo que hace suponer que la variable de educación tiene un efecto importante en la toma de decisión de una mujer de utilizar o no algún método anticonceptivo. En términos globales para el año de 1997 la probabilidad promedio ponderada de uso para una mujer que tiene un nivel nulo de educación es de 0.6126 contra una probabilidad de uso de 0.7560 que corresponde a una mujer que tiene un nivel de educación profesional o postgrado.

Asimismo los resultados revelan un contraste marcado en la probabilidad promedio de uso de métodos anticonceptivos por lugar de residencia. Es decir, en el medio urbano al existir alguna institución educativa o alguna institución de salud, así como fuente de trabajo, entre otras, favorece el hecho que una mujer utilice algún método anticonceptivo aumentando por lo tanto la probabilidad de uso. En contraste con el medio rural, en el cual no se tienen estas condiciones, estas tienden a no incrementar la probabilidad de uso de métodos anticonceptivos. Lo anterior se puede ver reflejado en las probabilidades siguientes; para el lugar de residencia rural la probabilidad de uso es de 0.5433827 y para el lugar de residencia urbano es de 0.7253437.

Las condiciones de aseguramiento suponen una fuerte relación con el uso de métodos anticonceptivos. Ya que una mujer que tiene algún sistema de aseguramiento puede recibir algunas pláticas de planificación familiar, además de poder adquirir algún método anticonceptivo sin costo alguno, es decir, no se presentan barreras de accesibilidad a éstos. Lo anterior supone una relación positiva entre tener algún sistema de aseguramiento y utilizar algún método anticonceptivo. Lo cual se ve

reflejado en las probabilidades de uso para las categorías de aseguramiento. Ya que la probabilidad de uso para una mujer que no tiene aseguramiento es de 0.6157424 contra 0.7618631 para una mujer que tiene algún esquema de aseguramiento.

La mujer dentro de una relación conyugal, ya sea el matrimonio o unión libre es considerada la responsable de la planificación familiar, además del hecho que los métodos anticonceptivos modernos están hechos para las mujeres. Lo anterior es la justificación de haber utilizado variables relacionadas con las características de la mujer, como son edad, educación, tipo de relación conyugal, condición de trabajo, entre otras, las cuales se supone que son capaces de predecir la posible utilización.

Las variables incluidas en el modelo son consideradas posibles determinantes que afectan la probabilidad de uso de servicios de planificación familiar. Aunque no se descarta la posibilidad de que algunas de estas variables estén mal reportadas o se tengan problemas al codificar la información en el centro de recolección de la misma. Lo anterior se podría reflejar en la variable de educación, la cual presenta para los estados de alta marginación, una mayor tasa de mujeres que no responden o no saben el grado de educación que poseen, con lo anterior se limita la posibilidad de realizar inferencias más concretas para dichos estados.

Las consecuencias de la utilización de métodos anticonceptivos traen consigo beneficios para las mujeres, como por ejemplo; la mayoría de las mujeres están convencidas de que la practica de la planificación familiar y tener familias menos numerosas proporcionan beneficios económicos y de salud; la planificación familiar puede liberar a las parejas del temor de que ocurra un embarazo no planificado y pueda mejorar la vida sexual de las parejas, así como la relación de éstas; cuando hay

posibilidades de trabajar, las usuarias de la planificación familiar suelen tener más probabilidades de aprovechar las oportunidades de trabajo que las no usuarias.[18]

Con lo anterior se ha mostrado una aplicación de los modelos lineales generalizados, utilizando una distribución y una función liga específica. Sin embargo, este tipo de modelos tienen una gran cantidad de ligas y funciones que se pueden asociar a una gran variedad de conjunto de datos en diferentes situaciones, lo que hace que estos modelos sean importantes en la resolución de diferentes problemas.

[18] Según artículo de Family Health Internacional de Nancy Williamson: Como Influye el uso de la planificación familiar en la vida de las mujeres, 1998.

APÉNDICE A

Se tiene la función de log - verosimilitud dada por

$$l(\theta; y) = \left\{ \sum y_i b(\theta_i) + \sum c(\theta_i) + \sum d(y_i) \right\} \tag{A.1}$$

P.D. $E(Y_i) = \mu_i = \frac{-c'(\theta_i)}{b'(\theta_i)}$ y $Var(Y_i) = \frac{b''(\theta_i)c'(\theta_i) - b'(\theta_i)c''(\theta_i))}{[b'(\theta_i)]^3}$

Demostración.

Derivando (A.1) se obtiene

$$U_i = \frac{\partial l_i}{\partial \theta} = y_i b'(\theta_i) + c'(\theta_i) \tag{A.2}$$

Y dado que $\frac{\partial l_i}{\partial \theta} = 0$ por el apéndice B y calculando la esperanza a y_i se tiene

$$E(y_i) = \frac{-c'(\theta_i)}{b'(\theta_i)} = \mu_i \tag{A.3}$$

Por definición la varianza de Y_i aplicada a (A.2), i.e. $Var(cy) = c^2 Var(y)$

$$Var(U) = (b'(\theta))^2 Var(Y_i) \tag{A.4}$$

$Var(U) = E(U^2) - E^2(U)$

Pero por el apéndice B se tiene que $E(U) = 0$ y $E(-U') = E(U^2)$

Por lo que

$$Var(U) = E(-U') \tag{A.5}$$

Por lo tanto derivando la ecuación (A.2) se tiene

$U_j' = \frac{\partial^2 l_i}{\partial \theta} = y_i b''(\theta_i) + c''(\theta_i)$ tomando la esperanza sobre y_i se tiene

$$E(-U_j') = -E(y_i)b''(\theta_i) - c''(\theta_i) \tag{A.6}$$

Y sustituyendo (A.3) en (A.6) se tiene

$$\begin{aligned} E(-U_j') &= \frac{c'(\theta_i)}{b'(\theta)} b''(\theta_i) - c''(\theta_i) \\ &= \frac{c'(\theta_i)b''(\theta_i) - c''(\theta_i)b'(\theta)}{b'(\theta)} \end{aligned} \tag{A.7}$$

Por lo tanto despejando $Var(Y_i)$ en (A.4) y aplicando (B.6) del apéndice B en

$$\begin{aligned} Var(U_j') &= E(-U_j') \\ &= (b'(\theta))^2 Var(Y_i) \end{aligned}$$

Por lo que

$$Var(Y_i) = \frac{E(-U_j')}{(b'(\theta))^2} \tag{A.8}$$

Sustituyendo (A.7) en (A.8) se tiene

$$Var(Y_i) = \frac{-c'(\theta_i)b''(\theta_i) + c''(\theta_i)b'(\theta)}{(b'(\theta))^3} \tag{A.9}$$

APÉNDICE B

Caso Univariado

1) P.D. $E(U) = 0$

Considérese una función variable aleatoria continua Y con función de densidad f(y;θ) que depende de un solo parámetro. La función de log-verosimil es el logaritmo de $f(y;\theta)$ considerada como una función de θ, es decir

$$l(\theta;y) = \log f(y;\theta)$$

Se tiene que la primera derivada de la función *l* es llamada la función puntaje(o score), la cual está dada por

$$U = \frac{dl}{d\theta} \tag{B.1}$$

Para encontrar la derivada se tiene que

$$\frac{d\log f(y;\theta)}{d\theta} = \frac{1}{f(y;\theta)}\frac{df(y;\theta)}{d\theta} \tag{B.2}$$

Calculando la esperanza a (B.2) se tiene y aplicando $E[g(y)]=\int g(y)f_y(y;\theta)dy$ se obtiene

$$\begin{aligned} E(U) &= \int \frac{d\log f(y;\theta)}{d\theta} f(y;\theta)\, dy \\ &= \int \frac{1}{\cancel{f(y;\theta)}}\frac{df(y;\theta)}{d\theta}\cancel{f(y;\theta)}\, dy \\ &= \int \frac{df(y;\theta)}{d\theta} dy \end{aligned}$$

Bajo condiciones de regularidad la ecuación anterior queda como

$$\int \frac{df(y;\theta)}{d\theta} dy = \frac{d}{d\theta}\int f(y;\theta) dy$$

Pero por definición $\int f(y;\theta) = 1$ quedando así

$$\frac{d}{d\theta}1 = 0$$

Por lo que $E(U) = 0$ (B.3)

2) P.D. $E(U') = E(U^2)$

Derivando la ecuación (B.2) y tomando esperanza, además de intercambiar los operadores bajo condiciones de regularidad se tiene:

$$\frac{d}{d\theta}\int \frac{d\log f(y;\theta)}{d\theta} f(y;\theta)dy = \frac{d^2}{d\theta}\int f(y;\theta)dy \qquad \text{(B.4)}$$

Igual que en el caso anterior el lado derecho de la ecuación es cero dado que

$$\int f(y;\theta) = 1$$

Es decir $\frac{d}{d\theta}\int \frac{d\log f(y;\theta)}{d\theta} f(y;\theta)dy = 0$ (B.5)

Ahora utilizando que

$(uv)' = u'v + v'u$ donde $u = \frac{d\log f(y;\theta)}{d\theta}$ y $v = f(y;\theta)$,

el lado derecho de la ecuación queda como:

$$\int \frac{d^2 \log f(y;\theta)}{d\theta^2} f(y;\theta)dy + \int \frac{d \log f(y;\theta)}{d\theta}\frac{df(y;\theta)}{d\theta}dy$$

Despejando de (B.2) a $\frac{df(y;\theta)}{d\theta}$ y sustituyendo se obtiene

$$\int \frac{d^2 \log f(y;\theta)}{d\theta^2} f(y;\theta)dy + \int \left[\frac{d \log f(y;\theta)}{d\theta}\right]^2 f(y;\theta)dy = 0$$

Pero lo anterior es equivalente

$$E\left[-\frac{d^2 \log f(y;\theta)}{d\theta^2}\right] = E\left[\left[\frac{d \log f(y;\theta)}{d\theta}\right]^2\right]$$

Renombrando ésta ecuación en términos de estadística puntaje (o score) queda como:

$$E(-U') = E(U^2) \qquad \text{(B.6)}$$

Donde U' es la derivada de U con respecto a θ.

Dado que la varianza esta dada por

$Var(U) = E(U^2) - (E(U))^2$

Y dado que E(U)=0, entonces

Var(U)=E(U^2) y por (B.6)

=E(-U')

Caso Multivariado

3) P.D. $E(U_i) = 0 \qquad \forall\ i$

En forma general, si se tienen n variables aleatorias independientes Y_1 , Y_2 , ..., Y_n, las cuales dependen de p parámetros, $\theta_1, \theta_2, \ldots, \theta_p$ donde $p \leq n$. Sea $l_i(\theta_i;y_i)$ que denota a la función logaritmo de la verosimilitud que depende del vector $\underline{\theta}$ con p parámetros. Entonces la función log-verosímil de Y_1 , Y_2 , ..., Y_n es

$l(\theta;y) = \sum_{i=1}^{n} l_i(\theta;y_i)$ donde $y=[Y_1,\ldots,Y_n]^T$.

El puntaje (score) total con respecto a θ es definido como

$$U_j = \frac{\partial l(\theta; y)}{\partial \theta_j} = \sum_{i=1}^{n} \frac{\partial l_i(\theta; y_i)}{\partial \theta_j} \qquad \text{(B.7)}$$

Por el argumento de (B.3) aplicado a cada uno de los elementos de la suma se tiene

$$E(U_j) = E\left[\sum_{i=1}^{n} \frac{\partial l_i(\theta; y_i)}{\partial \theta_j}\right] \qquad \text{para toda j}$$
$$= 0$$

4) P.D. $E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] = E\left[-\frac{\partial^2 l}{\partial \theta_j \partial \theta_k}\right]$

Demostración.
La matriz de información es definida como la matriz de varianzas y covarianzas de las U_j's, $\wp = E[UU^T]$ donde $U = [U_1, \ldots, U_p]$, asi que sus elementos están dados por

$$\wp_{jk} = E[U_j U_k] = E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] \qquad \text{(B.8)}$$

$E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] = E\left[-\frac{\partial^2 l}{\partial \theta_j \partial \theta_k}\right]$ Por lo que los elementos de la matriz de información están dados por

$$\wp_{jk} = E\left[-\frac{\partial^2 l}{\partial \theta_j \partial \theta_k}\right] \qquad \text{(B.9)}$$

Por lo que el objetivo es demostrar que

$$E\left[\frac{\partial l}{\partial \theta_j}\frac{\partial l}{\partial \theta_k}\right] = E\left[-\frac{\partial^2 l}{\partial \theta_j \partial \theta_k}\right]$$

Aplicando las condiciones de regularidad a la ecuación (B.4) para el caso multivariado, se tiene

$$\begin{aligned} U_j &= \frac{\partial l(\underline{\theta}; \underline{y})}{\partial \theta_j} \\ &= \frac{\partial \log f(\underline{y}; \underline{\theta})}{\partial \theta_j} = \sum_{i=1}^{n} \frac{\partial l_i(\underline{\theta}; \underline{y})}{\partial \theta_j} \qquad \text{(B.10)} \\ &= \sum_{i=1}^{n} \frac{1}{f(y_i; \underline{\theta})} \frac{\partial f(y_i; \underline{\theta})}{\partial \theta_j} \end{aligned}$$

$$\frac{d}{d\theta_k}E(U_j) = \frac{d}{d\theta_k}\int\cdots\int\left[\sum_{i=1}^{n}\frac{1}{f(y_i;\underline{\theta})}\frac{\partial f(y_i;\underline{\theta})}{\partial\theta_j}\right]f(y_1;\underline{\theta})f(y_2;\underline{\theta})\cdots f(y_n;\underline{\theta})dy_1\cdots dy_n$$

Desarrollando la suma

$$= \frac{d}{d\theta_k}\int\cdots\int\left\{\left[\frac{1}{\cancel{f(y_1;\underline{\theta})}}\frac{\partial f(y_1;\underline{\theta})}{\partial\theta_j}\cancel{f(y_1;\underline{\theta})}f(y_2;\underline{\theta})\cdots f(y_n;\underline{\theta})\right]+\cdots+\right.$$

$$\left.+\cdots+\left[\frac{1}{\cancel{f(y_n;\underline{\theta})}}\frac{\partial f(y_n;\underline{\theta})}{\partial\theta_j}f(y_1;\underline{\theta})f(y_2;\underline{\theta})\cdots\cancel{f(y_n;\underline{\theta})}\right]\right\}dy_1\cdots dy_n$$

$$= \frac{d}{d\theta_k}\left\{\int\cdots\int\left[\frac{\partial f(y_1;\underline{\theta})}{\partial\theta_j}f(y_2;\underline{\theta})\cdots f(y_n;\underline{\theta})dy_1\cdots dy_n\right]+\cdots+\right.$$

$$\left.+\cdots+\int\cdots\int\left[\frac{\partial f(y_n;\underline{\theta})}{\partial\theta_j}f(y_1;\underline{\theta})f(y_2;\underline{\theta})\cdots f(y_{n-1};\underline{\theta})dy_1\cdots dy_n\right]\right\}$$

Por las condiciones de regularidad se puede intercambiar el operador derivada por integral, por lo tanto se tiene

$$= \frac{d}{d\theta_k}\left\{\frac{\partial}{\partial\theta_j}\int\cdots\int[f(y_1;\underline{\theta})f(y_2;\underline{\theta})\cdots f(y_n;\underline{\theta})dy_1\cdots dy_n]+\cdots+\right.$$

$$\left.+\cdots+\frac{\partial}{\partial\theta_j}\int\cdots\int[f(y_1;\underline{\theta})f(y_2;\underline{\theta})\cdots f(y_{n-1};\underline{\theta})f(y_n;\underline{\theta})dy_1\cdots dy_n]\right\}$$

Pero por ser variables aleatorias cumplen que la suma sobre el intervalo es 1, por lo que la ecuacion anterior queda como

$$= \frac{d}{d\theta_k}\left\{\frac{\partial}{\partial\theta_j}(1)+\cdots+\frac{\partial}{\partial\theta_j}(1)\right\}$$

$$= \frac{d}{d\theta_k}\frac{\partial}{\partial\theta_j}(n)$$

$$= 0$$

Entonces se tiene

$$\frac{\partial E(U_j)}{\partial \theta_k} = \frac{\partial}{\partial \theta_k} \int \cdots \int \frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} f(\underline{y};\underline{\theta}) d\underline{y} = 0$$

Aplicando las condiciones de regularidad se tiene

$$\frac{\partial}{\partial \theta_k} \int \cdots \int \frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} f(\underline{y};\underline{\theta}) d\underline{y} = \int \cdots \int \frac{\partial}{\partial \theta_k} \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} f(\underline{y};\underline{\theta}) \right] d\underline{y} = 0$$

Aplicando que

$$(uv)' = u'v + v'u \text{ donde } u = \frac{d \log f(\underline{y};\underline{\theta})}{d\underline{\theta}} \quad \text{y} \quad v = f(\underline{y};\underline{\theta})$$

al segundo miembro de la igualdad se tiene

$$\int \cdots \int \frac{\partial}{\partial \theta_k} \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} f(\underline{y};\underline{\theta}) \right] d\underline{y} = \int \cdots \int \left\{ \frac{\partial^2 \log f(\underline{y};\underline{\theta})}{\partial \theta_k \partial \theta_j} \right\} f(\underline{y};\underline{\theta}) d\underline{y} + \int \cdots \int \left\{ \frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} \right\} \frac{\partial f(\underline{y};\underline{\theta})}{\partial \theta_k} d\underline{y} \tag{B.11}$$

en la ecuación (B.2) se tiene

$$\frac{d \log f(y;\theta)}{d\theta} = \frac{1}{f(y;\theta)} \frac{df(y;\theta)}{d\theta}$$

Por lo tanto despejando se tiene

$$\frac{df(y;\theta)}{d\theta} = \frac{d \log f(y;\theta)}{d\theta} f(y;\theta)$$

sustituyendo en (B.11)

$$\int \cdots \int \left\{ \frac{\partial^2 \log f(\underline{y};\underline{\theta})}{\partial \theta_k \partial \theta_j} \right\} f(\underline{y};\underline{\theta}) d\underline{y} + \int \cdots \int \left\{ \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} \right] \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_k} \right] \right\} f(\underline{y};\underline{\theta}) d\underline{y} = 0$$

Despejando el segundo elemento de la igualdad

$$\int \cdots \int \left\{ \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_j} \right] \left[\frac{\partial \log f(\underline{y};\underline{\theta})}{\partial \theta_k} \right] \right\} f(\underline{y};\underline{\theta}) d\underline{y} = -\int \cdots \int \left\{ \frac{\partial^2 \log f(\underline{y};\underline{\theta})}{\partial \theta_k \partial \theta_j} \right\} f(\underline{y};\underline{\theta}) d\underline{y}$$

Renombrando en términos de esperanzas se tiene

$$E\left[\frac{\partial l}{\partial \theta_j} \frac{\partial l}{\partial \theta_k} \right] = E\left[-\frac{\partial^2 l}{\partial \theta_j \partial \theta_k} \right] \qquad \text{q.e.d.}$$

APÉNDICE C

Los resultados presentados corresponden al modelo lineal generalizado con distribución Binomial y liga probit. Estos resultados no tienen el factor de expansión, esto se hizó con el fin de obtener significancia estadística.

Para este caso se tiene un total de 15,281 grupos.

```
Iteration 1 : deviance = -2951.3082
Iteration 2 : deviance = -2940.6951
Iteration 3 : deviance = -2939.5774
Iteration 4 : deviance = -2939.5733
Iteration 5 : deviance = -2939.5727
Iteration 6 : deviance = -2939.5727
Iteration 7 : deviance = -2939.5727

Generalized linear models                 No. of obs      =      15281
Optimization     : MQL Fisher scoring     Residual df     =      15272
(IRLS EIM)                                Scale param     =          1
Deviance         =  2939.572742           (1/df) Deviance =  0.1924812
Pearson          =  19894.41345           (1/df) Pearson  =   1.302672

Variance function: V(u) = u*(1-u/ni)      [Binomial]
Link function    : g(u) = invnorm(u/ni)   [Probit]
Standard errors  : EIM

BIC              =  2852.863452
```

sum_usa	Coef.	EIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
educ1	-0.0231434	0.0027083	-8.55	0	-0.0284516	-0.0178351
num_mel	0.1458792	0.0031028	47.02	0	0.1397978	0.1519605
gedad	0.0219693	0.0042544	5.16	0	0.0136308	0.0303078
edo_cvl	0.0080308	0.0027348	2.94	0.003	0.0026707	0.0133908
trab	0.0484208	0.0127739	3.79	0	0.0233845	0.0734572
n_hijos	0.0113787	0.0009548	11.92	0	0.0095073	0.01325
aseg	0.1949573	0.013425	14.52	0	0.1686449	0.2212698
ru	0.198793	0.0142336	13.97	0	0.1708957	0.2266902
_cons	-0.9831725	0.0287423	-34.21	0	-1.039506	-0.9268387

En donde se obtienen resultados muy buenos, la prueba de bondad de ajuste resulta ser positivo, es decir, el valor de la devianza es menor que el de una Ji-Cuadrada con N-p grados de libertad, teniéndose una Devianza de 2939.572742 contra el valor de 15560.601 que corresponde a una Ji-Cuadrada con 15,272 grados de libertad. Por lo que se acepta no se rechaza que el modelo con p-parámetros ajusta bien a los datos. Asimismo en cada una de las estimaciones de los parámetros se rechaza que sean cero en forma individual.

APÉNDICE D

P.D. $$E[U_j U_k] = E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k}\right] \quad \text{(D.1)}$$

Tomando la ecuación $U_j = \sum_{i=1}^{n} \frac{\partial l_i(\theta; y_i)}{\partial \theta_j}$ y desarrollando el lado izquierdo de la ecuación (D.1), se tiene

$$E[U_j U_k] = E\left[\left(\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\right)\left(\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_k}\right)\right] \quad \text{(D.2)}$$

Utilizando que

$$\left[\sum_{i=1}^{n} a_i\right]^2 = \left[\sum_{i=1}^{n} a_i^2\right] + \left[\sum_{i\neq j} a_i a_j\right] \quad \text{(D.3)}$$

Por lo que aplicando (D.3) a (D.2) se obtiene

$$E[U_j U_k] = E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k} + \sum_{i\neq l}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_l}{\partial \theta_k}\right]$$

$$= E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k}\right] + E\left[\sum_{i\neq l}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_l}{\partial \theta_k}\right]$$

Pero como las Y_i's son independientes se tiene $E(Y_j Y_k) = E(Y_j)E(Y_k)$

$$E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k}\right] + E\left[\sum_{i\neq l}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_l}{\partial \theta_k}\right] = E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k}\right] + \left[\sum_{i\neq l}^{n} E\left(\frac{\partial l_i}{\partial \theta_j}\right) E\left(\frac{\partial l_l}{\partial \theta_k}\right)\right]$$

Además por el Apéndice A se tiene que $E(U_i) = 0 \quad \forall\, i$, cancelándose por lo tanto el segundo miembro del lado derecho de la igualdad. Por lo que la ecuación final queda de la siguiente manera

$$E[U_j U_k] = E\left[\sum_{i=1}^{n} \frac{\partial l_i}{\partial \theta_j}\frac{\partial l_i}{\partial \theta_k}\right] \quad \text{q.e.d.}$$

BIBLIOGRAFÍA

McCullagh P, Nelder J.A., Generalized Linear Models 2da. Ed., Londres, Reino Unido, Chapman and Hall, 1989

Hardin J., Hilbe J., Generalized Linear Models and Extensions, Texas, EUA, STATA PRESS, 2001.

Dobson Annette, An Introduction to Generalized Linear Models, Londres, Reino Unido,Chapman & Hall, 1990.

Liao Tim Futing, Interpreting probability models; Logit, Probit and other Generalized Linear Models, EUA, SAGE Publications, 1994.

Hosmer D., Lemeshow S., Applied Logistic Regression, John Wiley & Sons, 1989.

Everitt B.S., The Analysis of Contingency Tables 2da. Ed., Londres, Reino Unido, Chapman and Hall, 1992.

Agresti Alan, Categorical Data Analysis, Londres, Reino Unido, John Wiley & Sons, 2000.

Ross Sheldon, A first course in probability 3a. Ed.. Singapur, 1989.

STATA Reference Manual Release 7, volume 1-4, Texas , EUA, STATA PRESS 2001

Swokowski E., Calculus with Analytic geometric, 2nd edition, Prindle, Weber & Schmidt, Boston, USA, 1979.

Artículos:

1. **Beltrán Arlette, Informe sobre investigaciones aplicadas secundarias no. 9;Utilización de los servicios de planificación familiar : El caso Peruano. 1999.**
2. **Williamson Nancy, Introducción: Como influye el uso de la planificación familiar en la vida de las mujeres, Family Health Internacional, 1998.**

Sitios de internet

1. **http://data.princeton.edu/wws509**
2. **http://www.stat.sfu.ca/~lockhart/richard /350/97_1/examples/lectures33/lecture33.html**
3. **http://www.stat.sfu.ca/~lockhart/richard/350/99_1/lectures/35/web.html**
4. **http://sepwww.stanford.edu/public/docs/sep61/gilles/paper_html/node4.html**
5. **http://opr.princeton.edu/archive/pcp/cr81.asp**
6. **http://cddhcu.gob.mx/ca,dip/comlvii/compyd/pnp02.htm**
7. **http://www.fhi.org/en/wsp/wspubs/concept.html**
8. **http://www.fhi.org/sp/networks/sv18-4/ns1841.html**
9. **http://members.tripod/Smilox/slsnewton-raphson.html**
10. **http://imf.kvl.dk/biomodel/biomodeller-k-97/notes/afsnit11/ afsnit11.html**
11. **www.conapo.org.mx**