UNIVERSIDAD DE BARCELONA
DIVISION DE CIENCIAS JURIDICAS ECONOMICAS Y SOCIALES
FACULTAD DE CIENCIAS ECONOMICAS Y EMPRESARIALES
DEPARTAMENTO DE MATEMATICA ECONOMICA, FINANCIERA Y ACTUARIAL

LA TEORIA DE LA CREDIBILIDAD

Y SU APLICACION A LOS SEGUROS COLECTIVOS

Tesis Doctoral presentada por:
Mª Angeles Pons Cardell
Director:
Dr. D. Antonio Alegre Escolano
Catedrático de Universidad.

Barcelona, Noviembre 1991.

CAPITULO III

MODELOS DE REGRESION

INDICE DEL CAPITULO

Desde la aparición de los modelos que hemos denominado clásicos la Teoría de la Credibilidad se ha bifurcado, según nuestro punto de vista, en tres direcciones, una de las cuales es la que a continuación analizaremos, y que hemos denominado la línea de los MODELOS DE REGRESION, ya que en ellos la utilización de la técnica de regresión es una de las notas dominantes, junto con el completo abandono de la hipótesis de homogeneidad en el tiempo, tanto en las observaciones esperadas como en la matriz de covarianzas.

Esta vertiente fue iniciada por HACHEMEISTER, C. (1975) con su Modelo de Regresión, que es una generalización del Modelo de BÜHLMANN-STRAUB, y seguida por De VYLDER, F. (1985) con su Modelo de Regresión No-Lineal, que es una generalización del Modelo de Regresión de HACHEMEISTER.

En el **Modelo de Regresión de HACHEMEISTER** se asume que las observaciones esperadas se ajustan a una función polinómica, mientras que en el **Modelo de Regresión No-Lineal de De VYLDER** la hipótesis de homogeneidad en el tiempo no es reemplazada por una necesariamente de tipo polinómica, sino por otra mucho más general, siendo el Modelo de Regresión Exponencial uno de los más utilizados.

El objetivo de ambos modelos sigue siendo estimar de la mejor manera posible las primas de riesgo individuales, pero el procedimiento utilizado en ambos es distinto. HACHEMEISTER, C. (1975) sigue utilizando el procedimiento de los mínimos cuadrados para hallar la mejor aproximación lineal para las primas de riesgo individuales, mientras que De VYLDER, F. (1985) utiliza el instrumental de las distancias. Lo que sí cabe destacar, es que los factores de credibilidad dejan de ser escalares para convertirse en matrices, denominadas **matrices de credibilidad.**

A diferencia de los modelos clásicos, estos modelos son aptos para detectar tendencias y tener en cuenta los efectos de la inflación.

3.1.- MODELO DE REGRESION DE HACHEMEISTER

El Modelo de Regresión de HACHEMEISTER, que es una generalización del Modelo de BÜHLMANN-STRAUB, fue presentado por primera vez en 1974, en una conferencia sobre credibilidad titulada "Actuarial Research Conference on Credibility" que tuvo lugar en la Universidad de Berkeley, California.

HACHEMEISTER, C. (1975) planteó la necesidad de establecer estimaciones para las tendencias que presentaban los datos de los montantes de las reclamaciones de los seguros de automóviles por estados, debido a los efectos de la inflación, ya que en su opinión la inflación se había convertido en un elemento importante a tener en cuenta en las tarificaciones. Disponiendo de los datos sobre los montantes de las reclamaciones por estados y por periodos, propuso utilizar un estinmador que fuese una ponderación entre el estimador individual para cada estado y el estimador para el conjunto del país, utilizando para ello un modelo de regresión lineal.

HACHEMEISTER, C. (1975) no formuló explícitamente las hipótesis de su modelo, aunque sí las dejó entrever, pero desarrolló el procedimiento para obtener los estimadores ajustados de credibilidad para las primas de riesgo individuales, siendo en realidad NORBERG, R. (1979), pp. 208, quien formuló propiamente las hipótesis del modelo, y se ocupó también del problema de la estimación de los parámetros estructurales.

El Modelo de Regresión de HACHEMEISTER es el primero de los dos modelos de regresión que apareció y está fuertemente inspirado en el Modelo de BÜHLMANN-STRAUB, en el cual, aunque las observaciones esperadas son homogéneas en el tiempo, ya se permite la variación en el tiempo de la $Var[X_{js}/\theta_j]$. En el Modelo de HACHEMEISTER no sólo se permite esta variación sino también en la $E[X_j/\theta_j]$, donde la hipótesis de homogeneidad en el tiempo es sustituida por otra de tipo polinómico, pudiéndose de este modo tener en cuenta los efectos de la inflación o trazar tendencias. Otra novedad del modelo es que incorpora la técnica de regresión para la obtención de los estimadores de credibilidad.

3.1.1.- Variables relevantes.

Las variables más relevantes de este modelo son:

- θ_j : Es el parámetro de riesgo, con $j = 1,2,...,k$.

- X_{js}: Es la variable aleatoria que nos indica la experiencia de reclamaciones, con $j = 1,2,...,k$ y $s = 1,2,...,t_j$, siendo k el número total de pólizas que integran la cartera y t_j el número de periodos observados para la póliza j-ésima.

Ambas variables ya han sido utilizadas en los dos modelos descritos anteriormente, pero en este caso se define además para cada contrato j:

- Y_j : Matriz dada de dimensión (t_j,n), de rango pleno n, que debe verificar $n < t_j$. La n indica el grado, más una unidad, de la mayor tendencia polinómica prevista para las distintas pólizas que integran la cartera.

- v_j : Matriz dada, en este caso de dimensión (t_j,t_j), y semidefinida positiva, con $j = 1,2,\dots,k$.

3.1.2.- Hipótesis del modelo.

Las hipótesis que se asumen en este modelo son:

H.1) Las pólizas $j = 1,2,\dots,k$, esto es, los pares $(\theta_1, \vec{X}_1)$, $(\theta_2, \vec{X}_2),\dots,(\theta_k, \vec{X}_k)$ son independientes, y las variables $\theta_1,\theta_2,\dots,\theta_k$ son independientes y están idénticamente distribuidas.

Como en los dos modelos anteriores, se sigue asumiendo la independencia entre las pólizas, y la equivalencia exterior de las mismas.

H.2) Para todo $j = 1,2,\dots,k$ y todo $r,s = 1,2,\dots,t_j$ se tiene:

a) $E[X_j/\theta_j] = Y_j \cdot \beta(\theta_j)$ $\qquad (t_j,1)$

donde $\beta(\theta_j)$ es un vector de regresión desconocido de dimensión $(n,1)$, Y_j es una matriz dada de dimensión (t_j,n), y X_j es un vector columna:

$$X_j = \begin{bmatrix} X_{j1} \\ X_{j2} \\ \vdots \\ X_{jt} \end{bmatrix}$$

La introducción de las matrices Y_j reemplaza la rígida asunción de homogeneidad en el tiempo, asumida hasta ahora, por otra de tipo polinómica, y en particular, cuando $n = 2$ por otra de tipo lineal, siendo posible por tanto en este modelo trazar tendencias, o tener en cuenta los efectos de la inflación. En el caso particular que $n = 2$ las matrices Y_j adoptan la siguiente forma:

$$Y_j = \begin{bmatrix} 1 & t_j \\ 1 & t_j - 1 \\ 1 & t_j - 2 \\ \vdots & \\ 1 & 1 \end{bmatrix} \quad \text{o alternativamente} \quad Y_j = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \\ 1 & t_j \end{bmatrix}$$

En general, utilizaremos la primera por ventajas desde el punto de vista de la predicción, ya que el próximo periodo sería $s = \emptyset$ con lo que del modelo polinómico o lineal, sólamente será significativo el término independiente o punto de corte.

Esta primera hipótesis constituye una ampliación de la considerada en el Modelo de BÜHLMANN-STRAUB que establece:

$$E[X_{js}/\theta_j] = \mu(\theta_j) \qquad s \in \{1,2,\dots,t_j\}$$

esto es, que las observaciones esperadas son homogéneas en el tiempo.

b) $\mathrm{Cov}[X_j/\theta_j] = \sigma^2(\theta_j)\cdot v_j \qquad (t_j,t_j)$

donde $\sigma^2(\theta_j)$ es una función escalar desconocida, y v_j con $j = 1,2,\dots,k$, son matrices dadas, semidefinidas positivas de dimensión (t_j,t_j), siendo $[\sigma^2(\theta_j)\cdot v_j]$ la matriz de varianzas y covarianzas correspodiente a la póliza j-ésima.

Si consideramos la hipótesis de covarianzas nulas y varianzas constantes en el tiempo, v_j sería la matriz identidad $I_{(t_j,t_j)}$. Otra hipótesis menos restrictiva, sería considerar nulas las covarianzas, pero las varianzas variables en el tiempo, siendo en este caso v_j una matriz diagonal, con todos sus elementos estrictamente positivos, siendo este el caso asumido en el Modelo de BÜHLMANN-STRAUB. En general, la condición que deberá verificar v_j es que sea una matriz semidefinida positiva, admitiéndose covarianzas significativas entre instantes temporales distintos.

En principio, esta hipótesis se diferencia de la de BÜHLMANN-STRAUB en la posibilidad de que existan covarianzas no nulas, pero la dispersión, que viene medida por $\sigma^2(\theta_j)$, no se hace variable en el tiempo.

NORBERG, R. (1979) cuando formuló por primera vez las hipótesis del Modelo de Regresión de HACHEMEISTER, no contempló la posibilidad de que existieran covarianzas no nulas, sino que estableció que v_j, con $j = 1,2,\ldots,k$, son matrices dadas, conocidas y diagonales. Posteriormente, el propio NORBERG, R. (1982) amplió esta hipótesis y estableció que v_j son matrices conocidas no aleatorias y semidefinidas positivas.

3.1.3.- Estimación de la prima de riesgo individual.

El objetivo de este modelo sigue siendo estimar, de la mejor manera posible, las primas de riesgo individuales.

HACHEMEISTER, C. (1975) señaló que existían dos alternativas para estimar dichas primas:

- Una es a través del análisis de las series temporales,
- y la otra, a través de un modelo de regresión.

En la primera, se pone el énfasis en la interdependencia de X_{js}/θ_j para varias j, mientras que en la segunda la esperanza condicionada, $E[X_{js}/\theta_j]$, se considera una combinación lineal de otras variables observadas.

Estas dos alternativas no son totalmente independientes, ya que es posible crear un modelo que contenga a ambos elementos, la interdependencia de las X_{js}/θ_j por un lado, y la esperanza condicionada como combinación lineal de otras variables observadas, por otro.

De las dos posibilidades que propuso para estimar las primas de riesgo individuales HACHEMEISTER, C. eligió la segunda, la utilización de un modelo de regresión, en el cual el problema fundamental que se presenta es principalmente de orden práctico, ya que se debe estimar la matriz de covarianza de X_{js} para diferentes s, y al mismo tiempo, estimar los coeficientes de regresión.

El planteamiento de HACHEMEISTER, C. (1975) fue el siguiente:

Estamos interesados en estimar la prima de riesgo individual para la póliza j-ésima en el periodo s-ésimo:

$$E[X_{js}/\theta_j] = Y'_{js}\cdot\beta(\theta_j) = \mu_{js}(\theta_j)$$

donde Y'_{js} el vector fila s-ésimo de la matriz Y_j anteriormente definida, siendo su dimensión (1,n).

Vamos a aceptar que el estimador $\hat{\mu}^*_{js}(\theta_j)$ es el estimador óptimo si verifica, de forma análoga a lo efectuado en el Modelo de BÜHLMANN, que:

$$E[[\mu_{js}(\theta_j) - \hat{\mu}^*_{js}(\theta_j)]^2] \leq E[[\mu_{js}(\theta_j) - \hat{\mu}_{js}(\theta_j)]^2]$$

para todo posible estimador $\hat{\mu}_{js}(\theta_j)$.

Siguiendo a BÜHLMANN, H. y STRAUB, E., consideró estimadores de tipo lineal:

$$\hat{\mu}_{js}(\theta_j) = \alpha_0 + \sum_{i=1}^{k} \sum_{r=1}^{t_i} \alpha_{ir} \cdot X_{ir} = \alpha_0 + X' \cdot A$$

donde X es un vector columna particionado de dimensión $(k \cdot t_i, 1)$:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_k \end{bmatrix} \quad \text{con} \quad X_i = \begin{bmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{ir} \\ \vdots \\ X_{it_i} \end{bmatrix}$$

que es un vector columna de dimensión $(t_i, 1)$, y A es también un vector columna particionado de dimensión $(k \cdot t_i, 1)$ definido como:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_k \end{bmatrix} \quad \text{con} \quad A_i = \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ir} \\ \vdots \\ \alpha_{it_i} \end{bmatrix}$$

que también es un vector columna de dimensión $(t_i, 1)$.

Si calculamos:

$$X' \cdot A = \begin{bmatrix} X'_1 \; X'_2 \cdots X'_i \cdots X'_k \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_k \end{bmatrix} =$$

$$= \sum_{i=1}^{k} X'_i \cdot A_i = \sum_{i=1}^{k} \begin{bmatrix} X'_{i1} \; X'_{i2} \cdots X'_{ir} \cdots X'_{it_i} \end{bmatrix} \cdot \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ir} \\ \vdots \\ \alpha_{it_i} \end{bmatrix} =$$

$$= \sum_{i=1}^{k} \sum_{r=1}^{t_i} \alpha_{ir}^{(js)} \cdot X_{ir}$$

La mejor aproximación lineal para la prima de riesgo individual la obtendremos hallando los valores de $\alpha_\emptyset$ y de α_{ir} con $i = 1,2,\ldots,k$ y $r = 1,2,\ldots,t_i$ tales que minimicen la siguiente esperanza total:

$$\phi_{js} = E_{Total}\left[\left[\mu_{js}(\theta_j) - \alpha_\emptyset - X' \cdot A\right]^2\right]$$

en la que aparecen $(k \cdot t_i + 1)$ parámetros.

Para facilitar la exposición del proceso de minimización vamos a introducir los siguientes elementos:

•) $\beta = E[\beta(\theta_j)]$ que es un vector de dimensión $(n,1)$.

•) La matriz de covarianza de los coeficientes de regresión, que se simboliza por Γ, es una matriz cuadrada de dimensión (n,n):

$$\Gamma = Cov[\beta(\theta_j),\ \beta'(\theta_j)] = E[\beta(\theta_j) \cdot \beta'(\theta_j)] - \beta \cdot \beta'$$

•) Hemos definido $\mu_{js}(\theta_j) = E[X_{js}/\theta_j] = Y'_{js}\cdot\beta(\theta_j)$ que no es más que la esperanza condicionada para la póliza j-ésima.

Si calculamos la esperanza de $\mu_{js}(\theta_j)$ respecto a los distintos parámetros de riesgo θ, tenemos:

$$E[\mu_{js}(\theta_j)] = E[Y'_{js}\cdot\beta(\theta_j)] = Y'_{js}\cdot E[\beta(\theta_j)] = Y'_{js}\cdot\beta$$

de manera que:

$$E_{Total}[X_{js}] = E\big[E[X_{js}/\theta_j]\big] = Y'_{js}\cdot\beta$$

siendo $Y'_{js}\cdot\beta$ un escalar.

Por otro lado, sabemos que:

$$E[X_j/\theta_j] = \mu_j(\theta_j) = Y_j\cdot\beta(\theta_j)$$

Si calculamos su esperanza respecto a θ:

$$E[\mu_j(\theta_j)] = E[Y_j\cdot\beta(\theta_j)] = Y_j\cdot E[\beta(\theta_j)] = Y_j\cdot\beta$$

siendo:

$$E_{Total}[X_j] = E\big[E[X_j/\theta_j]\big] = Y_j\cdot\beta$$

que es un vector de dimensión $(t_j,1)$.

Y por último:

$$E_{Total}[X] = E_{Total}\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_j \\ \vdots \\ X_k \end{bmatrix} = \begin{bmatrix} Y_1\cdot\beta \\ Y_2\cdot\beta \\ \vdots \\ Y_j\cdot\beta \\ \vdots \\ Y_k\cdot\beta \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_j \\ \vdots \\ Y_k \end{bmatrix}\cdot\beta = Y^*\cdot\beta = m$$

ya que:

$$E_{Total}[X_j] = E\big[E[X_j/\theta_j]\big] = E[\mu(\theta_j)] = Y_j\cdot\beta$$

donde Y^* es una matriz particionada de dimensión $(k\cdot t_j,n)$, siendo sus elementos a su vez matrices de dimensión (t_j,n). La matriz $m = Y^*\cdot\beta$ es una matriz de dimensión $(k\cdot t_j,1)$, esto es, un vector columna con $(k\cdot t_j)$ filas.

•) La matriz de covarianzas entre X_i y $\mu_{js}(\theta_j)$, para $i = 1,2,\ldots,j,\ldots,k$ es igual:

$$Cov[X_i,\ \mu_{js}(\theta_j)] = \delta_{ij}\cdot Cov[X_i,\ \mu_{js}(\theta_j)]$$

por la hipótesis de independencia asumida entre las pólizas. Si $i = j$ resulta que:

$$Cov[X_j,\ \mu_{js}(\theta_j)] = E\big[Cov[X_j,\ \mu_{js}(\theta_j)/\theta_j]\big] +$$

$$+ Cov\big[E[X_j/\theta_j],\ E[\mu_{js}(\theta_j)/\theta_j]\big] =$$

$$= 0 + Cov[\mu_j(\theta_j),\ \mu_{js}(\theta_j)] =$$

$$= Cov[Y_j \cdot \beta(\theta_j),\ Y'_{js} \cdot \beta(\theta_j)] =$$

$$= Y_j \cdot Cov[\beta(\theta_j)] \cdot Y_{js} = Y_j \cdot \Gamma \cdot Y_{js}$$

de manera que:

$$Cov[X,\ \mu_{js}(\theta_j)] = \begin{bmatrix} Y_1 \cdot \Gamma \cdot Y_{js} \cdot \delta_{1j} \\ Y_2 \cdot \Gamma \cdot Y_{js} \cdot \delta_{2j} \\ \vdots \\ Y_k \cdot \Gamma \cdot Y_{js} \cdot \delta_{kj} \end{bmatrix}$$

que es un vector de dimensión $(k \cdot t_j, 1)$, donde δ_{ij} es la delta de Kronecker.

•) La matriz de covarianzas entre los distintos valores de las variables aleatorias experiencia de reclamaciones vamos a simbolizarla por C, y viene definida del siguiente modo:

$$C = Cov[X,\ X'] = \begin{bmatrix} C_1 & \emptyset & \cdots & \emptyset \\ \emptyset & C_2 & \cdots & \emptyset \\ \vdots & \vdots & \ddots & \vdots \\ \emptyset & \emptyset & \cdots & C_k \end{bmatrix}$$

debido a la hipótesis de independencia asumida en el modelo, donde:

$$C_j = Cov[X_j,\ X_j'] = E\big[Cov[X_j,\ X_j'/\theta_j]\big] +$$

$$+ Cov\big[E[X_j/\theta_j],\ E[X_j'/\theta_j]\big] =$$

$$= E\big[Var[X_j/\theta_j]\big] + Cov[\mu_j(\theta_j),\ \mu_j'(\theta_j)] =$$

$$= E[\sigma^2(\theta_j)]\cdot v_j + Cov[Y_j\cdot\beta(\theta_j),\ \beta'(\theta_j)\cdot Y_j'] =$$

$$= S^2\cdot v_j + Y_j\cdot Cov[\beta(\theta_j)]\cdot Y_j' = V_j + Y_j\cdot\Gamma\cdot Y_j'$$

siendo $V_j = S^2\cdot v_j$ una matriz cuadrada de dimensión (t_j,t_j).

Una vez introducidos todos estos elementos pasemos a resolver el problema de minimización que nos interesa. Queremos hallar los valores de $\alpha_\emptyset$ y A que minimicen la siguiente esperanza total:

$$\phi_{js} = E_{Total}[[\mu_{js}(\theta_j) - \alpha_\emptyset - X'\cdot A]^2]$$

y de este modo hallar el estimador $\hat{\mu}_{js}(\theta_j)$ óptimo para la prima de riesgo individual.

Derivando parcialmente ϕ_{js} respecto a $\alpha_\emptyset$ y A, es decir, calculando $\dfrac{\delta\phi_{js}}{\delta\alpha_\emptyset}$ y $\dfrac{\delta\phi_{js}}{\delta A}$[10], e igualando a cero, obtenemos:

---

[10] Es la expresión matricial de la derivada respecto a los elementos del vector A, siendo por ejemplo:

$$\frac{\delta\phi_{js}}{\delta A_i} = -2\cdot E_{Total}[X_i\cdot[\mu_{js}(\theta_j) - \alpha_\emptyset - X_i'\cdot A_i]]$$

$$\frac{\partial \phi_{js}}{\partial \alpha_{0}} = -2 \cdot E_{Total}[[\mu_{js}(\theta_j) - \alpha_0 - X' \cdot A]] = 0 \qquad (1)$$

$$\frac{\partial \phi_{js}}{\partial A} = -2 \cdot E_{Total}[X \cdot [\mu_{js}(\theta_j) - \alpha_0 - X' \cdot A]] = 0 \qquad (2)$$

De la ecuación (1) resulta:

$$E_{Total}[\mu_{js}(\theta_j) - \alpha_0 - X' \cdot A] = 0 \qquad (3)$$

o lo que es lo mismo:

$$E_{Total}[\mu_{js}(\theta_j)] - \alpha_0 - E_{Total}[X'] \cdot A = 0 \qquad (4)$$

de donde:

$$\alpha_0 = E_{Total}[\mu_{js}(\theta_j)] - E_{Total}[X'] \cdot A \qquad (5)$$

Sustituyendo ambas esperanzas por su valor nos queda:

$$\alpha_0 = Y'_{js} \cdot \beta - \beta' \cdot Y^{*'} \cdot A \qquad (6)$$

Como $Y'_{js} \cdot \beta$ es un escalar, da lo mismo tener $Y'_{js} \cdot \beta$ que su transpuesto $\beta' \cdot Y_{js}$, de manera que:

$$\alpha_0 = \beta' \cdot Y_{js} - \beta' \cdot Y^{*'} \cdot A$$

siendo:

$$\boxed{\alpha_0 = \beta'[Y_{js} - Y^{*'} \cdot A]} \tag{7}$$

que es un escalar.

Por otro lado, de (2) obtenemos:

$$E_{Total}\left[X \cdot [\mu_{js}(\theta_j) - \alpha_0 - X' \cdot A]\right] = 0 \tag{8}$$

o lo que es lo mismo:

$$E_{Total}[X \cdot \mu_{js}(\theta_j)] - E_{Total}[X] \cdot \alpha_0 - E_{Total}[X \cdot X'] \cdot A = 0 \tag{9}$$

Sustituyendo α_0 por (5) resulta:

$$E_{Total}[X \cdot \mu_{js}(\theta_j)] - E_{Total}[X] \cdot E_{Total}[\mu_{js}(\theta_j)] +$$

$$+ E_{Total}[X] \cdot E_{Total}[X'] \cdot A - E_{Total}[X \cdot X'] \cdot A = 0 \tag{10}$$

Reordenando los términos y sacando factor común la ecuación (10) se puede escribir también como sigue:

$$Cov[X, \mu_{js}(\theta_j)] - Cov[X, X'] \cdot A = 0$$

o de forma análoga:

$$Cov[X, X'] \cdot A = Cov[X, \mu_{js}(\theta_j)] \tag{11}$$

Sustituyendo ambas covarianzas por sus respectivas expresiones resulta:

$$[V_i + Y_i \cdot \Gamma \cdot Y_i'] \cdot A_i = Y_i \cdot \Gamma \cdot Y_{js} \cdot \delta_{ij} \qquad (12)$$

$$i = 1,2,\dots,k$$

tratándose de un sistema de $k \cdot t_i$ ecuaciones con $k \cdot t_i$ incógnitas α_{ir}.

Si $i \neq j$ el sistema de ecuaciones anterior se convierte en:

$$[V_i + Y_i \cdot \Gamma \cdot Y_i'] \cdot A_i = \emptyset \qquad (13)$$

$$i = 1,2,\dots,k - \{j\}$$

resultando ser un sistema de ecuaciones homogéneo compatible determinado, ya que por la hipótesis de independencia asumida entre las pólizas todas las ecuaciones son independientes, siendo el rango de la matriz asociada al sistema igual al número de incógnitas.

Al tener solución trivial resulta que:

$$A_i = \emptyset \quad \text{para } i \neq j$$

lo que implica que todos los coeficientes α_{ir} son iguales a cero para $i = 1,2,\dots,k - \{j\}$ y $r = 1,2,\dots,t_i$

Si $i = j$ el sistema de ecuaciones (12) se transforma en:

$$[V_i + Y_i \cdot \Gamma \cdot Y_i'] \cdot A_i = Y_j \cdot \Gamma \cdot Y_{js} \qquad (14)$$

siendo un sistema con t_j ecuaciones con t_j incógnitas α_{js}, con $s = 1,2,\dots,t_j$.

Si a la ecuación (14) la premultiplicamos por $(Y_j' \cdot V_j^{-1})$ obtenemos:

$$[Y_j' \cdot V_j^{-1} \cdot V_j + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot Y_j'] \cdot A_j = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot Y_{js} \qquad (15)$$

pero como:

$$Y_j' \cdot V_j^{-1} \cdot V_j = Y_j' \cdot I_{(t,t)} = I_{(n,n)} \cdot Y_j'$$

podemos sacar factor común y despejar $Y_j' \cdot A_j$:

$$[I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma] \cdot Y_j' \cdot A_j = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot Y_{js} \tag{16}$$

$$Y_j' \cdot A_j = [I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} \cdot Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot Y_{js}$$

que es lo mismo que escribir:

$$Y_j' \cdot A_j = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} \cdot Y_{js} \tag{17}$$

ya que[11]:

$$Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} =$$

$$= [I + Y_j \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} \cdot Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma$$

Si definimos una nueva matriz Z_j, de dimensión (n,n), como:

---

11 Si simbolizamos por $B = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma$ resulta que:

$$[I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} \cdot Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma = (I + B)^{-1} \cdot B = (I + B)^{-1} \cdot (B^{-1})^{-1} =$$

$$= (B^{-1} \cdot (I + B))^{-1} = (B^{-1} + I)^{-1}$$

y a su vez:

$$Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} = B \cdot (I + B)^{-1} = (B^{-1})^{-1} \cdot (I + B)^{-1} =$$

$$= ((I + B) \cdot B^{-1})^{-1} = (B^{-1} + I)^{-1}$$

$$\boxed{Z_j = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I + Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1}} \tag{18}$$

la ecuación (17) se puede escribir también como:

$$Y_j' \cdot A_j = Z_j \cdot Y_{js} \tag{19}$$

y sustituyendo la ecuación (19) en la (14) nos queda:

$$V_j \cdot A_j + Y_j \cdot \Gamma \cdot Z_j \cdot Y_{js} = Y_j \cdot \Gamma \cdot Y_{js} \tag{20}$$

de donde:

$$A_j = V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I - Z_j] \cdot Y_{js} \tag{21}$$

Ahora bien, si en la ecuación (19) sustituimos A_j por la expresión obtenida en (21) se desprende que:

$$Z_j = Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma \cdot [I - Z_j] \tag{22}$$

siendo:

$$[I - Z_j] = (Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma)^{-1} \cdot Z_j \tag{23}$$

Si sustituimos (23) en la ecuación (18) obtenemos:

$$A_j = V_j^{-1} \cdot Y_j \cdot \Gamma \cdot (Y_j' \cdot V_j^{-1} \cdot Y_j \cdot \Gamma)^{-1} \cdot Z_j \cdot Y_{js} \tag{24}$$

que premultiplicada por X_j' resulta ser igual:

$$\boxed{X'_j \cdot A_j = \hat{\beta}'_j \cdot Z_j \cdot Y_{js}} \tag{25}$$

siendo:

$$\boxed{\hat{\beta}'_j = X'_j \cdot V_j^{-1} \cdot Y_j \cdot (Y'_j \cdot V_j^{-1} \cdot Y_j)^{-1}} \tag{26}$$

o lo que es lo mismo:

$$\boxed{\hat{\beta}'_j = X'_j \cdot v_j^{-1} \cdot Y_j \cdot (Y'_j \cdot v_j^{-1} \cdot Y_j)^{-1}} \tag{27}$$

pues $V_j = S^2 \cdot v_j$.

El vector $\hat{\beta}_j$, de dimensión $(n,1)$, no es más que un **estimador mínimo-cuadrático generalizado** para el modelo de regresión polinómico, basado en X_j.

Anteriormente hemos obtenido en la ecuación (7) que el coeficiente $\alpha_{\emptyset}$ era igual:

$$\alpha_{\emptyset} = \beta' \cdot [Y_{js} - Y^{*'} \cdot A] \qquad \text{escalar}$$

pero como hemos obtenido que $A_i = \emptyset$ para $i \neq j$, con $i = 1,2,\ldots,k$, resulta que:

$$Y^{*'} \cdot A = \sum_{i=1}^{k} Y'_i \cdot A_i = Y'_j \cdot A_j$$

siendo:

$$\boxed{\alpha_{\emptyset} = \beta' \cdot [Y_{js} - Y'_j \cdot A_j]} \tag{28}$$

y a su vez, si sustituimos $Y'_j \cdot A_j$ por (19) resulta que también podemos escribirlo del siguiente modo:

$$\boxed{\alpha_{\emptyset} = \beta' \cdot [I - Z_j] \cdot Y_{js}} \tag{29}$$

Una vez obtenidas las expresiones de $\alpha_{\emptyset}$ y de $X'_j \cdot A_j$, siendo:

$$\alpha_{\emptyset} = \beta' \cdot [I - Z_j] \cdot Y_{js} \tag{29}$$

y

$$X'_j \cdot A_j = \hat{\beta}'_j \cdot Z_j \cdot Y_{js} \tag{25}$$

estamos en condiciones de obtener el estimador óptimo para la prima de riesgo individual, $\hat{\mu}_{js}(\theta_j)$.

Para ello, recordemos que habíamos definido $\hat{\mu}_{js}(\theta_j)$ como:

$$\hat{\mu}_{js}(\theta_j) = \alpha_{\emptyset} + X' \cdot A = \alpha_{\emptyset} + \sum_{i=1}^{k} X'_i \cdot A_i$$

Sin embargo como en (13) hemos obtenido que:

$$A_i = \emptyset \qquad \text{para} \quad i \neq j$$

resulta que:

$$\hat{\mu}_{js}(\theta_j) = \alpha_{\emptyset} + X'_j \cdot A_j = \beta' \cdot [I - Z_j] \cdot Y_{js} + \hat{\beta}'_j \cdot Z_j \cdot Y_{js}$$

siendo:

$$\boxed{\hat{\mu}_{js}(\theta_j) = \left[\beta'\cdot[I - Z_j] + \hat{\beta}'_j\cdot Z_j \right]\cdot Y_{js}} \tag{30}$$

tratándose de un escalar.

El estimador ajustado de credibilidad lineal óptimo también podemos escribirlo como sigue:

$$\boxed{\hat{\mu}_{js}(\theta_j) = \hat{\beta}(\theta_j)\cdot Y_{js}} \tag{31}$$

donde:

$$\boxed{\hat{\beta}(\theta_j) = \beta'\cdot[I - Z_j] + \hat{\beta}'_j\cdot Z_j} \tag{32}$$

que es un vector de dimensión (1,n), y se le conoce con el nombre de **estimador de credibilidad del vector de regresión $\beta(\theta_j)$**, siendo:

$$\beta = E[\beta(\theta_j)] \qquad \text{dim } (n,1)$$

$$\hat{\beta}'_j = X'_j\cdot v_j^{-1}\cdot Y_j\cdot(Y'_j\cdot v_j^{-1}\cdot Y_j)^{-1} \qquad \text{dim } (1,n)$$

$$Z_j = Y'_j\cdot V_j^{-1}\cdot Y_j\cdot\Gamma\cdot[I + Y'_j\cdot V_j^{-1}\cdot Y_j\cdot\Gamma]^{-1} \qquad \text{dim } (n,n)$$

A la matriz Z_j, HACHEMEISTER, C. la denominó **matriz de credibilidad**, que también se suele escribir del siguiente modo:

$$Z_j = \Gamma\cdot(\Gamma + S^2\cdot\omega_j)^{-1} \tag{33}$$

siendo:

$$\omega_j = (Y_j' \cdot v_j^{-1} \cdot Y_j)^{-1} \tag{34}$$

Para obtener esta expresión alternativa para la matriz de credibilidad, Z_j, recordemos que hemos definido $V_j = S^2 \cdot v_j$. Si sustituimos en la ecuación (18) V_j por esta expresión resulta:

$$Z_j = Y_j' \cdot (S^2 \cdot v_j)^{-1} \cdot Y_j \cdot \Gamma \cdot [I + Y_j' \cdot (S^2 \cdot v_j)^{-1} \cdot Y_j \cdot \Gamma]^{-1}$$

siendo:

$$(Y_j' \cdot (S^2 \cdot v_j)^{-1} \cdot Y_j)^{-1} = S^2 \cdot (Y_j' \cdot v_j^{-1} \cdot Y_j)^{-1} = S^2 \cdot \omega_j$$

de donde[12]:

$$Z_j = (S^2 \cdot \omega_j)^{-1} \cdot \Gamma \cdot [I + (S^2 \cdot \omega_j)^{-1} \cdot \Gamma]^{-1} =$$

$$= \Gamma \cdot [(I + (S^2 \cdot \omega_j)^{-1} \Gamma) \cdot S^2 \omega_j]^{-1} =$$

$$= \Gamma \cdot (S^2 \cdot \omega_j + \Gamma)^{-1} = \Gamma \cdot (\Gamma + S^2 \cdot \omega_j)^{-1}$$

De manera que:

$$Z_j = \Gamma \cdot (\Gamma + S^2 \cdot \omega_j)^{-1}$$

12 $(S^2 \cdot \omega_j)^{-1} \cdot \Gamma = \Gamma \cdot (S^2 \cdot \omega_j)^{-1}$ ya que Γ y ω_j son matrices simétricas.

3.1.4.- Estimación de los parámetros estructurales.

Una vez obtenido el estimador para la prima de riesgo individual el siguiente paso consiste en estimar:

- $\beta = E[\beta(\theta_j)]$
- $S^2 = E[\sigma^2(\theta_j)]$
- $\Gamma = Cov[\beta(\theta_j)]$

A los parámetros $\boldsymbol{\beta}$, S^2 y Γ NORBERG, R. (1979) los denomina **parámetros estructurales,** y la estimación de los mismos es un prerrequisito indispensable para poder aplicar el estimador ajustado de credibilidad.

Entre los autores que han tratado este tema cabe destacar a HACHEMEISTER, C. (1975), De VYLDER, F. (1978, 1981a y 1984), NORBERG, R. (1979 y 1982) y a ZEHNWIRTH, B. (1984), entre otros.

Al estimar estos tres parámetros estructurales nos volvemos a encontrar con el problema de los pseudo-estimadores, al igual que en el Modelo de BÜHLMANN-STRAUB. Sin embargo, en este caso, la estimación de los parámetros a través de los pseudo-estimadores está más desarrollada. La idea básica consiste en definir una familia de pseudo-estimadores insesgados para cada uno de los parámetros estructurales, y después elegir, de entre ellos, el que tenga varianza mínima, ya que la insesgadez y la varianza mínima se siguen considerando las propiedades más deseables para los estimadores.

Dejando de lado los problemas que presentan los pseudo-estimadores, pasemos a la estimación de los tres parámetros estructurales.

3.1.4.1.- Estimación de $\beta = E[\beta(\theta_j)]$.

HACHEMEISTER, C. (1975), pp. 148, propuso estimar el parámetro β a través del estimador insesgado $\hat{\beta}$, definido como sigue:

$$\hat{\beta} = (Y^{*'} \cdot v^{-1} \cdot Y^{*})^{-1} \cdot Y^{*'} \cdot v^{-1} \cdot X \tag{35}$$

donde v una matriz diagonal, cuyos elementos de la diagonal principal son las matrices v_j, con $j = 1,2,\ldots,k$.

Este no ha sido el único estimador propuesto para β. De VYLDER, F. (1978) y NORBERG, R. (1979) propusieron estimar β del siguiente modo:

Al ser $\hat{\beta}_j$, definido anteriormente como:

$$\hat{\beta}_j = (Y_j' \cdot v_j^{-1} \cdot Y_j)^{-1} \cdot Y_j' \cdot v_j^{-1} \cdot X_j$$

un estimador insesgado[13] de β, ya que $E[\hat{\beta}_j] = \beta$, resulta que:

$$\hat{\beta}_F = \sum_{j=1}^{k} F_j \cdot \hat{\beta}_j \tag{36}$$

13 La insesgadez del estimador fue demostrada por De VYLDER, F. (1978), pp. 102-103.

es también un estimador insesgado de β, siendo F_j, con $j = 1,2,\ldots,k$, una serie de matrices cuadradas de dimensión (n,n), que verifican:

$$\sum_{j=1}^{k} F_j = I$$

siendo la matriz I la matriz identidad.

La relación (36) define una familia de pseudo-estimadores para β, de entre los cuales el que tiene varianza mínima, como demostró De VYLDER, F. (1978), pp. 108-109, se obtiene cuando F_j es la matriz de credibilidad Z_j prenormada, esto es, cuando:

$$F_j = \bar{Z}_j = \left[\sum_{j=1}^{k} Z_j\right]^{-1} \cdot Z_j$$

siendo el estimador colectivo para $E[\beta(\theta_j)]$ propuesto por De VYLDER, F.:

$$\boxed{\hat{\beta}_Z = \left[\sum_{j=1}^{k} Z_j\right]^{-1} \cdot \sum_{j=1}^{k} Z_j \cdot \hat{\beta}_j} \tag{37}$$

que será el estimador que nosotros utilizaremos en la posterior aplicación práctica.

3.1.4.2.- Estimación del parámetro $S^2 = E[\sigma^2(\theta_j)]$.

Según NORBERG, R. (1979) el estimador para el parámetro S^2 se obtiene a través de:

$$\phi_k = \sum_{j=1}^{k} S^2_{j,G} \tag{38}$$

donde:

$$S^2_{j,G} = (X_j - Y_j \cdot \hat{\beta}_j)' \cdot G_j \cdot (X_j - Y_j \cdot \hat{\beta}_j) \tag{39}$$

es una suma ponderada generalizada de los residuos cuadrados con G_j matrices de ponderación de dimensión (t_j, t_j), definidas positivas, simétricas y sujetas a la restricción:

$$\sum_{j=1}^{k} [\text{traza}(v_j \cdot G_j) - \text{traza}[(Y'_j \cdot v_j^{-1} \cdot Y_j)^{-1} \cdot (Y'_j \cdot G_j \cdot Y_j)]] = 1$$

Por su parte, De VYLDER, F. (1981) utiliza el siguiente razonamiento para hallar el estimador para S^2:

El estimador clásico insesgado[14] para S^2, basado sólamente en el vector X_j, viene dado por:

$$\hat{S}^2_j = \frac{1}{t_j - n} \cdot (X_j - Y_j \cdot \hat{\beta}_j)' \cdot v_j^{-1} \cdot (X_j - Y_j \cdot \hat{\beta}_j) \tag{40}$$

que no es más que un estimador mínimo-cuadrático generalizado.

---

[14] La insesgadez del estimador fue demostrada por De VYLDER, F. (1978), pp. 102-103.

Para todos los escalares G_j $(j = 1,2,\ldots,k)$, podemos definir S_G^2 como:

$$S_G^2 = \sum_{j=1}^{k} G_j \cdot \hat{S}_j^2 \qquad \text{con} \qquad \sum_{j=1}^{k} G_j = 1 \tag{41}$$

el cual define una familia de estimadores insesgados para S^2.

Introduciendo la hipótesis de normalidad para el vector X_j, es decir, si X_j se distribuye como $N(Y\cdot\beta(\theta_j),\ \sigma^2(\theta_j)\cdot v_j)$, resulta que el estimador de varianza mínima, como demostró De VYLDER, F. (1981), pp. 120-122, dentro de la familia de estimadores (41) es:

$$\hat{S}^2 = \frac{1}{k} \sum_{j=1}^{k} \hat{S}_j^2 \tag{42}$$

siendo la media aritmética de los estimadores parciales.

Sustituyendo $\hat{S}_j^2$ por su valor, se obtiene:

$$\boxed{\hat{S}^2 = \frac{1}{k\cdot(t_j - n)} \cdot \sum_{j=1}^{k} (X_j - Y_j\cdot\hat{\beta}_j)\cdot v_j^{-1}\cdot(X_j - Y_j\cdot\hat{\beta}_j)} \tag{43}$$

que es un escalar, y será el estimador que nosotros utilizaremos.

3.1.4.3.- Estimación de la matriz $\Gamma = \mathrm{Cov}[\beta(\theta_j)]$.

Se han propuesto distintos estimadores para la matriz de covarianza Γ, algunos de ellos de una gran complejidad.

HACHEMEISTER, C. (1975), pp. 149-151, propuso como estimador de Γ a la siguiente matriz:

$$\hat{\Gamma} = \frac{1}{2}\cdot(H + H') \tag{44}$$

siendo:

$$H = \pi^{-1}\cdot[G - (k - 1)\cdot(Y^{*'}\cdot v^{-1}\cdot Y^{*})^{-1}\cdot\hat{S}^2] \tag{45}$$

$$G = \sum_{j=1}^{k} (Y^{*'}\cdot v^{-1}\cdot Y^{*})^{-1}\cdot(Y_j'\cdot v_j^{-1}\cdot Y_j)\cdot(\hat{\beta}_j - \hat{\beta})\cdot(\hat{\beta}_j - \hat{\beta})' \tag{46}$$

$$\pi = I - \sum_{j=1}^{k} (Y^{*'}\cdot v^{-1}\cdot Y^{*})^{-1}\cdot(Y_j'\cdot v_j^{-1}\cdot Y_j)\cdot \\ \cdot(Y^{*'}\cdot v^{-1}\cdot Y^{*})\cdot(Y_j'\cdot v_j^{-1}\cdot Y_j) \tag{47}$$

$$\hat{\beta} = (Y^{*'}\cdot v^{-1}\cdot Y^{*})^{-1}\cdot Y^{*'}\cdot v^{-1}\cdot X \tag{48}$$

Se trata de un estimador insesgado, que no depende del parámetro que estamos estimando, representando una gran ventaja desde el punto de vista calculatorio. No obstante, no posee varianza mínima.

NORBERG, R. (1979), pp. 210, por su lado, considera el siguiente estimador para Γ:

$$\Lambda_k = \frac{1}{2}\cdot\sum_{j=1}^{k} \left[H_j\cdot\left[\hat{\beta}_j\cdot\hat{\beta}_j' - \phi_k\cdot(Y_j'\cdot v_j^{-1}\cdot Y_j)^{-1}\right] + \right.$$

$$+ \left[\hat{\beta}_j \cdot \hat{\beta}'_j - \phi_k \cdot (Y'_j \cdot v_j^{-1} \cdot Y_j)^{-1} \right] \cdot H'_j \Big] - \beta_F \cdot \beta'_F \qquad (49)$$

siendo H_j matrices de ponderación, de dimensión (n,n), sujetas a la restricción:

$$\sum_{j=1}^{k} H_j = I \qquad (50)$$

Y por último, cabe mencionar a De VYLDER, F. (1984) que propuso como estimador de la matriz Γ a $\hat{\Gamma}$, el cual sí tiene varianza mínima dentro de la clase de estimadores insesgados:

$$\boxed{\hat{\Gamma} = \frac{1}{k-1} \cdot \sum_{j=1}^{k} Z_j \cdot (\hat{\beta}_j - \hat{\beta}_Z) \cdot (\hat{\beta}_j - \hat{\beta}_Z)'} \qquad (51)$$

La matriz $\hat{\Gamma}$, de dimensión (n,n), debe ser una matriz simétrica, por lo que en cada iteración es sustituida por $\frac{\hat{\Gamma} + \hat{\Gamma}'}{2}$.

La matriz $\hat{\Gamma}$ es un pseudo-estimador, que contiene a través de Z_j el parámetro Γ que está siendo estimado. Estimadores como éste se solucionan mediante un proceso iterativo. En la práctica, se escoge un valor inicial para $\hat{\Gamma}$, con el que se calculan unos valores para Z_j y $\hat{\beta}_Z$. Con estos valores se obtiene un valor para $\hat{\Gamma}$ más perfeccionado, y el proceso se repite hasta que la matriz $\hat{\Gamma}$ se estabiliza. Al ser $\hat{\Gamma}$ una matriz simétrica, en cada una de las iteraciones sucesivas que se van haciendo, esta matriz es sustituida por $\frac{\hat{\Gamma} + \hat{\Gamma}'}{2}$.

Cabe destacar que los estimadores propuestos por NORBERG, R. (1979), como él indica, son consistentes bajo suaves restricciones, es decir:

$$(\hat{\beta}_F,\ \phi_k,\ \Lambda_k) \xrightarrow{\ P\ } (\beta,\ S^2,\ \Gamma) \text{ cuando } k \longrightarrow \infty$$

y que los estimadores propuestos por HACHEMEISTER, C. y De VYLDER, F. se obtienen tomando unas particulares matrices de ponderación en los estimadores dados por NORBERG, R.

Nosotros vamos a utilizar como estimadores de los parámetros estructurales β, S^2 y Γ los propuestos por De VYLDER, F. que vienen dados en (37), (43) y (51), pues son los que en general se han utilizado con mayor asiduidad, siendo los pasos a seguir para obtenerlos los siguientes:

Dados: $X_j\ (t_j,1)$; $v_j\ (t_j,t_j)$; $Y\ (t_j,n)$.

Primero se calcula $\hat{\beta}_j = (Y_j'\cdot v_j^{-1}\cdot Y_j)^{-1}\cdot Y_j'\cdot v_j^{-1}\cdot X_j$ y a continuación el escalar:

$$S^2 = \frac{1}{k\cdot(t-n)}\cdot\sum_{j=1}^{k} (X_j - Y_j\cdot\hat{\beta}_j)'\cdot v_j^{-1}\cdot(X_j - Y_j\cdot\hat{\beta}_j)$$

Para poder calcular el resto de los estimadores es necesario fijar un valor inicial para el parámetro Γ, por ejemplo, $\Gamma = \text{diag}(c,c,\ldots,c)$, matriz de dimensión (n,n), siendo c un número elevado.

Una vez fijado el valor de Γ se calcula Z_j, así como el valor de $\left[\sum_{j=1}^{k} Z_j \right]^{-1}$ y calculados ambos valores se puede obtener $\hat{\beta}_Z$, que es igual a:

$$\hat{\beta}_Z = \left[\sum_{j=1}^{k} Z_j \right]^{-1} \cdot \sum_{j=1}^{k} Z_j \cdot \hat{\beta}_j$$

estando ya en condiciones de calcular $\hat{\Gamma}$, pues todos los elementos que intervienen en su fórmula han sido obtenidos previamente. Una vez obtenido el valor de $\hat{\Gamma}$, se vuelve a iniciar el proceso a partir del cálculo de la Z_j, y es en el cálculo de Z_j cuando se utiliza en lugar de $\hat{\Gamma}$, $\dfrac{\hat{\Gamma} + \hat{\Gamma}'}{2}$ ya que debe ser una matriz simétrica. Este proceso se va repitiendo hasta que el valor de $\hat{\Gamma}$ se estabiliza.

Una vez conocidos los valores de Z_j, $\hat{\beta}_Z$ y $\hat{\beta}_j$ se puede calcular el estimador de credibilidad para la prima de riesgo individual:

$$\hat{\mu}_{js}(\theta_j) = \hat{\beta}(\theta_j) \cdot Y_{js} = \left[\hat{\beta}_j' \cdot Z_j + \hat{\beta}_Z' \cdot [I - Z_j] \right] \cdot Y_{js}$$

3.1.5.- Comentarios al modelo.

Como ya hemos indicado anteriormente, el Modelo de Regresión de HACHEMEISTER es una generalización del Modelo de BÜHLMANN-STRAUB, de modo que este último no es más que un caso particular del primero.

Si en dicho modelo consideramos el caso particular que:

$$n = 1$$

$$Y_j = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \text{ vector de dimensión } (t_j, 1)$$

$$\beta(\theta_j) = \mu(\theta_j) \quad \text{escalar}$$

$$v_j = \text{diag}\left(\frac{1}{w_{j1}}, \frac{1}{w_{j2}}, \ldots, \frac{1}{w_{jt_j}}\right)$$

resulta que:

$$E[X_j/\theta] = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \cdot \mu(\theta_j)$$

$$\text{Cov}[X_j/\theta_j] = \sigma^2(\theta_j) \cdot v_j = \text{Var}[X_j/\theta_j]$$

y los parámetros estructurales se convierten en:

$$\beta = E[\mu(\theta_j)] = m \qquad \text{escalar}$$

$$S^2 = E[\sigma^2(\theta_j)]$$

$$\Gamma = \text{Cov}[\mu(\theta_j)] = \text{Var}[\mu(\theta_j)] = a$$

siendo a su vez:

$$\hat{\beta}'_j = X'_j \cdot v_j^{-1} \cdot Y_j \cdot (Y'_j \cdot v_j^{-1} \cdot Y_j)^{-1} = \frac{1}{w_{j\cdot}} \cdot \sum_{s=1}^{t_j} w_{js} \cdot X_{js} = X_{jw}$$

$$Z_j = Y'_j \cdot V_j \cdot Y_j \cdot \Gamma \cdot [I + Y'_j \cdot V_j^{-1} \cdot Y_j \cdot \Gamma]^{-1} =$$

$$= a \cdot \frac{w_{j\cdot}}{S^2} \cdot \left[1 + a \cdot \frac{w_{j\cdot}}{S^2} \right]^{-1} = a \cdot w_{j\cdot} \cdot (S^2 + a \cdot w_{j\cdot})^{-1} =$$

$$= \frac{a \cdot w_{j\cdot}}{S^2 + a \cdot w_{j\cdot}}$$

En este caso, el estimador ajustado de credibilidad se convierte en:

$$\hat{\mu}_{js}(\theta_j) = [\, \beta'[I - Z_j] + \hat{\beta}'_j \cdot Z_j \,] \cdot Y_{js} = [1 - Z_j] \cdot m + X_{jw} \cdot Z_j$$

que es precisamente el estimador ajustado de credibilidad obtenido en el Modelo de BÜHLMANN-STRAUB para el caso en que cada póliza tenga t_j observaciones. Si consideramos además que $t_j = t \;\; \forall j$, el estimador de credibilidad que obtendremos será idéntico al de BÜHLMANN-STRAUB.

También hemos visto que una de las hipótesis del Modelo de Regresión de HACHEMEISTER es:

$$\mathrm{Cov}[X_j/\theta_j] = \sigma^2(\theta_j) \cdot v_j$$

donde $\sigma^2(\theta_j)$ es una función escalar desconocida (ya hemos visto como estimarla), y v_j es una matriz dada, de dimensión (t_j, t_j), semidefinida positiva.

Sin embargo, el propio HACHEMEISTER, C. (1975) cuando aplicó su modelo de regresión a un caso práctico introdujo la siguiente hipótesis:

$$v_j = \mathrm{diag}\left(\frac{1}{w_{j1}}, \frac{1}{w_{j2}}, \ldots, \frac{1}{w_{jt}}\right)$$

es decir, implícitamente para cada contrato asumió que las series temporales de las observaciones de la experiencia de reclamaciones están incorrelacionadas. En otras palabras, consideró que la cantidad media de reclamaciones del periodo r-ésimo no afecta al montante del periodo s-ésimo, siempre que $r \neq s$.

AL QUIRIN (1975), pp. 164-169, en la discusión que hizo del artículo de HACHEMEISTER, C. (1975), apuntó el hecho de que la no aceptación de la hipótesis de autocorrelación en la aplicación práctica de su modelo, junto con la asunción de que:

$$\mathrm{Var}[X_j] = \sigma^2(\theta_j) \qquad s \in \{1,2,\ldots,t_j\}$$

le sirvió para simplificar la obtención de sus resultados numéricos, y evitarse de este modo muchos de los problemas que representa estimar los parámetros que aparecen en su modelo de credibilidad, muchos de los cuales no son directamente observables, como es el caso de Z_j, que para conocer su valor es necesario estimar previamente los parámetros estructurales Γ y S^2.

Varios años más tarde, De VYLDER. F. (1981) propuso una simplificación del Modelo de Regresión de HACHEMEISTER, que consiste en utilizar como **pesos credibilísticos escalares** en lugar de matrices. Dicho autor justifica la utilización de escalares debido a que las matrices de credibilidad pueden contener elementos negativos, dando lugar a resultados que son de difícil interpretación en los casos prácticos. Esta modificación da lugar a estimadores menos precisos, pero el modelo adquiere una serie de ventajas. Al ser los pesos credibilísticos números comprendidos entre 0 y 1, ambos inclusives, los resultados son más atractivos en las situaciones prácticas.

Y por último, cabe mencionar a ZEHNWIRTH, B. (1984), que siguiendo a DUCAN y HORN (1972), expusó un **Modelo de Regresión Lineal Mixto.** Dicho modelo consiste en combinar, por un lado, un modelo lineal que describe los datos básicos, y por otro, unos estimadores mínimo-cuadráticos generalizados, siendo el modelo lineal que describe los datos básicos:

$$X^G = \hat{Y} \cdot \hat{\beta} + \hat{\epsilon}$$

donde $\hat{\epsilon}$ es el término de error de media cero y matriz de covarianza Σ, X^G es un vector de observaciones, y $\hat{Y}$ es la matriz de datos conocidos.

Además se supone que se poseen datos anteriores a los que se les aplica un modelo lineal con los mismos parámetros $\hat{\beta}$ que en el modelo anterior. Basándose en dichos datos se obtienen unos estimadores mínimo-cuadráticos generalizados:

$$\beta_0^G = \hat{\beta} + \hat{\delta}$$

donde $\hat{\delta}$ es un término de error de media cero, y con una matriz de covarianza Γ.

La utilización de ambos conjuntos, combinando las dos fórmulas anteriores, da lugar al Modelo de Regresión Lineal Mixto:

$$\begin{bmatrix} X^G \\ \beta_\emptyset^G \end{bmatrix} = \begin{bmatrix} \hat{Y} \\ I \end{bmatrix} \cdot \hat{\beta} + \begin{bmatrix} \hat{\epsilon} \\ \hat{\delta} \end{bmatrix}$$

donde $E[\hat{\epsilon}] = E[\hat{\delta}] = \emptyset$, y la matriz de $Cov(\hat{\epsilon},\hat{\delta})$ es una matriz diagonal con elementos Σ y Γ.

El estimador mínimo cuadrado generalizado para este modelo mixto es:

$$\hat{\beta}^a = (\hat{Y}' \cdot \Sigma^{-1} \cdot \hat{Y} + \Gamma^{-1})^{-1} \cdot (\Gamma^{-1} \cdot \hat{\beta}_\emptyset + \hat{Y}' \cdot \Sigma^{-1} \cdot X^G)$$

Si se define:

$$Z = (\hat{Y}' \cdot \Sigma^{-1} \cdot \hat{Y} + \Gamma^{-1}) \cdot \hat{Y}' \cdot \Sigma^{-1} \cdot \hat{Y}$$

y haciendo algunos cálculos previos se llega a:

$$\hat{\beta}^a = [I - Z] \cdot \hat{\beta}_\emptyset + Z \cdot \beta$$

donde:

$$\hat{\beta}_\emptyset = (\hat{Y}' \cdot \Sigma^{-1} \cdot \hat{Y})^{-1} \cdot \hat{Y}' \cdot \Sigma^{-1} \cdot X^G$$

siendo el estimador mínimo-cuadrático generalizado para el modelo en ausencia de información a priori de β^G.

Este modelo es fácilmente reducible al Modelo de BÜHLMANN, al de BÜHLMANN-STRAUB o al de HACHEMEISTER, con sólo introducir las hipótesis específicas de cada uno de estos modelos.

3.2.- MODELO DE REGRESION NO-LINEAL DE DE VYLDER

El Modelo de Regresión No-Lineal de De VYLDER no es más que una generalización del Modelo de Regresión de HACHEMEISTER. En este modelo, la hipótesis de homogeneidad en el tiempo es abandonada completamente y es sustituida por otra, no necesariamente de tipo polinómica, sino mucho más general.

Las observaciones esperadas pasan a ser función de un vector de regresión desconocido: $E[X_j/\theta_j] = f(\beta(\theta_j))$, función que puede adoptar distintas expresiones, de entre las cuales cabe destacar la de tipo exponencial, ya que el **Modelo de Regresión Exponencial** es uno de los más utilizados dentro de los modelo de regresión no-lineales.

El objetivo de este modelo sigue siendo, como en los modelos anteriores, hallar un estimador para la prima de riesgo individual, sin embargo en este caso, el procedimiento utilizado es distinto. De VYLDER, F. (1985), se plantea encontrar la mejor aproximación de tipo credibilístico para el vector de regresión $\beta(\theta_j)$, utilizando el instrumental de las distancias, abandonando el procedimiento de los mínimos cuadrados. Su primer paso fue hallar el estimador individual para el vector de regresión, para luego pasar a calcular la matriz de credibilidad óptima.

Debido a que no es posible conocer la expresión concreta que adopta el estimador individual de $\beta(\theta_j)$, sino tan sólo la ecuación que debe verificar, De VYLDER, F. (1985) se vió en la necesidad de hallar la **mejor aproximación polinómica para el modelo de regresión general**, para poder así hallar una expresión concreta para la matriz de credibilidad, y de este modo poder obtener la mejor aproximación de tipo credibilístico para el vector de regresión $\beta(\theta_j)$.

El Modelo de Regresión No-Lineal es el último de los modelos credibilísticos expuestos hasta el momento, según nuestra información disponible, y es el que se muestra como más adecuado cuando se trabaja con datos que incluyen los efectos de la inflación, en especial el Modelo de Regresión Exponencial.

3.2.1.- Variables relevantes.

Las variables más relevantes de este modelo son:

- θ_j : El ya conocido parámetro de riesgo, con $j = 1,2,\ldots,k$.

- X_{js}: Variable aleatoria que nos indica la experiencia de reclamaciones, con $j = 1,2,\ldots,k$ y $s = 1,2,\ldots,t$, siendo:

  $X_j = (X_{j1}, X_{j2},\ldots, X_{jt})'$

- Además se considera que para cada póliza existe una función f definida de la siguiente forma:

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

$$(b_1, b_2,\ldots, b_n) \longrightarrow f(b_1, b_2,\ldots, b_n)$$

siendo n la longitud del vector $\beta(\theta_j)$, vector de regresión desconocido que ya apareció en el modelo anterior, y donde t es la longitud de X_j.

Por razones de tipo matemático, se asume que todas las derivadas parciales $\frac{\delta f}{\delta b_i}$ existen, y simbolizaremos con $g(b_1, b_2, \ldots, b_n)$ la matriz Jacobiana de dimensión (t,n), formada por las derivadas parciales puestas en columnas, esto es:

$$g(b_1, b_2, \ldots, b_n) = \left[\frac{\delta f}{\delta b_1}, \frac{\delta f}{\delta b_2}, \ldots, \frac{\delta f}{\delta b_n} \right]$$

3.2.2.- Hipótesis del modelo.

Las hipótesis que se asumen en este modelo son:

D.1) Las pólizas j = 1,2,...,k , esto es, los pares $(\theta_1, \vec{X}_1)$, $(\theta_2, \vec{X}_2), \ldots, (\theta_k, \vec{X}_k)$ son independientes, y las variables $\theta_1, \theta_2, \ldots, \theta_k$ están idénticamente distribuidas.

Igual que en los tres modelos anteriores, se asume la independencia entre los riesgos, y la equivalencia exterior de los mismos.

D.2) Para cada $j = 1,2,\dots,k$:

a) $E[X_j/\theta_j] = f(\beta(\theta_j))$

donde $\beta(\theta_j)$ es un vector desconocido de dimensión n, y f es la función antes definida: $f\colon R^n \longrightarrow R^t$.

Como puede observarse, la hipótesis de homogeneidad en el tiempo es abandonada completamente, y es reemplazada por otra asunción, no necesariamente de tipo polinómica, como ocurría en el Modelo de Regresión de HACHEMEISTER, mostrándose como más adecuado cuando se está trabajando con datos que incluyen los efectos de la inflación.

b) $Cov[X_j/\theta_j] = \sigma^2(\theta_j)\cdot v_j$

donde $\sigma^2(\theta_j)$ es una variable aleatoria integrable desconocida, y v_j una matriz dada, de dimensión (t,t), semidefinida positiva y simétrica, con $j = 1,2,\dots,k$.

$Cov[X_j/\theta_j]$ es la matriz de varianzas y covarianzas correspondiente a la póliza j-ésima, y coincide con la matriz dada en el Modelo de Regresión de HACHEMEISTER.

Así, por ejemplo, en el caso particular que $n = 2$, $\beta_j = (b_{j1}, b_{j2})$, con $j = 1,2,\dots,k$, y definiendo:

$$f : R^2 \longrightarrow R^t$$

$$f(b_{j1}, b_{j2}) = \begin{bmatrix} b_{j1} \\ b_{j1} \cdot b_{j2} \\ b_{j1} \cdot b_{j2}^2 \\ \vdots \\ b_{j1} \cdot b_{j2}^{t-1} \end{bmatrix}$$

que es un vector de dimensión $(t,1)$, estamos considerando que:

$$E[X_{js}/\theta_j] = b_{j1} \cdot b_{j2}^{s-1} \quad \text{con } s = 1,2,\ldots,t$$

tratándose, a su vez, de un vector de dimensión $(t,1)$.

La matriz Jacobiana g, de las derivadas parciales, para la póliza j-ésima es en este caso una matriz de dimensión $(t,2)$:

$$g(b_{j1}, b_{j2}) = \begin{bmatrix} \dfrac{\delta f}{\delta b_{j1}} & \dfrac{\delta f}{\delta b_{j2}} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ b_{j2} & b_{j1} \\ b_{j2}^2 & 2 \cdot b_{j1} \cdot b_{j2} \\ \vdots & \\ b_{j2}^{t-1} & (t-1) \cdot b_{j1} \cdot b_{j2}^{t-2} \end{bmatrix}$$

Este modelo es conocido con el nombre de **Modelo de Regresión Exponencial**, y es uno de los más utilizados dentro de los modelos de regresión no-lineales.

3.2.3.- Estimación de la prima de riesgo individual.

Como en los modelos anteriores, el objetivo de este modelo sigue siendo estimar la prima de riesgo individual, que en este caso es:

$$E[X_j/\theta_j] = f(\beta(\theta_j))$$

Para el tratamiento de este problema vamos a seguir el artículo original de De VYLDER, F. (1985).

Para dicho autor, el problema reside en este caso, en encontrar una aproximación de tipo credibilístico para $\beta(\theta_j)$ del tipo:

$$\boxed{\hat{\beta}(\theta_j) = Z_j \cdot \hat{\beta}_j + [I - Z_j] \cdot \beta}$$

donde:

$\beta = E[\beta(\theta_j)]$ es un vector de dimensión (n,1).

$\hat{\beta}_j$ es el estimador individual para $\beta(\theta_j)$, siendo a su vez un vector de dimensión (n,1).

Z_j es la matriz de credibilidad, de dimensión (n,n).

A $\hat{\beta}(\theta_j)$ se le conoce con el nombre de **estimador ajustado de credibilidad**, y es un vector de dimensión (n,1).

Antes de adentrarnos en el problema, vamos a introducir una serie de conceptos y lemas, que después vamos a utilizar para la determinación del estimador ajustado de credibilidad.

•) $\boxed{\text{LEMA 1}}$:

Si a, x, b son matrices tales que el producto $a \cdot x \cdot b$ es un escalar, entonces:

$$a \cdot x \cdot b = 0 \quad \forall x \iff b \cdot a = 0$$

<u>DEMOSTRACION:</u>

$$a \cdot x \cdot b = tr(a \cdot x \cdot b) = tr(x \cdot c) = \sum_{ij} x_{ij} \cdot c_{ij}$$

donde $c = b \cdot a$, x_{ij} son los elementos de la matriz x y c_{ij} los de la matriz c.

Si $a \cdot x \cdot b = 0 \quad \forall x \iff c_{ij} = 0 \quad \forall i,j \iff c = 0$

Esto significa que si $a \cdot x \cdot b = 0$ $\forall x$, un caso particular podría ser aquel en que todos los elementos de x tuviesen el mismo signo, y fuesen distintos de cero, con lo que todas las c_{ij} deberán ser nulas, o sea $c = 0$.

•) $\boxed{\text{Distancia en } R^n}$:

La distancia usual entre dos puntos $a,b \in R^n$ es:

$$[(b - a)'\cdot I\cdot(b - a)]^{1/2}$$

donde I es la matriz identidad.

Con más generalidad, con una matriz simétrica, definida positiva, de dimensión (n,n), p, se puede asociar una distancia en R^n definida por:

$$[(b - a)'\cdot p\cdot(b - a)]^{1/2}$$

Esta distancia es la correspondiente al producto escalar definido por:

$$\langle a,b\rangle = a'\cdot p\cdot b$$

siendo la relación de ortogonalidad:

$$a \perp b \iff a'\cdot p\cdot b = 0$$

•) $\boxed{\text{Distancia de optimización en } R^n}$:

Sea $a \in R^n$ un punto fijo, y x un punto variable sobre una superficie S en R^n fijada. La condición necesaria para que el punto $x \in S$ minimice la distancia al punto a es tal que:

$$(x - a) \perp dx$$

donde dx es la diferencial de x.

•) Matrices aleatorias :

Son matrices cuyos elementos son variables aleatorias.

Vamos a simbolizar por L_2^n el espacio de columnas aleatorias con n componentes cuadrado-integrables.

LEMA 2 :

Si A, B son matrices aleatorias y x una matriz tal que $A \cdot x \cdot B$ es una matriz de dimensión (1,1), entonces:

$$E(A \cdot x \cdot B) = \emptyset \quad \forall x \iff E(B \cdot A) = \emptyset$$

<u>DEMOSTRACION:</u>

$$E(A \cdot x \cdot B) = E\ tr(A \cdot x \cdot B) = E\ tr(x \cdot B \cdot A) =$$
$$= tr\ E(x \cdot B \cdot A) = tr(x \cdot E(B \cdot A))$$

Si $E(B \cdot A) = \emptyset$, entonces $E(A \cdot x \cdot B) = \emptyset$

Por otro lado, si $E(A \cdot x \cdot B) = \emptyset \quad \forall x$, entonces se sigue que $E(B \cdot A) = \emptyset$.

•) Distancias en L_2^n :

A una matriz p de dimensión (n,n), simétrica y definida positiva, vamos a asociar un producto escalar en L_2^n definido por:

$$\langle x,y \rangle = E(x' \cdot p \cdot y) \quad \text{siendo} \quad x,y \in L_2^n$$

La relación de ortogonalidad correspondiente se define como:

$$x \perp y \iff E(x' \cdot p \cdot y) = \emptyset \qquad x,y \in L_2^n$$

Y por último, la distancia asociada entre los puntos $x,y \in L_2^n$ es:

$$E[(y - x)' \cdot p \cdot (y - x)]^{1/2}$$

•) $\boxed{\text{Distancia de optimización en } L_2^n}$:

Sea A un punto fijado en L_2^n , y x un punto que se extiende sobre una superficie fijada S en L_2^n . Entonces el punto $x \in S$ que minimiza la distancia a A cumple:

$$(x - A) \perp dx$$

para todo desplazamiento infinitesimal dx en S.

3.2.3.1.- Cálculo del estimador individual para $\beta(\theta_j)$.

Una vez introducidos los elementos anteriores pasemos al problema que nos ocupa. De VYLDER, F. (1985) antes de atacar el problema de la aproximación credibilística para $\beta(\theta_j)$, centró su atención en el cálculo de su estimador individual, que nosotros hemos simbolizado por $\hat{\beta}_j$, y que es uno de los elementos que aparece en la fórmula del estimador ajustado de credibilidad.

El problema que plantea es el siguiente:

Sea $x \in R^t$ y sea f una función $f: R^n \longrightarrow R^t$, con $n \leq t$. El objetivo es encontrar un punto $b \in R^n$ tal que f(b) sea lo más próximo posible al punto dado x. La distancia es medida a través de una matriz arbitraria p, de dimensión (t,t), pero simétrica y definida positiva.

El punto óptimo $b = h(x)$ depende de x, siendo h una función definida como $h: R^t \longrightarrow R^n$, de manera que:

$$h[f(b)] = b \qquad \forall b \in R^n$$

ya que el punto $b_\emptyset$ tal que $f(b_\emptyset)$ es el más próximo a $x = f(b)$ es el propio b.

La solución que da De VYLDER, F. a este problema es el siguiente:

Sea $b = (b_1, b_2, \ldots, b_n)'$. Vamos a asumir que las derivadas parciales de $f(b)$, con $b \in R^n$, respecto a b_i $(i = 1,2,\ldots,n)$ existen. Entonces:

$$df(b) = \sum_{i=1}^{n} \frac{\delta f}{\delta b_i} \cdot db_i = g(b) \cdot db \qquad (b \in R^n)$$

donde $g(b)$ es la matriz Jacobiana de f, de las derivadas parciales puestas en columnas, de dimensión (t,n).

A continuación se introduce el siguiente teorema:

TEOREMA 1 :

El punto $b = h(x) \in R^n$ tal que $f(b)$ es el más próximo a $x \in R^t$ es tal que:

$$\boxed{g'(b) \cdot p \cdot x = g'(b) \cdot p \cdot f(b)}$$

Esto es, es el punto que hace que la distancia entre $g(b)$ y x coincida con las distancia entre $g(b)$ y $f(b)$.

DEMOSTRACION:

Según la definición de distancia de optimización en R^n, el punto b es tal que:

$$x - f(b) \perp df(b) \qquad \forall db$$

$$x - f(b) \perp g(b) \cdot db \qquad \forall db$$

es decir, tal que:

$$db' \cdot g'(b) \cdot p \cdot (x - f(b)) = 0 \qquad \forall db$$

Según el LEMA 1:

$$a \cdot x \cdot b = 0 \quad \forall x \quad \Longleftrightarrow \quad b \cdot a = 0$$

Si consideramos que $a = I_{(n,n)}$ resulta que:

$$db' \cdot g'(b) \cdot p \cdot (x - f(b)) = 0 \qquad \forall db$$

se verifica, si:

$$g'(b) \cdot p \cdot (x - f(b)) = 0$$

Si adaptamos este problema a nuestro caso particular, es decir, al cálculo del estimador individual para $\beta(\theta_j)$, que es el vector $\hat{\beta}_j$, tenemos:

Dado $X_j \in R^t$, y $f: R^n \longrightarrow R^t$ con $n \leq t$, estamos interesados en hallar un punto $\hat{\beta}_j \in R^n$ tal que $f(\hat{\beta}_j)$ sea lo más próximo posible a X_j. En este caso, la distancia es medida a través de la matriz v_j^{-1}, que es elegida por razones de tipo credibilístico, siendo una matriz simétrica, definida positiva, y de dimensión (t,t).

El punto $\hat{\beta}_j = h(X_j) \in R^n$ tal que $f(\hat{\beta}_j)$ es el más próximo a $X_j \in R^t$ es tal que:

$$g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot X_j = g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot f(\hat{\beta}_j)$$

o dicho de otra manera, el estimador individual para $\beta(\theta_j)$ es el vector $\hat{\beta}_j$ de dimensión $(n,1)$, tal que verifica:

$$\boxed{g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot (X_j - f(\hat{\beta}_j)) = 0}$$

3.2.3.2.- Cálculo del estimador individual para $\beta(\theta_j)$ en el caso particular del Modelo de Regresión Exponencial.

Como ya hemos dicho anteriormente, el Modelo de Regresión Exponencial es uno de los más utilizados dentro de los modelos de regresión no-lineales, y aparece en el caso particular en que $n = 2$. En este caso, la función f viene definida del siguiente modo:

$$f : R^2 \longrightarrow R^t$$

$$f(b_{j1}, b_{j2}) = \begin{bmatrix} b_{j1} \\ b_{j1} \cdot b_{j2} \\ b_{j1} \cdot b_{j2}^2 \\ \vdots \\ b_{j1} \cdot b_{j2}^{t-1} \end{bmatrix}$$

siendo la prima de riesgo individual para la póliza j-ésima:

$$E[X_j/\theta_j] = b_{j1} \cdot b_{j2}^{s-1} \quad \text{con } s = 1,2,\ldots,t$$

que es un vector de dimensión $(t,1)$.

Nuestro objetivo es estimar la prima de riesgo individual, hallando una aproximación para $\beta(\theta_j) = (b_{j1}, b_{j2})$ de tipo credibilístico.

Como ya hemos dicho, el primer paso consiste en hallar el estimador individual para $\beta(\theta_j)$, que es el vector $\hat{\beta}_j$, tal que verifica:

$$g'(\hat{\beta}_j) \cdot v_j^{-1} \cdot (X_j - f(\hat{\beta}_j)) = 0$$

En este caso:

$$\hat{\beta}_j = (\hat{b}_{j1}, \hat{b}_{j2})$$

$$f(\hat{\beta}_j) = \hat{b}_{j1} \cdot \hat{b}_{j2}^{s-1} \quad \text{con } s = 1,2,\ldots,t$$

$$g(\hat{b}_{j1}, \hat{b}_{j2}) = \left[\frac{\delta f}{\delta \hat{b}_{j1}} \quad \frac{\delta f}{\delta \hat{b}_{j2}} \right] =$$

$$= \begin{bmatrix} 1 & \emptyset \\ \hat{b}_{j2} & \hat{b}_{j1} \\ \hat{b}_{j2}^{2} & 2\cdot\hat{b}_{j1}\cdot\hat{b}_{j2} \\ \vdots & \\ \hat{b}_{j2}^{t-1} & (t-1)\cdot\hat{b}_{j1}\cdot\hat{b}_{j2}^{t-2} \end{bmatrix}$$

Sustituyendo cada uno de estos elementos en:

$$g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot(X_j - f(\hat{\beta}_j)) = \emptyset$$

obtenemos un sistema de dos ecuaciones con dos incógnitas, sistema que se puede simplificar considerablemente si asumimos que:

$$v_j = \text{diag}\left(\frac{1}{w_{j1}}, \frac{1}{w_{j2}}, \ldots, \frac{1}{w_{jt}}\right)$$

esto es, si asumimos que no existe correlación entre las observaciones referentes a periodos distintos de tiempo, y que la varianza es inversamente proporcional a las ponderaciones para cada periodo.

El sistema de ecuaciones que resulta es:

$$\sum_{s=1}^{t} X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{s-1} = \hat{b}_{j1}\cdot\sum_{s=1}^{t} w_{js}\cdot\hat{b}_{j2}^{2\cdot(s-1)} \qquad (1)$$

$$\hat{b}_{j1}\cdot\sum_{s=1}^{t}(s-1)\cdot X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{\,s-2}=\hat{b}_{j1}^{\,2}\cdot\sum_{s=1}^{t}(s-1)\cdot w_{js}\cdot b_{j2}^{\,2s-3} \qquad (2)$$

De la primera ecuación podemos despejar $\hat{b}_{j1}$, que es igual:

$$\boxed{\hat{b}_{j1}=\frac{\sum_{s=1}^{t}X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{\,s}}{\sum_{s=1}^{t}w_{js}\cdot\hat{b}_{j2}^{\,2s-1}}}$$

Si en la segunda ecuación simplificamos, obtenemos:

$$\sum_{s=1}^{t}(s-1)\cdot X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{\,s}=\hat{b}_{j1}\cdot\sum_{s=1}^{t}(s-1)\cdot w_{js}\cdot\hat{b}_{j2}^{\,2s-1}$$

y sustituyendo $\hat{b}_{j1}$ por lo que vale:

$$\sum_{s=1}^{t}(s-1)\cdot X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{\,s}=$$

$$=\frac{\sum_{s=1}^{t}X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{\,s}}{\sum_{s=1}^{t}w_{js}\cdot\hat{b}_{j2}^{\,2s-1}}\cdot\left[\sum_{s=1}^{t}(s-1)\cdot w_{js}\cdot\hat{b}_{j2}^{\,2s-1}\right]$$

resultando:

$$\left[\sum_{s=1}^{t} w_{js}\cdot\hat{b}_{j2}^{2s-1}\right]\cdot\left[\sum_{s=1}^{t}(s-1)\cdot X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{s}\right]=$$

$$=\left[\sum_{s=1}^{t} X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{s}\right]\cdot\left[\sum_{s=1}^{t}(s-1)\cdot w_{js}\cdot\hat{b}_{j2}^{2s-1}\right]$$

Si multiplicamos los dos miembros de la igualdad anterior por $\hat{b}_{j2}$ nos queda:

$$\left[\sum_{s=1}^{t} w_{js}\cdot\hat{b}_{j2}^{2s}\right]\cdot\left[\sum_{s=1}^{t}(s-1)\cdot X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{s}\right]=$$

$$=\left[\sum_{s=1}^{t} X_{js}\cdot w_{js}\cdot\hat{b}_{j2}^{s}\right]\cdot\left[\sum_{s=1}^{t}(s-1)\cdot w_{js}\cdot\hat{b}_{j2}^{2s}\right]$$

de donde obtendremos $\hat{b}_{j2}$.

3.2.3.3.- Cálculo del estimador ajustado de credibilidad para $\beta(\theta_j)$.

Una vez hallado el estimador individual para $\beta(\theta_j)$ pasemos a calcular su aproximación credibilística.

Según De VYLDER, F. (1985), lo que se pretende es mejorar la aproximación dada por $\hat{\beta}_j$ para $\beta(\theta_j)$, reemplazándola por su aproximación credibilística $\hat{\beta}(\theta_j)$:

$$\hat{\beta}(\theta_j) = Z_j \cdot \hat{\beta}_j + [I - Z_j] \cdot \beta$$

donde $\hat{\beta}(\theta_j)$ es el llamado **estimador ajustado de credibilidad.**

El problema reside en **hallar la matriz Z_j, que es la llamada matriz de credibilidad**, de dimensión (n,n), matriz que deberá ser fijada a través de un camino óptimo.

Para obtener dicha matriz De VYLDER, F. (1985) introduce el siguiente teorema:

TEOREMA 2 :

La matriz Z_j, de dimensión (n,n) tal que hace que:

$$\hat{\beta}(\theta_j) = Z_j \cdot \hat{\beta}_j + [I - Z_j] \cdot \beta$$

sea el estimador credibilístico más próximo a $\beta(\theta_j)$ es:

$$\boxed{Z_j = E[\beta^o(\theta_j) \cdot \hat{\beta}^{o\prime}_j] \cdot E^{-1}[\hat{\beta}^o_j \cdot \hat{\beta}^{o\prime}_j]}$$

donde:

$$\hat{\beta}^o_j = \hat{\beta}_j - \beta$$

$$\beta^o(\theta_j) = \beta(\theta_j) - \beta$$

DEMOSTRACION:

Según la definición de distancia de optimización en L^n_2 , la matriz óptima Z_j es aquella tal que para todo dZ_j:

$$Z_j \cdot \hat{\beta}_j + [I - Z_j] \cdot \beta - \beta(\theta_j) \perp d(Z_j \cdot \beta_j + [I - Z_j] \cdot \beta)$$

o sea:

$$Z_j \cdot \hat{\beta}_j^o - \beta(\theta_j) \perp dZ_j \cdot \hat{\beta}_j^o$$

es decir:

$$E[\hat{\beta}_j^{o\prime} \cdot dZ_j' \cdot q \cdot (Z_j \cdot \hat{\beta}_j^o - \beta^o(\theta_j))] = 0$$

donde q es una matriz simétrica, definida positiva, de dimensión (n,n), que utilizamos para medir las distancias.

Utilizando el LEMA 2, resulta que esta última condición equivale a decir que, para todo dZ_j:

$$E[q \cdot (Z_j \cdot \hat{\beta}_j^o - \beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime})] = 0$$

Como la matriz q es una matriz conocida, resulta que:

$$q \cdot E[(Z_j \cdot \hat{\beta}_j^o - \beta^o(\theta_j)) \cdot \hat{\beta}_j^{o\prime}] = 0$$

o lo que es lo mismo:

$$q \cdot E[Z_j \cdot \hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime} - \beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime}] = 0$$

Al ser la esperanza de una resta, diferencia de esperanzas:

$$q \cdot E[Z_j \cdot \hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime}] - q \cdot E[\beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime}] = \emptyset$$

o de forma análoga:

$$E[Z_j \cdot \hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime}] = E[\beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime}]$$

Como estamos calculando esperanzas respecto a un parámetro de riesgo determinado, resulta que Z_j es una constante, verificándose:

$$Z_j \cdot E[\hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime}] = E[\beta^o(\theta_j) \cdot \hat{\beta}^{o\prime}].$$

de donde:

$$Z_j = E[\beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime}] \cdot E^{-1}[\hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime}]$$

quedando de este modo probado el teorema anterior.

Cabe destacar que la obtención de la matriz de credibilidad óptima, Z_j, no depende de la matriz q, que es la matriz que hemos utilizado para medir las distancias.

Como acabamos de ver, la matriz de credibilidad Z_j, de dimensión (n,n), que hace que $\hat{\beta}(\theta_j)$ sea el estimador credibilístico más próximo a $\beta(\theta_j)$ es:

$$\boxed{Z_j = E[\beta^o(\theta_j) \cdot \hat{\beta}_j^{o\prime}] \cdot E^{-1}[\hat{\beta}_j^o \cdot \hat{\beta}_j^{o\prime}]}$$

o lo que es lo mismo:

$$Z_j = E[(\beta(\theta_j) - \beta)\cdot(\hat{\beta}_j - \beta)']\cdot E^{-1}[(\hat{\beta}_j - \beta)\cdot(\hat{\beta}_j - \beta)']$$

que a su vez resulta ser:

$$\boxed{Z_j = \mathrm{Cov}[\beta(\theta_j),\ \hat{\beta}_j]\cdot \mathrm{Cov}^{-1}[\hat{\beta}_j]}$$

Si calculamos las dos covarianzas por separado, obtenemos por un lado:

$$\mathrm{Cov}[\beta(\theta_j),\ \hat{\beta}_j] = E[\mathrm{Cov}[\beta(\theta_j),\ \hat{\beta}_j/\theta_j]] +$$

$$+ \mathrm{Cov}[E[\beta(\theta_j)/\theta_j],\ E[\hat{\beta}_j/\theta_j]] = \mathrm{Cov}[\beta(\theta_j)]$$

ya que $\mathrm{Cov}[\beta(\theta_j),\ \hat{\beta}_j/\theta_j] = 0$, puesto que para un parámetro de riesgo fijo $\beta(\theta_j)$ es una constante, y $E[\hat{\beta}_j/\theta_j] = \beta(\theta_j)$, pues $\hat{\beta}_j$ es un estimador insesgado de $\beta(\theta_j)$, siendo $\mathrm{Cov}[\beta(\theta_j)]$ uno de los parámetros estructurales de este modelo, que más tarde estimaremos y simbolizaremos por Γ.

Y por otro lado:

$$\mathrm{Cov}[\hat{\beta}_j] = E[\mathrm{Cov}[\hat{\beta}_j/\theta_j]] + \mathrm{Cov}[E[\hat{\beta}_j/\theta_j]] =$$

$$= E[\mathrm{Cov}[\hat{\beta}_j/\theta_j]] + \mathrm{Cov}[\beta(\theta_j)]$$

Una vez que hemos llegado a este punto, resulta que no podemos calcular $\mathrm{Cov}[\hat{\beta}_j/\theta_j]$, ya que no conocemos la expresión concreta que adopta el estimador individual para $\beta(\theta_j)$, sólo sabemos que $\hat{\beta}_j$ verifica la ecuación:

$$g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot(X_j - f(\hat{\beta}_j)) = 0$$

y de la misma no podemos despejar $\hat{\beta}_j$, excepto en el caso que adoptemos la hipótesis:

$$f[\beta(\theta_j)] = Y_j\cdot\beta(\theta_j)$$

que es el caso particular del Modelo de Regresión de HACHEMEISTER. Con esta asunción se obtiene que $\hat{\beta}_j$ es un estimador mínimo-cuadrático generalizado, definido como sigue:

$$\hat{\beta}_j = (Y_j'\cdot v_j^{-1}\cdot Y_j)\cdot Y_j'\cdot v_j^{-1}\cdot X_j$$

Para solucionar este problema, De VYLDER, F. (1985) propuso **buscar la mejor aproximación polinómica al modelo de regresión general.**

3.2.3.4.- Aproximación polinómica al Modelo de Regresión No-Lineal.

El objetivo de De VYLDER, F. (1985) es hallar la mejor aproximación polinómica para el modelo general, es decir, la mejor aproximación del tipo:

$$E[X_j/\theta_j] = X_0 + Y\cdot\beta(\theta_j)$$

donde:

$X_{\emptyset}$: es un vector de dimensión (t,1), de términos independientes que actúan como un cambio de origen en la medida de las variables X_{js}.

Y : es la matriz que ya apareció en el Modelo de Regresión de HACHEMEISTER, de dimensión, en este caso, (t,n).

$\beta(\theta_j)$: es el vector de coeficientes de regresión desconocido de dimensión (n,1).

Para hallar la mejor aproximación polinómica se introduce el siguente teorema:

TEOREMA 3 :

En el modelo polinómico más próximo al modelo de regresión general, tenemos que:

$$\boxed{Y = \mathrm{Cov}[f(\beta(\theta_j)),\ \beta'(\theta_j)]\cdot \mathrm{Cov}^{-1}[\beta(\theta_j)] = d\cdot \Gamma^{-1}}$$

siendo:

$$\boxed{X_{\emptyset} = E[f(\beta(\theta_j))] - Y\cdot E[\beta(\theta_j)]}$$

DEMOSTRACION:

Estamos interesados en hallar el vector aleatorio $X_{\emptyset} + Y\cdot\beta(\theta_j)$ más próximo a $f(\beta(\theta_j))$ en L_2^t. Para ello vamos a medir las distancias con una matriz arbitraria, p, de dimensión (t,t).

Según el concepto de distancia de optimización en L_2^t visto anteriormente, los $X_\emptyset$ e Y óptimos son tales que para todo $dX_\emptyset$, dY:

$$X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)) \perp d(X_\emptyset + Y\cdot\beta(\theta_j))$$

o lo que es lo mismo:

$$X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)) \perp dX_\emptyset + dY\cdot\beta(\theta_j)$$

es decir:

$$E[(\beta'(\theta_j)\cdot dY' + dX_\emptyset')\cdot p\cdot(X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)))] = \emptyset$$

Como $[\beta'(\theta_j)\cdot dY' + dX_\emptyset']$, y $[X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j))]$ son matrices aleatorias y p es una matriz tal que:

$$[\beta'(\theta_j)\cdot dY' + dX_\emptyset')\cdot p\cdot(X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j))]$$

es un escalar, por el LEMA 2 sabemos que:

$$E[(\beta'(\theta_j)\cdot dY + dX_\emptyset')\cdot p\cdot(X_\emptyset + Y\cdot\beta(\theta_j)] = \emptyset \qquad \forall p \quad \Longleftrightarrow$$

$$\Longleftrightarrow \quad E[(X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j))\cdot(\beta'(\theta_j)\cdot dY' + dX_\emptyset)] = \emptyset$$

Si consideramos $dY = \emptyset$ en la esperanza anterior, resulta:

$$E[(X_\emptyset + Y\cdot\beta(\theta_j) - f(\beta(\theta_j))\cdot dX_\emptyset] = \emptyset$$

y al ser dX'_0 una constante respecto a θ_j:

$$E[(X_0 + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)] = 0$$

o lo que es lo mismo:

$$X_0 + Y\cdot E[\beta(\theta_j)] = E[f(\beta(\theta_j))]$$

siendo:

$$\boxed{X_0 = E[f(\beta(\theta_j))] - Y\cdot E[\beta(\theta_j)]}$$

Por otro lado, si consideramos $dX_0 = 0$ en la esperanza anterior obtenemos:

$$E[(X_0 + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)))\cdot\beta'(\theta_j)\cdot dY'] = 0$$

Como dY es una constante respecto de θ_j:

$$E[(X_0 + Y\cdot\beta(\theta_j) - f(\beta(\theta_j)))\cdot\beta'(\theta_j)] = 0$$

es decir:

$$\boxed{E[(X_0 + Y\cdot\beta(\theta_j))\cdot\beta'(\theta_j)] = E[f(\beta(\theta_j))\cdot\beta'(\theta_j)]}$$

Si multiplicamos $[X_0 + Y\cdot E[\beta(\theta_j)]]$ por $E[\beta'(\theta_j)]$, y restamos esta expresión de la anterior nos queda:

$$E[(X_0 + Y\cdot\beta(\theta_j))\cdot\beta'(\theta_j)] - E[X_0 + Y\cdot\beta(\theta_j)]\cdot E[\beta'(\theta_j)] =$$

$$= \mathrm{Cov}[X_\emptyset + Y\cdot\beta(\theta_j),\ \beta'(\theta_j)]$$

A su vez, esta covarianza es igual:

$$\mathrm{Cov}[X_\emptyset + Y\cdot\beta(\theta_j),\ \beta'(\theta_j)] = E[f(\beta(\theta_j))\cdot\beta(\theta_j)] -$$

$$- E[f(\beta(\theta_j))]\cdot E[\beta'(\theta_j)] = \mathrm{Cov}[f(\beta(\theta_j)),\ \beta'(\theta_j)]$$

ya que:

$$E[(X_\emptyset + Y\cdot\beta(\theta_j))\cdot\beta'(\theta_j)] = E[f(\beta(\theta_j))\cdot\beta'(\theta_j)]$$

y

$$E[X_\emptyset + Y\cdot\beta(\theta_j)] = E[f(\beta(\theta_j))]$$

Si aplicamos las propiedades de la covarianza resulta que:

$$\mathrm{Cov}[X_\emptyset + Y\cdot\beta(\theta_j),\ \beta'(\theta_j)] = \mathrm{Cov}[Y\cdot\beta(\theta_j),\ \beta'(\theta_j)] =$$

$$= Y\cdot\mathrm{Cov}[\beta(\theta_j),\ \beta'(\theta_j)] = Y\cdot\mathrm{Cov}[\beta(\theta_j)]$$

Por otra parte, como:

$$\mathrm{Cov}[X_\emptyset + Y\cdot\beta(\theta_j),\ \beta'(\theta_j)] = \mathrm{Cov}[f(\beta(\theta_j),\ \beta'(\theta_j)]$$

obtenemos que:

$$\boxed{Y = \mathrm{Cov}[f(\beta(\theta_j)),\ \beta'(\theta_j)]\cdot\mathrm{Cov}^{-1}[\beta(\theta_j)] = d\cdot\Gamma^{-1}}$$

siendo:

$$d = \mathrm{Cov}[f(\beta(\theta_j)),\ \beta'(\theta_j)]$$

$$\Gamma = \mathrm{Cov}[\beta'(\theta_j)]$$

quedando demostrado el teorema anterior.

A partir del sistema de ecuaciones anterior podemos hallar los coeficientes $\beta(\theta_j)$, ya que los hemos elegido de tal manera que:

$$\mathrm{Cov}[f(\beta(\theta_j)),\ \beta'(\theta_j)]\cdot\mathrm{Cov}^{-1}[\beta(\theta_j)]$$

sea igual a la matriz Y dada.

Una vez obtenida la mejor aproximación polinómica para el modelo de regresión general, podemos retomar el problema que nos ocupa, la estimación de la matriz de credibilidad Z_j.

Si en el TEOREMA 1 introducimos la mejor aproximación polinómica, esto es si consideramos que:

$$f(\beta(\theta_j)) = X_\emptyset + Y\cdot\beta(\theta_j)$$

verificándose el sistema de ecuaciones:

$$Y = d\cdot\Gamma^{-1}$$

con:

$$X_\emptyset = E[f(\hat{\beta}_j)] - Y\cdot E[\hat{\beta}_j]$$

obtenemos que $\hat{\beta}_j$ es el estimador mínimo-cuadrático generalizado:

$$\hat{\beta}_j = (Y' \cdot p \cdot Y)^{-1} \cdot Y' \cdot p \cdot X_j$$

o de forma análoga:

$$\hat{\beta}_j = (Y' \cdot v_j^{-1} \cdot Y)^{-1} \cdot Y' \cdot v_j^{-1} \cdot X_j$$

si la matriz arbitraria que utilizamos para medir las distancias es la matriz v_j^{-1}.

Recordemos que la matriz de credibilidad Z_j, es aquella que verifica:

$$Z_j = \mathrm{Cov}[\beta(\theta_j), \hat{\beta}_j] \cdot \mathrm{Cov}^{-1}[\hat{\beta}_j]$$

siendo:

$$\mathrm{Cov}[\beta(\theta_j), \hat{\beta}_j] = \mathrm{Cov}[\beta(\theta_j)] = \Gamma$$

$$\mathrm{Cov}[\hat{\beta}_j] = E[\mathrm{Cov}[\hat{\beta}_j/\theta_j]] + \Gamma$$

En este caso, sí que podremos calcular la $\mathrm{Cov}[\hat{\beta}_j/\theta_j]$ puesto que conocemos la expresión que adopta el estimador $\hat{\beta}_j$.

$$\mathrm{Cov}[\hat{\beta}_j/\theta_j] = \mathrm{Cov}[(Y' \cdot v_j^{-1} \cdot Y)^{-1} \cdot Y \cdot v_j^{-1} \cdot X_j/\theta_j] =$$

$$= (Y' \cdot v_j^{-1} \cdot Y)^{-1} \cdot Y' \cdot v_j \cdot \mathrm{Cov}[X_j/\theta_j] \cdot v_j^{-1} \cdot Y \cdot (Y' \cdot v_j^{-1} \cdot Y)^{-1} =$$

$$= (Y' \cdot v_j^{-1} \cdot Y)^{-1} \cdot Y' \cdot v_j \cdot \sigma^2(\theta_j) \cdot v_j \cdot v_j^{-1} \cdot Y \cdot (Y' \cdot v_j^{-1} \cdot Y)^{-1} =$$

$$= \sigma^2(\theta_j)\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

y su esperanza:

$$E[Cov[\hat{\beta}_j/\theta_j]] = S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

Sustituyendo las expresiones de $Cov[\beta(\theta_j),\ \hat{\beta}_j]$ y $Cov[\hat{\beta}_j]$ en la fórmula de la matriz de credibilidad Z_j, que al considerar su aproximación polinómica simbolizaremos por Z_j^A, nos queda:

$$Z_j^A = \Gamma\cdot(\Gamma + S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1})^{-1}$$

Pero a su vez, como $Y = d\cdot\Gamma^{-1}$:

$$Z_j^A = \Gamma\cdot\left[\Gamma + S^2\cdot\left[\left[d\cdot\Gamma^{-1}\right]'\cdot v_j^{-1}\cdot d\cdot\Gamma^{-1}\right]^{-1}\right]^{-1} =$$

$$= \Gamma\cdot\left[\Gamma + S^2\cdot\left[(\Gamma')^{-1}\cdot d'\cdot v_j^{-1}\cdot d\cdot\Gamma^{-1}\right]^{-1}\right]^{-1} =$$

$$= \Gamma\cdot\left[\Gamma + S^2\cdot\Gamma\cdot\left[d'\cdot v_j^{-1}\cdot d\right]^{-1}\cdot\Gamma'\right]^{-1}$$

Si simbolizamos por $d_j = (d'\cdot v_j\cdot d)^{-1}$ resulta que:

$$Z_j^A = \Gamma\cdot\left[\Gamma + S^2\cdot\Gamma\cdot d_j\cdot\Gamma'\right]^{-1}$$

y como Γ es una matriz simétrica ($\Gamma = \Gamma'$) obtenemos:

$$Z_j^A = \Gamma\cdot\left[\Gamma + S^2\cdot\Gamma\cdot d_j\cdot\Gamma'\right]^{-1} = \Gamma\cdot\left[(I + S^2\cdot\Gamma\cdot d_j\cdot)\cdot\Gamma\right]^{-1} =$$

$$= \Gamma\cdot\Gamma^{-1}\cdot\left[I + S^2\cdot\Gamma\cdot d_j\right]^{-1}$$

siendo:

$$\boxed{Z_j^A = \left[I + S^2\cdot\Gamma\cdot d_j\right]^{-1}}$$

donde:

$$S^2 = E[\sigma^2(\theta_j)]$$

$$\Gamma = Cov[\beta'(\theta_j)]$$

$$d_j = (d'\cdot v_j\cdot d)^{-1}$$

Es importante remarcar que para obtener una expresión para la matriz de credibilidad ha sido necesario utilizar la aproximación polinómica para el modelo de regresión general, tratándose, por lo tanto, de una aproximación.

3.2.3.4.1.- Matrices de credibilidad diagonales.

Por razones prácticas, según indica De VYLDER, F. (1985), puede ser útil restringir las matrices de credibilidad, y él las restringió a ser diagonales. Para poder calcular la matriz de credibilidad diagonal, que simbolizaremos por Z_j^d, que hace que $\hat{\beta}(\theta_j)$ sea lo más próximo a $\beta(\theta_j)$, es necesario introducir dos nuevos lemas que aparecen en el artículo del citado autor.

Vamos a simbolizar por a_{ij} el elemento de la fila i y de la columna j de la matriz a. Si a es una matriz cuadrada:

$$a^d = (a_{11},\ldots,a_{nn})'$$

simboliza el vector columna formado por los elementos de la diagonal principal de a. Y por último, si a y b son matrices con la misma dimensión, entonces $c = a*b$ simboliza la matriz con elementos:

$$c_{ij} = a_{ij} \cdot b_{ij}$$

tratándose del producto de matrices calculado elemento a elemento.

Los dos nuevos lemas que introduce De VYLDER, F. (1985), son:

$\boxed{\text{LEMA 3}}$:

Si $a \cdot x \cdot b$ es una matriz cuadrada y x una matriz diagonal, entonces:

$$(a \cdot x \cdot b)^d = (a*b') \cdot x^d$$

<u>DEMOSTRACION</u>:

El elemento i-ésimo de $(a \cdot x \cdot b)^d$ es igual:

$$(a \cdot x \cdot d)_i^d = (a \cdot x \cdot b)_{ii} = \sum_j \sum_k a_{ij} \cdot x_{jk} \cdot b_{ki}$$

Pero al ser x diagonal, queda simplemente:

$$(a\cdot x\cdot d)_i^d = \sum_j a_{ij}\cdot x_{jj}\cdot b_{ji} = \sum_j a_{ij}\cdot b'_{ij}\cdot x_{jj} =$$

$$= \sum_j (a*b')_{ij}\cdot x_j^d = ((a*b')\cdot x^d)_i$$

quedando demostrado el lema anterior.

$\boxed{\text{LEMA 4}}$:

Si A, B son matrices aleatorias y x es una matriz tal que $A\cdot x\cdot B$ es una matriz de dimensión (1,1), entonces:

$$E(A\cdot x\cdot B) = 0 \quad \forall x \text{ diagonal} \quad \Longleftrightarrow \quad E^d(B\cdot A) = 0$$

<u>DEMOSTRACION:</u>

Por las propiedades de la traza de una matriz:

$$E(A\cdot x\cdot B) = E\ tr(A\cdot x\cdot B) = E\ tr\ (x\cdot B\cdot A)$$

Como la esperanza de una suma es suma de esperanzas resulta:

$$E\ tr(x\cdot B\cdot A) = tr(x\cdot E(B\cdot A)) = tr(x\cdot c)$$

siendo $c = E(B\cdot A)$

A su vez:

$$tr(x\cdot c) = \sum_i (x\cdot c)_{ii} = \sum_i \sum_j x_{ij}\cdot c_{ij} = \sum_i x_{ii}\cdot c_{ii}$$

De manera que:

$$E(A \cdot x \cdot B) = \emptyset \qquad \forall x \text{ diagonal si } c_{ii} = \emptyset \qquad \forall i,$$

o lo que es lo mismo si $E^d(B \cdot A) = \emptyset$

Una vez introducidos estos dos lemas, De VYLDER, F. (1985) propone el siguiente teorema:

$\boxed{\text{TEOREMA 4}}$:

La matriz diagonal Z_j^d, de dimensión (n,n), tal que:

$$\hat{\beta}(\theta_j) = Z_j^d \cdot \hat{\beta}_j + [I - Z_j^d] \cdot \beta$$

es la más próxima a $\beta(\theta_j)$, es aquella que verifica:

$$\boxed{Z_j^d = \left[q * E[\hat{\beta}_j^o \cdot \hat{\beta}_j^o{}']\right]^{-1} \cdot \left[q \cdot E[\beta^o(\theta_j) \cdot \hat{\beta}_j^o{}']\right]^d}$$

<u>DEMOSTRACION:</u>

Por la definición de distancia de optimización en L_2^n, la matriz óptima Z_j^d es aquella que para todo dZ_j:

$$Z_j^d \cdot \hat{\beta}_j + [I - Z_j^d] \cdot \beta - \beta(\theta_j) \perp d(Z_j^d \cdot \hat{\beta}_j + [I - Z_j^d] \cdot \beta)$$

o

$$Z_j^d \cdot \hat{\beta}_j^o - \beta^o(\theta_j) \perp dZ_j^d \cdot \hat{\beta}_j^o$$

o

$$E[\hat{\beta}_j^o{}' \cdot dZ_j^d{}' \cdot q \cdot (Z_j^d \cdot \hat{\beta}_j^o - \beta^o(\theta_j))] = \emptyset$$

Por el LEMA 4, esto es lo mismo que:

$$\left[E\left[p\cdot(Z_j^d\cdot\hat{\beta}_j^o - \beta^o(\theta_j))\right]\cdot\hat{\beta}_j^{o\prime}\right]^d = 0$$

o que:

$$[q\cdot Z_j^d\cdot E[\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]^d = [q\cdot E^o[\beta(\theta_j)\cdot\hat{\beta}_j^{o\prime}]]^d$$

ya que para un parámetro de riesgo determinado Z_j^d es una constante.

Como $[q\cdot Z_j^d\cdot E[\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]$ es una matriz cuadrada de dimensión (n,n), y Z_j^d es una matriz diagonal (recordemos que hemos restringido las matrices de credibilidad a ser diagonales), podemos aplicar el LEMA 3, verificándose:

$$[q\cdot Z_j^d\cdot E[\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]^d = [q*E]\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]\cdot Z_j^d$$

y a su vez:

$$[q*E[\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]\cdot Z_j^d = [q\cdot E[\beta^o(\theta_j)\cdot\hat{\beta}_j^{o\prime}]]^d$$

de donde Z_j^d resulta que es igual a:

$$Z_j^d = [q*E[\hat{\beta}_j^o\cdot\hat{\beta}_j^{o\prime}]]^{-1}\cdot[q\cdot E[\beta^o(\theta_j)\cdot\hat{\beta}_j^{o\prime}]]^d$$

quedando demostrado el TEOREMA 3.

Cabe destacar, que cuando imponemos que la matriz de credibilidad sea diagonal, entonces dicha matriz Z_j^d depende de la matriz q que hemos introducido para medir las distancias, cosa que no ocurría al no imponer en el caso general la restricción de diagonalidad en la matriz de credibilidad.

Como:

$$\hat{\beta}_j^o = \hat{\beta}_j - \beta$$

$$\beta^o(\theta_j) = \beta(\theta_j) - \beta$$

la matriz Z_j^d la podemos también escribir del siguiente modo:

$$Z_j^d = [q*(E[(\hat{\beta}_j - \beta)\cdot(\hat{\beta}_j - \beta)']]^{-1}\cdot[q\cdot E[(\beta(\theta_j) - \beta)\cdot(\hat{\beta}_j - \beta)']]^d$$

$$\boxed{Z_j^d = [q*Cov[\hat{\beta}_j]]^{-1}\cdot[q\cdot Cov[\beta(\theta_j),\ \hat{\beta}_j)]^d}$$

siendo:

$$Cov[\beta(\theta_j),\ \hat{\beta}_j] = E[Cov[\beta(\theta_j),\ \hat{\beta}_j/\theta_j]] +$$

$$+ Cov[E(\beta(\theta_j)/\theta_j),\ E(\hat{\beta}_j/\theta_j)] = \emptyset + Cov[\beta(\theta_j)] = \Gamma$$

tratándose de uno de los parámetros estructurales del modelo, y

$$Cov[\hat{\beta}_j] = E[Cov[\hat{\beta}_j/\theta_j]] + Cov[E(\hat{\beta}_j/\theta_j)] =$$

$$= E[Cov[\hat{\beta}_j/\theta_j]] + Cov[\beta(\theta_j)] =$$

$$= E[Cov[\hat{\beta}_j/\theta_j]] + \Gamma$$

Como en el caso anterior, nos volvemos a encontrar con la problemática del cálculo de la $Cov[\hat{\beta}_j/\theta_j]$, imprescindible para hallar la matriz de credibilidad Z_j^d.

Una aproximación de la misma la podemos obtener si utilizamos la mejor aproximación polinómica al modelo de regresión general. Por el TEOREMA 3, sabemos que en el modelo polinómico más próximo al modelo de regresión general:

$$Y = Cov[f(\beta(\theta_j)),\ \beta'(\theta_j)]\cdot Cov^{-1}[\beta(\theta_j)] = d\cdot\Gamma^{-1}$$

y

$$X_\emptyset = E[f(\beta(\theta_j))] - Y\cdot E[\beta(\theta_j)]$$

Repitiendo el mismo proceso que en el caso de la obtención de una aproximación para Z_j, obtenemos a través del TEOREMA 1 que:

$$\hat{\beta}_j = (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot X_j$$

Sustituyendo $\hat{\beta}_j$ para el cálculo de la $E[Cov[\hat{\beta}_j/\theta_j]]$, que es el único elemento que nos falta por conocer en la fórmula de Z_j^d, obtenemos:

$$E[Cov[\hat{\beta}_j/\theta_j]] = S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

siendo:

$$Cov[\hat{\beta}_j] = S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1} + \Gamma$$

como en el caso anterior.

Por último si en $Z_j^d = [q*Cov[\hat{\beta}_j]]^{-1}\cdot[q\cdot Cov[\beta(\theta_j),\ \hat{\beta}_j)]^d$ sustituimos cada covarianza por su valor, nos queda que la aproximación polinómica para la matriz de credibilidad diagonal, que simbolizaremos por Z_j^{dA}, es la siguiente:

$$Z_j^{dA} = [q*[S^2 \cdot (Y' \cdot v_j^{-1} \cdot Y) + \Gamma]]^{-1} \cdot (q \cdot \Gamma)^d$$

Pero a su vez, como $Y = d \cdot \Gamma^{-1}$, resulta que:

$$(Y' \cdot v_j^{-1} \cdot Y)^{-1} = [(d \cdot \Gamma^{-1})' \cdot v_j^{-1} \cdot d \cdot \Gamma^{-1}]^{-1} =$$

$$= [(\Gamma')^{-1} \cdot d' \cdot v_j^{-1} \cdot d \cdot \Gamma^{-1}]^{-1} =$$

$$= \Gamma \cdot (d' \cdot v_j^{-1} \cdot d)^{-1} \cdot \Gamma' = \Gamma \cdot d_j \cdot \Gamma' = \Gamma \cdot d_j \cdot \Gamma$$

pues Γ es una matriz simétrica.

Sustituyéndolo en Z_j^{dA} obtenemos:

$$\boxed{Z_j^{dA} = [q*[S^2 \cdot \Gamma \cdot d_j \cdot \Gamma + \Gamma]]^{-1} \cdot (q \cdot \Gamma)^d}$$

y si denominamos:

$$\boxed{m_j = \Gamma + S^2 \cdot \Gamma \cdot d_j \cdot \Gamma}$$

Z_j^{dA} también se puede escribir como sigue:

$$\boxed{Z_j^{dA} = (q*m_j)^{-1} \cdot (q \cdot \Gamma)^d}$$

De VYLDER, F. (1985) considera que la matriz q más adecuada para este caso es:

$$q = Cov^{-1}[\hat{\beta}_j]$$

Al considerar la mejor aproximación polinómica para el modelo de regresión general resulta que:

$$Cov[\hat{\beta}_j] = \Gamma + S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1} = \Gamma + S^2\cdot\Gamma\cdot d_j\cdot\Gamma = m_j$$

siendo, por lo tanto:

$$\boxed{q = m_j^{-1}}$$

y sustituyendo q en Z_j^{dA}, resulta:

$$\boxed{Z_j^{dA} = (m_j^{-1} * m_j)^{-1}\cdot(m_j^{-1}\cdot\Gamma)^d}$$

3.2.3.4.2.- Resumen de los resultados obtenidos.

El estimador individual para $\beta(\theta_j)$ es el vector $\hat{\beta}_j$, que verifica:

$$g'(\hat{\beta}_j)\cdot v_j^{-1}\cdot(X_j - f(\hat{\beta}_j)) = 0$$

y el estimador ajustado de credibilidad $\hat{\beta}(\theta_j)$, es un vector de dimensión

$(n,1)$, que viene definido como sigue:

$$\hat{\beta}(\theta_j) = Z_j \cdot \hat{\beta}_j + [I - Z_j] \cdot \beta$$

siendo:

β el estimador colectivo para $E[\beta(\theta_j)]$.

Z_j la matriz de credibilidad.

En el caso general, la matriz de credibilidad es:

$$Z_j = Cov[\beta(\theta_j),\ \hat{\beta}_j]^{-1} \cdot Cov^{-1}[\beta_j]$$

y en el caso de restringir dicha matriz a ser diagonal:

$$Z_j^d = [q * Cov[\hat{\beta}_j]]^{-1} \cdot [q \cdot Cov[\beta(\theta_j),\ \hat{\beta}_j]]^d$$

Si en el cálculo de las matrices de credibilidad utilizamos la mejor aproximación polinómica para el modelo de regresión general, resulta que las aproximaciones para Z_j y Z_j^d son:

$$Z_j^A = (I + S^2 \cdot \Gamma \cdot d_j)^{-1}$$

siendo:

$$d_j = (d' \cdot v_j^{-1} \cdot d)^{-1}$$

y

$$Z_j^{dA} = (m_j^{-1} * m_j)^{-1} \cdot (m_j^{-1} \cdot \Gamma)^d$$

donde:

$$m_j = \Gamma + S^2 \cdot \Gamma \cdot d_j \cdot \Gamma$$

3.2.4.- Estimación de los parámetros estructurales.

Los parámetros estructurales que aparecen en este modelo son:

- $\beta = E[\beta(\theta_j)]$ vector de dimensión (n,1).

- $S^2 = E[\sigma^2(\theta_j)]$ que es un escalar.

- $\Gamma = Cov[\beta(\theta_j)]$ matriz simétrica de dimensión (n,n).

- $d = Cov[f(\beta(\theta_j)), \beta'(\theta_j)]$ matriz de dimensión (n,t).

parámetros, que como puede observarse, son independientes de j.

Para el cálculo de estos parámetros se sigue utilizando la aproximación polinómica al modelo de regresión general. En este caso, los estimadores propuestos coinciden con los señalados en el Modelo de Regresión de HACHEMEISTER, aunque con algunos pequeños cambios, que son las adaptaciones al modelo actual. Los estimadores son los siguientes:

- El estimador para $\boldsymbol{\beta} = E[\boldsymbol{\beta}(\theta_j)]$ es:

$$\boxed{\hat{\beta}_Z = \left[\sum_{j=1}^{k} Z_j \right]^{-1} \cdot \sum_{j=1}^{k} Z_j \cdot \hat{\beta}_j}$$

- El estimador para $S^2 = E[\sigma^2(\theta_j)]$ es:

$$\boxed{\hat{S}^2 = \frac{1}{k\cdot(t-n)}\cdot\sum_{j=1}^{k}(X_j - f(\hat{\beta}_j))'\cdot v_j^{-1}\cdot(X_j - f(\hat{\beta}_j))}$$

- El estimador para $\Gamma = Cov[\beta(\theta_j)]$ es:

$$\boxed{\hat{\Gamma} = \frac{1}{k-1}\cdot\sum_{j=1}^{k} Z_j\cdot(\hat{\beta}_j - \hat{\beta}_Z)\cdot(\hat{\beta}_j - \hat{\beta}_Z)'}$$

Al tratarse de un pseudo-estimador, para su cálculo se sigue un proceso iterativo, como ya hemos visto en el Modelo de Regresión de HACHEMEISTER, y en cada iteración es sustituida por $\frac{\hat{\Gamma} + \hat{\Gamma}'}{2}$, ya que debe ser una matriz simétrica.

- Y el estimador para $d = Cov[f(\beta(\theta_j)),\ \beta'(\theta_j)]$ es:

$$\boxed{\hat{d}' = \frac{1}{k-1}\cdot\sum_{j=1}^{k} Z_j\cdot(\hat{\beta}_j - \hat{\beta}_Z)\cdot(f(\hat{\beta}_j) - f(\hat{\beta}_Z))'}$$

Cabe indicar, que las matrices de credibilidad que aparecen en estos cuatro estimadores son las matrices generales, no diagonales, es decir:

$$Z_j^A = (I + S^2 \cdot \Gamma \cdot d_j)^{-1}$$

Para poder obtener estos cuatro estimadores el procedimiento a seguir es el siguiente:

Se toma como punto de partida unos valores iniciales para $\hat{\beta}_Z$, $\hat{\Gamma}$ y $\hat{d}$, por ejemplo, los correspondientes para $Z_j = I$. Con ellos, se calculan unos nuevos valores para $Z_1, Z_2, \ldots, Z_k$, $\hat{\beta}_Z$, $\hat{\Gamma}$ y $\hat{d}$, y el proceso se va repitiendo hasta que los valores de los parámetros se estabilizan. Excepto en casos excepcionales, que a continuación indicaremos, el proceso converge y los valores finales obtenidos no dependen de los valores inicialmente tomados.

3.2.5.- Caso degenerado.

Cuando hemos hallado la mejor aproximación polinómica para el modelo de regresión general hemos obtenido el siguiente sistema de ecuaciones:

$$Y = d \cdot \Gamma^{-1}$$

Como es de suponer, esta expresión no tiene sentido si la matriz de covarianzas Γ no tiene inversa.

En este caso, el proceso iterativo antes descrito resulta divergente, ya que en las iteraciones los valores característicos de la matriz tienden a cero, siendo preciso en esta circunstancia utilizar una aproximación diferente para el modelo de regresión general, es decir, una aproximación polinómica menos próxima.

Esta aproximación se obtiene a través de la Fórmula de Taylor, limitada a los dos primeros términos en el punto β:

$$E[X_j/\theta_j] = f(\beta(\theta_j)) \cong f(\beta) + g(\beta)\cdot(\beta(\theta_j) - \beta) =$$

$$= X_\emptyset + Y\cdot\beta(\theta_j)$$

donde:

$$X_\emptyset = f(\beta) - g(\beta)\cdot\beta$$

$$Y = g(\beta)$$

En este caso la aproximación para la matriz de credibilidad general (no diagonal) Z_j^A es:

$$\boxed{Z_j^A = \Gamma\cdot(\Gamma + S^2\cdot\alpha_j)^{-1}}$$

siendo:

$$\alpha_j = (g'(\beta)\cdot v_j^{-1}\cdot g(\beta))^{-1}$$

Esta expresión se obtiene sustituyendo en la fórmula de Z_j para el caso de utilizar la aproximacón polinómica, la matriz Y por lo que vale, es decir, sustituyendo en:

$$Z_j^A = \Gamma\cdot(\Gamma + S^2\cdot(Y'\cdot v_j\cdot Y)^{-1})^{-1}$$

Y por $g(\beta)$.

Del mismo modo, para obtener la aproximación para la matriz de credibilidad diagonal, basta sustituir en la expresión de Z_j^{dA}, obtenida anteriormente en el caso de utilizar una aproximación polinómica para el modelo de regresión general, la matriz Y por su valor:

$$Z_j^{dA} = [Cov^{-1}[\hat{\beta}_j]*Cov[\hat{\beta}_j]]^{-1}\cdot[Cov^{-1}(\beta_j)\cdot Cov[\beta(\theta_j),\ \hat{\beta}_j]]^d$$

siendo:

$$Cov[\hat{\beta}_j] = \Gamma + S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

$$Cov[\beta(\theta_j),\ \hat{\beta}_j] = Cov[\beta(\theta_j)] = \Gamma$$

Si sustituimos $Y = g(\beta)$ resulta que:

$$Cov[\hat{\beta}_j] = \Gamma + S^2\cdot[g'(\beta)\cdot v_j^{-1}\cdot g(\beta)]^{-1} = \Gamma + S^2\cdot\alpha_j$$

y sustituyéndola en Z_j^d nos queda:

$$Z_j^{dA} = [(\Gamma + S^2\cdot\alpha_j)^{-1}*(\Gamma + S^2\cdot\alpha_j)]^{-1}\cdot[(\Gamma + S^2\cdot\alpha_j)^{-1}\cdot\Gamma]^d$$

A su vez, si simbolizamos por $\hat{m}_j = \Gamma + S^2\cdot\alpha_j$, resulta que Z_j^{dA} también se puede escribir como sigue:

$$\boxed{Z_j^{dA} = (\hat{m}_j^{-1}*m_j)^{-1}\cdot(m_j\cdot\Gamma)^d}$$

3.2.6.- Obtención del Modelo de Regresión de HACHEMEISTER como caso particular del Modelo de Regresión No-Lineal.

Como ya hemos dicho, el Modelo de Regresión No-Lineal de De VYLDER, no es más que una generalización del Modelo de Regresión polinómico de HACHEMEISTER, o dicho de otra manera, el Modelo de Regresión de HACHEMEISTER es un caso particular del Modelo de De VYLDER que surge cuando $f(\beta(\theta_j)) = Y\cdot\beta(\theta_j)$, es decir cuando consideramos f una función polinómica.

Para este caso particular, la segunda hipótesis del modelo de De VYLDER, se escribe del siguiente modo:

D.2) Para todo $j = 1,2,\dots,k$:

a) $E[X_j/\theta_j] = Y\cdot\beta(\theta_j)$

siendo la función f una función:

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

$$f(\beta(\theta_j)) = Y\cdot\beta(\theta_j)$$

donde Y es una matriz dada, de dimensión (t,n), y $\beta(\theta_j)$ es un vector de regresión desconocido de dimensión $(n,1)$

En este caso, $g(\beta(\theta_j)) = Y$ siendo un vector de dimensión (t,n).

b) $Cov[X_j/\theta_j] = \sigma^2(\theta_j)\cdot v_j$

que es una matriz de dimensión (t,t).

Esta hipótesis coincide exactamente con la segunda hipótesis del Modelo de Regresión de HACHEMEISTER.

Para obtener el estimador individual para $\beta(\theta_j)$ aplicamos el TEOREMA 1 del Modelo de De VYLDER:

$\boxed{\text{TEOREMA 1}}$:

El punto $\hat{\beta}_j \in R^n$ tal que $f(\hat{\beta}_j) = Y\cdot\hat{\beta}_j$ es el más próximo a $X_j \in R^t$ es:

$$g'(\hat{\beta}_j)\cdot p\cdot X_j = g'(\hat{\beta}_j)\cdot p\cdot f(\hat{\beta}_j)$$

$$Y'\cdot p\cdot X_j = Y'\cdot p\cdot Y\cdot\hat{\beta}_j$$

$$\hat{\beta}_j = (Y'\cdot p\cdot Y)^{-1}\cdot Y'\cdot p\cdot X_j$$

Lo habitual es considerar la matriz $p = v_j^{-1}$, siendo $\hat{\beta}_j$ igual a:

$$\boxed{\hat{\beta}_j = (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot X_j}$$

donde $\hat{\beta}_j$ es el estimador mínimo-cuadrático generalizado que habíamos obtenido en el Modelo de HACHEMEISTER.

Del mismo modo, si aplicamos el TEOREMA 2 para obtener la matriz de credibilidad Z_j tenemos:

$\boxed{\text{TEOREMA 2}}$:

La matriz Z_j, de dimensión (n,n), tal que $Z_j \cdot \hat{\beta}_j + [I - Z_j]\cdot\beta$ es el más próximo a $\beta(\theta_j)$ es:

$$Z_j = Cov[\beta(\theta_j),\hat{\beta}_j]\cdot Cov^{-1}[\hat{\beta}_j]$$

Para este caso particular:

$$Cov[\beta(\theta_j),\ \hat{\beta}_j] = E[Cov[\beta(\theta_j),\ \hat{\beta}_j/\theta_j]] +$$

$$+ Cov[E[\beta(\theta_j)/\theta_j],\ E[\hat{\beta}_j/\theta_j]] = 0 + Cov[\beta(\theta_j)] = \Gamma$$

y

$$Cov[\hat{\beta}_j] = E[Cov[\hat{\beta}_j/\theta_j]] + Cov[E[\hat{\beta}_j/\theta_j]] = E[Cov[\hat{\beta}_j/\theta_j]] + \Gamma$$

donde:

$$Cov[\hat{\beta}_j/\theta_j] = Cov[\hat{\beta}_j,\ \hat{\beta}'_j/\theta_j] =$$

$$= (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot Cov[X_j/\theta_j]\cdot v_j^{-1}\cdot Y\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

pues como ya hemos visto $\hat{\beta}_j = (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot X_j$, siendo v_j^{-1} una matriz simétrica definida positiva.

Como $Cov[X_j/\theta_j] = \sigma^2(\theta_j)\cdot v_j^{-1}$, si lo sustituimos en $Cov[\hat{\beta}_j/\theta_j]$ nos queda:

$$Cov[\hat{\beta}_j/\theta_j] = (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot\sigma^2(\theta_j)\cdot v_j\cdot v_j^{-1}\cdot Y\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1} =$$

$$= \sigma^2(\theta_j)\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot I\cdot Y\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1} =$$

$$= \sigma^2(\theta_j)\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

y si calculamos:

$$E[Cov[\hat{\beta}_j/\theta_j]] = E[\sigma^2(\theta_j)\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}] = S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

Por último, sustituyendo cada uno de estos resultados en la fórmula de la matriz de credibilidad, obtenemos:

$$Z_j = \Gamma\cdot[\Gamma + S^2\cdot(Y'\cdot v_j^{-1}\cdot Y)^{-1}]^{-1}$$

y si simbolizamos por:

$$\omega_j = (Y'\cdot v_j^{-1}\cdot Y)^{-1}$$

como hicimos en el Modelo de HACHEMEISTER, resulta que:

$$\boxed{Z_j = \Gamma\cdot(\Gamma + S^2\cdot\omega_j)^{-1}}$$

expresión que coincide con la obtenida en el Modelo de Regresión de HACHEMEISTER.

Si en los estimadores de estos parámetros estructurales consideramos:

$$f(\hat{\beta}_j) = Y\cdot\hat{\beta}_j \quad \text{y} \quad \hat{\beta}_j = (Y'\cdot v_j^{-1}\cdot Y)^{-1}\cdot Y'\cdot v_j^{-1}\cdot X_j$$

obtenemos también los mismos estimadores para los parámetros estructurales que los obtenidos en el Modelo de Regresión de HACHEMEISTER.

CAPITULO IV

MODELOS SEMILINEALES

INDICE DEL CAPITULO

En este capítulo vamos a analizar los dos modelos semilineales expuestos por De VYLDER, F. (1976), denominados respectivamente **MODELO SEMILINEAL DE De VYLDER** y **MODELO SEMILINEAL OPTIMO DE De VYLDER**, siendo el segundo una generalización del primero.

Ambos modelos, configuran una de las vertientes en la cual se ha bifurcado la Teoría de la Credibilidad desde la aparición del Modelo de BÜHLMANN, hallándose en su misma línea en cuanto a que siguen asumiendo la existencia de homogeneidad en el tiempo. La diferencia fundamental respecto al Modelo de BÜHLMANN es que en estos modelos no se utilizan directamente los datos obtenidos de la experiencia de reclamaciones, sino que se trabaja con los datos cambiados, transformados, a través de unas funciones. En el Modelo Semilineal se considera que existen P funciones dadas cuadrado-integrables, mientras que en el Modelo Semilineal Optimo solamente se asume la existencia de una única función, siendo uno de sus objetivos la búsqueda de la misma.

En cuanto a los estimadores ajustados de credibilidad para las primas de riesgo individuales son lineales, pero en este caso, no respecto a las variables aleatorias observadas, experiencia de reclamaciones, sino respecto a las funciones o función antes citadas.

Al ser el Modelo Semilineal Optimo una generalización del Modelo Semilineal, este último se puede obtener como un caso particular del primero, puesto que las funciones introducidas en el Modelo Semilineal pueden ser una de las posibles soluciones que hallemos en la búsqueda de la función óptima. Por otro lado, el estimador obtenido a través del Modelo Semilineal Optimo es por lo menos tan bueno como el hallado en el Modelo Semilineal. A pesar de ello, se sigue utilizando la aproximación dada por el Modelo Semilineal ya que en este modelo no es demasiado difícil hallar los estimadores ajustados de credibilidad, pero se complica en gran medida en el Modelo Semilineal Optimo, puesto que un prerrequisito para obtenerlos es estimar previamente la distribución conjunta de los pares (X_r, X_s), con $r,s = 1,2,\ldots,t$.

4.1.- MODELO SEMILINEAL DE DE VYLDER

El Modelo Semilineal fue el primero de los dos modelos semilineales que expuso De VYLDER, F. (1976). Se trata de una extensión del Modelo de BÜHLMANN, en el cual se sigue asumiendo la existencia de homogeneidad en el tiempo.

La principal innovación que introduce De VYLDER, F. (1976) es que ya no trabaja directamente con los datos de la experiencia de reclamaciones, sino con dichos datos transformados. Asume que existen p funciones dadas, con $p = 1,2,...,P$, cuyo objetivo es transformar los datos de la experiencia de reclamaciones, dotando de este modo de mayor flexibilidadal modelo. Si las funciones son cuidadosamente elegidas pueden servir, entre otros fines, para detectar observaciones atípicas y eliminarlas.

En este modelo, los estimadores de las primas de riesgo individuales son lineales, pero no necesariamente respecto a las variables aleatorias observadas, sino respecto a las p funciones dadas. Para la obtención de dichos estimadores se utiliza el procedimiento de los mínimos cuadrados, y en cada estimador aparecen tantos factores de credibilidad como funciones dadas han sido utilizadas.

En la mayor parte de las aplicaciones prácticas no se trabaja con las p funciones dadas, sino que normalmente existen buenas razones para utilizar una única función. En este caso, la obtención del estimador ajustado de credibilidad se simplifica apareciendo un único factor de credibilidad.

4.1.1.- Variables relevantes.

Las variables más relevantes de este modelo son:

- θ_j : El ya conocido parámetro de riesgo, con $j = 1,2,\dots,k$.

- X_{js}: Variable aleatoria que nos indica la experiencia de reclamaciones, con $j = 1,2,\dots,k$ y $s = 1,2,\dots,t$, siendo:

  $X_s = (X_{1s}, X_{2s},\dots, X_{ks})'$

  un vector de dimensión $(k,1)$.

- Además se asume que: $f_0,f_1,f_2,\dots,f_P$ son funciones dadas, tales que $f_p(X_s)$, con $p = 0,1,2,\dots,P$ y $s = 1,2,\dots,t$, son cuadrado-integrables. Estas funciones no son más que transformaciones de las variables aleatorias observables, siendo:

$$f_p: R^k \longrightarrow R^k$$

$$X_s = \begin{bmatrix} X_{1s} \\ X_{2s} \\ \vdots \\ X_{ks} \end{bmatrix} \longrightarrow f_p(X_s) = \begin{bmatrix} f_p(X_{1s}) \\ f_p(X_{2s}) \\ \vdots \\ f_p(X_{ks}) \end{bmatrix}$$

donde $f_p(X_{js})$ es el elemento j-ésimo del vector $f_p(X_s)$ de dimensión $(k,1)$.

Generalmente se considera que $f_\emptyset$ es la función identidad siendo: $f_\emptyset(x) = x$ para todo $x \in R$.

En este modelo, De VYLDER, F. (1976) no está interesado en la búsqueda de las funciones óptimas $f_p(X_s)$, sino que las considera como funciones dadas, como ya hemos dicho.

La introducción de estas funciones proporciona toda una serie de ventajas, aunque también algunos inconvenientes. En primer lugar, podemos destacar que si se eligen cuidadosamente pueden servir para detectar observaciones no deseadas. Así, por ejemplo, en una situación donde hallemos de forma esporádica algunas observaciones, X_{js}, que estén por encima de una cierta cantidad, M_p, da lugar a considerar, de forma intuitiva, a estos puntos como **puntos outliers**. Una definición adecuada podría ser en este caso:

$$f_p(X_s) = \min(M_p,\ X_s) = X_s^t$$

donde X_s^t es un vector de dimensión (k,1) truncado, en el cual se han sustituido las observaciones originales del vector X_s que están por encima del nivel M_p, por dicho valor.

En general, estas funciones hacen más flexible al modelo, en comparación con los modelos anteriores, sin embargo es imposible incorporar diferentes precisiones para las observaciones, y además no es apto para detectar tendencias, como sí lo son el Modelo de Regresión de HACHEMEISTER, y el Modelo de Regresión No-Lineal de De VYLDER.

4.1.2.- Hipótesis del modelo.

Las hipótesis de este modelo son:

D.1) Para un parámetro de riesgo fijo, las variables condicionadas: X_1/θ, $X_2/\theta, \ldots, X_t/\theta$ son independientes y están idénticamente distribuidas.

Esta hipótesis es idéntica a la de los modelos anteriores.

D.2) Hay unas funciones dadas, $f_p(X_s)$, con $p = 1,2,\ldots,P$ y $s = 1,2,\ldots,t$ que son cuadrado-integrables, tales que:

$$E[f_p(X_s)/\theta] = \mu_p(\theta)$$

donde $\mu_p(\theta)$ es un vector de dimensión $(k,1)$, siendo su elemento j-ésimo:

$$\mu_p(\theta_j) = E[f_p(X_{js})/\theta_j]$$

En esta segunda hipótesis cabe destacar que el valor esperado, $\mu_p(\theta)$, no depende del tiempo. De forma idéntica al Modelo de BÜHLMANN se sigue asumiendo que las observaciones esperadas son homogéneas en el tiempo, que hay estacionariedad, la diferencia está en que en este modelo no se utilizan los datos que se obtienen directamente de la experiencia de reclamaciones, sino que se utilizan estos datos transformados, cambiados, a través de las funciones dadas f_p.

4.1.3.- Estimación de la prima de riesgo individual.

En el Modelo de BÜHLMANN[15] comprobamos que cuando se siguen manteniendo las hipótesis iniciales del modelo para las observaciones del periodo siguiente, esto es, para X_{t+1}, se verifica para el caso particular de la póliza j-ésima que:

$$E_{Total}[\mu(\theta_j)/X_{j1}, X_{j2},\dots, X_{jt}] =$$
$$= E_{Total}[X_{j,t+1}/X_{j1}, X_{j2},\dots, X_{jt}]$$

Si trasladamos dicha igualdad al Modelo Semilineal de De VYLDER, ésta se puede escribir del siguiente modo:

$$E_{Total}[\mu_{\emptyset}(\theta_j)/f_1(X_{j1}),\dots,f_1(X_{jt}),\dots,f_p(X_{j1}),\dots,f_p(X_{jt})] =$$

$$= E_{Total}[f_{\emptyset}(X_{j,t+1})/f_1(X_{j1}),\dots,f_1(X_{jt}),\dots,f_p(X_{j1}),\dots,f_p(X_{jt})]$$

que expresada en forma vectorial es como sigue:

15 Véase para más detalles el apartado 2.1.5. del presente trabajo.

$$E_{Total}[\mu_{\emptyset}(\theta)/f_1(X_1),\dots,f_1(X_t),\dots,f_P(X_1),\dots,f_P(X_t)] =$$

$$= E_{Total}[f_{\emptyset}(X_{t+1})/f_1(X_1),\dots,f_1(X_t),\dots,f_P(X_1),\dots,f_P(X_t)]$$

tratándose de dos vectores de dimensión (k,1).

Naturalmente, somos libres de elegir la función $f_{\emptyset}(X_{t+1})$ que mejor se adapte a cada situación, sin embargo, en la práctica, lo usual es considerar:

$$f_{\emptyset}(X_{t+1}) = X_{t+1}$$

ya que lo que nos interesa es estimar la esperanza condicionada de la experiencia de reclamaciones para el periodo t+1.

Al tratarse de una extensión del Modelo de BÜHLMANN, el estimador de la prima de riesgo individual para la póliza j-ésima, $\mu(\theta_j)$, sigue siendo lineal, no respecto a las variables aleatorias observables, experiencia de reclamaciones:

$$c_{j\emptyset} + \sum_{s=1}^{t} c_{js}\cdot X_{js} = c_{j\emptyset} + c_{j1}\cdot X_{j1} + \dots + c_{jt}\cdot X_{jt} \qquad (1)$$

sino respecto a las transformaciones de las mismas a través de las funciones dadas f_p, con p = 1,2,...,P, esto es:

$$a_{j\emptyset} + \sum_{p=1}^{P}\sum_{s=1}^{t} a_{jps}\cdot f_p(X_{js}) \qquad (2)$$

que expresado de forma vectorial es:

$$\boxed{a_{\emptyset} + \sum_{p=1}^{P} \sum_{s=1}^{t} a_{ps} \cdot f_p(X_s)} \tag{3}$$

Nuestro objetivo es, por lo tanto, determinar la combinación lineal del vector $1_{(k,1)}$ y de los vectores $f_p(X_s)$, con $p = 1,2,\ldots,P$ y $s = 1,2,\ldots,t$, más próxima al vector $\mu_{\emptyset}(\theta)$, en el sentido de los mínimos cuadráticos. Esto equivale, a su vez, a determinar la combinación lineal anterior más próxima al vector $f_{\emptyset}(X_{t+1})$.

Antes de pasar a la obtención de la combinación lineal del tipo (3) más próxima a $\mu_{\emptyset}(\theta)$ o a $f_{\emptyset}(X_{t+1})$, vamos a introducir la siguiente notación para los parámetros estructurales que aparecen en dicho modelo. Sean:

$$m_p = E[\mu_p(\theta)] = E[f_p(X_s)] \qquad p = \emptyset,1,2,\ldots,P$$

$$a_{pq} = E[Cov[f_p(X_s),\ f_q(X_s)/\theta]] \qquad p,q = \emptyset,1,2,\ldots,P$$

$$b_{pq} = Cov[\mu_p(\theta),\ \mu_q(\theta)] \qquad p,q = \emptyset,1,2,\ldots,P$$

$$c_{pq} = Cov[f_p(X_s),\ f_q(X_s)] \qquad p,q = \emptyset,1,2,\ldots,P$$

$$d_{pq} = Cov[f_p(X_r),\ f_q(X_s)] \qquad \begin{cases} p,q = \emptyset,1,2,\ldots,P \\ r \neq s \end{cases}$$

Estas expresiones no dependen de $r,s = 1,2,\ldots,t$ y para todos los valores de p y q están relacionadas del siguiente modo:

$$c_{pq} = a_{pq} + b_{pq}$$

$d_{pq} = b_{pq} = Cov[f_p(X_s),\ \mu_q(\theta)]$ debido a la hipótesis de homogeneidad en el tiempo.

Para obtener la combinación lineal de $1_{(k,1)}$ y de los vectores $f_p(X_s)$, de dimensión $(k,1)$, más próxima a $\mu_\emptyset(\theta)$ o a $f_\emptyset(X_{t+1})$, vamos a seguir el artículo de De VYLDER, F. y GOOVAERTS, M. (1985). Estos dos autores proponen el siguiente teorema para solucionar el problema planteado:

TEOREMA 1 :

La combinación lineal de $1_{(k,1)}$ y de las variables aleatorias $f_p(X_s)$, con $p = 1,2,\ldots,P$ y $s = 1,2,\ldots,t$, más próxima a $\mu_\emptyset(\theta)$ y a $f_\emptyset(X_{t+1})$ en el sentido de los mínimos cuadrados es igual a:

$$\boxed{\hat{\mu}_p(\theta) = \sum_{p=1}^{P} Z_p \sum_{s=1}^{t} \frac{1}{t} f_p(X_s) + m_\emptyset - \sum_{p=1}^{P} Z_p \cdot m_p}$$

donde:

$m_\emptyset = E[\mu_\emptyset(\theta)]$ vector de dimensión $(k,1)$

$m_p = E[\mu_p(\theta)]$ vector de dimensión $(k,1)$

$Z_1, Z_2, \ldots, Z_P$ son P escalares[16], solución del sistema lineal de P ecuaciones con P incógnitas:

$$\sum_{p=1}^{P} [c_{pq} + (t-1)\cdot d_{pq}]\cdot Z_p = t\cdot d_{\emptyset q} \qquad q = 1,2,\ldots,P$$

o del sistema lineal equivalente:

$$\sum_{p=1}^{P} (a_{pq} + t\cdot b_{pq})\cdot Z_p = t\cdot b_{\emptyset q} \qquad q = 1,2,\ldots,P$$

A su vez, el estimador ajustado de credibilidad para el elemento j-ésimo se suele escribir como sigue:

$$\boxed{\hat{\mu}_p(\theta_j) = \sum_{p=1}^{P} Z_p\cdot M_{pj} + M_\emptyset - \sum_{p=1}^{P} Z_p\cdot M_p}$$

o de forma alternativa:

$$\boxed{\hat{\mu}_p(\theta_j) = M_\emptyset + \sum_{p=1}^{P} Z_p\cdot (M_{pj} - M_p)}$$

donde:

16 En este modelo, a diferencia con el de BÜHLMANN, los factores de credibilidad no son necesariamente números comprendidos entre Ø y 1, sino que pueden ser cualquier número perteneciente a los R.

$$M_{pj} = \frac{1}{t} \cdot \sum_{s=1}^{t} f_p(X_{js}) \qquad p = \emptyset,1,2,\ldots,P$$

$$M_p = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{pj} \qquad p = \emptyset,1,2,\ldots,P$$

siendo M_{pj} el elemento j-ésimo del vector de estimadores de m_p.

$$M_\emptyset = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{\emptyset j} \qquad p = \emptyset,1,2,\ldots,P$$

siendo $M_{\emptyset j}$ el elemento j-ésimo del vector de estimadores de $m_\emptyset$.

DEMOSTRACION:

$\hat{\mu}_p(\theta)$ serà el estimador más próximo a $\mu_\emptyset(\theta)$ si verifica:

$$E[\hat{\mu}_p(\theta)] = E[\mu_\emptyset(\theta)]$$

y

$$\mathrm{Cov}[\hat{\mu}_p(\theta),\ f_p(X_s)] = \mathrm{Cov}[f_\emptyset(X_{t+1}),\ f_p(X_s)]$$

para $p = 1,2,\ldots,P$ y $s = 1,2,\ldots,t$.

Y del mismo modo, $\hat{\mu}_p(\theta)$ es el estimador más próximo a $f_\emptyset(X_{t+1})$ si verifica:

$$E[\hat{\mu}_p(\theta)] = E[f_\emptyset(X_{t+1})]$$

y

$$Cov[\hat{\mu}_p(\theta),\ f_p(X_s)] = Cov[f_{\emptyset}(X_{t+1}),\ f_p(X_s)]$$
$$\text{para}\quad p = 1,2,\dots,P \quad y \quad s = 1,2,\dots,t.$$

Lo único que hay que hacer es comprobar que sí se verifican las cuatro igualdades anteriores. Nosotros lo haremos para un elemento en particular de estos vectores, el j-ésimo.

En primer lugar vamos a comprobar que:

$$E[\hat{\mu}_p(\theta_j)] = E[\mu_{\emptyset}(\theta_j)]$$

Como ya sabemos:

$$\hat{\mu}_p(\theta_j) = \sum_{p=1}^{P} Z_p \cdot M_{pj} + M_{\emptyset} - \sum_{p=1}^{P} Z_p \cdot M_p$$

siendo su esperanza:

$$E[\hat{\mu}_p(\theta_j)] = \sum_{p=1}^{P} Z_p \cdot E[M_{pj}] + E[M_{\emptyset}] - \sum_{p=1}^{P} Z_p \cdot E[M_p]$$

$$\cdot)\ E[M_{pj}] = E\left[\frac{1}{t}\cdot\sum_{s=1}^{t} f_p(X_{js})\right] = \frac{1}{t}\cdot\sum_{s=1}^{t} E[f_p(X_{js})] =$$

$$= \frac{1}{t}\cdot\sum_{s=1}^{t} E[\mu_p(\theta_j)] = E[\mu_p(\theta_j)]$$

•) Por otro lado, como $M_p = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{pj}$ resulta que:

$$E[M_p] = E\left[\frac{1}{k} \cdot \sum_{j=1}^{k} M_{pj}\right] = \frac{1}{k} \cdot \sum_{j=1}^{k} E[M_{pj}] =$$

$$= \frac{1}{k} \cdot \sum_{j=1}^{k} E[\mu_p(\theta_j)] = \frac{k}{k} \cdot E[\mu_p(\theta_j)] = E[\mu_p(\theta_j)]$$

•) Y por último, como $M_\emptyset$ es el estimador colectivo de $\mu_\emptyset(\theta_j)$, se verifica que $E[M_\emptyset] = E[\mu_\emptyset(\theta_j)]$. Esto es:

$$E[M_\emptyset] = \frac{1}{k} \cdot \sum_{j=1}^{k} \frac{1}{t} \cdot \sum_{s=1}^{t} E[f_\emptyset(X_{js})] =$$

$$= \frac{1}{k} \sum_{j=1}^{k} \frac{1}{t} \cdot \sum_{s=1}^{t} E[\mu_\emptyset(\theta_j)] =$$

$$= \frac{1}{k} \sum_{j=1}^{k} \frac{t}{t} E[\mu_\emptyset(\theta_j)] = \frac{1}{k} \sum_{j=1}^{k} E[\mu_\emptyset(\theta_j)] =$$

$$= \frac{k}{k} \cdot E[\mu_\emptyset(\theta_j)] = E[\mu_\emptyset(\theta_j)]$$

Sustituyendo en $E[\hat{\mu}_p(\theta_j)]$ estas tres esperanzas por su valor respectivo, nos queda:

$$E[\hat{\mu}_p(\theta_j)] = \sum_{p=1}^{P} Z_p \cdot E[\mu_p(\theta_j)] + E[\mu_\emptyset(\theta_j)] - \sum_{p=1}^{P} Z_p \cdot E[\mu_p(\theta_j)] =$$

$$= E[\mu_\emptyset(\theta_j)]$$

quedando demostrado que $E[\hat{\mu}_p(\theta_j)] = E[\mu_\emptyset(\theta_j)]$.

En segundo lugar debemos comprobar que:

$$Cov[\hat{\mu}_p(\theta_j),\ f_p(X_{js})] = Cov[\mu_\emptyset(\theta_j),\ f_p(X_{js})]$$

para $p = 1,2,\ldots,P$ y $s = 1,2,\ldots t$.

Para ello aplicaremos la definición de covarianza, siendo:

$$Cov[\hat{\mu}_p(\theta_j),\ f_p(X_{js})] = E[Cov[\hat{\mu}_p(\theta_j),\ f_p(X_{js})/\theta_j]] +$$

$$+ Cov[E[\hat{\mu}_p(\theta_j)/\theta_j],\ E[f_p(X_{js})/\theta_j]]$$

•) $Cov[\hat{\mu}_p(\theta_j),\ f_p(X_{js})/\theta_j] = \emptyset$, ya que para un determinado θ_j, el estimador ajustado de credibilidad es una constante.

•) $E[f_p(X_{js})/\theta_j] = \mu_p(\theta_j)$ por la hipótesis D.2) del modelo.

•) $E[\hat{\mu}_p(\theta_j)/\theta_j] = \sum_{p=1}^{P} Z_p \cdot E[M_{pj}/\theta_j] + E[M_\emptyset/\theta_j] -$

$$- \sum_{p=1}^{P} Z_p \cdot E[M_p/\theta_j] = \sum_{p=1}^{P} Z_p \cdot \mu_p(\theta_j) + \mu_\emptyset(\theta_j) -$$

$$- \sum_{p=1}^{P} Z_p \cdot \mu_p(\theta_j) = \mu_\emptyset(\theta_j)$$

De ello se desprende:

$$\mathrm{Cov}[E[\hat{\mu}_p(\theta_j)/\theta_j], E[f_p(X_{js})/\theta_j]] = \mathrm{Cov}[\mu_\emptyset(\theta_j), \mu_p(\theta_j)] = b_{\emptyset p}$$

siendo, por lo tanto:

$$\mathrm{Cov}[\hat{\mu}_p(\theta_j), f_p(X_{js})] = b_{\emptyset p} = \mathrm{Cov}[\mu_\emptyset(\theta_j), \mu_p(\theta_j)]$$

Pero como:

$$d_{pq} = b_{pq} = \mathrm{Cov}[f_p(X_{js}), \mu_p(\theta_j)]$$

resulta que:

$$b_{\emptyset p} = \mathrm{Cov}[\mu_\emptyset(\theta_j), f_p(X_{js})]$$

siendo:

$$\mathrm{Cov}[\hat{\mu}_p(\theta_j), f_p(X_{js})] = b_{\emptyset p} = \mathrm{Cov}[\mu_\emptyset(\theta_j), f_p(X_{js})]$$

que es lo que queríamos comprobar.

En cuanto a las dos últimas igualdades que nos falta por comprobar del TEOREMA 1, se haría de idéntica forma a lo efectuado para las dos primeras.

El estimador de credibilidad que hemos obtenido, a diferencia con el del Modelo de BÜHLMANN, ya no es una combinación lineal convexa entre el estimador colectivo y el estimador individual de la prima de riesgo individual, sino que en este modelo adopta, para la póliza j-ésima, la siguiente expresión:

$$\hat{\mu}_p(\theta_j) = M_\emptyset + \sum_{p=1}^{P} Z_p \cdot (M_{pj} - M_p)$$

Como puede apreciarse, dicho estimador se obtiene sumando la media colectiva de las observaciones de la cartera transformadas a través de la función dada $f_\emptyset(X_s)$ más p sumandos, tantos como funciones $f_p(X_s)$ hayamos considerado, siendo $p = 1,2,\ldots,P$. A su vez, cada uno de estos sumandos es el resultado de multiplicar el factor de credibilidad correspondiente a la función dada $f_p(X_s)$, por la diferencia entre la media individual y la media colectiva de las observaciones transformadas a través de dicha función.

Normalmente se considera que $f_\emptyset(X_s)$ es la función identidad $(f_\emptyset(X_s) = X_s)$ por lo que en la mayoría de los casos $M_\emptyset$ está definido del siguiente modo:

$$M_\emptyset = \frac{1}{k\cdot t}\cdot\sum_{j=1}^{k}\sum_{s=1}^{t} X_{js}$$

siendo la media colectiva de las observaciones para la cartera en su totalidad, coincidiendo con el estimador colectivo para la prima de riesgo individual obtenido en el Modelo de BÜHLMANN.

4.1.4.- Estimación de los parámetros estructurales.

A partir del TEOREMA 1 hemos obtenido el estimador ajustado de credibilidad para $\mu_\emptyset(\theta)$ o para $f_\emptyset(X_{t+1})$, que es:

$$\hat{\mu}_p(\theta) = \sum_{p=1}^{P} Z_p\cdot\sum_{s=1}^{t}\frac{1}{t}\cdot f_p(X_s) + m_\emptyset - \sum_{p=1}^{P} Z_p\cdot m_p$$

y en particular para el elemento j-ésimo:

$$\hat{\mu}_p(\theta_j) = \sum_{p=1}^{P} Z_p\cdot M_{pj} + M_\emptyset - \sum_{p=1}^{P} Z_p\cdot M_p$$

donde:

$$M_{pj} = \frac{1}{t}\cdot\sum_{s=1}^{t} f_p(X_{js})$$

$$M_p = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{pj} \qquad \text{con } p = \emptyset,1,2,\dots,P$$

$Z_1, Z_2, \dots, Z_P$ son P escalares, solución del sistema de ecuaciones lineal:

$$\sum_{p=1}^{P} (a_{pq} + t \cdot b_{pq}) \cdot Z_p = t \cdot b_{\emptyset q} \qquad q = \emptyset,1,2,\dots,P$$

Al ser $f_\emptyset$, $f_1, \dots, f_P$ funciones dadas, los valores de M_{pj}, M_p y $M_\emptyset$ son conocidos o de fácil cálculo, y lo único que nos falta para poder obtener el estimador ajustado de credibilidad para la póliza j-ésima, $\hat{\mu}_p(\theta_j)$, son los valores de los factores de credibilidad: $Z_1, Z_2, \dots, Z_P$, factores, que como ya hemos dicho, son la solución del sistema de ecuaciones lineal:

$$\sum_{p=1}^{P} (a_{pq} + t \cdot b_{\emptyset p}) \cdot Z_p = t \cdot b_{\emptyset q} \qquad q = \emptyset,1,2,\dots,P$$

Para poder solucionar este sistema de ecuaciones, debemos estimar previamente los valores de los parámetros estructurales a_{pq} y b_{pq}, que son respectivamente:

$$a_{pq} = E[Cov[f_p(X_s), f_q(X_s)]] \qquad p,q = \emptyset,1,2,\dots P$$

$$b_{pq} = Cov[\mu_p(\theta), \mu_q(\theta)] \qquad p,q = \emptyset,1,2,\dots P$$

Los estimadores propuestos para ambos parámetros estructurales son $\hat{a}_{pq}$ y $\hat{b}_{pq}$, que vienen definidos del siguientes modo:

$$\boxed{\hat{a}_{pq} = \frac{1}{k\cdot(t-1)}\cdot\sum_{j=1}^{k}\sum_{s=1}^{t}(f_p(X_{js}) - M_{pj})\cdot(f_q(X_{js}) - M_{qj})}$$

que la media aritmética de las k covarianzas individuales.

$$\boxed{\begin{array}{l}\hat{b}_{pq} = \dfrac{1}{k-1}\cdot\displaystyle\sum_{j=1}^{k}(M_{pj} - M_p)\cdot(M_{qj} - M_q) - \dfrac{1}{t}\cdot\hat{a}_{pq} \\ \qquad\qquad\qquad\qquad\qquad p,q = 0,1,2,\ldots P\end{array}}$$

que es la media aritmética de las k covarianzas de M_{pj}.

Estos dos estimadores son insesgados.

4.1.5.- Casos particulares.

1) En la mayoría de las aplicaciones prácticas a menudo existen buenas razones para utilizar el Modelo Semilineal de De VYLDER, con **una única función, $f(X_s)$.**

En este caso, los parámetros estructurales se suelen escribir del siguiente modo:

$$\mu_f(\theta) = E[f(X_s)/\theta]$$

$$m_f = E[\mu_f(\theta)] = E[f(X_s)]$$

$$a_{fg} = E[Cov[f(x_s), g(X_s)/\theta]]$$

$$b_{fg} = Cov[\mu_f(\theta), \mu_g(\theta)]$$

$$c_{fg} = Cov[f(X_s), g(X_s)]$$

$$d_{fg} = Cov[f(X_r), g(X_s)]$$

donde:

$$g(X_s) = f(X_s) \quad o \quad g(X_s) = f_\emptyset(X_s)$$

Evidentemente se verifica:

$$c_{fg} = a_{fg} + b_{fg}$$
$$d_{fg} = b_{fg}$$

Un elemento importante a tener en cuenta es que en este caso los parámetros estructurales dejan de ser matrices para convertirse en escalares.

Al utilizar una única función, $f(X_s)$, resulta que el estimador ajustado de credibilidad que hemos obtenido en el TEOREMA 1:

$$\hat{\mu}_p(\theta) = \sum_{p=1}^{P} Z_p \cdot \sum_{s=1}^{t} \frac{1}{t} f_p(X_s) + m_\emptyset - \sum_{P=1}^{P} Z_p \cdot m_p$$

pasa a ser:

$$\hat{\mu}_f(\theta) = Z \cdot \sum_{s=1}^{t} f(X_s) + m_\emptyset - Z \cdot m_f$$

donde:

$$m_f = E[\mu_f(\theta)]$$

$$m_\emptyset = E[\mu_\emptyset(\theta)]$$

siendo $\hat{\mu}_f(\theta)$, $f(X_s)$, $m_\emptyset$ y m_f vectores de dimensión $(k,1)$.

En este caso, no tenemos p factores de credibilidad sino un único factor[17] Z, que se obtiene a partir de la ecuación:

$$(a_{fg} + t \cdot b_{fg}) \cdot Z = t \cdot b_{\emptyset f}$$

de donde:

$$\boxed{Z = \frac{t \cdot b_{\emptyset f}}{a_{ff} + t \cdot b_{ff}} = \frac{t \cdot d_{\emptyset f}}{c_{ff} + (t-1) \cdot d_{ff}}}$$

---

[17] En este caso particular, aunque el factor de credibilidad sea único no tiene porque estar comprendido entre 0 y 1, sino que puede ser cualquier número perteneciente a R, como se desprende de su definición.

siendo los estimadores de los parámetros estructurales que aparecen en dicha expresión los siguientes:

$$\hat{a}_{ff} = \frac{1}{k\cdot(t-1)}\cdot\sum_{j=1}^{k}\sum_{s=1}^{t}(f(X_{js}) - M_{fj})^2$$

$$\hat{a}_{\emptyset f} = \frac{1}{k\cdot(t-1)}\cdot\sum_{j=1}^{k}\sum_{s=1}^{t}(f(X_{js}) - M_{fj})\cdot(f_{\emptyset}(X_{js}) - M_{\emptyset j})$$

$$\hat{b}_{ff} = \frac{1}{k-1}\cdot\sum_{j=1}^{k}(M_{fj} - M_{f})^2 - \frac{1}{t}\cdot\hat{a}_{ff}$$

$$\hat{b}_{\emptyset f} = \frac{1}{k-1}\cdot\sum_{j=1}^{k}(M_{fj} - M_{f})\cdot(M_{\emptyset j} - M_{\emptyset}) - \frac{1}{t}\cdot\hat{a}_{\emptyset f}$$

Para el caso particular de la póliza j-ésima, el estimador ajustado de credibilidad es:

$$\boxed{\hat{\mu}_f(\theta) = Z\cdot M_{fj} + M_{\emptyset} + Z\cdot M_f}$$

o de forma alternativa:

$$\boxed{\hat{\mu}_f(\theta) = M_{\emptyset} + Z\cdot(M_{fj} - M_f)}$$

donde:

$$M_{fj} = \frac{1}{t} \cdot \sum_{s=1}^{t} f(X_{js})$$

$$M_f = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{fj}$$

$$M_\emptyset = \frac{1}{k} \cdot \sum_{j=1}^{k} M_{\emptyset j} = \frac{1}{k \cdot t} \cdot \sum_{j=1}^{k} \sum_{s=1}^{t} f_\emptyset(X_{js})$$

2) Si consideramos además la hipótesis:

$$f(X_s) = g(X_s) = f_\emptyset(X_s)$$

resulta que el estimador ajustado de credibilidad para la póliza j-ésima es:

$$\boxed{\hat{\mu}_f(\theta) = Z \cdot M_{fj} + [1 - Z] \cdot M_f}$$

ya que en este caso $M_\emptyset = M_f$, y a su vez:

$$a_{fg} = a_{ff} \quad \text{y} \quad b_{fg} = b_{\emptyset f} = b_{ff}$$

siendo el factor de credibilidad Z, igual a:

$$Z = \frac{t \cdot b_{ff}}{a_{ff} + t \cdot b_{ff}} = \frac{t \cdot d_{ff}}{c_{ff} + (t - 1) \cdot d_{ff}}$$

En este caso, el estimador ajustado de credibilidad es ya una combinación lineal convexa de la media individual y de la media colectiva, al igual que en el Modelo de BÜHLMANN, pero con la salvedad que se siguen utilizando los datos transformados a través de la función dada $f(X_s)$, siendo el factor de credibilidad $0 \leq Z \leq 1$.

3) Si además de considerar una única función $f(X_s)$ introducimos la hipótesis:

$$f_0(X_s) = f(X_s) = X_s$$

resulta que el estimador ajustado de credibilidad es:

$$\hat{\mu}_f(\theta) = Z \cdot M_{fj} + [1 - Z] \cdot M_f$$

donde:

$$M_{fj} = \frac{1}{t} \cdot \sum_{s=1}^{t} X_{js}$$

$$M_f = \frac{1}{k \cdot t} \cdot \sum_{j=1}^{k} \sum_{s=1}^{t} X_{js}$$

$$Z = \frac{t \cdot a_{ff}}{a_{ff} + t \cdot b_{ff}}$$

$$\hat{a}_{ff} = \frac{1}{k \cdot (t - 1)} \cdot \sum_{j=1}^{k} \sum_{s=1}^{t} (X_{js} - M_{fj})^2$$

$$\hat{b}_{ff} = \frac{1}{k - 1} \cdot \sum_{j=1}^{k} (M_{fj} - M_f)^2 - \frac{1}{t} \cdot \hat{a}_{ff}$$

que coincide con el obtenido en el Modelo de BÜHLMANN, aunque con distinta nomenclatura, siendo $M_f = M_\emptyset$, $\hat{a}_{ff} = \hat{S}^2$ y $\hat{b}_{ff} = \hat{a}$. En realidad, coincide con el estimador ajustado de credibilidad de BÜHLMANN, una vez que en éste $m = E[X_{js}]$ es sustituido por su estimador $M_\emptyset$.

Para concluir, diremos que el Modelo de BÜHLMANN se puede considerar como un caso particular del Modelo Semilineal de De VYLDER, que se obtiene cuando se considera que existe una única función dada y además que $f(X_s) = f_\emptyset(X_s) = X_s$.

4.2.- MODELO SEMILINEAL OPTIMO DE DE VYLDER

Se trata de una generalización del Modelo Semilineal de De VYLDER, de manera que se sigue asumiendo la existencia de homogeneidad en el tiempo. Como en el Modelo Semilineal no se utilizan directamente los datos que se obtienen de la experiencia de las reclamaciones, sino que se trabaja con los datos transformados, cambiados, a través de una función.

La diferencia fundamental que existe entre el Modelo Semilineal y el Modelo Semilineal Optimo es que, mientras que en el primero se considera que existen p funciones dadas, cuadrado-integrables, con $p = 1,2,\ldots,P$, en el segundo se asume que existe una **única función**, siendo uno de los objetivos del modelo la búsqueda de la misma, utilizando para ello el procedimiento de los mínimos cuadrados.

El estimador ajustado de credibilidad sigue siendo lineal, pero respecto a la función óptima que se está buscando, complicándose considerablemente su obtención, ya que un paso previo es estimar la distribución conjunta de los pares (X_r, X_s), con $r,s = 1,2,\ldots,t$.

4.2.1.- Variables relevantes.

Las variables más relevantes de este modelo son:

- θ_j : El ya conocido parámetro de riesgo, con $j = 1,2,\ldots,k$.

- X_{js}: Variable aleatoria que nos indica la experiencia de reclamaciones, con $j = 1,2,\dots,k$ y $s = 1,2,\dots,t$, siendo
$X_s = (X_{1s}, X_{2s}, \dots, X_{ks})'$
un vector de dimensión $(k,1)$.

- También se asume que existe una función $f(X_s)$, tal que:

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

$$X_s = \begin{bmatrix} X_{1s} \\ X_{2s} \\ \vdots \\ X_{ks} \end{bmatrix} \longrightarrow f(X_s) = \begin{bmatrix} f(X_{1s}) \\ f(X_{2s}) \\ \vdots \\ f(X_{ks}) \end{bmatrix}$$

siendo $f(X_{js})$ el elemento j-ésimo del vector $f(X_s)$, de dimensión $(k,1)$.

4.2.2.- Hipótesis del modelo.

Las hipótesis de este modelo coinciden con las del Modelo Semilineal de De VYLDER[18] para el caso particular que $p = 1$. Esto es:

---

[18] Véase para más detalles el apartado 4.1.2. del presente capítulo.

D.1) Para un parámetro de riesgo fijo, las variables aleatorias condicionadas: X_1/θ, $X_2/\theta, \ldots,$ X_t/θ, son independientes y están idénticamente distribuidas.

D.2) Hay una función $f(X_s)$, con $s = 1,2,\ldots,t$, que es cuadrado-integrable, tal que:

$$E[f(X_s)/\theta] = \mu_f(\theta)$$

donde $\mu_f(\theta)$ un vector de dimensión $(k,1)$, siendo su elemento j-ésimo:

$$\mu_f(\theta_j) = E[f(X_{js})/\theta_j]$$

Las observaciones esperadas son homogéneas en el tiempo, como en el Modelo de BÜHLMANN, y en el Modelo Semilineal de De VYLDER.

4.2.3.- Estimación de la prima de riesgo individual.

En el Modelo Semilineal de De VYLDER hemos considerado un estimador de tipo lineal para $\mu_\emptyset(\theta)$ o para $f_\emptyset(X_{t+1})$, cuya forma vectorial es la siguiente:

$$a_{\emptyset} + \sum_{p=1}^{P} \sum_{s=1}^{t} a_{ps} \cdot f_p(X_s)$$

En este modelo, seguimos estando interesados en hallar estimadores de tipo lineal para las primas de riesgo individuales, pero al haber asumido la existencia de una única función, $f(X_s)$, resulta que en este caso adoptará la siguiente forma:

$$a_{\emptyset} + \sum_{s=1}^{t} a_s \cdot f(X_s)$$

A su vez, al no existir limitaciones en cuanto a la estructura de la función $f(X_s)$, podemos eliminar los coeficientes $a_{\emptyset}$ y a_s, con $s = 1,2,\ldots,t$, ya que podemos asumir que están incorporados en la propia función $f(X_s)$. En consecuencia, sin pérdida de generalidad, podemos considerar estimadores del siguiente tipo:

$$\sum_{s=1}^{t} f(X_s)$$

para cuya obtención vamos a utilizar el mismo procedimiento que en el Modelo de BÜHLMANN: Minimizar la esperanza total del error cuadrático medio cometido.

El estimador ajustado de credibilidad obtenido en el Modelo Semilineal de De VYLDER, cuya expresión es:

$$\hat{\mu}_p(\theta) = \sum_{p=1}^{P} Z_p \cdot \sum_{s=1}^{t} \frac{1}{t} \cdot f_p(X_s) + m_\emptyset - \sum_{p=1}^{P} Z_p \cdot m_p$$

se puede también expresar como sigue:

$$\hat{\mu}_p(\theta) = f(X_1) + f(X_2) + \ldots + f(X_t)$$

donde:

$$f(X) = \frac{1}{t} \cdot \sum_{p=1}^{P} Z_p \cdot f_p(X) + \frac{1}{t} \cdot m_\emptyset - \frac{1}{t} \cdot \sum_{p=1}^{P} Z_p \cdot m_p \qquad (X \in R)$$

Ello muestra que las funciones introducidas en el Modelo Semilineal pueden ser una de las posibles soluciones que hallemos en la búsqueda de la función óptima, al aplicar el Modelo Semilineal Optimo.

La aproximación credibilística que proporciona la Teoría de la Credibilidad semilineal óptima, para $f_\emptyset(X_{t+1})$ o para $\mu_\emptyset(\theta)$, es por lo menos tan buena como la hallada a través del Modelo Semilineal, pero si la comparamos con la prima lineal usual, la del Modelo de BÜHLMANN, la diferencia que existente entre ambas es considerable.

Todo ello nos lleva a preguntarnos el motivo por el cual se sigue utilizando la aproximación dada por el Modelo Semilineal. La respuesta es fácil. En el Modelo Semilineal de De VYLDER para calcular los factores de credibilidad Z_1, $Z_2, \ldots, Z_P$, que son la solución del sistema de ecuaciones lineal:

$$\sum_{p=1}^{P} (a_{pq} + t \cdot b_{pq}) \cdot Z_p = t \cdot b_{\emptyset q} \qquad q = 1,2,\ldots,P.$$

es suficiente conocer los valores de los parámetros a_{pq}, b_{pq} y $b_{\emptyset q}$.

En general, no es demasiado difícil construir estimadores insesgados para los mismos basados en las observaciones de la cartera en consideración. Sin embargo, la determinación del estimador semilineal óptimo es una tarea más difícil y que consume más tiempo, ya que para su obtención es necesario estimar la distribución conjunta de los pares (X_r, X_s), con $r,s = 1,2,\dots,t$.

4.2.3.1.- Obtención del estimador ajustado de credibilidad para $\mu_{\theta}(\theta)$ o para $f_{\theta}(X_{t+1})$.

Para su obtención vamos a seguir el artículo de De VYLDER, F. y BALLEGEER, Y. (1979), pp. 131-148.

Sea:

- $m_\theta = E[X_1/\theta]$ la prima de riesgo para cada año[19].

- Para cada función f, de una variable, vamos a simbolizar por f_θ la variable aleatoria:

$$f_\theta = E[f(X_1)/\theta]$$

$\boxed{\text{LEMA}}$:

•) Para cada par de funciones f,g de una variable:

$$E[f(X_1)\cdot g(X_2)] = E[f_\theta\cdot g(X_2)] = E[f(X_1)\cdot g_\theta] = E[f_\theta\cdot g_\theta] \quad (1)$$

19 En este caso y en otras situaciones similares, el subíndice 1 puede ser reemplazado por otro indistintamente, ya que las variables $X_1, X_2,\dots, X_t$ son intercambiables en el sentido de De FINETTI.

•) Para cada función f, de una variable, y para cada función ς, de t variables:

$$E[\varsigma(X_1,\ldots,X_t)\cdot f(X_{t+1})] = E[\varsigma(X_1,\ldots,X_t)\cdot f_\theta] \tag{2}$$

•) Para cada función f, de una variable:

$$E[f(X_{t+1})/\ X_1,X_2,\ldots,X_t] = E[f_\theta/\ X_1,X_2,\ldots,X_t] \tag{3}$$

TEOREMA 1:

Sea la f^* la solución de:

$$\boxed{E[X_2/X_1] = f^*(X_1) + (t-1)\cdot E[f^*(X_2)/X_1]} \tag{4}$$

Entonces, para cada función f:

$$E[m_\theta - f^*(X_1) - \ldots - f^*(X_t)]^2 \le$$
$$\le E[m_\theta - f(X_1) - \ldots - f(X_t)]^2 \tag{5}$$

El error cuadrático medio cometido en la aproximación de m_θ por $f^*(X_1),\ldots,f^*(X_t)$ viene dado por:

$$E[m_\theta - f^*(X_1) - \ldots - f^*(X_t)]^2 =$$
$$= E[X_1\cdot X_2] - t\cdot E[X_1\cdot f^*(X_2)] \tag{6}$$

Si g^* también satisface:

$$E[X_2/X_1] = g^*(X_1) + (t - 1)\cdot E[g^*(X_2)/X_1] \tag{7}$$

entonces:

$$f^*(X_1) = g^*(X_1) \tag{8}$$

Este teorema nos indica que existe la función óptima f^* que es la solución de la ecuación (4), y a su vez que es UNICA.

$\boxed{\text{COROLARIO}}$:

Sea la f^* la solución de la ecuación (4). Entonces, para cada f:

$$E[X_{t+1} - f^*(X_1) - \dots - f^*(X_t)]^2 \leq$$
$$\leq E[X_{t+1} - f(X_1) - \dots - f(X_t)]^2 \tag{9}$$

En este modelo, la prima de credibilidad semilineal óptima para el año t+1 se define y se simboliza del siguiente modo:

$$\boxed{E[X_{t+1}/X_1,\dots,X_t] = f^*(X_1) + \dots + f^*(X_t)} \tag{10}$$

tratándose de un vector de dimensión (k,1), y donde f^* es la solución de la ecuación (4), esto es:

$$E[X_2/X_1] = f^*(X_1) + (t - 1)\cdot E[f^*(X_2)/X_1] \tag{4}$$

TEOREMA 2 :

$$E\Big[E[X_{t+1}/X_1,\dots,X_t]\Big] = E[X_{t+1}] \tag{11}$$

Del mismo modo, podemos definir el estimador ajustado de credibilidad o la prima de credibilidad semilineal óptima del elemento j-ésimo de la siguiente manera:

$$\hat{\mu}_f(\theta_j) = \sum_{s=1}^{t} f^*(X_{js}) = f^*(X_{j1}) + \dots + f^*(X_{jt}) \tag{12}$$

donde f^* es la solución de las ecuaciones:

$$E[f_\emptyset(X_{jr})/X_{js}] = f^*(X_{js}) + (t-1)\cdot E[f^*(X_{jr})/X_{js}] \tag{13}$$

$$j = 1,2,\dots,k; \quad r,s = 1,2,\dots,t$$

Lo usual es que $f_\emptyset(X_{j,t+1}) = X_{j,t+1}$

Si retomamos el vector de la prima de credibilidad semilineal óptima, y consideramos que las variables X_1 y X_2 tienen una función de densidad conjunta $p(x,y)$, resulta que la ecuación (4) se puede también escribir como sigue:

$$\boxed{\int y\cdot p(x,y)\cdot dy = f^*(x)\cdot\int p(x,y)\cdot dy + (t-1)\cdot\int f^*(y)\cdot p(x,y)\cdot dy} \tag{14}$$

ya que:

$$E[X_2/X_1] = \int y \cdot \frac{p(x,y)}{p(x)} \cdot dy$$

$$f^*(X_1) = f^*(X) \cdot \int p(y) \cdot dy = f^*(X) \cdot \int \frac{p(x,y)}{p(x)} \cdot dy$$

$$E[f^*(X_2)/X_1] = \int f^*(y) \cdot \frac{p(x,y)}{p(x)} \cdot dy$$

Esta ecuación es una ecuación integral del tipo FREDHOM para una función f^* desconocida. Aunque existen métodos numéricos para solucionarla, en la práctica no se utilizan, sino que su cálculo se lleva a cabo, a menudo, mediante distintas aproximaciones.

Si asumimos que X_1 toma un número finito de valores, que es lo más usual, sea $0,1,\ldots,n$, entonces la ecuación (4) se convierte en el sistema de ecuaciones lineal siguiente:

$$\boxed{\sum_{l=0}^{n} l \cdot p_{ml} = f_m^* \cdot \sum_{l=0}^{n} p_{ml} + (t-1) \cdot \sum_{l=0}^{n} f_l^* \cdot p_{ml} \qquad m = 0,1,\ldots,n} \tag{15}$$

donde:

$$p_{ml} = P(X_1 = m,\ X_2 = l) \tag{16}$$

$$f_m^* = f^*(m) \tag{17}$$

Las ecuaciones (14) y (15), tanto sirven para llevar a cabo investigaciones teóricas como para obtener los cálculos numéricos de las primas de credibilidad semilineales óptimas, para lo cual, sólo es necesario conocer la distribución conjunta de X_1 y X_2.

4.2.3.2.- Estimación de la distribución conjunta del par (X_r, X_s).

Estimar la distribución conjunta del par genérico (X_r, X_s), con $r,s = 1,2,\dots,t$, es un prerrequisito para poder hallar el estimador ajustado de credibilidad:

$$\mu_f(\theta) = \sum_{s=1}^{t} f^*(X_s)$$

ya que para obtener $f^*(X_s)$ es preciso solucionar el sistema de ecuaciones:

$$\sum_{l=0}^{n} l\cdot p_{ml} = f_m^* \cdot \sum_{l=0}^{n} p_{ml} + (t-1)\cdot \sum_{l=0}^{n} f_l^* \cdot p_{ml} \qquad (15)$$

$$m = 0,1,2,\dots,n$$

en el cual hemos asumido que las variables X_s toman un número finito de valores, y donde p_{ml} son las probabilidades conjuntas, probabilidades que se deben estimar previamente para poder solucionar este sistema de ecuaciones. A la matriz cuadrada de dimensión $(n+1, n+1)$, cuyos elementos son las probabilidades conjuntas antes citadas, se la suele denominar **matriz de credibilidad**, y la simbolizaremos por P.

Estimar dichas probabilidades no es tarea fácil. De VYLDER, F. y GOOVAERTS, M. (1984), pp. 179-188, hicieron un estudio exahustivo de las condiciones que debe verificar una matriz para poderla considerar una matriz de credibilidad, para el caso particular en que las variables: $X_1, X_2, \dots, X_t$ tomen un número finito de valores.

Dichos autores probaron que para **$n+1 \leq 4$**, una matriz P es una matriz de credibilidad si satisface las siguientes condiciones:

- P es no-negativa, es decir, si $p_{ml} \geq 0$.

- P es normalizada: $\sum_{m=0}^{n} \sum_{l=0}^{n} p_{ml} = 1$.

- P es simétrica y semidefinida positiva.

Sin embargo, si **$n+1 \geq 5$** las condiciones anteriores no son suficientes para asegurar que la matriz P sea una matriz de credibilidad de orden n+1, y en estos casos propusieron el siguiente TEOREMA:

> Cualquier matriz de orden n+1 es una combinación convexa de no más de $\frac{1}{2}\cdot(n+1)\cdot(n+2)$ matrices de credibilidad simples.

Su inconveniente, es que su aplicación no es nada práctica, debido al gran número de valores desconocidos involucrados en los cálculos.

Por lo general, considerar $n+1 \leq 4$ no presenta grandes dificultades en la mayoría de las aplicaciones prácticas. Así, por ejemplo, en los seguros de automóviles asumir que la probabilidad de tener más de 3 accidentes en un año es cero, no es difícil de aceptar. Si la variable X_s, en lugar de ser el número de reclamaciones fuese el importe de las mismas, considerar $n+1 \leq 4$, sí supone una limitación en principio, limitación que se puede solventar agrupando los importes de las reclamaciones en tres intervalos, pudiéndose de este modo calcular, de manera más fácil, la matriz P.

Un camino a seguir para hallar los valores de las probabilidades conjuntas, p_{ml}, es construir un estimador para la matriz P, que verifique las tres condiciones anteriores y que a la vez sea insesgado. GOOVAERTS, H. y HOOGSTAD, W. (1987), pp. 81, propusieron el siguiente estimador:

Sea I_q una función indicador, definida como sigue:

$$I_q = \begin{cases} 1 & \text{si q es verdaero} \\ 0 & \text{si q es falso} \end{cases}$$

entonces el estimador insesgado para p_{ml} , $\hat{p}_{ml}$, es:

$$\boxed{\hat{p}_{ml} = \frac{1}{k \cdot t \cdot (t-1)} \cdot \sum_{\substack{r=1 \\ r \neq s}}^{t} \sum_{s=1}^{t} \sum_{j=1}^{k} I_{X_{jr} = m,\ X_{js} = 1} \qquad l,m = 0,1,2,,..,n}$$

Este estimador tiene una limitación, es demasiado sensible a pequeños cambios en los datos, sin embargo, su cálculo es sencillo ya que si disponemos de una tabla de contingencias[20], la matriz $\hat{P}$, de dimensión (n+1, n+1), cuyos elementos son las probabilidades conjuntas estimadas

20 Una tabla de contingencias no es más que una tabla que nos proporciona la información concerniente al número de pólizas con l reclamaciones el primer año y m reclamciones durante el segundo año, siendo l,m = 0,1,2,...n.

$\hat{p}_{ml}$, coincide con $\frac{\pi + \pi'}{2 \cdot k}$, siendo π la matriz asociada a la tabla de contingencias.

Otro camino para hallar la matriz de credibilidad, es utilizar el **Procedimiento de la Máxima Verosimilitud.** Se trata de un camino alternativo al de los mínimos cuadráticos, y en este caso parece más adecuado, ya que su objetivo es hallar un estimador para la matriz P, de tal manera que, los resultados empíricos obtenidos tengan la probabilidad máxima de verificarse cuando el parámetro estimado tomase el valor asignado por el estimador.

Supongamos que tenemos una cartera que ha sido observada durante t años, y que está compuesta por k pólizas independientes e idénticamente distribuidas, siguiendo las hipótesis de nuestro modelo. En este caso, los subíndices 1 y 2 representan dos años cualesquiera observados. Vamos a considerar que los valores observados son el número de reclamaciones, siendo $0,1,2,\ldots,n$, de tal manera que la variable X_{js}, con $j = 1,2,\ldots,k$ y $s = 1,2$, nos indica el número de reclamaciones del contrato j durante el periodo s.

En este caso:

$$p_{ml} = P(X_{j1} = m,\ X_{j2} = l) \qquad \text{con} \quad m,l = 0,1,2,\ldots,n$$

probabilidad que no depende de j.

Por otro lado, vamos a simbolizar con N_{ml} el número de pólizas con m reclamaciones durante el primer año y l reclamaciones durante el segundo.

Si $n_{ml} \geq 0$, son números dados, cuya suma es k, entonces la probabilidad de que ocurra el evento:

$$N_{ml} = n_{ml}$$

de acuerdo con la ley multinomial, ya que parece la más adecuada, es:

$$P(N_{ml} = n_{ml}) = \frac{1}{\prod_{m=0}^{n} \prod_{l=0}^{n} n_{ml}!} \cdot \prod_{m=0}^{n} \prod_{l=0}^{n} p_{ml}^{\,n_{ml}}$$

Para estimar las probabilidades, p_{ml}, vamos a utilizar el procedimiento de la máxima verosimilitud, de tal manera que dichas probabilidades maximicen la probabilidad de obtener $P(N_{ml} = n_{ml})$, lo cual equivale a maximizar la función:

$$L = \log\left[\prod_{m=0}^{n} \prod_{l=0}^{n} p_{ml}^{\,n_{ml}}\right] = \sum_{m=0}^{n} \sum_{l=0}^{n} n_{ml} \cdot \log\, p_{ml}$$

ya que $\dfrac{1}{\prod_{m=0}^{n} \prod_{l=0}^{n} n_{ml}!}$ es una constante respecto a la probabilidad que deseamos estimar, y por lo tanto, no interviene en el proceso de maximización.

Según De VYLDER, F. y GOOVAERTS, M. (1984), pp. 184-185, la matriz P, de dimensión (n+1, n+1), cuyos elementos son las probabilidades p_{ml} se debe construir de tal manera que sea de la forma $P = b' \cdot b$, donde b es también una matriz de dimensión (n+1, n+1), no-negativa y triangular superior, siendo además necesario que se verifique:

$$\sum_{m=0}^{n} \sum_{l=0}^{n} p_{ml} = 1.$$

Para hallar los valores de las probabilidades p_{ml} que maximizan la función L:

$$L = \sum_{m=0}^{n} \sum_{l=0}^{n} n_{ml} \cdot \log p_{ml}$$

sujeta a las restricciones anteriores, De VYLDER, F. y GOOVAERTS, M. (1985) proponen utilizar el algoritmo que aparece en la subrutina EO4UAF de la NAGLIB: 1442-0:Mk 9.

Como punto de partida para este algoritmo podría ser útil utilizar la matriz $\hat{P}$, cuyos elementos $\hat{p}_{ml}$, son los estimadores de las probabilidades conjuntas, p_{ml}, obtenidos a partir de la expresión:

$$\hat{p}_{ml} = \frac{1}{k \cdot t \cdot (t-1)} \cdot \sum_{\substack{r=1 \\ r \neq s}}^{t} \sum_{s=1}^{t} \sum_{j=1}^{k} I_{X_{jr}=m,\ X_{js}=l}$$

$$l,m = 0,1,2,\dots,n$$

ya que la matriz $\hat{P}$ es un estimador insesgado de la matriz P.

CAPITULO V

MODELOS JERARQUICOS

INDICE DEL CAPITULO

Paralelamente a la aparición del Modelo de Regresión de HACHEMEISTER, JEWELL, W. (1975e), presentó su modelo conocido con el nombre de **MODELO JERARQUICO DE JEWELL**, tratándose de otra generalización del Modelo de BÜHLMANN-STRAUB pero en otra dirección, siendo a su vez el precursor de otra de las vertientes en la cual se ha desarrollado la Teoría de la Credibilidad desde la aparición del Modelo de BÜHLMANN

Mientras que HACHEMEISTER, C. (1975) generalizó el Modelo de BÜHLMANN-STRAUB abandonando la hipótesis de homogeneidad en el tiempo, en cuanto a las observaciones esperadas y la sustituyó por otra de tipo polinómica, e incorporó además la técnica de regresión, JEWELL, W. (1975e) asumió que cada cartera puede estar dividida en un cierto número de subcarteras, donde cada una de ellas está caracterizada por un parámetro de riesgo desconocido que describe cómo una subcartera difiere de las otras. Es decir, consideró un modelo jerárquico con parámetros aleatorios a dos niveles: en el nivel de las pólizas y en el nivel de las subcarteras.

En los modelos que hemos descrito hasta el momento, las características inobservables de los riesgos individuales están representados por unas variables aleatorias independientes e idénticamente distribuidas, los parámetros de riesgo, que describen cómo una póliza difiere de otras similares pertenecientes a la misma cartera. Los estimadores para las primas de riesgo individuales, están basados en la experiencia propia del riesgo, ya que se considera que los distintos riesgos contenidos en la cartera son independientes, y los parámetros de riesgo son también independientes y están generados por la misma función de distribución.

Todos estos modelos son apropiados cuando la cartera se puede considerar como una muestra aleatoria de una población de riesgos. En ellos, la información colateral sobre otros riesgos de la cartera no se tiene en cuenta de forma expresa, aunque intuitivamente parece que puede contener información relevante respecto al riesgo que se está considerando. No obstante, la información colateral acaba siendo introducida a través de los estimadores de los parámetros estructurales.

Como indica JEWELL, W. (1975e), pp. 7-8, en su modelo jerárquico los conceptos de variables aleatorias de riesgos individuales, parámetros de riesgo y cohorte de riesgos elegidos dentro de en un colectivo, se siguen utilizando, pero ahora consideramos que nuestro colectivo no es necesariamente representativo de otros posibles colectivos, los cuales constituyen un universo de colectivos. Según JEWELL, W. (1975e) ello significa que existe un hiperparámetro colectivo, el cual describe cómo los posibles colectivos pueden variar de unos a otros. Una vez que este hiperparámetro es elegido, y las características del colectivo estan definidas, entonces se eligen los parámetros de riesgo para cada uno de los miembros de nuestra cohorte, independientes e idénticamente distribuidos.

Para JEWELL, W. (1975e), su modelo tiene una interpretación práctica clara. Para ello basta que imaginemos una compañía de seguros en la cual cada riesgo individual es una póliza de seguro individual, y el colectivo es la cartera de coberturas similares dentro de nuestra compañía. Está claro que las carteras pueden variar de una compañía a otra, y que nuestra cartera puede ser mejor o peor que la media nacional. El universo de colectivos corresponde a la unión de todos los posibles contratos de riesgo de este tipo en la nación, esto es, está formado por los colectivos o carteras parecidas de todas las compañías del país, para las cuales, como indica JEWELL, W., podemos asumir que disponemos de estadísticos adecuados.

Desde su aparición, el Modelo Jerárquico de JEWELL ha despertado un gran interés, hasta tal punto que el modelo jerárquico no ha qrdado limitado a dos niveles, sino que se ha extendido para el caso de múltiples niveles, dando lugar a la aparición del **MODELO JERARQUICO CON MULTIPLES NIVELES**, siendo su idea básica la siguiente: los riesgos pertenecen a cohortes, los cohortes pertenecen a las compañías, las compañías pertenecen a una región determinada, etc... Los autores artífices de este modelo han sido TAYLOR, G. (1979), el cual utilizó las técnicas del espacio Hilberiano, BÜHLMANN, H. y JEWELL, W. (1987) y GOOVAERTS, M.; KAAS, R.; HEERWAARDEN, A. van y BAUWELINCKX, T. (1990).

Por otro lado, siguuiendo dentro de la linea de la creibilidad jerárquica a dos niveles, SUNDT, B. (1978, 1979a, 1980) presentó su modelo conocido con el nombre de Modelo de Regresión Jerárquico de SUNDT. Se trata de una combinación del Modelo Jerárquico de JEWELL y del Modelo de Regresión de HACHEMEISTER, en el cual se considera, al igual que en el Modelo de JEWELL, que la cartera se halla dividida en un cierto número de subcarteras, estando cada póliza caracterizada por dos parámetros de riesgo, sin embargo, la hipótesis de homogeneidad en el tiempo se abandona completamente, y es sustituida por otra de tipo polinómica para las observaciones esperadas, como se hizo en el Modelo de Regresión de HACHEMEISTER, al mismo tiempo que se utiliza la técnica de la regresión.

Del mismo modo que el Modelo Jerárquico de JEWELL se ha extendido para más de dos niveles, ha aparecido el **Modelo de Regresión Jerárquico con parámetros aleatorios para n niveles** que no es más que una generalización del Modelo de Regresión Jerárquico de SUNDT. Entre los autores que han tratado este tema están el propio SUNDT, B. (1980) y NORBERG, R. (1986), el cual centra su atención principalmente en la obtención de los estimadores lineales de credibilidad, al mismo tiempo que propone estimadores para los parámetros estructurales.

5.1.- MODELO JEARARQUICO DE JEWELL.

El Modelo Jerárquico de JEWELL se puede considerar como una generalización[21] del Modelo de BÜHLMANN-STRAUB, y aparece paralelamente al Modelo de Regresión de HACHEMEISTER, pero su óptica es completamente distinta.

En este modelo, se considera que la cartera existente puede ser dividida en un cierto número, p, de **subcarteras** o sectores, donde cada una de ellas está caracterizada por un parámetro de riesgo, que describe las diferencias existentes entre las distintas subcarteras, ya que cada subcartera está formada por un cierto número de pólizas que han sido agrupadas por poseer determinadas características básicas comunes. Sin embargo, cada póliza posee unas características específicas que la diferencian de las demás pólizas dentro de la subcartera, características específicas que vienen cuantificadas por otro parámetro de riesgo.

Se trata, por lo tanto, de un **Modelo Jerárquico a dos niveles**, donde cada póliza tiene asociados dos parámetros de riesgo, uno en el nivel de pólizas y otro en el nivel de las subcarteras.

Al igual que en el modelo de BÜHLMANN-STRAUB, las observaciones esperadas son homogéneas en el tiempo, mientras que la varianza depende del período considerado vía los pesos naturales.

---

21 En realidad, JEWELL, W. (1975e) considera que su modelo jerárquico es una generalización del Modelo de BÜHLMANN pero, al ser el Modelo de BÜHLMANN un caso particular del Modelo de BÜHLMANN-STRAUB, se puede también considerar como una generalizacion del Modelo de BÜHLMANN-STRAUB.

El objetivo de este modelo sigue siendo hallar los mejores estimadores de credibilidad lineales para las primas de riesgo individuales utilizando el procedimiento de los mínimos cuadrados, sin embargo, en este caso al haber estructurado la cartera en dos niveles, su obtención está ligada a la de los estimadores de credibilidad para las primas de riesgo de cada subcartera.

5.1.1.- Variables relevantes.

Las variables más relevantes de este modelo son:

- θ_p : Es una nueva variable que se introduce en este modelo. Es el parámetro de riesgo que caracteriza a la subcartera p, con p = 1,2,...,P, siendo P el número de subcarteras en las que se ha dividido la cartera.
  Al tratarse de un parámetro de riesgo, es una variable aleatoria inobservable.

- θ_{pj} : Es el parámetro de riesgo que describe las características de la póliza j-ésima perteneciente a la subcartera p, con $j = 1,2,\ldots,k_p$, siendo k_p el número de pólizas que constituyen la subcartera p. Es el mismo parámetro de riesgo que aparece en los otros modelos, pero en este caso le hemos añadido un nuevo subíndice, p, para indicar a que subcartera pertenece la póliza que estamos considerando.

- X_{pjs}: Es la variable aleatoria observable que nos indica la experiencia de reclamaciones para la póliza j-ésima, perteneciente a la subcartera p, en el periodo s-ésimo, con $p = 1,2,\dots,P$; $j = 1,2,\dots,k_p$ y $s = 1,2,\dots,t_{pj}$, siendo t_{pj} el número de periodos observados para la póliza j-ésima perteneciente a la subcartera p.

- w_{pjs}: Son las ponderaciones o pesos naturales de las variables observables X_{pjs}. Son números positivos conocidos.

En este modelo, la **subcartera p** viene caracterizada por el siguiente conjunto de variables:

$$(\theta_p, \theta_{pj}, X_{pjs}) \qquad j = 1,2,\dots,k_p;\ s = 1,\ 2,\dots,t_{pj}$$

y el **contrato pj** por:

$$(\theta_{pj}, X_{pjs}) \qquad s = 1,2,\dots,t_{pj}$$

Un esquema clarificador y representativo de este modelo puede ser el siguiente:

Subcarteras

	$p = 1$				$p = P$		
	θ_1			...	θ_P		
Variables estructurales	θ_{11}	θ_{12}	...	...	θ_{P1}	θ_{P2}	...
Variables observables	X_{111}	X_{121}	...	...	X_{P11}	X_{P21}	...
	(w_{111})	(w_{121})	...	...	(w_{P11})	(w_{P21})	...
(Ponderaciones naturales)	X_{112}	X_{121}	...	...	X_{P12}	X_{P21}	...
	(w_{112})	(w_{122})	...	...	(w_{P12})	(w_{P22})	...
	$\vdots$	$\vdots$			$\vdots$	$\vdots$	
Contratos	11	12	...		P1	P2	...

5.1.2.- Hipótesis del modelo.

Las hipótesis de este modelo son las siguientes:

J.1) Las subcarteras $p = 1,2,\dots,P$ son independientes, es decir, los pares $(\theta_p, \theta_{pj}, X_{pjs})$, con $j = 1,2,\dots,k_p$ y $s = 1,2,\dots,t_{pj}$, son independientes.

J.2) Para cada $p = 1,2,\dots,P$ y para cada θ_p dado, los contratos $pj = p1,p2,\dots,pk_p$, es decir, los pares (θ_{pj}, X_{pjs}), con $s = 1,2,\dots,t_{pj}$, son condicionalmente independientes.

J.3) Para cada $p = 1,2,\dots,P$; $j = 1,2,\dots,k_p$ y para cada (θ_p, θ_{pj}) dado, las observaciones condicionales: $X_{pj1}/\theta_p,\theta_{pj}$; $X_{pj2}/\theta_p,\theta_{pj}$;...; $X_{pjt_{pj}}/\theta_p,\theta_{pj}$ son independientes.

J.4) Todos los pares de variables (θ_p, θ_{pj}), con $p = 1,2,\dots,P$ y $j = 1,2,\dots,k_p$ están idénticamente distribuidos.

J.5) Para todo p, j, y s:

a) $E[X_{pjs}/\theta_p,\theta_{pj}] = \mu(\theta_p, \theta_{pj})$

b) $Var[X_{pjs}/\theta_p,\theta_{pj}] = \frac{1}{w_{pjs}}\cdot\sigma^2(\theta_p, \theta_{pj})$

donde $\mu(\theta_p, \theta_{pj})$ y $\sigma^2(\theta_p, \theta_{pj})$ no dependen de los subíndices p, j y s.

En este modelo las observaciones esperadas son homógeneas en el tiempo, mientras que la varianza depende del periodo considerado vía los pesos naturales, como ocurre en el Modelo de BÜHLMANN-STRAUB.

Vamos a definir $\mu(\theta_p)$ como:

$$\mu(\theta_p) = E[\mu(\theta_p, \theta_{pj})/\theta_p] = E[X_{pjs}/\theta_p]$$

que no es más que el valor esperado para todos los elementos de la subcartera p en el tiempo, o dicho de otra manera, es la prima de riesgo para la subcartera p, con $p = 1,2,\dots,P$.

5.1.3.- Estimadores de credibilidad lineales.

El Modelo Jerárquico de JEWELL nos proporciona, en última instancia, estimadores de credibilidad lineales para $\mu(\theta_p)$, en el nivel de las subcarteras, y para $\mu(\theta_p, \theta_{pj})$, en el nivel de las pólizas. Gráficamente podríamos indicar las esperanzas condicionadas para el modelo jerárquico del siguiente modo:

Subcarteras	1	...	p	...	P
Pólizas	(p,1)	...	(p,j)	...	(p,k_p)
Variables estructurales			θ_p $\theta_{p1} \ldots \theta_{p2} \ldots \theta_{pk_p}$		
Esperanzas de las subcarteras			$\mu(\theta_p) = E[X_{pjs}/\theta_p] = E[\mu(\theta_p, \theta_{pj})/\theta_p)]$		
Esperanzas de las pólizas	$\mu(\theta_p, \theta_{p1})$	...	$\mu(\theta_p, \theta_{pj})$	...	$\mu(\theta_p, \theta_{pk_p})$

Antes de obtener los estimadores de credibilidad vamos a definir los parámetros estructurales que aparecen en este modelo, que son los siguientes:

•) $m_p = \mu(\theta_p) = E[\mu(\theta_p, \theta_{pj})/\theta_p] = E[X_{pjs}/\theta_p]$

Es el valor esperado para las pólizas de la subcartera p, con $p = 1,2,\ldots,P$.

•) $m = E[\mu(\theta_p)] = E_{Total}[\mu(\theta_p, \theta_{pj})] = E_{Total}[X_{pjs}]$

Es la esperanza conjunta para la cartera en su totalidad.

•) $S^2 = E[\sigma^2(\theta_p, \theta_{pj})]$

Mide la heterogeneidad esperada en el tiempo de la experiencia de reclamaciones.

•) $a = E\left[Var[\mu(\theta_p, \theta_{pj})/\theta_p]\right]$

Mide ahora el grado de variabilidad esperado dentro de las subcarteras, o la heterogeneidad esperada dentro de las subcarteras.

•) $b = Var[\mu(\theta_p)]$

Mide la heterogeneidad entre las distintas subcarteras, siendo por lo tanto, una novedad de este modelo.

Por otro lado, la notación que vamos a utilizar no es más que una ampliación de la ya usada en el Modelo de BÜHLMANN-STRAUB. En este caso:

•) $w_{p\cdot} = \sum_{j=1}^{k_p} w_{pj\cdot} = \sum_{j=1}^{k_p} \sum_{s=1}^{t_{pj}} w_{pjs}$

•) Z_{pj}: Es el factor de credibilidad en el nivel de los contratos, con $p = 1,2,\dots,P$ y $j = 1,2,\dots,k_p$ siendo igual :

$$Z_{pj} = \frac{a\cdot w_{pj\cdot}}{S^2 + a\cdot w_{pj\cdot}}$$

factor que coincide con el obtenido en el Modelo de BÜHLMANN-STRAUB.

•) Z_p: Es el factor de credibilidad pero en el nivel de la subcarteras con $p = 1,2,\dots,P$, que viene definido de la siguiente manera como más adelante probaremos:

$$Z_p = \frac{b\cdot Z_{p\cdot}}{a + b\cdot Z_{p\cdot}}$$

En la definición de Z_{pj} aparecen como pesos, los pesos naturales w_{pjs}, mientras que en la definición de Z_p aparecen los pesos de credibilidad acumulados, siendo:

$$Z_{p\cdot} = \sum_{j=1}^{k_p} Z_{pj}$$

Además vamos a introducir las siguientes medias ponderadas:

$$X_{pjw} = \sum_{s=1}^{t_{pj}} \frac{w_{pjs} \cdot X_{pjs}}{w_{pj\cdot}}$$

$$X_{pzw} = \sum_{j=1}^{k_p} \frac{Z_{pj} \cdot X_{pjw}}{Z_{p\cdot}}$$

$$X_{zzw} = \sum_{p=1}^{P} \frac{Z_p \cdot X_{pzw}}{Z_\cdot}$$

siendo: $$Z_\cdot = \sum_{p=1}^{P} Z_p$$

Cabe destacar que en las medias que utilizaremos para las subcarteras y para la cartera, aparecen como ponderaciones naturales los factores de credibilidad acumulados en lugar de los pesos naturales w_{pjs}.

Según GOOVAERTS, M.; KAAS, R.; HEERWAARDEN, A. van, y BAUWELINCKX, T. (1990), pp. 162-164, bajo las hipótesis del Modelo de JEWELL se pueden obtener los siguientes resultados para las covarianzas:

- $\mathrm{Cov}[\mu(\theta_p, \theta_{pj}), X_{qis}] = \delta_{pq}(\delta_{ij} \cdot a + b)$

- $\mathrm{Cov}[\mu(\theta_p), X_{qjw}] = \delta_{pq} \cdot b$

- $\mathrm{Cov}[X_{pjs}, X_{pjs'}] = \dfrac{\delta_{ss'} \cdot S^2}{w_{pjs}} + a + b$

$$- \operatorname{Cov}[X_{pjs}, X_{pj's'}] = \delta_{jj'} \cdot \left[\frac{\delta_{ss'} \cdot S^2}{w_{pjs}} + a\right] + b$$

$$- \operatorname{Cov}[X_{pjs}, X_{qj's'}] = \emptyset \qquad p \neq q$$

$$- \operatorname{Cov}[X_{pjs}, X_{pj'w}] = \operatorname{Cov}[X_{pjw}, X_{pj'w}] = b + \delta_{jj'} \cdot \frac{a}{Z_{pj}}$$

$$- \operatorname{Cov}[X_{pjw}, X_{pzw}] = \operatorname{Cov}[X_{pzw}, X_{pzw}] = \frac{b}{Z_p} = b + \frac{a}{Z_{p\cdot}}$$

$$- \operatorname{Cov}[X_{pzw}, X_{zzw}] = \operatorname{Cov}[X_{zzw}, X_{zzw}] = \frac{b}{Z_{\cdot}}$$

Una vez introducidos todos estos elementos, vamos a pasar al problema que nos ocupa.

Estamos interesados en hallar los estimadores de credibilidad lineales, no-homogéneos, para las primas de riesgo individuales y para su obtención vamos a centrarnos en una póliza concreta, la pj.

Para obtener dicho estimador vamos a seguir el mismo procedimento que en el Modelo de BÜHLMANN-STRAUB: Seleccionar la mejor prima lineal dentro de la clase restringida de las primas lineales, utilizando el procedimiento de los mínimos cuadrados. Sin embargo, en este caso deberemos aplicar este procedimiento dos veces, como ya veremos.

Al tratarse de un Modelo Jerárquico a dos niveles, la póliza pj forma parte de la cartera considerada a través de la subcartera p a la cual pertenece, de manera que debemos obtener el estimador de credibilidad para $\mu(\theta_p, \theta_{pj})$, restringiendo su pertenencia a la subcartera de la cual forma parte, siendo en este caso el problema que debemos considerar el siguiente:

$$\text{Min } E\left[\left[\mu(\theta_p, \theta_{pj}) - c_0 - \sum_{i=1}^{k_p}\sum_{s=1}^{t_{pi}} c_{pis}\cdot X_{pis}\right]^2 / \theta_p\right] \qquad (1)$$

Este proceso de minimización tiene fácil solución ya que consiste simplemente en aplicar el Modelo de BÜHLMANN-STRAUB pero restringido a la subcartera p, siendo el resultado del mismo:

$$\mu^*(\theta_p, \theta_{pj}) = Z_{pj}\cdot X_{pjw} + [1 - Z_{pj}]\cdot E[\mu(\theta_p, \theta_{pj})/\theta_p] \qquad (2)$$

que a su vez es lo mismo que escribir:

$$\mu^*(\theta_p, \theta_{pj}) = Z_{pj}\cdot X_{pjw} + [1 - Z_{pj}]\cdot \mu(\theta_p) \qquad (3)$$

donde:

Z_{pj} : Es el factor de credibilidad para la póliza j-ésima perteneciente a la subcartera p, siendo:

$$Z_{pj} = \frac{a\cdot w_{pj\bullet}}{S^2 + a\cdot w_{pj\bullet}}$$

X_{pjw}: Es el estimador individual para $\mu(\theta_p, \theta_{pj})$ dentro de la subcartera p, siendo:

$$X_{pjw} = \sum_{s=1}^{t_{pj}} \frac{w_{pjs}\cdot X_{pjs}}{w_{pj\bullet}}$$

$m_p = \mu(\theta_p) = E[\mu(\theta_p, \theta_{pj})/\theta_p] = E[X_{pjs}/\theta_p]$: es el valor esperado para todas las pólizas de la subcartera p, o dicho de otra manera, es el estimador colectivo de $\mu(\theta_p, \theta_{pj})$.

Para poder obtener el estimador ajustado de credibilidad para la póliza pj, es necesario que estimemos $\mu(\theta_p)$, esto es, la prima de riesgo para la subcartera p, ya que ha aparecido en nuestro proceso de minimización.

Para obtener el estimador ajustado de credibilidad lineal, no-homogéneo, para $\mu(\theta_p)$ deberemos utilizar toda la información relativa a las demás subcarteras, ya que la subcartera p es una de las p subcarteras, con $p = 1,2,\ldots,P$, en las que hemos dividido nuestra cartera. Se trata, por lo tanto, de aplicar otra vez el Modelo de BÜHLMANN-STRAUB, pero ahora para estimar la prima de riesgo para la subcartera p, perteneciente a una cartera dada. En este caso, el problema de minimización que debemos considerar es el siguiente:

$$\text{Min } E_{\text{Total}}\left[\left[\mu(\theta_p) - c_\emptyset - \sum_{q=1}^{P}\sum_{j=1}^{k_q}\sum_{s=1}^{t_{qj}} c_{qjs}\cdot X_{qjs}\right]^2\right] \qquad (4)$$

El objetivo sigue siendo hallar los coeficientes $c_\emptyset$ y c_{qjs} tales que minimicen la esperanza total anterior. En este caso, vamos a desarrollar el proceso de minimización debido a la importancia de algunos resultados intermedios. El procedimiento a seguir es el siguiente:

Si derivamos parcialmente respecto a $c_\emptyset$ e igualamos a cero, resulta que:

$$c_\emptyset = E[\mu(\theta_p)] - \sum_{q,j,s} c_{qjs}\cdot E[X_{qjs}] \qquad (5)$$

o lo que es lo mismo:

$$c_{\emptyset} = m - \sum_{q,j,s} c_{qjs} \cdot m \tag{6}$$

Sustituyendo (5) en (4) tenemos:

$$\operatorname*{Min}_{c_{q'j's'}} E\left[\left[\mu(\theta_p) - E[\mu(\theta_p)] + \sum_{qjs} c_{qjs} \cdot E[X_{qjs}] - \sum_{qjs} c_{qjs} \cdot X_{qjs}\right]^2\right] \tag{7}$$

Derivando parcialmente respesto a $c_{q'j's'}$ e igualando a cero resulta:

$$E\left[\left[\mu(\theta_p) - E[\mu(\theta_p)] - \sum_{qjs} c_{qjs}(X_{qjs} - E[X_{qjs}])\right] \cdot X_{q'j's'}\right] = 0 \tag{8}$$

$$q' = 1,2,\ldots,P; \quad j' = 1,2,\ldots,k_{q'}; \quad s' = 1,2,\ldots,t_{qj'}$$

esperanza que también podemos escribir:

$$E\left[[\mu(\theta_p) - E[\mu(\theta_p)]] \cdot X_{q'j's'}\right] = \sum_{qjs} c_{qjs} \cdot \left[E\left[[X_{qjs} - E[X_{qjs}]] \cdot X_{q'j's'}\right]\right] \tag{9}$$

$$q' = 1,2,\ldots,P; \quad j' = 1,2,\ldots,k_{q'}; \quad s' = 1,2,\ldots,t_{qj'}$$

Aplicando la relación de la covarianza[22] esta igualdad es análoga:

$$\mathrm{Cov}[\mu(\theta_p),\ X_{q'j's'}] = \sum_{qjs} c_{qjs}\cdot \mathrm{Cov}[X_{qjs},\ X_{q'j's'}] \qquad (10)$$

$$q' = 1,2,\ldots,P;\quad j' = 1,2,\ldots,k_{q'};\quad s' = 1,2,\ldots,t_{q'j'}$$

que a su vez es igual:

$$\delta_{pq'}\cdot b = \sum_{qjs} c_{qjs}\cdot\delta_{qq'}\cdot\left[\delta_{jj'}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{qjs}} + a\right] + b\right] \qquad (11)$$

$$q' = 1,2,\ldots,P;\quad j' = 1,2,\ldots,k_{q'};\quad s' = 1,2,\ldots,t_{q'j'}$$

tratándose de un sistema de tantas ecuaciones como incógnitas c_{qjs} tenemos, esto es $P \times k_p \times t_{pj}$.

Si **p ≠ q'** el sistema de ecuaciones anterior se convierte en:

$$0 = \sum_{qjs} c_{qjs}\cdot\delta_{qq'}\cdot\left[\delta_{jj'}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{qjs}} + a\right] + b\right] \qquad (12)$$

$$q' =\{1,2,\ldots,P\}-\{p\};\quad j' = 1,2,\ldots,k_{q'};\quad s' = 1,2,\ldots,t_{q'j'}$$

tratándose de un sistema de ecuaciones homogéneo, compatible determinado, ya que por la hipótesis de independencia asumidas en este modelo resulta que todas las ecuaciones son independientes, siendo el rango de la matriz

22 Recordemos que $\mathrm{Cov}[x,\ y] = E[(x - E[X])\cdot(y - E[y])] =$
$= E[(x - E[x])\cdot y] - E[(x - E[X])]\cdot E[y] =$
$= E[(x - E[x])\cdot y]$

asociada al sistema igual al número de incógnitas. Al tener solución trivial resulta que todos los coeficientes $c_{q'j's'}$ son iguales a cero para todo j' y s'. Esto es $c_{q'j's'} = \emptyset$ para $q' = \{1,2,\ldots,P\} - \{p\}$.

Si $p = q'$ tenemos para $j' = 1,2,\ldots,k_p$; $s' = 1,2,\ldots,t_{pj'}$:

$$b = \sum_{qjs} c_{qjs}\cdot\delta_{qp}\cdot\left[\delta_{jj'}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{qjs}} + a\right] + b\right] \tag{13}$$

$$j' = 1,2,\ldots,k_p;\quad s' = 1,2,\ldots,t_{pj'}$$

que a su vez es igual:

$$\boxed{b = \sum_{js} c_{pjs}\cdot\left[b + \delta_{jj'}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{pjs}} + a\right]\right] \qquad j' = 1,2,\ldots,k_p;\; s' = 1,2,\ldots,t_{pj'}} \tag{14}$$

o también:

$$b = \sum_{js} c_{pjs}\cdot b + \sum_{js} c_{pjs}\cdot\delta_{jj'}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{pjs}} + a\right] \tag{15}$$

$$j' = 1,2,\ldots,k_p;\quad s' = 1,2\ldots,t_{pj'}$$

expresión que a su vez se reduce:

$$b = \sum_{js} c_{pjs}\cdot b + \sum_{s=1}^{t_{pj'}} c_{pj's}\cdot\left[\delta_{ss'}\cdot\frac{S^2}{w_{pj's}} + a\right] \tag{16}$$

$$j' = 1,2,\ldots,k_p;\quad s' = 1,2,\ldots,t_{pj'}$$

o de forma análoga:

$$b = \sum_{js} c_{pjs} \cdot b + \sum_{s=1}^{t_{pj'}} c_{pj's} \cdot a + \frac{c_{pj's'}}{w_{pj's'}} \cdot S^2 \tag{17}$$

$$j' = 1,2,\ldots,k_p;\; s' = 1,2,\ldots,t_{pj'}$$

Operando resulta que:

$$b = c_{p\cdot\cdot} \cdot b + a \cdot c_{pj'\cdot} + S^2 \cdot \frac{c_{pj's'}}{w_{pj's'}} \tag{18}$$

$$j' = 1,2,\ldots,k_p;\; s' = 1,2,\ldots,t_{pj'}$$

que es un sistema de tantas ecuaciones como incógnitas $k_p \times t_{pj'}$ tenemos, y con la peculiaridad de que todas las ecuaciones son iguales a la constante b.

Para cada valor de j', tenemos $t_{pj'}$ ecuaciones iguales, que en lo único que se diferencian es en el subíndice s' de $\frac{c_{pj's'}}{w_{pj's'}}$ ya que $s' = 1,2,\ldots,t_{pj'}$. Si consideramos el caso particular $j' = j$, obtenemos un sistema de t_{pj} ecuaciones con $k_p \times t_{pj}$ incógnitas.

$$\left\{\begin{array}{l} b = c_{p\cdot\cdot} \cdot b + a \cdot c_{pj\cdot} + S^2 \cdot \dfrac{c_{pj1}}{w_{pj1}} \\ b = c_{p\cdot\cdot} \cdot b + a \cdot c_{pj\cdot} + S^2 \cdot \dfrac{c_{pj2}}{w_{pj2}} \\ \vdots \qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ b = c_{p\cdot\cdot} \cdot b + a \cdot c_{pj\cdot} + S^2 \cdot \dfrac{c_{pjt_{pj}}}{w_{pjt_{pj}}} \end{array}\right. \tag{19}$$

Cogiendo dos ecuaciones cualesquiera, por ejemplo, para s y s' resulta que al tener el mismo término independiente se verifica:

$$c_{p\cdot\cdot}\cdot b + a\cdot c_{pj\cdot} + S^2\cdot\frac{c_{pjs}}{w_{pjs}} = c_{p\cdot\cdot}\cdot b + a\cdot c_{pj\cdot} + S^2\cdot\frac{c_{pjs'}}{w_{pjs'}}$$

o lo que es lo mismo:

$$\boxed{\frac{c_{pjs}}{w_{pjs}} = \frac{c_{pjs'}}{w_{pjs'}}} \tag{20}$$

De manera que los coeficientes c_{pjs} son proporcionales a las ponderaciones naturales w_{pjs}, para cada valor de j fijado.

Al darse esta relación de proporcionalidad entre los coeficientes $c_{pj's}$, y $w_{pj's}$, para cada valor de j' dado, el sistema de ecuaciones (18) de $k_p \times t_{pj'}$ ecuaciones y $k_p \times t_{pj'}$ incógnitas, se reduce a un sistema de k_p ecuaciones con k_p incógnitas, siendo las incógnitas: c_{p1}, $c_{p2}, \ldots, c_{pk_p}$.

Para obtener este sistema de ecuaciones, consideremos de nuevo el sistema de ecuaciones (19). Al verificarse la relación de proporcionalidad (20), podemos considerar que:

$$\boxed{\frac{c_{pj's}}{w_{pj's}} = \frac{c_{pj'}}{w_{pj'\cdot}} \quad ====\rangle \quad c_{pj's} = \frac{c_{pj'}\cdot w_{pj's}}{w_{pj'\cdot}} \qquad s = 1,2,\ldots,t_{pj'}} \tag{21}$$

Al verificarse (21), el sistema de ecuaciones (19) queda reducido a una sola ecuación, y si en ella sustituimos $c_{pj's}$ por su valor, nos queda:

$$b = b \cdot c_{p\bullet} + a \cdot c_{pj'} + S^2 \cdot \frac{c_{pj'}}{w_{pj'\bullet}} \tag{23}$$

y sacando factor común c_{pj}:

$$b = b \cdot c_{p\bullet} + \left[a + \frac{S^2}{w_{pj'\bullet}} \right] \cdot c_{pj'} \tag{24}$$

Si la expresión de dentro del corchete la multiplicamos y dividimos por a, obtenemos:

$$b = b \cdot c_{p\bullet} + a \cdot \frac{c_{pj'}}{Z_{pj'}} \tag{25}$$

resultando que para cada valor de $j' = 1,2,\ldots,k_p$, tenemos un sistema de $t_{pj'}$ ecuaciones como en (19), y cada uno de ellos, a su vez, queda reducido a una sola ecuación, al verificarse la relación de proporcionalidad (20) para cada $j' = 1,2,\ldots,k_p$. De manera, que el sistema de ecuaciones inicial (18) se convierte en un sistema de k_p ecuaciones con k_p incógnitas, siendo su expresión general:

$$\boxed{\begin{array}{c} b = b \cdot c_{p\bullet} + a \cdot \dfrac{c_{pj'}}{Z_{pj'}} \\ j' = 1,2,\ldots,k_p \end{array}} \tag{26}$$

A su vez, este sistema de ecuaciones es simétrico respecto a $\frac{c_{pj'}}{Z_{pj'}}$, de manera que:

$$\frac{c_{pj'}}{Z_{pj'}} = R = \text{CONSTANTE} \qquad \text{para} \quad j' = 1,2,\ldots,k_p \tag{27}$$

constante que puede ser determinada como sigue:

$$b = b\cdot c_{p\cdot} + a\cdot R = a\cdot R + b\cdot R\cdot \sum_{j'=1}^{k_p} Z_{pj'} =$$

$$= R\cdot(a + b\cdot Z_{p\cdot})$$

siendo:

$$\boxed{R = \frac{b}{a + b\cdot Z_{p\cdot}}} \tag{28}$$

Sustituyendo este resultado en (27), obtenemos para cada $j' = 1,2,\ldots,k_p$:

$$c_{pj'} = R\cdot Z_{pj'} = \frac{b\cdot Z_{pj'}}{a + b\cdot Z_{p\cdot}} = \frac{Z_{pj'}}{Z_{p\cdot}}\cdot\frac{b\cdot Z_{p\cdot}}{a + b\cdot Z_{p\cdot}} \tag{29}$$

y si definimos una nueva variable Z_p como:

$$\boxed{Z_p = \frac{b\cdot Z_{p\cdot}}{a + b\cdot Z_{p\cdot}}} \tag{30}$$

resulta que la ecuación (29) se puede también escribir:

$$\boxed{\begin{array}{l} c_{pj'} = Z_p \cdot \dfrac{Z_{pj'}}{Z_{p\cdot}} \\ \qquad j' = 1,2,\ldots,k_p \end{array}} \tag{31}$$

Una vez que hemos obtenido los valores de los coeficientes c_{pjs}, la obtención del coeficiente $c_\emptyset$ a partir de la ecuación (6) es inmediata, ya que al verificarse que todos los coeficientes $c_{qjs} = \emptyset$ para $q = \{1,2,\ldots,P\} - \{p\}$ y todo j y s, resulta:

$$c_\emptyset = m - \sum_{q,j,s} c_{qjs} \cdot m = m - \sum_{j=1}^{k_p} \sum_{s=1}^{t_{pj}} c_{pjs} \cdot m$$

Susutituyendo c_{pjs} por (21), y luego c_{pj} por (31) obtenemos:

$$c_\emptyset = m - \sum_{j=1}^{k_p} c_{pj} \cdot m = m - \sum_{j'=1}^{k_p} Z_p \cdot \frac{Z_{pj'}}{Z_{p\cdot}} \cdot m =$$

$$= m - Z_{p\cdot} \cdot m \cdot \sum_{j'=1}^{k_p} \frac{Z_{pj'}}{Z_{p\cdot}} = m - Z_p \cdot m$$

siendo:

$$\boxed{c_\emptyset = [1 - Z_p] \cdot m} \tag{32}$$

y al mismo tiempo como todos los coeficientes $c_{qjs} = \emptyset$ para $q = \{1,2,\ldots,P\} - \{p\}$ y para todo j y s, resulta que el estimador ajustado de credibilidad para la subcartera p, que habíamos definido como:

$$\widehat{\mu(\theta_p)} = c_\emptyset + \sum_{q=1}^{p} \sum_{j=1}^{k_q} \sum_{s=1}^{t_{qj}} c_{qjs} \cdot X_{qjs}$$

se convierte en:

$$\widehat{\mu(\theta_p)} = c_\emptyset + \sum_{j=1}^{k_p} \sum_{s=1}^{t_{pj}} c_{pjs} \cdot X_{pjs}$$

que, a su vez, se puede también escribir del siguiente modo si sustituimos el coeficiente c_{pjs} por (21):

$$\widehat{\mu(\theta_p)} = c_\emptyset + \sum_{j=1}^{k_p} c_{pj} \cdot X_{pjw}$$

Y por último, si sustituimos $c_\emptyset$ y c_{pj} por sus respectivos valores, dados en (32) y (31), obtenemos:

$$\widehat{\mu(\theta_p)} = [1 - Z_p] \cdot m + Z_p \cdot \sum_{j=1}^{k} \frac{Z_{pj} \cdot X_{pjw}}{Z_{p\cdot}}$$

siendo:

$$\boxed{\widehat{\mu(\theta_p)} = [1 - Z_p] \cdot m + Z_p \cdot X_{pzw}} \qquad (33)$$

donde:

$m = E[\mu(\theta_p)]$: Es la esperanza conjunta de la cartera en su totalidad, que es uno de los parámetros estructurales del modelo.

Z_p: Es el factor de credibilidad para la subcartera p, cuya expresión es la siguiente:

$$Z_p = \frac{b \cdot Z_{p\cdot}}{a + b \cdot Z_{p\cdot}}$$

X_{pzw}: Es el estimador individual para $\mu(\theta_p)$, donde:

$$X_{pzw} = \sum_{j=1}^{k_p} \frac{Z_{pj} \cdot X_{pjw}}{Z_{p\cdot}}$$

Una vez obtenido el estimador ajustado de credibilidad para $\mu(\theta_p)$, lo sustituimos en (3) y obtenemos, de este modo, el estimador ajustado de credibilidad para la prima de riesgo individual pj, siendo:

$$\widehat{\mu(\theta_p, \theta_{pj})} = Z_{pj} \cdot X_{pjw} + [1 - Z_{pj}] \cdot [[1 - Z_p] \cdot m + Z_p \cdot X_{pzw}] =$$

$$= Z_{pj} \cdot X_{pjw} + [1 - Z_{pj}] \cdot \left[\frac{a \cdot m + b \cdot Z_p \cdot X_{pzw}}{a + b \cdot Z_{p\cdot}} \right]$$

Pero como:

$$X_{pzw} = \sum_{j=1}^{k_p} \frac{Z_{pj} \cdot X_{pjw}}{Z_{p\cdot}}$$

resulta que:

$$Z_{p\cdot} \cdot X_{p\ w} = \sum_{j=1}^{k_p} Z_{pj} \cdot X_{pjw}$$

siendo:

$$\widehat{\mu(\theta_p, \theta_{pj})} = Z_{pj} \cdot X_{pjw} + \\ + [1 - Z_{pj}] \cdot \left[\frac{a \cdot m + b \cdot \sum_{j=1}^{k_p} Z_{pj} \cdot X_{pjw}}{a + b \cdot Z_{p\bullet}} \right] \tag{34}$$

Este estimador coincide prácticamente con el obtenido por JEWELL, W. (1975e), pp. 9. La única diferencia está en que JEWELL, W. no trabajó con los pesos naturales y nosotros si.

Una vez obtenido el estimador ajustado de credibilidad lineal, no-homogéneo para la prima de riesgo individual pj, el siguiente paso es comprobar que $\widehat{\mu(\theta_p, \theta_{pj})}$ **es el mejor estimador de credibilidad lineal.** Para ello es necesario que verifique, como indica, entre otros, SUNDT, B. (1980), pp 26-27, las dos condiciones siguientes:

a) $E[\mu(\theta_p, \theta_{pj})] = E[\widehat{\mu(\theta_p, \theta_{pj})}]$

b) $Cov[\mu(\theta_p, \theta_{pj}), X_{pzw}] = Cov[\widehat{\mu(\theta_p, \theta_{pj})}, X_{pzw}]$

Comprobar que $E[\mu(\theta_p, \theta_{pj})] = E[\widehat{\mu(\theta_p, \theta_{pj})}]$ no presenta demasiadas dificultades, ya que por un lado, ya sabemos que:

$$E[\mu(\theta_p, \theta_{pj})] = E[X_{pjs}] = E[\mu(\theta_p)] = m$$

y por otro :

$$E[\widehat{\mu(\theta_p, \theta_{pj})}] = Z_{pj}\cdot E[X_{pjw}] + + [1 - Z_{pj}]\cdot \left[\frac{a\cdot m + b\cdot \sum_{j=1}^{k_p} Z_{pj}\cdot E[X_{pjw}]}{a + b\cdot Z_{p\cdot}}\right]$$

siendo:

$$E[X_{pjw}] = \sum_{s=1}^{t_{pj}} \frac{w_{pjs}\cdot E[X_{pjs}]}{w_{pj\cdot}} = \frac{m}{w_{pj\cdot}}\cdot \sum_{s=1}^{t_{pj}} w_{pjs} = m$$

de manera que:

$$E[\widehat{\mu(\theta_p,\theta_{pj})}] = Z_{pj}\cdot m + [1 - Z_{pj}]\cdot \frac{a\cdot m + b\cdot m\cdot Z_{p\cdot}}{a + b\cdot Z_{p\cdot}} =$$

$$= Z_{pj}\cdot m + [1 - Z_{pj}]\cdot m = m$$

En cuanto a la igualdad entre las covarianzas, debemos comprobar que se verifica:

$$\mathrm{Cov}[\mu(\theta_p, \theta_{pj}), X_{pzw}] =$$

$$= \mathrm{Cov}\left[\mu(\theta_p, \theta_{pj}), \sum_{j=1}^{k_p} \frac{Z_{pj}}{Z_{p\cdot}}\cdot \sum_{s=1}^{t_{pj}} \frac{w_{pjs}}{w_{pj\cdot}}\cdot X_{pjs}\right] =$$

$$= \mathrm{Cov}[\widehat{\mu(\theta_p, \theta_{pj})}, X_{pzw}]$$

Como ya sabemos:

$$\mathrm{Cov}[\mu(\theta_p, \theta_{pj}), X_{qir}] = \delta_{pq}\cdot(\delta_{ij}\cdot a + b) = \delta_{pq}\cdot\delta_{ij}\cdot a + \delta_{pq}\cdot b$$

de manera que:

$$\mathrm{Cov}[\mu(\theta_p,\ \theta_{pj}),\ X_{pzw}] = \sum_{j=1}^{k_p} \frac{Z_{pj}}{Z_{p\bullet}} \cdot \sum_{s=1}^{t_{pj}} \frac{w_{pjs}}{w_{pj\bullet}} \cdot b\ +$$

$$+ \sum_{j=1}^{k_p} \frac{Z_{pj}}{Z_{p\bullet}} \cdot \sum_{s=1}^{t_{pj}} \frac{w_{pjs}}{w_{pj\bullet}} \cdot a = b + \frac{Z_{pj}}{Z_{p\bullet}} \cdot a$$

A su vez:

$$\mathrm{Cov}[\widehat{\mu(\theta_p,\ \theta_{pj})},\ X_{pzw}] = Z_{pj} \cdot \mathrm{Cov}[X_{pjw},\ X_{pzw}]\ +$$

$$+ [1 - Z_{pj}] \cdot \frac{b \cdot \sum_{j=1}^{k_p} Z_{pj} \cdot \mathrm{Cov}[X_{pjw},\ X_{pzw}]}{a + b \cdot Z_{p\bullet}}$$

Como:

$$\mathrm{Cov}[X_{pjw},\ X_{pzw}] = \frac{b}{Z_p} = b + \frac{a}{Z_{p\bullet}}$$

resulta que:

$$\mathrm{Cov}[\widehat{\mu(\theta_p,\ \theta_{pj})},\ X_{pzw}] = \frac{Z_{pj}}{Z_p} \cdot b\ +$$

$$+ [1 - Z_{pj}] \cdot \frac{\frac{b^2}{Z_p} \cdot \sum_{j=1}^{k_p} Z_{pj}}{a + b \cdot Z_{p\bullet}} =$$

$$= \frac{Z_{pj}}{Z_p} \cdot b + [1 - Z_{pj}] \cdot \frac{\frac{b^2}{Z_p} \cdot Z_{p\bullet}}{a + b \cdot Z_{p\bullet}} =$$

$$= \frac{Z_{pj}}{Z_p} \cdot b + [1 - Z_{pj}] \cdot \frac{b}{Z_p} \cdot \frac{b \cdot Z_{p\bullet}}{a + b \cdot Z_{p\bullet}} =$$

$$= \frac{Z_{pj}}{Z_p} \cdot b + [1 - Z_{pj}] \cdot b = a \cdot \frac{Z_{pj}}{Z_p} + b$$

siendo:

$$\mathrm{Cov}[\mu(\theta_p, \theta_{pj}), X_{pzw}] = \mathrm{Cov}[\widehat{\mu(\theta_p, \theta_{pj})}, X_{pzw}] = a \cdot \frac{Z_{pj}}{Z_p} + b$$

Como $\widehat{\mu(\theta_p, \theta_{pj})}$ verifica las dos condiciones antes citadas, resulta que se trata del mejor estimador de credibilidad lineal, no-homogéneo, para la prima de riesgo individual.

En lugar de plantearnos la obtención de los estimadores de credibilidad lineales no-homogéneos, nos hubiésemos podido plantear su obtención en el caso de que fuesen **homogéneos**, esto decir, en el caso de que no existiese la constante c_0. En esta circunstancia, bajo las hipótesis del modelo considerado, los estimadores de credibilidad lineales homogéneos para la prima de riesgo individual pj, y para la subcartera p, son respectivamente:

$$\widehat{\mu(\theta_p, \theta_{pj})} = [1 - Z_{pj}] \cdot \widehat{\mu(\theta_p)} + Z_{pj} \cdot X_{pjw}$$

$$\widehat{\mu(\theta_p)} = [1 - Z_p] \cdot X_{zzw} + Z_p \cdot X_{pzw}$$

donde:

$$Z_{pj} = \frac{b \cdot w_{pj\cdot}}{S^2 + a \cdot w_{pj\cdot}}$$

$$Z_p = \frac{b \cdot Z_{p\cdot}}{a + b \cdot Z_{p\cdot}}$$

5.1.3.1.- Un caso particular.

Si consideramos que **θ_p es constante para $\forall$ p, y que $m_p = m$,** es decir, que la esperanza conjunta de la cartera en su totalidad coincide con el valor esperado para las pólizas de cada subcartera resulta que al ser

$$E[\mu(\theta_p, \theta_{pj})/\theta_p] = E[\mu(\theta_p)] = m$$

el estimador de credibilidad lineal para la prima de riesgo individual pj es igual:

$$\widehat{\mu(\theta_p, \theta_{pj})} = Z_{pj} \cdot X_{pjw} + [1 - Z_{pj}] \cdot m$$

no siendo preciso en este caso volver a aplicar el Modelo de BÜHLMANN-STRAUB para estimar $E[\mu(\theta_p, \theta_{pj})/\theta_p]$.

Se trata de una hipótesis muy restrictiva, ya que estamos asumiendo que todas las subcarteras tienen idéntica prima y que, a su vez, coincide con la esperanza conjunta de la cartera. En realidad, introduciendo esta hipótesis hemos suprimido la jerarquización propia del Modelo de JEWELL.

5.1.4.- Estimación de los parámetros estructurales.

Como ya hemos visto, el estimador de credibilidad lineal, no-homogéneo, para la prima de riesgo individual $\mu(\theta_p, \theta_{pj})$ es:

$$\widehat{\mu(\theta_p, \theta_{pj})} = Z_{pj}\cdot X_{pjw} + \\ + [1 - Z_{pj}]\cdot\left[\frac{a\cdot m + b\cdot\sum_{j=1}^{k_p} Z_{pj}\cdot X_{pjw}}{a + b\cdot Z_{p\bullet}}\right]$$

siendo:

m: la esperanza conjunta de la cartera en su totalidad.

$$Z_{pj} = \frac{b\cdot w_{pj\bullet}}{S^2 + a\cdot w_{pj\bullet}}$$

Por otro lado, el estimador ajustado de credibilidad lineal que hemos obtenido para $\mu(\theta_p)$ es:

$$\widehat{\mu(\theta_p)} = N_p^a = [1 - Z_p]\cdot m + Z_p\cdot X_{pzw}$$

siendo:

$$Z_p = \frac{b\cdot Z_{p\bullet}}{a + b\cdot Z_{p\bullet}}$$

Para poder aplicar ambos estimadores es preciso que estimemos previamente los parámetros estructurales que aparecen en los mismos: m, S^2, a y b, parámetros que ya hemos definido anteriormente.

Los estimadores insesgados propuestos para estos parámetros son los siguientes:

•) $\hat{m} = N_{\emptyset} = X_{zzw}$ para la esperanza conjunta de la cartera en su totalidad, siendo:

$$\boxed{\hat{m} = \sum_{p=1}^{P} \frac{Z_p \cdot X_{pzw}}{Z_{\bullet}}}$$

•) $\hat{S}^2$ para S^2, que mide la heterogeneidad esperada de la experiencia de reclamaciones en el tiempo:

$$\boxed{\hat{S}^2 = \frac{\displaystyle\sum_{p=1}^{P}\sum_{j=1}^{k_p}\sum_{s=1}^{t_{pj}} w_{pjs}\cdot[X_{pjs} - X_{pjw}]^2}{\displaystyle\sum_{p=1}^{P}\sum_{j=1}^{k_p}(t_{pj} - 1)} \qquad \text{con } t_{pj} \geq 2}$$

•) $\hat{a}$ para a, que mide la heterogeneidad esperada dentro de las subcarteras:

$$\hat{a} = \frac{\sum_{p=1}^{P} \sum_{j=1}^{k_p} Z_{pj} \cdot [X_{pjw} - X_{pzw}]^2}{\sum_{p=1}^{P} (k_p - 1)} \qquad \text{con } k_p \geq 2$$

•) $\hat{b}$ para b, que mide la heterogeneidad entre las distintas subcarteras:

$$\hat{b} = \frac{\sum_{p=1}^{P} Z_p \cdot [X_{pzw} - X_{zzw}]^2}{P - 1}$$

La demostración de la insesgadez de estos estimadores puede hallarse en el libro de GOOVAERTS, M.; KAAS, R.; HEERWAARDEN, A. van, y BAUWELINCKX, T. (1990), pp. 168-169.

Como puede apresciarse, el estimador para S^2 se puede calcular directamente a través de las variables observables, experiencia de reclamaciones. Sin embargo, para poder aplicar la fórmula para calcular $\hat{a}$, es necesario que conozcamos los valores de a y S^2, ya que ambos aparecen en la definición de Z_{pj} y X_{pzw}. En realidad, $\hat{a}$ es un **pseudo-estimador** ya que contiene, a través de Z_{pj} y X_{pzw}, el propio parámetro que estamos estimando. Estimadores como este se solucionan, como ya hemos visto, mediante un proceso iterativo. Esto es, escogemos un valor inicial para $\hat{a}$, con el que calcularemos unos valores para Z_{pj} y

X_{pzw}. Con estos valores obtendremos un nuevo valor para $\hat{a}$, más perfeccionado, e iremos repitiendo el proceso hasta que $\hat{a}$ se estabilice.

Para calcular $\hat{b}$, nos encontramos con el mismo problema que para $\hat{a}$, ya que dicho estimador es un **pseudo-estimador**, que contiene a través de Z_p y X_{zzw} el propio parámetro b que estamos estimando, y a través de X_{pzw} el parámetro a. Para hallar el valor de $\hat{b}$, introduciremos el valor de $\hat{a}$ obtenido en el proceso iterativo anterior, y tomaremos un valor inicial para $\hat{b}$ iniciándose un nuevo proceso iterativo, a través del cual obtendremos el valor de $\hat{b}$ a utilizar.

En general, la utilización de un proceso iterativo proporciona una solución, sin embargo en algunas ocasiones converge a cero. Así, por ejemplo, si $\hat{a} = 0$, indicaría que no hay heterogeneidad dentro de las subcarteras. En este caso, deberíamos pensar en otra subdivisión para las subcarteras.

Una vez obtenidos los valores de $\hat{S}^2$, $\hat{a}$ y $\hat{b}$ podremos hallar los factores de credibilidad Z_{pj} y Z_p, así como X_{pzw}, pudiendo de este modo calcular $\hat{m}$.