Los métodos y las aplicaciones del *credit-scoring*

Gilbert Saporta*

La técnica de discriminación sobre variables cualitativas propuesta por el autor [5], fue inicialmente concebida para calcular puntajes de riesgos bancarios. Desde entonces y con la implementación computacional del proceso para calcular los puntajes, más de una centena de estudios han sido realizados en Francia.

En una primera parte se presenta un resumen de las diferentes técnicas estadísticas utilizables y evaluamos su interés. En la segunda parte se presenta un resumen de los trabajos operacionales. En particular se muestra que surgen numerosos problemas (muestreo, reajuste, acceso a los datos, robustez, fiabilidad, ...) y que la realización de un puntaje operacional depende no sólo de la escogencia de un buen algoritmo, sino además, de un *savoir faire* adquirido gracias a la experiencia. Algunos perfeccionamientos estadísticos son igualmente evocados.

1 Metodología estadística

1.1 Planteamiento del problema

Esquematicamente un problema de *credit-scoring* se puede describir como la búsqueda de un procedimento de separación entre dos grupos: los "malos pagadores" y los "buenos pagadores", conociendo un conjunto de descriptores. Las variables explicativas son casi siempre cualitativas (categoría profesional, estado matrimonial, geotipo, etc.) o se transforman en cualitativas por descomposición en clases (antigüedad en el empleo, por ejemplo). Esta situación excluye el uso clásico de los procedimientos de análisis discriminante. Los datos son por tanto constituidos por dos muestras de tamañ generalmente elevado (varias centenas de observaciones) descritas por p variables cualitativas X_i con m_i modalidades; $i = 1, 2, \ldots, p$ donde p es geenralmente del orden de varias decenas.

1.2 Técnicas estadísticas utilizables

El objetivo es prever la calidad de un expediente a partir de las variables explicativas. Dos grandes categorías de procedimientos se pueden proponer: los fundamentados en la estimación de la probabilidad de pertenencia a una de las clases y los fundamentados en la estimación de una regla numérica de clasificación: "el puntaje". A continuación se describen algunas de esas técnicas estadísticas.

*Département de Mathématiques et Informatique, Centre National d'Arts et Métiers, París

1.2.1 Estimación de la probabilidad

El esquema multinomial es utilizable puesto que hay $2q\Pi_{i=1}^{p} m_i$ casos posibles y este número es en general excesivo. Se debe, por tanto, reducir el nivel de las interacciones útiles y tomar por ejemplo unos modelos del tipo loglineal (ver [2,4] por ejemplo). Se puede también utilizar un método de segmentación. Sin embargo, estas técnicas son muy pesadas para implementar y mal adaptadas al tamaño de las bases de datos que tratamos.

1.2.2 Función de puntaje

Esta enfoque es el más usado: se atribuye a cada categoría de cada variable un valor numérico. El puntaje es la suma de los valores de las categorías a las cuales pertenece un individuo y es la variable numérica de decisión: se puede fijar ya sea un umbral de aceptación o bien decidir en función de las curvas dando los porcentajes de la demanda y el porcentaje de malos expedientes en función del puntaje. La regla es de muy fácil aplicación e interpretación.

El puntaje no es más que la variable discriminate en el sentido de Fisher, combinación lineal de las $\sum_i^p m_i$ variables indicatrices de las categorías de las p variables. El método del puntaje es por lo tanto un análisis discriminante sobre las variables cualitativas, y una forma de cuantificar las variables cualitativas.

El conjunto de las variables indicatrices no es de rango pleno puesto que la suma de las indicatrices de una misma variable cualitativa vale 1. Soluciones técnicas elementales como la supresión de la última categoría de cada variable, permitirían utilizar un programa ordinario de análisis discriminante. Sin embargo, esta no es la técnica que hemos propuesto [5] y que es utilizada bajo el nombre de Disqual.

1.2.3 Disqual

Este método consiste en hacer un análisis factorial (de correspondencias múltiples) del conjunto de las variables cualitativas explicativas. El conjunto de todos los factores engendra el mismo espacio que el conjunto de las indicatrices. Se eliminan los factores de poca inercia y los que no separan suficientemente los dos grupos. (Una de las ventajas de este procedimiento es que permite usar un métrica diagonal ya que los factores son dos a dos ortogonales. Además los resultados de un análisis de correspondencias múltiples son muy robustos, lo que da una seguridad).

Las k coordenadas factoriales retenidas tienen la función de variables explicativas numéricas para un análisis discriminante. Se calcula luego la función de Fisher, que se expresa como una combinación lineal de las indicatrices, lo que da la función de puntaje (una variante de este método, pero sin la selección de los factores, figura en el programa SPAD, etapa DIS2G).

En efecto, si d es la función de Fisher; $d = (d_1 \quad d_k)$ donde $d_j = \dfrac{\overline{z}_1^j - \overline{z}_2^j}{\lambda_j}$ y $\overline{z}_i^j$ es la media de las coordenadas del grupo i sobre el eje factorial j de varianza λ_j.
El puntaje s de un individuo se obtiene por: $s = \sum_{j=1}^{k} d_j z^j$.

Como $z^j = \sum_{l=1}^{m} \alpha_l^j x_l$ se tiene $s = \sum_{j=1}^{k} \sum_{l=1}^{m} d_j \alpha_l^j x_l$ o sea

$$s = \sum_l (\sum_j d_j \alpha_l^j) x_l$$

donde x_l es la indicatriz de la categoría l. l valor del puntaje de esta categoría vale por lo tanto $\sum_{j=1}^{k} d_j \alpha_l^j$.

Por construcción de la función de puntaje, esto corresponde a un modelo puramente aditivo sin interacción. Tomando en cuenta las interacciones entre las variables explicativas mejoraría claramente el comportamiento del puntaje. Esto se puede hacer de la manera siguiente: se detecta por medio de un modelo log-lineal o por medio de una segmentación, que X_1 y X_2 interacctúan. Se reemplazan estas variables por la variable $X_1 \times X_2$ con $m_1 m_2$ categorías y se aplica el método Disqual. Si $m_1 m_2$ es muy grande (normalmente lo es) se pueden agrupar categorías.

2 Las aplicaciones del *credit-scoring*: riesgos de fracaso y condiciones de éxito

2.1 Algunos problemas metodológicos

Mencionamos aquí algunos aspectos de la implementación del método mencionado al *credit-scoring*. Aparecen varias dificultades ligadas ya sea a problemas metodológicos o bien a problemas prácticos. Para que un puntaje sea operacional y utilizable, no sólo hay que escoger un buen algoritmo aino que también se debe adquirir un gran *savoir faire*.

2.1.1 La población, el muestreo y los reajustes

El primer problema es quizás el más delicado para resolver. Es el que lleva a un gran número de fracasos. Para que un puntaje sea operacional debe ser aplicable al conjunto de la demanda. Esta es asimilada a una población numerable P. Para construir un puntaje lo ideal sería observar el comportamiento pagador de una muestra aleatoria representativa de P. Ahora bien, en la práctica, se realiza una preselección y algunos expedientes son rechazados. Consecuentemente la muestra observada M_1, es sacada de una parte P_1 de P, compuesta de los expedientes aceptables. Si no se tienen los cuidados necesarios, el puntaje se aplicará por tanto a P_1 y no a P.
Durante la implementación del procedimiento ello significa que los expedientes deberán ser primero seleccionados con la ayuda de los métodos tradicionales, y en segundo término, con la ayuda del puntaje. Pero esto, en la práctica, es inaceptable.

Para solucionar el inconveniente, varios enfoques son posibles. Se les puede clasificar en dos categorías las cuales hacen referencia a una muestra M_2 de expedientes rechazados sacados de $P_2 = P - P_1$. En el primer caso, se busca rehacer M_1 para darle la estructura de P sobre las principales variables de control. En el segundo caso, se simula el comportamiento pagador (desconocido) de las observaciones pertenecientes a M_2.

Un segundo problema de muestreo se presenta al tener en cuenta la dimensión temporal de la demanda. Se observan frecuentemente fenómenos estacionales. Por ejemplo, en Francia junio es más malo que octubre. Además, la observación del comportamiento pagador de

un expediente rechazado se debe hacer en un período fijo (por ejemplo 24 meses), teniendo en cuenta las leyes de aparición de las cuentas morosas. Los expedientes técnicamente utilizables tienen cuando menos 24 meses de antigüedad. Mientras tanto la estructura de la demanda, la reglamentación y las condiciones de aceptación pueden evolucionar y el puntaje corre el riesgo de ser obsoleto antes de ser utilizado. Un plan de sondeo bien construido permite crear una muestra completa de expedientes suficientemente antiguos, además de una muestra representativa de la demanda reciente sobre la cual buscará ajustarse.

En conclusión, dos tipos de reajustes se imponen. El primero permite corregir la estructura de la submuestra observada para volverle a dar la estructura de la demanda. El segundo permite corregir los efectos de envejecimiento de la muestra para volverle a dar la estructura de la demanda reciente.

2.1.2 La escogencia de las variables y de las categorías

Otros problemas surgen cuando se trata de escoger las variables que se van a introducir en el modelo y su descomposición en categorías. Un algoritmo eficaz de selección que no necesitaría realizar el análisis discriminante en cada paso, no ha sido aun implementado. El método que he propuesto con la ayuda de los coeficientes parciales de Chuprow, no ha aportado los resultados deseados. Una dificultad proviene del hecho que las variables indicatrices de una variable cualitativa no pueden ser disociadas y por tanto se debe proceder por bloques de variables.

Sin embargo, los métodos automáticos de selección del tipo "stepwise" conducen usualmente a soluciones inaceptables para el usuario. Una variable "clásica" como la tasa de endeudamiento, por ejemplo, puede ser rechazada puesto que su poder explicativo es ligeramente inferior que una variable "exótica" y poco familiar para el usuario. De la misma manera un método de descomposición optimal en categorías o de reagrupamiento optimal, puede conducir a la construcción de clases que sorprenden al usuario. Además ciertas escogencias "optimales" sobre una muestra pueden ser engañosas. A este nivel, el verdadero *savoir faire* no reside en la escogencia de las técnicas de optimización, sino esencialmente en el conocimiento del campo y la acumulación de experiencias. Los métodos de optimización proveen una primera solución "estandarizada", luego es necesario buscar las soluciones operacionales.

Lo mismo ocurre con la escogencia de las variables cruzadas. La experiencia nos ha mostrado que, frecuentemente resulta preferible sustituir las dos variables cualitativas por una sola variable obtenida por producto cartesiano. En este caso, es además, necesario reagrupar entre ellas algunas categorías de la variable cruzada. La búsqueda de mejores cruzamientos se puede hacer con la ayuda de modelos tales como el análisis de la varianza cualitativo o el modelo loglineal. Este tipo de modelo puede conducir también hacia soluciones artificiosas, y, a menudo, la experiencia guiará al análisis en sus escogencias *a priori*, de cruzamientos por probar.

2.1.3 Estabilidad e interpretación de los coeficientes

Las técnicas de puntaje se parecen a las de los modelos lineales generalizados: son proclives a aumentar artificialmente el número de variables y de categorías en el modelo. Ello permite, en apariencia, mejorar la calidad de la discriminación, pero, el problema de la estabilidad

de los estimadores de los puntajes es crucial. Además no se dispone en este caso del arsenal de tests utilizados en el caso del modèlo lineal. Los métodos de remuestreo del tipo "bootstrap" deberían permitir determinar intervalos de confianza para los puntajes de las categorías y probar las hipótesis concernientes en ausencia de toda hipótesis distribucional. En la práctica es deseable poder interpretar los pesos de los diferentes puntajes en relación con la tasa de no pago en la categoría. Una inversión, aún cuando sea estadísticamente justificada es, para el usuario, difícil de explicar. Un modelo más "económico" en variables y en reagrupamientos juiciosos de categorías permitiría eliminar tales inversiones.

2.1.4 Estimación de las probabilidades a posteriori

Si las distribuciones condicionales de las coordenadas factoriales son normales se pueden aplicar los resultados usuales del análisis discriminante. Esta hipótesis no se verifica siempre y una estimación precisa de las probabilidades se puede efectuar por medio de un método no paramétrico del tipo núcleos de Parzen [3]. Su integración en un programa del tipo Disqual está por hacer.

2.1.5 Desviaciones entre las previsiones y las realizaciones

Una vez el modelo elaborado su utilización es doble. Por supuesto, las políticas de selección fundadas sobre la red de puntaje son implementadas. Por otra parte, el modelo es utilizado para hacer previsiones de no pagadores, en número y en volumen, teniendo en cuenta las políticas de selección escogidas. Rápidamente pueden aparecer desviaciones entre las previsiones y las realizaciones. Su análisis es un punto cardinal y permite orientar las acciones. Las desviaciones pueden surgir por causa de una evolución del ambiente: contexto reglamentario o política comercial del organismo. También pueden deberse a una evolución de la demanda, lo que se detecta facilmente siguiendo el perfil de la clientela y comparándolo con el de la muestra. Los sesgos no controlados en el momento de la construcción de la muestra también pueden contribuir a las desviaciones.
La detección de todas estas desviaciones requiere la puesta a punto de herramientas de control. ¿Cómo combatir el problema? Normalmente un simple ajuste de umbrales de aceptación permite corregir la desviación. Si ello no es suficiente y si las causas son estructurales, la única solución es ajustar un nuevo modelo sobre una muestra actualizada.

2.2 Algunos problemas prácticos

2.2.1 El acceso a los datos pertinentes

Usualmente las informaciones contenidads en los expedientes aceptados y rechazados no son almacenados sobre soporte informático. Los expedientes rechazados normalmente no son conservados. Las estadísticas de demanda y de aceptación que permiten elaborar el plan de sondeo no siempre existen. Es por lo tanto necesario realizar una fase previa de conteo, de codificación y de digitación. En lo sucesivo se deben implementar los procedimientos para la adquisición de datos de manera sistemática, que permita alimentar una base de datos. Esta base servirá para realizar los trabajos de seguimiento de las desviaciones y sacar las muestras de futuros estudios de puntaje.

Otro problema que se presenta frecuentemente es que lo histórico no es utilizable, ya sea porque se trata de una actividad nueva o reciente, o bien porque las reglas de selección eran tan severas que ningún expediente deriva en moroso. Es entonces posible implementar técnicas totalmente diferentes, haciendo simulaciones de la demanda, de la aceptación y del comportamiento.

2.2.2 La modificación de los procedimentos

Otros problemas están ligados a las modificaciones de los procedimentos administrativos e informáticos, y éstos son aún más delicados. Para resolverlos es necesario evolucionar las aplicaciones informáticas e implementar acciones de formación y motivación de los usuarios. Por el momento no desarrollamos estos aspectos, solamente llamamos la atención sobre esas dificultades cuya subevaluación lleva, la mayor parte de las veces, a un fracaso.

3 Conclusión

Pese a todas las dificultades metodológicas y prácticas y a los riesgos de fracaso derivados, 10 años de consultoría y de realización me han permitido poner en evidencia las ventajas decisivas del *credit-scoring*. Comparado con otros procedimentos tradicionales, el puntaje posee tres cualidades fundamentales:

1. La **simplicidad**: la recolección de un número mínimo de informes permite tomar una decisión rápida.
2. La **homogeneidad**: la política de aceptación es aplicada uniformente en la red.
3. La **flexibilidad**: la política de selección está basada sobre bases objetivas. Toda inflexión puede ser decidida rápidamente midiendo las consecuencias de las nuevas escogencias.

Aparte de estas ventajas, específicas del método, citamos las principales consecuencias de un puntaje:

1. La **disminución del número de morosos**: una selección optimizada permite reducir el volumen de los morosos.
2. La **previsión de los morosos**: el modelo permite simular con precisión los futuros no pagadores y por tanto optimizar el cálculo de las provisiones.
3. La **productividad**: solamente los expedientes con riesgo son objeto de un examen cuidadoso. En la mayoría delos casos la decisión puede ser delegada y el análisis acelerado.

Para obtener todas estas ventajas, los bancos y los organismos financieros deben ser conscientes que la implementación de un puntaje necesita un conjunto de competencias estadísticas y prácticas y que, por consiguiente, se trata de una decisión estratégica cuya realización necesita tiempo e inversión.

Bibliografía

[1] Bouroche J.M., Saporta G., Tenenhaus M. (1977) *Some methods of qualitative data analysis*, in Recent Developments in Statistics, J.R. Barra (ed.), pp 749-755, North Holland, Amsterdam.

[2] Daudin J.J. (1980) *Régression qualitative: choix de l'espace prédicteur*, in Data Analysis and Informatics, E. Diday (ed.), pp 329-345, North Holland, Amsterdam.

[3] Gautier J.M., Saporta G. (1984) *Méthodes non paramétriques en analyse discriminante; quelques propositions nouvelles*, in Data Analisis and Informatics III, E. Diday et al. (eds.), pp 591-605, North Holland, Amsterdam.

[4] Goldstein M., Dillon W.R. (1978) *Discrete Discriminant Analysis*, John Wiley & Sons, Nueva York.

[5] Saporta G. (1976) *Discriminant Analysis when the variables are nominal*, Spring Meeting of the Psychometric Society, Murray Hill, Nueva York.

[6] Saporta G. (1977) *Une méthode et un programme d'analyse discriminante pas à pas sur variables qualitatives*, 1eres Journées Analyse des Donées et Informatique, INRIA, pp 201-210, París.

[7] Sireyjol N. (1987) *Les apports du crédit-scoring*, Banque, N^0 457, pp 788-794.