IX Congreso de Ingeniería de Organización
Gijón, 8 y 9 de septiembre de 2005

Desempeño estadístico en modelos de preferencias y relación con el diseño de experimentos no lineal

J. Nicolás Ibáñez[1], Luis Onieva[1], Jesús Muñuzuri[1], José Guadix[1]

[1] Grupo de Ingeniería de Organización. Escuela Superior de Ingenieros de Sevilla. Avda. Descubrimientos, s/n, 41092 Sevilla. juannicolas@us.es, onieva@us.es, munuzuri@esi.us.es, guadix@esi.us.es

Palabras clave: Diseño de experimentos, Ajuste de modelos, Matrices de covarianza

Resumen

En este trabajo se abordan los modelos de preferencias de tipo no lineal, entendiendo por ellos los que forman la familia de modelos de elección discreta, los cuales se emplean para cuantificar factores que definan las preferencias observadas de un grupo de individuos. En este trabajo se exponen resultados relacionados con el diseño de encuestas para recoger las preferencias de los individuos y aplicar modelos no lineales (diseños no lineales de experimentos). Se hace especial hincapié en los tres siguientes aspectos: la diferente naturaleza con respecto al caso lineal, explicar las interacciones de este diseño con el resto de fases más conocidas de un estudio de preferencias y en presentar el origen de la componente estadística de los modelos. De esta explicación se deriva la formulación del problema más general de diseño no lineal y una aplicación numérica que matiza resultados existentes en la literatura sobre la eficiencia de una metodología concreta de diseño.

1. *INTRODUCCIÓN AL DISEÑO DE EXPERIMENTOS*

El diseño de un experimento consiste en la selección de un grupo de perfiles o combinaciones de atributos para pasarlos a una serie de individuos y pedirles que los valoren. Se trata por tanto de construir una matriz de datos **X** en la que se agrupen por filas cada uno de los perfiles. La relación del diseño de experimentos con los modelos de preferencias es que éstos se aplican sobre los datos recogidos mediante aquéllos para establecer así un modelo de cómo deciden los individuos encuestados.

Ejemplos de empleo de diseño de experimentos son las aplicaciones de investigación y desarrollo de productos donde se requiera de una evaluación empírica de la importancia de las características de estos productos sobre su aceptación entre el público. Una característica (atributo) del producto puede ser el tiempo de entrega desde la compra por el cliente, y tres configuraciones o niveles posibles de este atributo pueden ser entrega inmediata, semanal o mensual. Por lo general, el número de características distintas y sus diferentes configuraciones dan lugar a un número elevado de posibles perfiles de productos, haciendo impracticable la observación de valoraciones de los individuos para todos ellos. Si por ejemplo se dispone de tres atributos con dos niveles y tres atributos con tres niveles, existen hasta $2^3 \times 3^3 = 216$ perfiles distintos; en modelos no lineales, el número de combinaciones es incluso mayor, pues hay que agrupar a los perfiles en conjuntos de decisión, existiendo para el caso de ejemplo hasta 2795 posibilidades si los conjuntos de decisión están formados por dos alternativas. Es en este contexto donde aparece la necesidad de realizar diseño de experimentos, para seleccionar las combinaciones de atributos a considerar en el estudio atendiendo a criterios de recogida eficiente de información.

2. *ADAPTACIÓN DEL CASO DE DISEÑO LINEAL DE EXPERIMENTOS*

En el presente trabajo se analiza el proceso de construcción de cuestionarios/encuestas a partir de los cuales obtener información acerca de las preferencias de los individuos mediante metodología de elección discreta. Este hecho implica la necesidad, no siempre observada en la literatura, de efectuar cambios en el proceso tradicional (lineal) de construcción de las encuestas, para así tener en cuenta las especificidades de la metodología de elección discreta en el análisis de dichas encuestas. La literatura que aborda el diseño o construcción de las encuestas a partir de las cuales recoger las preferencias de los individuos es considerablemente más escasa que la dedicada a la propia especificación de los modelos de preferencias; este escasez es aún incluso más acuciada para el caso de diseños para aplicar modelos no lineales.

2.1. Criterio para diseñar un experimento y problema de optimización asociado

La construcción razonada de las encuestas es útil principalmente porque permite incrementar la fiabilidad de los resultados que se obtengan al aplicar modelos de preferencias a las respuestas de los individuos a estas encuestas. La fiabilidad es un concepto derivado del desempeño estadístico de los modelos, y se traduce en la forma y tamaño de las matrices de covarianza de los parámetros que definen a estos modelos. Mientras más diagonal sea esta matriz, menos solapamiento existe entre el poder explicativo de los parámetros, y cuánto más pequeña sea[1], más se acercan los parámetros estimados a ser representativos del comportamiento de la población total subyacente. Obsérvese pues la estrecha relación que existe entre el objetivo que persigue el diseño de experimentos y la evaluación estadística de los modelos de preferencias.

El criterio para diseñar una matriz de datos $\mathbf{X}$ es que la matriz de covarianza de $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ resultante sea lo mínima posible. Esta matriz depende del tipo de modelo de preferencias que se considere y es una de las siguientes,

Diseño lineal

$$\mathrm{var}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)=\sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\ ,\ \sigma^2\simeq\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)/\left(T-1\right)$$

Diseño de elección discreta

$$\begin{aligned}\mathrm{var}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)&=\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}\middle|\mathbf{y},\mathbf{X}\right)\right]^{-1}E\left(\left[\nabla w\left(\boldsymbol{\beta}\middle|\mathbf{y},\mathbf{X}\right)\right]'\left[\nabla w\left(\boldsymbol{\beta}\middle|\mathbf{y},\mathbf{X}\right)\right]\right)\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}\middle|\mathbf{y},\mathbf{X}\right)\right]^{-1}\\&\simeq\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}\middle|\mathbf{y},\mathbf{X}\right)\right]^{-1}\mathbf{BHHH}_{L\times L}\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}\middle|\mathbf{y},\mathbf{X}\right)\right]^{-1}\end{aligned}$$

donde $w\left(\boldsymbol{\beta}\middle|\mathbf{y},\mathbf{X}\right)$ es el logaritmo de la función de verosimilitud y **BHHH** el gradiente de las probabilidades observado a través de toda la muestra disponible,

$$w\left(\boldsymbol{\beta}\middle|\mathbf{y},\mathbf{X}\right)=\sum_{m=1}^{M}\sum_{j\in m}\left[\delta_{mj}\ln g_m\left(d_m\left(\boldsymbol{\beta},\mathbf{x}_j^m,\boldsymbol{\varepsilon}^m\right);\forall j\in m\right)\right]$$

$$\mathbf{BHHH}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p\middle|\mathbf{y},\mathbf{X}\right)=\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{C_m}\left(\delta_{mj}^p\nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)'\nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)\right)$$

El problema que se ha de resolver por tanto cuando se habla de diseñar una encuesta es el siguiente,

---

[1] La medida más frecuente que se emplea del tamaño de una matriz en diseño de experimentos es el determinante de dicha matriz (*D_Error*)

$$\begin{array}{ll} MIN & f\left(\operatorname{var}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right)=\begin{cases} f\left(\mathbf{X},\mathbf{y},\hat{\boldsymbol{\beta}}\right) & \textit{Diseño lineal} \\ f\left(\mathbf{X},\boldsymbol{\delta},\hat{\boldsymbol{\beta}},\boldsymbol{\varepsilon}\right) & \textit{Diseño no lineal} \end{cases} \\ s.a. & \mathbf{X}\in D\left(\mathbf{X}\right) \end{array}$$

La matriz **X** es la única variable a calcular en el problema cuando se diseña un experimento. $D\left(\mathbf{X}\right)$ representa a todas las combinaciones posibles de perfiles de atributos y la función *f* representa una medida del tamaño de la matriz de covarianza. La medida elegida en este trabajo (si bien no la única) es el determinante de la matriz, obteniendo por tanto los llamados diseños *D_Óptimos*.

2.2. *Naturaleza iterativa del proceso de diseño de elección discreta*

Se puede observar de la definición de las matrices de covarianza hecha justo más arriba que el criterio de diseño depende precisamente de las valoraciones de los individuos a la encuesta que se está tratando de diseñar, o en otras palabras, se pretende elegir un diseño **X** de tal forma que se minimice una expresión que depende directamente de las preferencias mostradas por los individuos (vectores **y** en el caso lineal y **δ** en el no lineal) y con la restricción añadida de que éstas preferencias sólo se obtiene cuando **X** ya ha sido definida.

Esta dependencia da lugar a un teórico carácter iterativo para el proceso de cálculo de la mejor encuesta: suponer unos valores iniciales para las preferencias $\left(\mathbf{y}^0,\boldsymbol{\delta}^0\right)$, diseñar la mejor encuesta $\left(\mathbf{X}^1\right)$, pasar la encuesta a los individuos y obtener $\left(\mathbf{y}^1,\boldsymbol{\delta}^1\right)$, volver a diseñar la encuesta, … y así sucesivamente hasta obtener una convergencia en las preferencias de los individuos. Como en la práctica no se podrá disponer de la población objetivo tantas veces como sea necesario, se han de aproximar a priori cuáles son las preferencias de los individuos. Esta aproximación es ciertamente una fuente de incertidumbre no completamente controlable, si bien esta sigue siendo mejor que no realizar diseño alguno.

Esta fuente de incertidumbre es común a los dos tipos de diseños, si bien es menor para el caso lineal (aunque no nula), pues la dependencia del criterio de diseño de las valoraciones reales de los individuos es a través de σ^2, una constante de proporcionalidad de la matriz de datos **X** diseñada. Para el caso no lineal esta dependencia es más compleja, y la teórica naturaleza iterativa del proceso de diseño de la mejor encuesta más evidente.

Descendiendo más en detalle, se tiene que en lugar de suponer unos valores para **y** y **δ**, lo que se hace es suponer unos valores de los parámetros para $\hat{\boldsymbol{\beta}}$ (llamados $\hat{\boldsymbol{\beta}}_p$ o parámetros piloto) y derivar las aproximaciones a las preferencias reales mediante $\mathbf{y}_p=\mathbf{X}\hat{\boldsymbol{\beta}}_p$ o $\boldsymbol{\delta}^p=g\left(d\left(\hat{\boldsymbol{\beta}}_p,\mathbf{X},\boldsymbol{\varepsilon}\right)\right)=\mathbf{P}\left(\hat{\boldsymbol{\beta}}_p,\mathbf{X}\right)$. La elección de estos $\hat{\boldsymbol{\beta}}_p$ es conveniente que esté basada en información previa disponible de aplicaciones similares, pues en tal caso se podría garantizar con mayor seguridad que las matrices de covarianza evaluadas en los parámetros piloto y en los parámetros que realmente caracterizan el comportamiento de los individuos no son tan diferentes.

Una vez establecidas todas estas anotaciones el proceso de diseño queda formulado de la siguiente manera,

Diseño lineal

$$\begin{array}{ll} MIN & f\left(\mathbf{X}'\mathbf{X}\right)^{-1} \\ s.a. & \mathbf{X}\in D\left(\mathbf{X}\right) \end{array}$$

Diseño no lineal

$$
\begin{aligned}
&MIN \quad f\left(\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\right]^{-1} \mathbf{BHHH}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\right]^{-1}\right) \\
&s.a. \quad w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right) = \sum_{m=1}^{M} \sum_{j\in m}\left[\delta_{mj}^{p} \ln g_m\left(d_m\left(\hat{\boldsymbol{\beta}}_p, \mathbf{x}_j^m, \boldsymbol{\varepsilon}^m\right); \forall j \in m\right)\right] \\
&\qquad \mathbf{BHHH}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right) = \frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{C_m}\left(\delta_{mj}^{p}\nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)'\nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)\right) \\
&\qquad \mathbf{x}_j^m \in D(\mathbf{X}), \boldsymbol{\delta}^p = g\left(d\left(\hat{\boldsymbol{\beta}}_p, \mathbf{X}, \boldsymbol{\varepsilon}\right)\right)
\end{aligned}
$$

Para el caso lineal no se ha incluido σ^2 en la función objetivo, debido a que la utilización de las $\mathbf{y}_p$ en el caso lineal llevaría a la situación no deseable de $\mathrm{var}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)=\mathbf{0}_{L\times L}$, por lo que en este caso sólo participa la minimización de $\left(\mathbf{X}'\mathbf{X}\right)^{-1}$.

El empleo de parámetros piloto nulos ($\hat{\boldsymbol{\beta}}_p = \mathbf{0}_{L\times L}$) deshace la dependencia de la función objetivo del diseño del experimento respecto a la propia matriz $\mathbf{X}$ diseñada; este resultado sin embargo no ha de esconder la artificialidad de la independencia conseguida, ya que la matriz de covarianza obtenida una vez se pase la encuesta no estará calculada con parámetros nulos.

En la Figura siguiente se resume gráficamente la participación de las propiedades y criterio de diseño de una encuesta y se representa la interacción con el resto de fases de un estudio de las preferencias de un grupo de individuos. Se puede observar también la naturaleza iterativa característica del proceso de diseño y cómo el tipo de modelo de preferencias y de expresión de la utilidad no sólo condiciona el proceso de estimación, sino también el proceso de diseño. El propósito de las secciones siguientes es precisamente considerar diferentes modelos de preferencias y expresiones de la utilidad cuando se diseña un experimento.

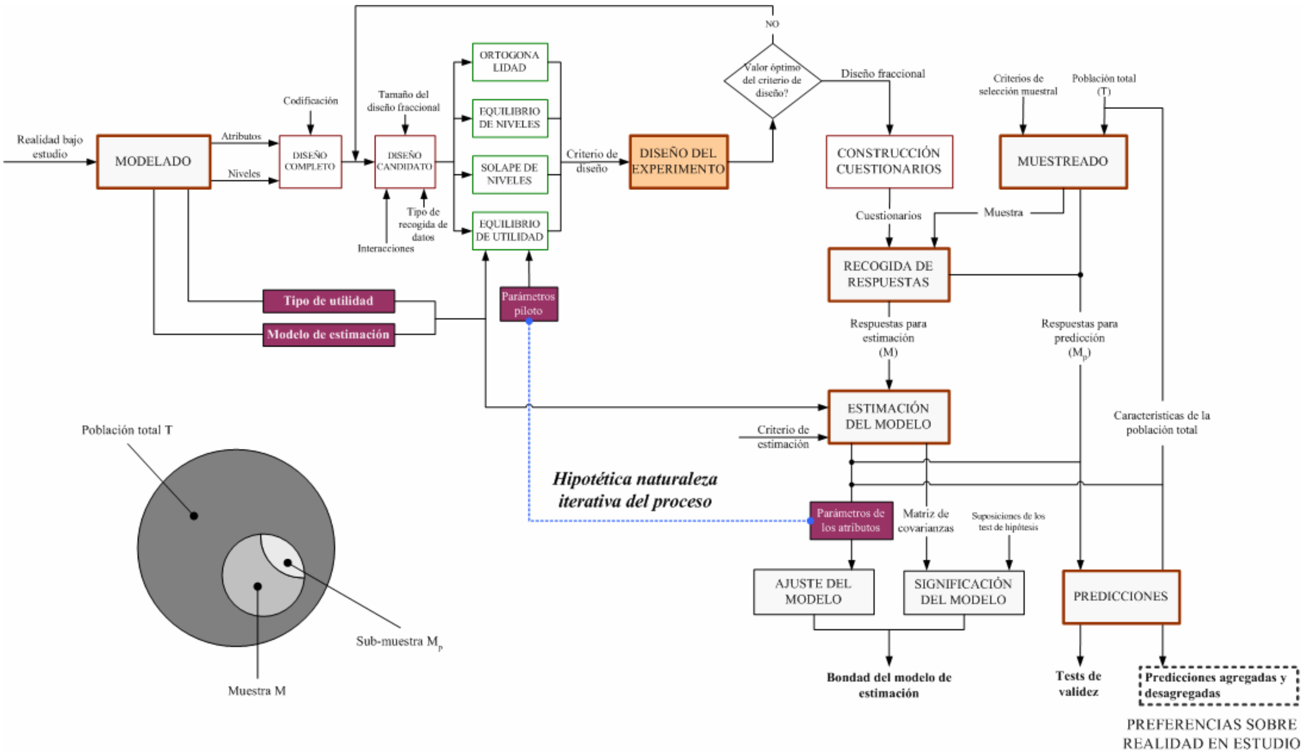

Figura 1 Esquema para analizar las preferencias de una muestra de individuos utilizando diseño de experimentos de elección discreta

3. *AMPLIACIONES DE DISEÑO DE EXPERIMENTOS DE ELECCIÓN DISCRETA*

El carácter iterativo argumentado con anterioridad para los diseños no lineales introduce una dependencia directa del diseño **X** construido no sólo con respecto a los parámetros piloto elegidos, sino también con respecto al modelo no lineal de preferencias utilizado, es decir, respecto a las funciones *g* y *d* que se escojan para calcular $\boldsymbol{\delta}^p = g\left(d\left(\hat{\boldsymbol{\beta}}_p, \mathbf{X}, \boldsymbol{\varepsilon}\right)\right) = \mathbf{P}\left(\hat{\boldsymbol{\beta}}_p, \mathbf{X}\right)$. Existen por tanto al menos dos razones que hacen que el proceso de búsqueda de la encuesta diseñada sea diferente dependiendo del modelo de preferencias específico que se vaya a emplear (incluso si ambos modelos son no lineales):

- La diferente función objetivo a minimizar en cada caso, debido a que las matrices de covarianza dependen del modelo de preferencias considerado, y
- La diferente aproximación de las valoraciones reales que harán los individuos (valoraciones piloto), debido a que éstas también dependen del modelo de preferencias que se considere.

3.1.1. Empleo de utilidades lineales en elección discreta simple

Al principio de este apartado sobre diseño de experimentos se ha expuesto cómo la construcción de diseños de elección en base a una disciplina propia, y no limitándose sólo a extensiones del caso lineal, implica un significativo avance en la teoría de construcción de encuestas; no es menos cierto tampoco que este avance se ha originado en la literatura casi en su totalidad bajo la asunción general de considerar que el modelo de preferencias a aplicar a los datos recogidos será de tipo **logit simple** y con utilidades de las alternativas **lineales** en los parámetros (en apartados posteriores se formula el proceso de cálculo de las encuestas cuando se emplean modelos más generales).

El modelo de preferencias de tipo logit simple con utilidades lineales se define construyendo una parte determinista de la utilidad en los parámetros y una parte aleatoria $\boldsymbol{\varepsilon}^m \to Gumbel\left(\mathbf{0}_{C_m \times 1}, \mu \mathbf{I}_{C_m \times C_m}\right)$, haciendo que la probabilidad de elegir la alternativa *j* del conjunto de decisión *m* sea la siguiente,

$$P_j^m\left(\hat{\boldsymbol{\beta}}_p, \mathbf{X}\right) = g_m\left(U_1^m, \ldots, U_{C_m}^m\right) = \frac{\exp\left(\mathbf{x}_j^m \hat{\boldsymbol{\beta}}_p\right)}{\sum_{j=1}^{C_m} \exp\left(\mathbf{x}_j^m \hat{\boldsymbol{\beta}}_p\right)}, \ j = 1, \ldots, C_m$$

donde $\mathbf{x}_j^m$ es una fila de la matriz de diseño **X** en la que se acumulan las características de la alternativa *j* del conjunto *m*, y $\hat{\boldsymbol{\beta}}$ los parámetros estimados para el modelo.

En este caso más simple, y debido a la relativa sencillez de las ecuaciones de cálculo de la probabilidad, la función objetivo para el cálculo de los parámetros (logaritmo de la función de verosimilitud) y su hessiano son los siguientes,

$$w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right) = \sum_{m=1}^{M}\left(\sum_{j \in C_m}\left[\delta_{mj}^p \mathbf{x}_j^m \hat{\boldsymbol{\beta}}_p\right] - \ln \sum_{j \in C_m} \exp\left(\mathbf{x}_j^m \hat{\boldsymbol{\beta}}_p\right)\right)$$

$$\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right) = -\mu^2 \sum_{m=1}^{M} \sum_{j \in C_m}\left\{\left[\mathbf{x}_j' - \sum_{k \in C_m}\left(\mathbf{x}_k' P_k^m\right)\right] P_j^m \left[\mathbf{x}_j - \sum_{k \in C_m}\left(\mathbf{x}_k P_k^m\right)\right]\right\} = -\mu^2 \mathbf{Z}'\mathbf{P}\mathbf{Z}$$

donde las matrices implicadas se definen como,

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}^1 & \cdots & \mathbf{z}^M \end{bmatrix}, \ \mathbf{z}^m = \mathbf{x}^m - \sum_{k \in C_m}\left(\mathbf{x}_k' P_k^m\right), \ \mathbf{P} = \begin{bmatrix} \mathbf{P}^1 & \cdots & \mathbf{P}^M \end{bmatrix}', \ \mathbf{P}^m = \begin{bmatrix} P_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_{C_m}^m \end{bmatrix}$$

Haciendo empleo de estas ecuaciones de probabilidad y de resultados de la teoría de distribuciones asintóticas, se establece que la matriz de covarianza de los parámetros $\hat{\boldsymbol{\beta}}$ estimados mediante este modelo logit simple se es igual al negativo de la inversa del hessiano de la función de verosimilitud expuesta más arriba, de tal forma que se tiene que,

$$\operatorname{var}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}_p\right)=-\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\right]^{-1}=\left(\mu^2\mathbf{Z}'\mathbf{P}\mathbf{Z}\right)^{-1}$$

por lo que el problema de búsqueda del mejor diseño para un logit simple con utilidades lineales es el siguiente,

$$\begin{array}{ll} MIN & \left(\det\left(\mu^2\mathbf{Z}'\mathbf{P}\mathbf{Z}\right)^{-1}\right)^{1/L} \\ s.a & \mathbf{x}_j^m \in D\left(\mathbf{X}\right), \forall j \in C_m, m=1,\ldots,M \end{array}$$

Las únicas variables del problema de optimización sin restricciones son los perfiles de atributos $\mathbf{x}_j^m$ que han de componer la matriz de datos. La función objetivo no es cóncava, aunque sea negativa definida la matriz $\left(\mu^2\mathbf{Z}'\mathbf{P}\mathbf{Z}\right)^{-1}$, quedando así aún más clara la diferencia entre los problemas (independientes para deshacer la naturaleza iterativa del proceso de diseño) de construcción de la matriz **X** (condicionado en $\hat{\boldsymbol{\beta}}_p$) y de estimación de los parámetros $\hat{\boldsymbol{\beta}}$.

Obsérvese de la formulación del problema que las valoraciones piloto $\boldsymbol{\delta}^p$ no participan directamente, haciendo no necesario establecer una regla para su cálculo en función de los parámetros piloto en un modelo logit simple. Por lo general $\boldsymbol{\delta}^p$ se aproxima por las probabilidades de elección ($\delta_{mj}^p = P_j^m = g_m\left(d_m\left(\hat{\boldsymbol{\beta}}_p,\mathbf{x}^m,\boldsymbol{\varepsilon}^m\right)\right)$), pero tal aproximación no es necesaria para un logit simple. Aún siendo esto válido, el proceso de construcción del mejor diseño teórico sigue teniendo carácter iterativo, pues las probabilidades dependen de los parámetros piloto $\hat{\boldsymbol{\beta}}_p$, y las probabilidades sí que participan del modelo.

3.1.2. Empleo de elección discreta no simple. Problema general de diseño.

El siguiente paso hacia la formulación del problema más general de diseño de un experimento es no partir ni de asunciones previas sobre la parte aleatoria de las utilidades ni sobre la expresión de la parte determinista de éstas. La única asunción que se mantiene con respecto a las anteriores es que las partes aleatorias y deterministas de las utilidades son independientes (poco restrictivo).

Una vez se haya especificado el tipo de parte aleatoria y la expresión para la parte determinista de las utilidades de las alternativas, y mediante aplicación de los postulados RUM, se construyen las expresiones de las probabilidades,

$$P_j^m = g_m\left(d_m\left(r_m\left(\hat{\boldsymbol{\beta}}_p,\mathbf{X}\right),\boldsymbol{\varepsilon}\right);\forall j \in m\right)$$

Al igual que para los dos casos anteriores, las únicas variables del problema son las T combinaciones de atributos a construir $\mathbf{x}_j^m \in D\left(\mathbf{X}\right), \forall j \in C_m, m=1,\ldots,M$, debiendo ser el resto de relaciones especificadas según el tipo de modelo concreto que se considere. Este problema es por tanto el más general para construir una encuesta de valoración de preferencias,

$$
\begin{aligned}
&MIN \quad f\left(\mathbf{\Omega}_{L\times L}\right)\\
&s.a.\\
&\qquad \mathbf{\Omega}_{L\times L} = \left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\right]^{-1} \mathbf{BHHH}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right)\right]^{-1}\\
&\qquad w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right) = \ln \prod_{m=1}^{M} \prod_{j=1}^{C_m} \left(P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)^{\delta_{mj}^p}\\
&\qquad \mathbf{BHHH}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y},\mathbf{X}\right) = \frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{C_m}\left(\delta_{mj}^p \nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)' \nabla\left(\ln P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X})\right)\right)\\
&\qquad P_j^m(\hat{\boldsymbol{\beta}}_p,\mathbf{X}) = g_m\left(d_m\left(r_m\left(\hat{\boldsymbol{\beta}}_p,\mathbf{x}^m\right),\boldsymbol{\varepsilon}^m\right)\right); V_j^m = r_m\left(\hat{\boldsymbol{\beta}}_p,\mathbf{X}\right); \boldsymbol{\delta}^p = g\left(d\left(r_m\left(\hat{\boldsymbol{\beta}}_p,\mathbf{X}\right),\boldsymbol{\varepsilon}\right)\right)\\
&\qquad \mathbf{x}_j^m \in D\left(\mathbf{X}\right), \forall j \in C_m, m = 1,\ldots,M
\end{aligned}
$$

4. *MATIZACIONES DE RESULTADOS EN LA LITERATURA EXISTENTE*

Una vez que se ha generalizado el proceso de diseño de una encuesta se exponen ciertas matizaciones sobre resultados que a este respecto se han publicado en la literatura reciente sobre construcción de experimentos para aplicar modelos de elección discreta. En concreto, hacemos referencia a resultados repetidos en Sándor (2001) y, Sándor & Wedel (2002, 2005). En todos estos trabajos se plantean problemas similares a los recogidos más arriba, pero con diferencias sustanciales en al menos los dos siguientes aspectos: el origen de la aleatoriedad en el modelo y la derivación de las expresiones de las matrices de covarianza. En las dos sub-secciones siguientes se abordan cada uno de estos aspectos, y por último se comparan resultados obtenidos bajo nuestra argumentación con algunos de los presentados en la literatura.

4.1.1. Aleatoriedad de las frecuencias de elección

En los trabajos mencionados se añaden nuevos elementos al carácter de las probabilidades de elección de las alternativas, en concreto, se establece que si f_j es la frecuencia observada de elección de la alternativa *j* y *M* es el número total de conjuntos de decisión, entonces $M\left[f_1 \ \cdots \ f_J\right]$ se comporta como una distribución multinomial de parámetros $\left[P_1(\hat{\boldsymbol{\beta}},\mathbf{X}) \ \cdots \ P_J(\hat{\boldsymbol{\beta}},\mathbf{X})\right]$ (Sándor, 2001 p. 17-18).

Obsérvese que se está suponiendo un comportamiento aleatorio para datos del problema en estudio, las frecuencias con las que los individuos eligen entre las *J* alternativas disponibles, específicamente que estas frecuencias aleatorias $\varsigma_1,\ldots,\varsigma_J$ siguen una distribución de tipo multinomial,

$$
f_{\varsigma_1,\ldots,\varsigma_J}\left(x_1,\ldots,x_J \middle| P_1(\hat{\boldsymbol{\beta}},\mathbf{X}),\ldots,P_J(\hat{\boldsymbol{\beta}},\mathbf{X})\right) = \Pr\begin{pmatrix} \varsigma_1 = x_1 \\ \vdots \\ \varsigma_J = x_J \end{pmatrix} = C\cdot\prod_{k=1}^{J}\left(P_k(\hat{\boldsymbol{\beta}},\mathbf{X})\right)^{M\cdot x_k}
$$

Bajo esta óptica, el criterio de cálculo de los parámetros $\hat{\boldsymbol{\beta}}$ es hacer que el siguiente valor, $f_{\varsigma_1,\ldots,\varsigma_J}\left(f_1,\ldots,f_J \middle| P_1(\hat{\boldsymbol{\beta}},\mathbf{X}),\ldots,P_J(\hat{\boldsymbol{\beta}},\mathbf{X})\right)$, sea lo máximo posible, donde $f_1,\ldots,f_J$ son datos obtenidos de las preferencias mostradas por los individuos.

Esta argumentación de la introducción de la aleatoriedad en modelos de preferencias es bien distinta a las presentadas en Ibáñez (2002) sobre consistencia de las decisiones observadas y representatividad de los resultados con una muestra finita y cuenta con al menos la siguiente limitación, que cada conjunto de decisión ha de constar de las mismas *J* alternativas y de que estás no sean genéricas.

4.1.2. Funciones objetivo para diseñar un experimento

La segunda de las diferencias está relacionada con la función objetivo y las matrices de covarianza que se utilizan en le proceso de diseño de un experimento. La derivación de la **matriz de covarianza** se realiza mediante inversión de la matriz de información asintótica de Fisher ($\mathbf{\Phi}_{L\times L}$). Esta matriz se define en estos trabajos como la covarianza del gradiente de la función objetivo del problema de cálculo de los parámetros $\hat{\boldsymbol{\beta}}$, es decir,

$$\mathbf{\Omega}_{L\times L}^{-1} = \mathbf{\Phi}_{L\times L}\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right) = E\left(\left[\nabla w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right)\right]' \left[\nabla w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right)\right]\right)$$

En lo que respecta al **valor esperado**, éste se calcula en base a la distribución multinomial supuesta para las frecuencias y no en base a la aleatoriedad del vector de parámetros reales,

$$w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right) = \ln\left(K \cdot \prod_{j=1}^{J}\left(P(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)^{M\cdot f_j^p}\right) \Rightarrow \nabla w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right) = M\mathbf{f}^{p\prime} \cdot \nabla\left(\ln \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)$$

$$\mathbf{\Omega}_{L\times L}^{-1} = M \cdot \nabla\left(\ln \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)' \left[\mathbf{P}_d(\hat{\boldsymbol{\beta}}_p, \mathbf{X}) + (M-1)\cdot \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X}) \cdot \mathbf{P}'(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right] \nabla\left(\ln \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)$$

Obsérvese cómo para el cálculo de este valor esperado se necesita que los conjuntos de decisión sean del mismo tamaño y que las alternativas no sean genéricas, requisitos no necesarios en nuestro enfoque más general, ya que se hace uso de

$$E\left(\mathbf{f}\mathbf{f}'\right) = \left(\mathbf{P}_d + (M-1)\mathbf{P}\mathbf{P}'\right) \big/ M$$

Bajo esta óptica, las derivadas involucradas en el gradiente se hacen con respecto a los parámetros del modelo, pero el valor esperado se hace con respecto a la distribución de las frecuencias. La matriz de covarianzas $\mathbf{\Omega}_{L\times L}$ calculada de esta forma siempre es positiva definida.

Así pues descartamos que la función objetivo en el proceso de diseño de un experimento esté basada en la distribución aleatoria de las frecuencias de decisión, por la limitación de la tipología de casos aplicables y porque en tal caso la aleatoriedad es introducida por la distribución de las frecuencias de elección de las alternativas y no por medio del carácter aleatorio de los parámetros que se pretenden estimar. Como consecuencia, las matrices de covarianza de modelos de preferencias no lineales cuyo uso proponemos son una de las dos siguientes:

$$\mathbf{\Omega}_{L\times L} = \left\{ \begin{array}{l} -\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right)\right]^{-1} \\ +\left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right)\right]^{-1} \mathbf{BHHH}_{L\times L} \left[\nabla^2 w\left(\hat{\boldsymbol{\beta}}_p \middle| \mathbf{y}, \mathbf{X}\right)\right]^{-1} \end{array} \right\} \neq M^2 \nabla\left(\ln \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)' E\left(\mathbf{f}^p \mathbf{f}^{p\prime}\right) \nabla\left(\ln \mathbf{P}(\hat{\boldsymbol{\beta}}_p, \mathbf{X})\right)$$

4.1.3. Comparación de matrices de covarianza para logit simple

En esta sección se compara el desempeño de diseños construidos utilizando el enfoque de Sándor & Wedel y el enfoque de estimador robusto o tipo sándwich de las matrices de covarianza expuesto en este trabajo.

La matriz de datos $\mathbf{X}$ que se quiere construir es de dimensión 240×9 ($T = 240; M = 60; L = 9$) y contiene los atributos de 240 alternativas de transporte interurbano; los niveles de estos nueve atributos son los que caracterizan a cada alternativa. Para construir el mejor diseño posible se parte de un vector de parámetros piloto basado en estudios similares de transporte. Sabemos que una vez que pasemos la encuesta a los individuos obtendremos unos valores de los parámetros diferentes a estas aproximaciones piloto. El objetivo de nuestra aplicación numérica es pues construir una matriz $\mathbf{X}$ de dimensiones 240×9 agrupada en conjuntos de cuatro alternativas y con empleo de los siguientes parámetros piloto,

Tiempo de viaje	Coste de viaje	Tiempo en el vehículo	Coste generalizado	Constante de Avión	Constante de Tren	Constante de Bus	Ingresos avión	No.pasajeros avión
-0.071	-0.027	-0.009	0.018	2.520	3.611	-26.546	0.027	-0.285

Tabla 1 Parámetros piloto utilizados para diseñar la encuesta

Se decide aplicar un modelo de preferencias de tipo logit simple para evaluar las preferencias que los individuos hacen de las alternativas contenidas en las filas de la matriz **X** diseñada; por ello, las matrices de covarianza que se minimiza en el proceso de diseño bajo cada enfoque son las derivadas con anterioridad,

$$\mathbf{\Omega}_{paper} = \left(\mathbf{Z}'\mathbf{P}\mathbf{Z}\right)^{-1} \;,\; \mathbf{\Omega}_{sandor} = \left[M \cdot \mathbf{X}'\left(\mathbf{P}_d - \mathbf{P}\mathbf{P}'\right)\mathbf{X}\right]^{-1}$$

Las medidas que se toman del tamaño de las matrices es su determinante ponderado por el número de parámetros que se estiman: $f\left(\mathbf{\Omega}\right) = \left(\det \mathbf{\Omega}\right)^{1/L}$.

El empleo del tamaño de la matriz $\mathbf{\Omega}_{paper}$ (y no el de $\mathbf{\Omega}_{sandor}$.) es el seguido también en trabajos como Huber & Zwerina (1996) y Zwerina (1997), y en paquetes de software como LIMDEP.

En la figura siguiente se muestra el proceso de minimización de cada una de las matrices. No se obtiene el mismo diseño ni medidas similares de las matrices de covarianza en los parámetros piloto,

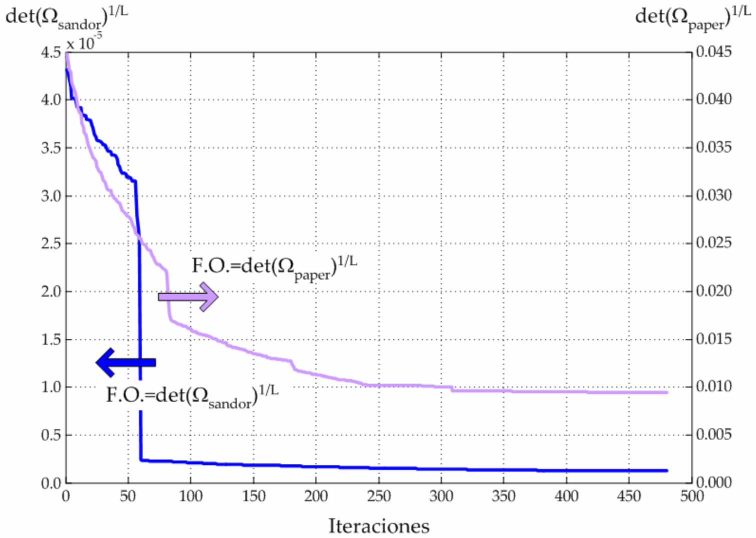

Figura 2 Proceso de búsqueda algorítmica con diferentes funciones objetivo (Ibáñez, 2002)

Se realiza ahora otra comparación entre enfoques que consiste en representar $f\left(\mathbf{\Omega}_{paper}\right)$ frente a $f\left(\mathbf{\Omega}_{sandor}\right)$, manteniendo el diseño **X** construido con anterioridad, pero en función de una batería de valores posibles de los parámetros del modelo, y no sólo para los parámetros piloto de la Tabla 1; en concreto se realiza la comparación para 5000 betas obtenidas de la distribución normal $N\left(\hat{\boldsymbol{\beta}}_p, \hat{\boldsymbol{\beta}}_p/2\right)$ y de la distribución uniforme $\left(-2\hat{\boldsymbol{\beta}}_p, 2\hat{\boldsymbol{\beta}}_p\right)$. Al diseño obtenido mediante la minimización de $f\left(\mathbf{\Omega}_{sandor}\right)$ se le denomina $\mathbf{X}_{sandor}$ y al obtenido minimizando $f\left(\mathbf{\Omega}_{paper}\right)$ por $\mathbf{X}_{paper}$.

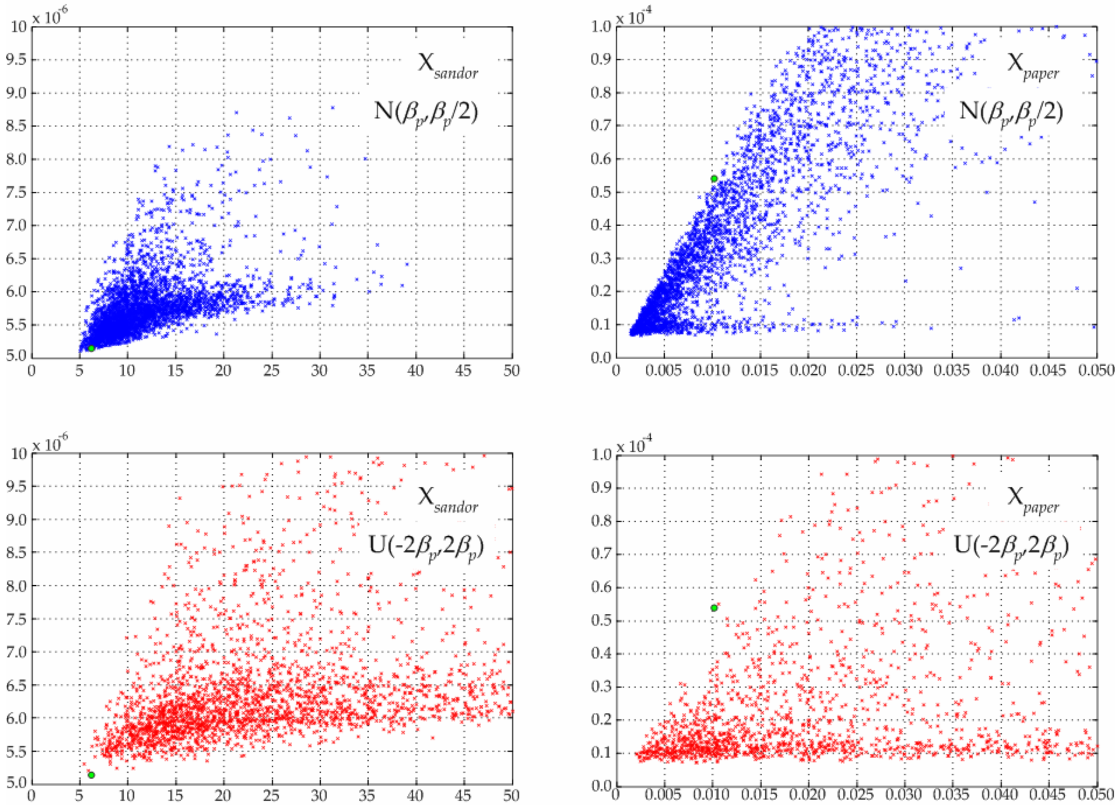

Figura 3 Comportamiento de los diseños propuestos por Sándor y los de este trabajo para diferentes baterías de parámetros beta

En la figura puede observarse cómo existe relación entre ambos enfoques, pero cómo también los resultados obtenidos con cada uno no son directamente comparables. En los trabajos de Sándor & Wedel se explicita que se obtienen mejores resultados si se utiliza su proceso de diseño de experimentos que para otros trabajos en la literatura como Huber & Zwerina (1996). A nuestro juicio este mejor comportamiento no está relacionado con el proceso de diseño del experimento, sino con la elección anterior de la función objetivo que se quiere minimizar en tal proceso. Las escogidas en Sándor & Wedel están basadas en las frecuencias de decisión, propiciando una componente de bondad artificial en sus modelos que habrá de ser tenida en cuenta a la hora de comparar el desempeño del proceso de diseño que proponen con otras estrategias existentes en la literatura.

5. *APORTACIONES DEL TRABAJO Y CONCLUSIONES*

En este trabajo se ha debatido sobre las implicaciones que los siguientes tres aspectos tienen sobre el diseño de elección discreta del experimento construido y en cómo difiere de lo actualmente publicado sobre utilización de modelos logit simples con utilidades lineales en los parámetros: Primero, la utilización de un modelo de elección discreta no simple (expresiones más generales para P_j^m), segundo, las generalizaciones de las expresiones lineales para la utilidad de las alternativas (U_j^m) y finalmente, la consideración de parámetros $\hat{\boldsymbol{\beta}}_{piloto}$. Las aportaciones principales de este trabajo son haber formulado el problema general de diseño de experimentos no lineal y haber concluido el rechazo parcial de una metodología propuesta en la literatura para diseñar experimentos debido al número más limitado de casos a la que es aplicable y una parte de bondad artificial para los resultados que de su aplicación se derivan.

6. *REFERENCIAS*

Ibáñez (2003). Modelos de Elección Discreta para el Diseño de Sistemas Logísticos en Entornos Urbanos. Informes, Estudios, Trabajos y Dictámenes.

KUHFELD, W. F., R. D. TOBIAS, and M. GARRATT (1994): "Efficient Experimental Design with Marketing Research Applications," Journal of Marketing Research, 21, 545-57.

Lazari, A.G., Anderson, D.A. (1994). "Design of Discrete Choice Set Experiments for Estimating Both Attribute and Availability Cross Effects", Journal of Marketing Research, 31: 375-383.

LOUVIERE, J. J. (1988): Analyzing Decision Making : Metric Conjoint Analysis. Newbury Park, CA: Sage Publications.

LOUVIERE, J. J., D. A. HENSHER, and J. D. SWAIT (2000): Stated Choice Methods: Analysis and Application. Cambridge University Press.

ZWERINA, K. (1997): Discrete Choice Experiments in Marketing. Heidelberg: Physica-Verlag.

Sándor, Z. (2001): Computation, Efficiency and Endogeneity in Discrete Choice Models. University of Groningen.

Sandor, Z., M. Wedel (2005): Heterogeneous Conjoint Choice Designs. Journal of Marketing Research, 2005, 42 (2), 210-218.

Sandor, Z., M. Wedel (2001): 'Designing Conjoint Choice Experiments using Managers' Prior Beliefs', Journal of Marketing Research, 38, 430-444.