31

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"TECNICAS ESTADISTICAS EN EL ANALISIS DE SOLVENCIA"

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

GABRIELA ESPINO HERNANDEZ

DIRECTOR DE TESIS: DR. MANUEL MENDOZA RAMIREZ

MEXICO, D.F. 2001

UNIVERSIDAD NACIONAL
AVENTA DE
MEXICO

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:

"Técnicas Estadisticas en el Análisis de Solvencia".

realizado por Gabriela Espino Hernández

con número de cuenta 9332794-3 , pasante de la carrera de Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	Dr. Manuel Mendoza Ramírez
Propietario	M. en C. José Antonio Flores Díaz
Propietario	M. en A. P. María del Pilar Alonso Reyes
Suplente	Act. Francisco Sánchez Villarreal
Suplente	Act. Jaime Vázquez Alamilla

Consejo Departamental de Matemáticas

Papá y Mamá, les agradezco mucho todo el amor y dedicación que me han brindado durante toda mi vida, han sido una gran enseñanza para mí. Los amo.

Rofa y Lulú, gracias por estar conmigo, los quiero muchísimo.

Mamá Jovita, gracias por apoyarme y escucharme siempre. Has sido una gran amiga para mí. Te quiero mucho.

Rodrigo, muchísimas gracias. Te quiero mucho.

A mis grandes amigas, Carmen, Liliano, Alejandra y Lisette, por ser mis confidentes, por superar juntas esas etapas difíciles de la vida que nos han unido más, por compartir tantas locuras y momentos felices, y por enseñarme las grandes teorías de la vida. Las quiero muchísimo.

A mis mejores amigos, Omar, Rodrigo, Pablo, Oswaldo, Carlos y Rubén, gracias por todo el apoyo, alegría, cariño y protección que me han dado. Son un gran tesoro para mí.

Agradezco especialmente a mi Director de Tesis

Dr. Manuel Mendoza Ramírez

por compartir conmigo sus conocimientos, tiempo y buenos momentos.

Gracias a todos mis profesores de la Facultad de Ciencias.

ÍNDICE

INTRODUCCIÓN

Existe una gran variedad de factores que pueden influir en el desempeño de una unidad económica. En un ambiente de competencia y relativo libre mercado, una empresa puede producir distintos resultados dependiendo de las variaciones que se presentan en su entorno tanto a nivel micro como macro. Cambios en la actitud o preferencia de sus clientes, en la agresividad de sus competidores, en la eficiencia de sus proveedores, en la habilidad de sus administradores, o la capacidad de sus empleados; cambios en la estructura del sector económico en el que se desempeña, en las disposiciones regulatorias de las autoridades que la supervisan, o en el entorno económico global, pueden originar resultados completamente distintos de los esperados.

Ahora bien, los resultados que reflejan el efecto de éstos y otros factores pueden ser juzgados desde diferentes perspectivas. Como por ejemplo, la rentabilidad que la empresa ofrece a sus accionistas, o la capacidad para enfrentar los compromisos que ha contraído con sus clientes.

Especialmente en el medio financiero, las agencias gubernamentales que supervisan las empresas de ese sector, persiguen, como uno de sus objetivos principales, la vigilancia con el fin de evitar, en el mayor grado posible, el incumplimiento con los usuarios de sus servicios. Un caso particular, muy importante, es el que se presenta con las compañías aseguradoras que precisamente ofrecen servicios que pretenden eliminar riesgos económicos y ofrecer protección a sus clientes (asegurados). Con este tipo de empresa, el objetivo de la supervisión se orienta a impedir que las aseguradoras dejen de cumplir las obligaciones que tienen con los asegurados.

Para evitar que una empresa aseguradora se encuentre en una situación de incumplimiento o insolvencia, es necesario tomar diversas medidas cuya oportunidad o conveniencia sólo puede ser establecida si se cuenta con un sistema que permita realizar pronósticos sobre la salud financiera de la compañía. De esta manera, es necesario observar el desempeño de la empresa y a partir de esa información aventurar un juicio sobre su equilibrio económico en un tiempo futuro preestablecido.

Ese tipo de procedimientos son complejos y necesariamente involucran incertidumbre, por ello, el problema que se pretende resolver es eminentemente estadístico. Así, el objetivo de esta tesis es presentar y examinar algunas de las técnicas estadísticas más comunes con las cuales se puede predecir oportunamente la insolvencia. En particular, se plantea este problema como uno de clasificación estadística; asimismo, se muestran algunos de los resultados obtenidos al aplicar estas técnicas estadísticas en el sector asegurador de Estados Unidos y México.

Así, en el capítulo 1 se abunda sobre la importancia de que las compañías se mantengan en estado solvente; asimismo, en forma general, se presentan los esfuerzos por parte de los organismos reguladores para detectar oportunamente la posible insolvencia de las compañías aseguradoras, y por último, se comenta la estructura de este problema para poder resolverlo desde una perspectiva estadística.

En el capítulo 2, se presentan los elementos teóricos más relevantes de la técnica estadística conocida como análisis discriminante, el cual es utilizado para resolver problemas de clasificación estadística. Asimismo, se comentan dos artículos en los que investigadores utilizan esta técnica multivariada para realizar pronósticos de la situación financiera de las compañías aseguradoras que operan en los Estados Unidos.

En el capítulo 3, se muestran los modelos de respuesta cualitativa, en particular el modelo logístico el cual es una herramienta estadística que puede ser utilizada para estimar probabilidades de insolvencia o calcular intervalos de confianza para esta probabilidad; asimismo, se presentan algunos de los resultados más relevantes que se han obtenido al utilizar esta técnica estadística en el análisis de solvencia de compañías aseguradoras de los Estados Unidos.

Una tercer técnica estadística, conocida como análisis de supervivencia se presenta en el capítulo 4, al igual que algunos de los resultados obtenidos al utilizar esta técnica en estudios de solvencia de compañías aseguradoras.

Por último, en el capítulo 5 se reporta una aplicación del análisis discriminante y un modelo de respuesta cualitativa, para poder inferir de forma anticipada la situación financiera de las compañías aseguradoras que operan en México, utilizando datos disponibles de la Comisión Nacional de Seguros y Fianzas.

Asimismo, se presentan las conclusiones y algunos resultados que se encuentran en los anexos.

CAPÍTULO 1
El problema de solvencia

1.1 Solvencia

La solvencia se define, en términos financieros, como la capacidad de una empresa para hacer frente oportunamente a sus obligaciones de corto y largo plazo.

La detección oportuna de patrones de inestabilidad económica o de problemas que influyan en la solvencia de empresas es una tarea primordial que se debe de realizar de manera eficaz y oportuna, con la finalidad de que las empresas mantengan una situación financiera sana, que les permita hacer frente a sus obligaciones; de tal manera que puedan ofrecer confianza y tranquilidad a sus clientes lo mismo que a sus socios e inversionistas. En el caso de las empresas del sector financiero, los clientes son los usuarios de servicios financieros, como por ejemplo ahorradores, asegurados, pensionados, etc.

En el caso de las empresas de servicios financieros, se han creado organismos reguladores encargados de supervisar la situación contable y financiera de las empresas del ramo. En particular, en el sector asegurador de los Estados Unidos, cada estado cuenta con un organismo regulador, los cuales, en 1871 crearon la NAIC (National Association of Insurance Commissioners), organización que ayuda a identificar problemas financieros en las compañías aseguradoras y, de esa manera, proteger a los asegurados de pérdidas debidas a insolvencia.

En los años 70, la NAIC desarrolló el IRIS (Insurance Regulatory Information System), un sistema utilizado para predecir la insolvencia de compañías aseguradoras que operan en los Estados Unidos, el cual hace uso de distintos índices contables y financieros, que se registran periódicamente de cada una de las instituciones que operan el seguro de daños-responsabilidad civil y el seguro de vida. Si el sistema detecta, para una empresa aseguradora, cuatro o más de estos índices fuera del rango establecido por la NAIC, se considera en peligro de ser insolvente.

El IRIS ha funcionado durante varios años; sin embargo, algunos investigadores consideran que no es confiable para prevenir la insolvencia de las compañías aseguradoras, debido a que no proporciona alertas tempranas sobre problemas financieros.

En México, la Secretaría de Hacienda y Crédito Público es la responsable para regular el sector financiero. La supervisión de la solvencia de las compañías aseguradoras, la lleva a cabo a través de la Comisión Nacional de Seguros y Fianzas (CNSF) que quedó constituida en 1990 como un organismo desconcentrado de la Secretaría de Hacienda y Crédito Público.

La CNSF, con las facultades y atribuciones que le confieren la Ley General de Instituciones y Sociedades Mutualistas de Seguros, así como las demás leyes, reglamentos y disposiciones administrativas aplicables al mercado asegurador mexicano, se encarga de revisar de forma periódica, a través de sus áreas de inspección y vigilancia, la situación contable y financiera de las compañías aseguradoras, con el fin de que éstas se mantengan de manera permanente en niveles adecuados que les permitan hacer frente al cumplimiento de las obligaciones adquiridas con los asegurados, y de esa manera se consoliden como empresas estables.

Por otra parte, en junio de 1994, supervisores del seguro provenientes de más de 100 jurisdicciones alrededor del mundo fundaron la IAIS (International Association of Insurance Supervisors), un organismo que estimula y promueve la cooperación internacional en materia de supervisión de la industria del seguro.

La IAIS cuenta con un Subcomité de Solvencia que analiza continuamente los requerimientos de solvencia que se establecen en el sector asegurador de distintos países y estudia diversas estrategias que se podrían implantar en la supervisión futura.

Sin embargo, y a pesar de estos esfuerzos, la insolvencia de las empresas del sector asegurador es un problema que en los Estados Unidos se ha incrementado con el paso del tiempo. Por ello, se han realizado una gran variedad de estudios, cuyo objetivo es utilizar técnicas estadísticas para detectar de manera oportuna y eficaz la posible insolvencia de las instituciones aseguradoras. En un número muy importante de tales trabajos, el problema de detección de la insolvencia se ha planteado, desde una perspectiva estadística, como un problema de clasificación.

1.2 El problema de clasificación

En muchas ocasiones, las personas o los organismos se enfrentan a la necesidad de saber a que población pertenece un objeto o individuo, de acuerdo a ciertas características que le fueron observadas. Este problema tiene sentido cuando existen al menos dos posibles poblaciones de las que pudiera proceder el elemento a clasificar. Es aquí, cuando a las personas interesadas en realizar la clasificación les surge la pregunta ¿Cómo decidir a qué población pertenece el individuo u objeto observado?.

Cuando en una situación de este tipo se tiene, además, que en cada una de las poblaciones relevantes las características objeto de observación se comportan como variables aleatorias, entonces se tiene lo que se conoce como un problema de clasificación estadística, en donde el individuo a clasificar sólo podrá pertenecer a una y sólo a una de las posibles poblaciones.

Para la solución de este tipo de problemas se puede recurrir a distintas técnicas estadísticas, como por ejemplo, el análisis discriminante y los modelos de respuesta cualitativa.

La idea original de los procedimientos de clasificación consiste en construir una regla δ, de manera que si un individuo presenta las características $\underline{X}$, entonces $\delta(\underline{X})$ tome un valor en el conjunto $\{1,2,\ldots,p\}$ (si hay p poblaciones en el problema) de manera que ese valor identifique la población en la que debe clasificarse al sujeto.

Una variante más general consiste en producir una regla δ que, aplicada al vector de características $\underline{X}$, determine una distribución de probabilidades $\{P(1|\underline{X}), P(2|\underline{X}),\ldots,P(p|\underline{X})\}$ la cual describa la probabilidad que tiene el sujeto de pertenecer a cada una de la poblaciones. Por supuesto, a partir de esta distribución es posible proceder a la clasificación.

El análisis discriminante tiene modalidades de cada uno de los dos tipos mencionados, mientras que los modelos de respuesta cualitativa claramente son del segundo tipo. En cualquier caso, estos procedimientos se consideran estáticos puesto que un individuo con características fijas $\underline{X}$ es clasificado en una población, o recibe una asignación de probabilidades, que no cambia como función del tiempo.

Existen, por otra parte, técnicas estadísticas que asignan a un individuo de características $\underline{X}$ una distribución de probabilidades $\{P_t(1|\underline{X}), P_t(2|\underline{X}),\ldots,P_t(p|\underline{X})\}$ que depende tanto de $\underline{X}$ como del tiempo.

Un ejemplo es el siguiente: Suponga que a un grupo de individuos con un padecimiento terminal se les administra un tratamiento que debiera prolongar su vida. En esta situación, cada sujeto tiene, al principio del estudio, una serie de características $\underline{X}$ que son relevantes para su supervivencia (sexo, edad, sintomatología, etc.) y resulta de interés observar el tiempo, a partir de que se aplica el tratamiento, que cada paciente se mantiene con vida.

Un análisis estadístico de la información recabada (las características y los tiempos de supervivencia) pueden dar lugar a una regla que asigne a un individuo con características $\underline{X}$ una probabilidad de mantenerse con vida al tiempo t (transcurrido desde la aplicación del tratamiento) $P_t(1|\underline{X})$. En este ejemplo, la distribución se completaría con la probabilidad $P_t(2|\underline{X})$, la probabilidad de morir al tiempo t.

Los modelos de este tipo reciben el nombre genérico de supervivencia y como puede observarse guardan relación con los procedimientos de clasificación. De hecho, pueden considerarse una generalización dinámica de los mismos.

Algunos ejemplos en los que se tiene un problema de clasificación con dos poblaciones, son los siguientes:

Ejemplo 1.1

Una universidad está interesada en decidir si un individuo debe ser aceptado o rechazado para estudios de posgrado, clasificándolo como buen o mal estudiante. Para ello, se debe de definir a que alumnos se consideran buenos estudiantes; un ejemplo podrían ser aquellos que en su primer semestre de estudios presentan un

promedio superior a ocho, además de que no reprueban ninguna materia. Dos grupos de alumnos, uno de los considerados como buenos y otro que no cumple con esta condición, son seleccionados para revisar algunas de las características que presentaban al solicitar el ingreso a la universidad, como por ejemplo, promedio en la carrera, nivel de inglés, si contaba o no con experiencia laboral, resultado de un examen de aceptación, etc. La idea entonces, consiste en recibir la solicitud de ingreso de un prospecto de alumno para el cual se revisarán estas mismas variables con el objeto de decidir, por medio de una técnica de clasificación estadística, si se considera como un buen o un mal estudiante y por consiguiente la universidad podrá tomar la decisión de aceptarlo o rechazarlo para los estudios de posgrado.

Ejemplo 1.2

Una institución bancaria desea clasificar a un individuo con el fin de saber si es un cliente al que es conveniente otorgarle o no una tarjeta de crédito; es decir, le interesa clasificarlo como un buen pagador o como una persona morosa. Para ello, el banco puede revisar, de aquellas personas que han pagado todos sus recibos a tiempo durante tres años (definición operativa de buen pagador), la información registrada en su solicitud de tarjeta como: sexo, edad, antigüedad en el trabajo donde labora actualmente, salario, número de dependientes económicos, situación en otros bancos, etc. Si se hace lo mismo con un grupo de clientes morosos, en tal caso, al recibir una nueva solicitud y analizar esta misma información, se podrá clasificar al individuo como buen pagador o como moroso, y así tomar la decisión de aceptarlo como un nuevo tarjetahabiente o no.

Ejemplo 1.3

Un organismo regulador, como la Comisión Nacional de Seguros y Fianzas, quiere clasificar de forma anticipada a una compañía aseguradora como sana en sentido financiero (es decir, solvente), o con problemas financieros que le podrían

ocasionar la insolvencia. Para esto, el organismo regulador debe definir con precisión lo que se entiende por una empresa solvente y puede, entonces, hacer uso de algunos índices financieros, los cuales revisará en un conjunto de compañías solventes lo mismo que en otro de compañías insolventes. Si se desea clasificar de forma anticipada a una empresa como sana o con problemas financieros, se deberán revisar estos mismos índices y por medio de una técnica estadística de clasificación, como el análisis discriminante, clasificarla como perteneciente a una de las dos posibles poblaciones.

Ejemplo 1.4

Haciendo referencia al ejemplo anterior, en el que se desea clasificar a una compañía de seguros como sana (solvente) o con problemas financieros, cabe mencionar que este problema de clasificación también se puede resolver utilizando análisis de supervivencia. En este caso, el organismo regulador puede estudiar, por ejemplo, en un grupo de compañías aseguradoras que hace dos años se mantenían en estado solvente, algunas de las características $\underline{X}$ que reflejen su situación contable y financiera, al inicio del estudio (es decir, hace dos años). Asimismo, el organismo regulador deberá observar y registrar, a partir del inicio del estudio y a lo largo de un periodo de dos años, los tiempos que estas compañías solventes se mantienen en ese estado. Entonces, después de realizar un análisis estadístico de la información anterior y obtener la distribución de los tiempos en que las compañías permanecen en estado solvente, si se desea clasificar a una nueva empresa aseguradora como solvente o con problemas financieros, se deberá estudiar el valor de las características que presentaba dos años previos al estudio o a la clasificación, para así poder conocer la distribución de su tiempo de vida, dado que al día del estudio se encuentra operando.

En cualesquiera de estos ejemplos, es posible clasificar al individuo de manera errónea. Esto puede suceder con mayor frecuencia cuando el vector con las

características, utilizado para clasificar, tiene un comportamiento semejante en las distintas poblaciones. Debido a esto, es necesario poner especial atención en la selección de las variables que se utilizarán, con el fin de incluir las características que mejor distinguen las poblaciones.

CAPÍTULO 2
Análisis discriminante

En este capítulo se presentan los elementos estadísticos de la técnica de clasificación conocida como análisis discriminante.

2.1 Regla de Fisher

En 1936, R.A. Fisher propuso una regla de clasificación que se basa en la combinación lineal $Y = \underline{\beta}'\underline{X}$, donde $\underline{\beta}$ es un vector de coeficientes por determinar, y $\underline{X}$ en R^k es el vector de características que se observan en el individuo objeto de la clasificación. Se supone que $\underline{X}$ tiene la misma matriz de varianzas y covarianzas Σ, (y distintos vectores de medias) en las diferentes poblaciones.

La idea que subyace el procedimiento, en el caso de dos poblaciones, es la siguiente:

Dado un vector $\underline{\beta}$ (diferente de $\underline{0}$), entonces

$$Y = \underline{\beta}'\underline{X}, \qquad (2.1)$$

es una variable aleatoria (real).

Suponga que la media de Y en la población i (i = 1,2) es μ_{iY}. Por otra parte y puesto que la matriz Σ no cambia entre poblaciones, la varianza de Y también es constante, esto es $Var(Y)= \sigma^2$; i = 1,2.

Entonces, el procedimiento consiste en determinar el vector de coeficientes $\underline{\beta}$ que maximice una distancia estandarizada entre las medias de Y asociadas a las poblaciones, es decir, que maximice

$$\Delta = \frac{(\mu_{1Y}-\mu_{2Y})^2}{\sigma^2_Y} \qquad (2.2)$$

Para esto, si se considera

$\underline{\mu}_1 = E(\underline{X} \mid \Pi_1)$ la media de la población Π_1,
$\underline{\mu}_2 = E(\underline{X} \mid \Pi_2)$ la media de la población Π_2, y
$\Sigma = Var(\underline{X})$ la matriz de varianzas y covarianzas de $\underline{X}$ (k x k).

entonces, al hacer uso de la ecuación (2.1), se tiene que:

$\mu_{iY} = E(Y \mid \Pi_i) = E[\underline{\beta}'\underline{X} \mid \Pi_i] = \underline{\beta}' E[\underline{X} \mid \Pi_i] = \underline{\beta}' \underline{\mu}_i$, para i = 1,2; y

$$\sigma^2_Y = Var(Y) = Var(\underline{\beta}'\underline{X}) = E[(\underline{\beta}'\underline{X} - E(\underline{\beta}'\underline{X})) (\underline{\beta}'\underline{X} - E(\underline{\beta}'\underline{X}))']$$
$$= E[\underline{\beta}'(\underline{X} - E(\underline{X})) (\underline{X} - E(\underline{X}))'\underline{\beta}] = \underline{\beta}'E[(\underline{X} - E(\underline{X})) (\underline{X} - E(\underline{X}))'] \underline{\beta}$$
$$= \underline{\beta}' Var(\underline{X}) \underline{\beta} = \underline{\beta}' \Sigma \underline{\beta}.$$

Así, la ecuación (2.2), queda de la forma

$$\Delta = \frac{(\beta'\mu_1 - \beta'\mu_2)^2}{\beta'\Sigma\beta} = \frac{(\beta'(\mu_1 - \mu_2))^2}{\beta'\Sigma\beta},$$

$$\Delta = \frac{\beta'(\mu_1 - \mu_2)(\mu_1 - \mu_2)'\beta}{\beta'\Sigma\beta},$$

$$\Delta = \frac{(\beta'\delta)^2}{\beta'\Sigma\beta},$$

donde, $\underline{\delta} = (\underline{\mu}_1 - \underline{\mu}_2)$.

Debido a que Σ es una matriz definida positiva y $\underline{\beta} \neq \underline{0}$, se puede demostrar (Apéndice A.3) que el máximo de Δ está dado por

$$\underset{\beta}{\text{Max}} \frac{(\beta'\delta)^2}{\beta'\Sigma\beta} = \underline{\delta}'\,\Sigma^{-1}\,\underline{\delta},$$

que se alcanza si $\underline{\beta} = c\,\Sigma^{-1}\,\underline{\delta} = c\,\Sigma^{-1}\,(\underline{\mu}_1 - \underline{\mu}_2)$, con c una constante arbitraria y distinta de cero.

Por lo tanto, sin pérdida de generalidad al elegir c = 1, se tiene que la combinación lineal que maximiza el cociente Δ es,

$$Y = \underline{\beta}'\underline{X} = (\underline{\mu}_1 - \underline{\mu}_2)'\,\Sigma^{-1}\,\underline{X},$$

esta combinación lineal es conocida como la función lineal discriminante de Fisher, que tiene la propiedad de convertir una observación multivariada $\underline{X}$, en una univariada Y, de manera que se maximiza la distancia estandarizada entre las medias de las dos poblaciones.

La función discriminante de Fisher sirve para clasificar a un individuo en una de dos poblaciones, de la siguiente forma:

Considere un individuo I, con el vector de observaciones $\underline{x}_0$, el cual, al utilizar la función discriminante de Fisher toma la siguiente forma univariada

$$y_0 = (\underline{\mu}_1 - \underline{\mu}_2)' \, \Sigma^{-1} \, \underline{x}_0,$$

entonces, la regla consiste en clasificar al individuo en la población Π_i si μ_{iY} es la media más cercana a y_0.

Equivalentemente, si μ es el punto medio entre μ_{1Y} y μ_{2Y}, y además, se tiene que $\mu_{2Y} < \mu_{1Y}$, entonces el individuo I, con vector de observaciones $\underline{x}_0$ se clasifica en la población

$$\Pi_1 \quad \text{si} \quad y_0 > \mu,$$
$$\Pi_2 \quad \text{si} \quad y_0 < \mu.$$

Esto es

$$I \in \Pi_1, \quad \text{si} \quad (\underline{\mu}_1 - \underline{\mu}_2)' \, \Sigma^{-1} \, \underline{x}_0 > \mu, \text{ ó}$$
$$I \in \Pi_2, \quad \text{si} \quad (\underline{\mu}_1 - \underline{\mu}_2)' \, \Sigma^{-1} \, \underline{x}_0 < \mu.$$

donde,

$$\mu = \frac{1}{2}(\mu_{1Y} + \mu_{2Y}) = \frac{1}{2}\left(\underline{\beta}'\underline{\mu}_1 + \underline{\beta}'\underline{\mu}_2\right) = \frac{1}{2}\underline{\beta}'(\underline{\mu}_1 + \underline{\mu}_2) = \frac{1}{2}(\underline{\mu}_1 - \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 + \underline{\mu}_2).$$

Este procedimiento se puede generalizar para el caso de p (mayor a dos) poblaciones.

2.2 Criterio de máxima verosimilitud

La regla de clasificación de Fisher es básicamente geométrica, sin embargo, esta regla tiene una interpretación probabilística dada a través del criterio de máxima verosimilitud. Para ello, considere que bajo Π_i, el vector $\underline{X}$ de k variables aleatorias con las características de interés del individuo a clasificar, tiene una función de densidad de probabilidad $P(\underline{X} \mid \Pi_i) = f_i(\underline{x})$ para $i = 1,\ldots,p$.

Entonces, el criterio de máxima verosimilitud establece que si un individuo I tiene un vector de características $\underline{x}_0$, se clasificará en la población cuya verosimilitud o función de densidad de probabilidad en $\underline{x}_0$ sea la mayor; es decir, se tiene que el individuo I con características $\underline{x}_0$ se clasificará en la población Π_i si y sólo si $f_i(\underline{x}_0) > f_j(\underline{x}_0)$ para toda $j \neq i$.

En particular, si bajo Π_i, $\underline{X}$ se distribuye normal multivariada con vector de medias $\underline{\mu}_i$ y matriz de varianzas y covarianzas $\Sigma > 0$, ambos conocidos para $i = 1,\ldots,p$; entonces, dado que

$$f_i(x) = (2\pi)^{-k/2} \mid \Sigma \mid^{-1/2} \exp\left(-\frac{1}{2}(x-\mu_i)' \Sigma^{-1}(x-\mu_i) \right),$$

se tiene que

$$f_i(\underline{x}_0) > f_j(\underline{x}_0),$$

si y sólo si

$$(\underline{x}_0 - \underline{\mu}_i)' \Sigma^{-1} (\underline{x}_0 - \underline{\mu}_i) < (\underline{x}_0 - \underline{\mu}_j)' \Sigma^{-1} (\underline{x}_0 - \underline{\mu}_j),$$

o equivalentemente, si

$$2\underline{x}_0' \ \Sigma^{-1} \ (\underline{\mu}_i - \underline{\mu}_j) \ > \ \underline{\mu}_i \ \Sigma^{-1} \underline{\mu}_i - \underline{\mu}_j \ \Sigma^{-1} \underline{\mu}_j,$$

esto es

$$\underline{x}_0' \ \Sigma^{-1} \ (\underline{\mu}_i - \underline{\mu}_j) \ > \ \frac{1}{2} \ (\underline{\mu}_i - \underline{\mu}_j)' \ \Sigma^{-1} \ (\underline{\mu}_i + \underline{\mu}_j). \qquad (2.3)$$

Por lo tanto, si se cumple la desigualdad (2.3) para toda $j \neq i$, entonces se clasificará al individuo I como perteneciente a Π_i.

Para el caso en que bajo las distintas poblaciones Π_i con $i = 1,\ldots,p$, las matrices de varianzas y covarianzas del vector $\underline{X}$ sean diferentes, es decir, $\Sigma_i \neq \Sigma_j$ para $i \neq j$, entonces se cumple que

$$f_i (\underline{x}_0) > f_j (\underline{x}_0),$$

si y sólo si

$$|\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x_0 - \mu_i)'\Sigma_i^{-1}(x_0 - \mu_i)\right) > |\Sigma_j|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x_0 - \mu_j)'\Sigma_j^{-1}(x_0 - \mu_j)\right)$$

o equivalentemente, si

$$\frac{1}{2}\ln\left(\frac{|\Sigma_j|}{|\Sigma_i|}\right) \ > \ \left\{ \begin{array}{l} \left[\frac{1}{2}\left(\underline{x}_0'\Sigma_i^{-1}\underline{x}_0 - 2x_0'\Sigma_i^{-1}\mu_i + \mu_i'\Sigma_i^{-1}\mu_i\right)\right] - \\ \left[\frac{1}{2}\left(\underline{x}_0'\Sigma_j^{-1}\underline{x}_0 - 2x_0'\Sigma_j^{-1}\mu_j + \mu_j'\Sigma_j^{-1}\mu_j\right)\right] \end{array} \right\}$$

es decir, si

$$\ln\left(\frac{|\Sigma_j|}{|\Sigma_i|}\right) - \left(\underline{\mu}_i'\Sigma_i^{-1}\mu_i - \underline{\mu}_j'\Sigma_j^{-1}\underline{\mu}_j\right) > \underline{x}_0'\left(\Sigma_i^{-1} - \Sigma_j^{-1}\right)x_0 - 2\underline{x}_0'\left(\Sigma_i^{-1}\mu_i - \Sigma_j^{-1}\mu_j\right).$$

Por lo tanto, el individuo con vector de características $\underline{x}_0$ se clasificará como proveniente de la población i, si se cumple la desigualdad anterior.

Lo que se ha hecho hasta el momento es encontrar la región R_i en R^k, tal que si el vector con las características observadas de un individuo I se localiza dentro de R_i, entonces el individuo será clasificado como proveniente de la población Π_i; lo que significa que habrá p posibles regiones R_i (i = 1,...,p) mutuamente excluyentes, tales que $R^k = \bigcup_i R_i$, de donde el individuo a clasificar podrá provenir.

2.3 Clasificación Bayesiana

En ciertos casos, se tiene información sobre la población a la que pertenece un individuo, la cual se expresa por medio de probabilidades que no utilizan la información del vector con las características del individuo. Por ejemplo, suponga que se sabe que de un total de compañías, el 90% son sanas y el 10% tienen problemas financieros. Si se desea clasificar a una nueva compañía, se puede considerar que inicialmente y en ausencia de otra fuente de datos, la probabilidad de que sea sana es 0.9 y de que tenga problemas financieros es 0.1, esto sin antes haber observado su vector de características $\underline{X}$.

Si tal información está disponible, puede ser incorporada al análisis mediante la regla de probabilidades condicionales. Para ello, si se sabe que un individuo pertenece a una de las poblaciones Π_i, con i = 1,...,p, entonces, la probabilidad a priori de que el individuo pertenezca a la población Π_i se denotará como π_i, es decir $P(\Pi_i) = \pi_i$, y $\pi_1 + \pi_2 + \ldots + \pi_p = 1$.

De la definición de probabilidad condicional, se puede establecer la siguiente forma del teorema de Bayes

$$P(\Pi_i \mid X) = \frac{P(X \mid \Pi_i)P(\Pi_i)}{P(X)},$$

por la fórmula de la probabilidad total se tiene que

$$P(\Pi_i \mid X) = \frac{P(X \mid \Pi_i)P(\Pi_i)}{\sum_{i=1}^{p} P(X \mid \Pi_i)P(\Pi_i)},$$

es decir

$$P(\Pi_i \mid X) = \frac{f_i(X)\pi_i}{\sum_{i=1}^{p} f_i(X)\pi_i}, \quad \text{para } i = 1,\ldots,p$$

Entonces, procediendo en similitud con el criterio de máxima verosimilitud, dado un vector de características $\underline{x}_0$ de un individuo I, una posibilidad es clasificarlo en la población cuya probabilidad sea mayor que la del resto. Es decir, se asignará I a Π_i si

$$P(\Pi_i \mid x_0) > P(\Pi_j \mid x_0), \quad \text{para toda } j \neq i$$

equivalentemente, el individuo I se asignará a Π_i si

$$\pi_i f_i(\underline{x}_0) > \pi_j f_j(\underline{x}_0), \quad \text{para toda } j \neq i$$

que es lo mismo que

$$\frac{f_i(\underline{x}_0)}{f_j(\underline{x}_0)} > \frac{\pi_j}{\pi_i}, \quad \text{para toda } j \neq i.$$

Se puede observar que si las probabilidades a priori son las mismas para todas las poblaciones, entonces se tiene el criterio de máxima verosimilitud.

Con el fin de ejemplificar, si se tiene que bajo Π_i (i = 1,...,p) $\underline{X}$ tiene una función de densidad de probabilidad normal multivariada con vector de medias $\underline{\mu}_i$ y matriz de varianzas y covarianzas $\Sigma > 0$, ambos conocidos, entonces a un individuo I con vector de características $\underline{x}_0$ se clasificará como perteneciente a Π_i si y sólo si

$$\frac{(2\pi)^{-k/2}\,|\Sigma|^{-1/2}\,\exp\left(-\frac{1}{2}(x-\mu_i)'\Sigma^{-1}(x-\mu_i)\right)}{(2\pi)^{-k/2}\,|\Sigma|^{-1/2}\,\exp\left(-\frac{1}{2}(x-\mu_j)'\Sigma^{-1}(x-\mu_j)\right)} > \frac{\pi_j}{\pi_i}, \quad \text{para toda } j \neq i,$$

es decir, si

$$-\frac{1}{2}\left[(x_0-\mu_i)'\Sigma^{-1}(x_0-\mu_i)-(x_0-\mu_j)'\Sigma^{-1}(x_0-\mu_j)\right] > \ln\left(\frac{\pi_j}{\pi_i}\right), \quad \text{para toda } j \neq i,$$

o equivalentemente, si

$$-\frac{1}{2}\left[\left(-2x_0'\Sigma^{-1}\mu_i+2x_0'\Sigma^{-1}\mu_j\right)+\left(\mu_i'\Sigma^{-1}\mu_i-\mu_j'\Sigma^{-1}\mu_j\right)\right] > \ln\left(\frac{\pi_j}{\pi_i}\right),$$

para toda $j \neq i$, esto es, si

$$\underline{x}_0'\Sigma^{-1}(\underline{\mu}_i-\underline{\mu}_j) > \frac{1}{2}(\underline{\mu}_i-\underline{\mu}_j)'\Sigma^{-1}(\underline{\mu}_i+\underline{\mu}_j)+\ln\left(\frac{\pi_j}{\pi_i}\right), \text{ para toda } j \neq i.$$

Para el caso en que la matriz de varianzas y covarianzas sea diferente para todas las poblaciones; es decir, $\Sigma_i \neq \Sigma_j$ para toda $j \neq i$, entonces se clasificará al individuo como proveniente de Π_i, si y sólo si para toda $j \neq i$ se tiene que

$$\frac{(2\pi)^{-k/2} |\Sigma_j|^{-1/2} \exp\left(-\frac{1}{2}(x-\mu_j)'\Sigma_j^{-1}(x-\mu_j)\right)}{(2\pi)^{-k/2} |\Sigma_i|^{-1/2} \exp\left(-\frac{1}{2}(x-\mu_i)'\Sigma_i^{-1}(x-\mu_i)\right)} \geq \frac{\pi_j}{\pi_i}, \text{ para toda } j \neq i,$$

si y sólo si

$$\frac{1}{2}\left[(x_0-\mu_i)'\Sigma_i^{-1}(x_0-\mu_i) - (x_0-\mu_j)'\Sigma_j^{-1}(x_0-\mu_j)\right] \geq \frac{1}{2}\ln\left(\frac{|\Sigma_j|}{|\Sigma_i|}\right) - \ln\left(\frac{\pi_j}{\pi_i}\right),$$

para toda $j \neq i$, es decir, si

$$\frac{1}{2}\left[x_0'\left(\Sigma_i^{-1}-\Sigma_j^{-1}\right)x_0 - 2x_0'\left(\Sigma_i^{-1}\mu_i - \Sigma_j^{-1}\mu_j\right)\right] \leq c - \ln\left(\frac{\pi_j}{\pi_i}\right), \text{ para toda } j \neq i,$$

con $c = \frac{1}{2}\ln\left(\frac{|\Sigma_j|}{|\Sigma_i|}\right) - \frac{1}{2}(\mu_i'\Sigma_i^{-1}\mu_i - \mu_j'\Sigma_j^{-1}\mu_j)$.

En este caso, se obtiene una función discriminante cuadrática, con la que se define la región en la que el individuo se debe de clasificar como perteneciente a la población Π_i.

2.4 Costos de una mala clasificación

El análisis discriminante, tal como se ha descrito hasta el momento, hace uso de la función de densidad del vector de características $\underline{X}$ para cada población, es decir $f_i(\underline{x})$ con $i = 1,\ldots,p$, así como de las probabilidades a priori de cada población. Sin embargo, otro aspecto que se puede considerar es que, como se mencionó en el capítulo anterior, al clasificar a un individuo se pueden cometer errores.

Cuando se trata de un problema de clasificación considerando dos poblaciones, existen dos tipos de errores que se pueden cometer; estos son, clasificar a un individuo de Π_1 como proveniente de Π_2, o clasificar a un individuo de Π_2 como de Π_1. A cada uno de estos errores se encuentra asignado un costo mayor a cero, sea $C(2|1)$ el costo relacionado al primer error y $C(1|2)$ el costo relacionado al segundo tipo de error. Generalmente, se supone que si no existe error al clasificar a un individuo, entonces no existe costo alguno.

Tabla 2.1 Costos de clasificación

		Población asignada	
		Π_1	Π_2
Población verdadera	Π_1	0	$C(2\|1)$
	Π_2	$C(1\|2)$	0

Cabe mencionar que, en general, los costos $C(j|i)$ para toda i, j con $j \neq i$, no son iguales. Por ejemplo, los riesgos monetarios de clasificar a una persona como merecedora de una tarjeta de crédito cuando realmente no lo es, son mayores a los derivados de no otorgar la tarjeta de crédito a aquella persona que sí lo era.

Entonces, con el fin de realizar una mejor clasificación, es posible tomar en cuenta los costos y las probabilidades de una mala clasificación, de tal forma que se minimicen.

Para ello, considere que la probabilidad de asignar un individuo a Π_j, dado que realmente pertenece a Π_i, está dada por

$$P(j \mid i) = P(\underline{X} \in R_j \mid \Pi_i) = \int_{R_j} f_i(x)dx, \quad \text{con } i, j = 1,2 \text{ para } i \neq j$$

Por lo tanto, la probabilidad de que una observación provenga de la población Π_i y que sea erróneamente clasificada como perteneciente a la población Π_j es

$$P(\underline{X} \in R_j, \Pi_i) = P(\underline{X} \in R_j \mid \Pi_i) * P(\Pi_i) = P(j|i)\, \pi_i.$$

En la siguiente tabla se muestran las probabilidades de clasificación para el caso de dos poblaciones.

Tabla 2.2 Probabilidades de clasificación

		Población asignada	
		Π_1	Π_2
Población verdadera	Π_1	$P(1\|1)\,\pi_1$	$P(2\|1)\,\pi_1$
	Π_2	$P(1\|2)\,\pi_2$	$P(2\|2)\,\pi_2$

Por lo tanto, como se mencionó, para realizar una clasificación óptima, lo que se busca es minimizar el costo esperado de una mala clasificación (CEMC), es decir, se busca minimizar

$$CEMC = C(2|1)\,P(2|1)\pi_1 + C(1|2)\,P(1|2)\pi_2$$

$$= C(2|1)\,\pi_1 \int_{R_2} f_1(x)dx + C(1|2)\,\pi_2 \int_{R_1} f_2(x)dx\,, \text{ con } R^k = R_1 \cup R_2$$

$$= C(2|1)\,\pi_1\left[1 - \int_{R_1} f_1(x)dx\right] + C(1|2)\,\pi_2 \int_{R_1} f_2(x)dx$$

$$= C(2|1)\pi_1 + \int_{R_1}[C(1|2)\pi_2 f_2(x) - C(2|1)\pi_1 f_1(x)]dx$$

En general, como los costos de una clasificación errónea son mayores a cero; entonces, minimizar el CEMC es equivalente a que la función a integrar sobre la región R_1 sea menor a cero (ver Seber, 1984), es decir

$$C(1|2)\pi_2 f_2(x) - C(2|1)\pi_1 f_1(x) < 0\,,$$

si y sólo si

$$C(2|1)\pi_1 f_1(x) > C(1|2)\pi_2 f_2(x)\,.$$

En caso de que las probabilidades a priori sean mayores a cero, esto es $\pi_i \neq 0$, se tiene que el CEMC es mínimo si y sólo si

$$\frac{f_1(\underline{x})}{f_2(\underline{x})} > \frac{C(1|2)\pi_2}{C(2|1)\pi_1}\,. \qquad (2.4)$$

Por lo tanto, si la condición (2.4) se cumple, y debido a que se está integrando sobre R_1, se clasificará al individuo como proveniente de la población Π_1.

Nuevamente observe que si los costos son iguales entre sí y lo mismo ocurre con las probabilidades a priori, se reproduce el criterio de máxima verosimilitud.

2.5 Parámetros desconocidos

En algunos problemas de clasificación, la distribución del vector de características $\underline{X}$ en las distintas poblaciones no se conoce en su totalidad. En particular, es frecuente suponer que sus parámetros son desconocidos. En ese caso, para realizar el análisis, los parámetros se deben de estimar con la información de una muestra.

Por ejemplo, considere que bajo Π_i, el vector de características $\underline{X}_i$ tiene una función de distribución normal multivariada con vector de medias $\underline{\mu}_i$ y matriz de varianzas y covariazas Σ_i, ambos desconocidos. Además, suponga que se tienen n_i observaciones del vector de variables aleatorias, $\underline{x}_i' = (x_{i1}, \ldots, x_{ik})$ bajo Π_i, cuya matriz de observaciones es $\mathbf{X}_i = (\underline{x}_{i1}', \ldots, \underline{x}_{in_i}')'$ con $i = 1,\ldots,p$, donde $\underline{x}_{ir}' = (x_{i1r}, x_{i2r}, \ldots, x_{ikr})$, para $r = 1,\ldots,n_i$.

Entonces, los parámetros de la distribución se pueden estimar de la siguiente forma:

$$\hat{\underline{\mu}}_i = \bar{\underline{x}}_i = (\bar{x}_{i1}, \bar{x}_{i2}, \ldots, \bar{x}_{ik})',$$

$$\hat{\Sigma}_i = S_i = \frac{1}{n_i - 1} \sum_{r=1}^{n_i} (\underline{x}_{ir} - \bar{\underline{x}}_i)(\underline{x}_{ir} - \bar{\underline{x}}_i)'$$

donde, $\bar{x}_{ij} = \frac{1}{n} \sum_{r=1}^{n_i} x_{ijr}$, para $i = 1,\ldots,p$; y $j = 1,\ldots,k$.

Si se supone que las poblaciones tienen la misma matriz de varianzas y covarianzas Σ; entonces, un posible estimador de Σ se obtiene al combinar los estimadores de Σ_i para toda i = 1,...,p, de la siguiente forma

$$\hat{\Sigma} = S = \frac{\sum_{i=1}^{p}(n_i - 1)S_i}{n - p},$$

donde n es la suma de las n_i.

Por lo tanto, la función de densidad de probabilidad (estimada) del vector de variables aleatorias $\underline{X}$ bajo la población Π_i, está dada por

$$\hat{f}_i(x) = (2\pi)^{-k/2} \mid \hat{\Sigma} \mid^{-1/2} \exp\left(-\frac{1}{2}(x - \hat{\mu}_i)' \hat{\Sigma}^{-1} (x - \hat{\mu}_i) \right).$$

Los procedimientos descritos previamente se pueden aplicar ahora con f_i sustituida por $\hat{f}_i$.

2.6 Pruebas estadísticas

En el análisis discriminante, la λ de Wilks (ver Apéndice B.1) es una estadística utilizada para probar diferencia de vectores de medias entre un conjunto de p grupos con datos multivariados, esto, con el fin de determinar si la diferencia entre los grupos es estadísticamente significativa. Para ello, se supone que los datos de cada grupo siguen una distribución normal multivariada con vector de medias $\underline{\mu}_i$ (i = 1,...,p), e igual matriz de varianzas y covarianzas Σ. Entonces, la hipótesis de interés H_0 se define como H_0: $\underline{\mu}_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_p$ y la hipótesis alternativa es H_a: $\underline{\mu}_i \neq \underline{\mu}_j$ para al menos dos grupos i, j, con $i \neq j$.

En particular, cuando se trabaja únicamente dos grupos (Π_0 y Π_1), y cualquier número de variables explicativas, la estadística λ puede aproximarse a una estadística con distribución F con k y n-k-1 grados de libertad, donde n es el número de observaciones considerando los dos grupos, y k es el número de variables en la función discriminante.

La regla de decisión de esta prueba consiste en rechazar la hipótesis nula H_0 al nivel de significancia α, si $F > F^{1-\alpha}_{(k,n-k-1)}$, es decir, se considera que existe una diferencia significativa entre los dos grupos.

2.7 Aplicación al análisis de solvencia

A partir de los años 70, la mayoría de los estudios para clasificar a las compañías aseguradoras como solventes o insolventes (Trieschmann y Pinches (1973, 1974, 1977), Hershbarger y Miller (1986), Ambrose y Seward (1988)) hicieron uso del análisis discriminante. Para tal fin, utilizaron datos contables y financieros de las compañías aseguradoras, de uno o dos años previos al estudio o a la declaración de insolvencia.

La introducción de la probabilidad a priori de insolvencia, así como los costos de una clasificación errónea no fueron incluidos en estos estudios, sino hasta 1975 por Cooley.

En este capítulo se comentan dos estudios recientes, realizados por BarNiv y Hershbarger en 1990 y por BarNiv y McDonald en 1992, cuyo objetivo es realizar un análisis de solvencia de las compañías aseguradoras, haciendo uso del análisis discriminante y otras técnicas estadísticas. Además, en estos artículos se recoge la experiencia de algunos estudios previos.

En las publicaciones citadas, se consideran dos grupos de información financiera y contable, uno de compañías insolventes y el otro de las solventes. La clasificación utilizando análisis discriminante se lleva a cabo mediante la siguiente regla

$$\underline{x}_0' \, \Sigma^{-1} (\underline{\mu}_1 - \underline{\mu}_2) > \frac{1}{2} (\underline{\mu}_1 - \underline{\mu}_2)' \, \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2),$$

donde $\underline{\mu}_i$ es el vector de medias de los datos del grupo i, i = 1,2,

Σ es la matriz de varianzas y covarianzas común de los dos grupos, y

$\underline{x}_0$ es un vector con la información de las variables explicativas de la compañía a clasificar.

De tal forma que si se cumple esta desigualdad la compañía analizada es clasificada como perteneciente al grupo 1.

Si además, se consideran las probabilidades a priori, así como los costos de clasificación errónea, se tiene la siguiente regla de clasificación

$$\underline{x}_0' \, \Sigma^{-1} (\underline{\mu}_1 - \underline{\mu}_2) > \frac{1}{2} (\underline{\mu}_1 - \underline{\mu}_2)' \, \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2) + \ln\left(\frac{c(2 \mid 1)\, \pi_1}{c(1 \mid 2)\, \pi_2}\right),$$

donde π_i es la probabilidad priori de la población i, con i = 1,2; y

$C(j \mid i)$ es el costo de clasificar erróneamente un elemento de Π_i, en Π_j.

Para hacer uso de esta regla, es necesario se cumpla el supuesto básico de que la matriz de varianzas y covarianzas debe ser la misma en cada una de las poblaciones. Por otra parte, tanto esta matriz como las medias de las poblaciones suelen ser desconocidas, de manera que es necesario estimar estos parámetros a partir de los datos disponibles.

En esta situación, y como consecuencia, los coeficientes en la combinación lineal de $\underline{x}_0$ que se emplean para clasificar, son también estimados y puede plantearse la conveniencia de probar significancia de algunos de los verdaderos coeficientes desconocidos.

Para llevar a cabo estas pruebas, se hace uso del supuesto de normalidad multivariada para la distribución de $\underline{X}$.

Adicionalmente, en ambos artículos los resultados del análisis discriminante tradicional son comparados con los que se puedan obtener con el llamado análisis discriminante no paramétrico. Esta otra técnica, utiliza la misma idea que da origen a la regla de Fisher, esto es, obtener la mejor combinación lineal $Z = \underline{X}'\underline{b}$, en el sentido de que las distribuciones de los dos grupos de datos se encuentren lo menos sobrepuestas. La diferencia estriba en que en el discriminante no paramétrico se define el siguiente índice de separación (IS)

$$IS(\underline{b}) = \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(Z_{1i} - Z_{2j})}{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}|Z_{1i} - Z_{2j}|} = \frac{Z_1 - Z_2}{\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}|Z_{1i} - Z_{2j}|},$$

donde $Z_{pl} = \underline{x}_l'\underline{b}$ para $l = i,j$; donde $i = 1,\ldots,n_1$ y $j = 1,\ldots,n_2$,

p se refiere al grupo, es decir toma los valores de 1 y 2,

n_p es el número de observaciones del grupo p,

$\bar{Z}_p$ es la media de las observaciones del grupo p.

Este índice de separación tiene la propiedad de tomar valores entre -1 y 1, y lo que se hace en el análisis discriminante no paramétrico es encontrar el vector de coeficientes de $\underline{b}$, tal que se maximice, en valor absoluto, el índice de separación. La maximización de este índice se lleva a cabo mediante el algoritmo de Zangwill

(1967/1968), el cual requiere de una condición inicial del vector de coeficientes, que puede establecerse a partir de los datos o bien, ser un vector de unos.

Lo que resta es seleccionar un punto de corte *cp*, de manera que la regla de clasificación del análisis discriminante no paramétrico consiste, por ejemplo, en asignar una nueva compañía al grupo 1, si

$$\underline{X}'\underline{b} \geq cp.$$

Los autores indican que el punto *cp* se elige de manera que el número de casos mal clasificado de la muestra sea mínimo.

Los mismos autores comentan que la ventaja de utilizar análisis discriminante no paramétrico es que la zona en que se traslapan las distribuciones de los dos grupos de observaciones es menor o igual a la zona obtenida con el análisis discriminante. Además, al igual que ocurre con el procedimiento de Fisher, no es necesaria una distribución para el vector de variables explicativas $\underline{X}$, el cual puede estar formado por variables cualitativas o cuantitativas. Por ello, el número de clasificaciones erróneas usando análisis discriminante no paramétrico suele ser menor.

En el primer artículo, BarNiv y Hershbarger (1990) utilizaron datos de aseguradoras de vida, de uno y dos años previos al estudio o a la declaración de insolvencia.

La muestra con la que trabajaron se formó con 28 compañías aseguradoras de vida que de 1975 a 1985 fueron declaradas insolventes, para las cuales se tenían datos disponibles y que habían sido registrados por la A. M. Best Company. Para los efectos del estudio consideraron que una compañía es insolvente cuando ha sido declarada como tal por el organismo regulador correspondiente, por ello, las compañías disueltas no fueron consideradas dentro del grupo de las insolventes, debido a que pueden incluir aquellas que se han disuelto de forma voluntaria.

Asimismo, en este artículo, se consideraron dos grupos de compañías solventes, el primero se formó por 28 compañías solventes, las cuales se obtuvieron de un apareamiento con las compañías insolventes, este apareamiento se realizó comparando el domicilio de las compañías, volumen de activos y tiempo de vida. El segundo grupo, formado por 49 compañías solventes, se obtuvo de un muestreo aleatorio sobre la población de compañías solventes reportadas por la A.M. Best Company para 1986, y teniendo como única restricción que el total de activos no rebasara los 60 millones de dólares que corresponde al total de activos más grande de las 28 compañías insolventes.

Con el propósito de realizar una validación cruzada, los autores obtuvieron un tercer grupo de compañías, formado por 31 compañías solventes de 1976 a 1982, y 31 compañías consideradas como prioritarias por el sistema de la NAIC, durante el periodo de 1979 a 1980.

Por otra parte, en el segundo artículo, BarNiv y McDonald (1992) utilizaron datos de compañías aseguradoras de daños-responsabilidad civil, de un año previo al estudio o a la declaración de la insolvencia.

BarNiv y McDonald formaron una muestra de 294 compañías, de las cuales 141 fueron declaradas insolventes dentro del periodo comprendido entre los años 1974 y 1988, para las cuales se tienen datos completos disponibles en Best's Insurance Reports, así como de un departamento de seguros, y las 153 compañía restantes correspondientes a solventes se eligieron al realizar un apareamiento con las compañías insolventes, para lo cual se consideró su volumen de primas y del total de activos, al igual que el tiempo para el cual los datos estaban disponibles. El número de compañías insolventes difiere del de las compañías solventes, debido a las limitaciones que se tienen en la disponibilidad de los datos.

Tanto el análisis discriminante como el análisis discriminante no paramétrico fueron aplicados por BarNiv y McDonald, utilizando datos de los años 1974 a 1983, mientras que el resto de los datos fueron utilizados para realizar una validación.

En ambos estudios, los autores realizaron una selección de las variables explicativas a utilizar en los modelos, mediante el procedimiento de selección stepwise.

Los autores investigaron el poder discriminatorio de las variables seleccionadas. En el primer artículo (BarNiv y Hershbarger, 1990), de 31 variables originalmente propuestas por los autores, 13 resultaron ser las más significativas y por consiguiente se utilizaron en el análisis. Mientras que, para el segundo artículo (BarNiv y McDonald, 1992) de 45 variables, 7 fueron las que para uno, dos y tres años previos al estudio o a la declaración de insolvencia resultaron más significativas y por consiguiente se utilizaron en la aplicación de las técnicas estadísticas mencionadas. Estas variables se definen de acuerdo a la práctica contable y financiera en los Estados Unidos, los conceptos a que se refieren no necesariamente tienen un equivalente en otros países. Por esta razón, en los Apéndices C.1 y C.2 se listan las variables con su definición en los términos originales tal como aparece en los artículos y se incluye, además, una colección aproximadamente equivalente producida a partir de la práctica contable en México.

En el primer artículo, BarNiv y Hershbarger analizaron de forma univariada los datos de las trece variables mencionadas para aseguradoras de vida solventes e insolventes, encontrando que existen diferencias considerables en el comportamiento de estos datos entre ambos grupos de compañías, uno y dos años antes de la insolvencia. Siendo la variable GP- *Gains to premium*, la que clasificó correctamente de forma univariada un mayor número de compañías.

Por otra parte, en el segundo artículo, BarNiv y McDonald analizaron el comportamiento de cada una de las siete variables explicativas, uno, dos y tres

años previos al estudio o a la declaración de la insolvencia, encontraron que las siete variables resultaron significativas de forma univariada, para distinguir entre compañías solventes e insolventes.

Finalmente, al realizar un estudio multivariado, BarNiv y Hershbarger aplicaron el análisis discriminante y el análisis discriminante no paramétrico, con datos de uno y dos años previos a la insolvencia o al estudio, para ambos grupos de muestras (muestra que consta de las compañías solventes obtenidas por un apareamiento con las compañías insolventes y muestra que consta de compañías solventes seleccionadas aleatoriamente).

Al utilizar la primera muestra, con datos de un año previo al estudio o a la declaración de la insolvencia, BarNiv y Hershbarger obtuvieron que el análisis discriminante clasificó correctamente un mayor número de compañías como solventes, mientras que el análisis discriminante no paramétrico clasificó correctamente un mayor número de compañías insolventes. Por otra parte, utilizando los datos de dos años previos al estudio o a la ocurrencia de la insolvencia, obtuvieron que el análisis discriminante no paramétrico clasificó correctamente un mayor número de compañías tanto solventes como insolventes, que el análisis discriminante.

Cuando los autores utilizaron la segunda muestra, es decir, aquella que incluye las compañías aseguradas solventes seleccionadas de forma aleatoria, obtuvieron que el análisis discriminante no paramétrico fue el que clasificó correctamente un mayor número de compañías insolventes, así como del total de compañías.

Al utilizar la muestra de las 31 compañías que dentro de los años 1979 y 1980 habían sido consideradas por la NAIC-IRIS como prioritarias a inspeccionar, y sin embargo, para 1985 las 31 compañías continuaban operando; se obtuvo que, únicamente fueron clasificadas como insolventes ocho compañías al utilizar el

análisis discriminante y seis compañías al utilizar el análisis discriminante no paramétrico.

Finalmente, BarNiv y Hershbarger concluyen que los modelos utilizados, además de disminuir el número de clasificación errónea de las compañía insolventes, también disminuyen las clasificaciones erróneas de las compañías solventes.

En el segundo artículo, BarNiv y McDonald (1992) realizaron un análisis multivariado en el que revisaron los supuestos básicos del análisis discriminante y encontraron que la matriz de varianzas y covarianzas no era la misma para ambos grupos de datos, por lo cual no se debería de usar análisis discriminante. Además, la distribución para cada grupo de datos no era normal multivariada.

Sin embargo, utilizando datos de un año previo al estudio o a la declaración de la insolvencia, los autores obtuvieron que el análisis discriminante y el análisis discriminante no paramétrico clasificaron correctamente el mismo número de compañías.

Por último, al utilizar datos de dos y tres años previos, los autores encontraron que el modelo obtenido mediante el análisis discriminante no paramétrico fue el que realizó la mejor clasificación de las compañías aseguradoras como solventes o insolventes.

Como conclusión de estos dos artículos, se tiene que el análisis discriminante no paramétrico parece producir una mejor clasificación que el análisis discriminante.

CAPÍTULO 3
Modelos de respuesta cualitativa

En este capítulo se presentan los elementos de los llamados modelos de respuesta cualitativa, así como los resultados más sobresalientes a que conducen.

La necesidad de describir la relación entre una variable de repuesta y una colección de variables explicativas, ha dado lugar a una gran variedad de modelos estadísticos. Los modelos más populares son los llamados modelos de regresión lineal que durante mucho tiempo se han utilizado en el análisis de datos. Estos modelos son de la forma

$$\underline{Y} = X\underline{\beta} + \underline{e} \qquad (3.1)$$

donde,

$\underline{Y}=(Y_1,\ldots,Y_n)'$ es un vector cuyos componentes son variables aleatorias independientes, conocidas como variables respuesta y que se distribuyen normal con media μ_i, para $i = 1, \ldots, n$,

$X=(\underline{X}_1,\ldots,\underline{X}_k)$ es la matriz formada por vectores de variables explicativas, en donde cada renglón de la matriz se refiere a una observación, y cada columna a una variable explicativa diferente,

$\underline{\beta}=(\beta_1,\ldots,\beta_k)'$ es el vector de parámetros desconocidos del modelo, cada parámetro esta asociada a una variable explicativa, y

$\underline{e}=(e_1,\ldots,e_n)'$ es el vector de errores aleatorios.

La matriz de variables explicativas, es una matriz de n x k formada por variables no aleatorias con valores fijos, que pueden representar medidas continuas en términos cuantitativos, en cuyo caso son conocidas como covariables; o por el contrario pueden representar categorías nominales o términos cualitativos, en cuyo caso reciben el nombre de factores.

El vector de errores aleatorios, en los modelos lineales clásicos, está formado por variables aleatorias independientes e idénticamente distribuidos cuya función de distribución es Normal con media cero y varianza σ^2 constante; es decir, Normal$(0,\sigma^2)$. Por consiguiente, las Y_i son variables aleatorias también independientes con una distribución Normal(μ_i, σ^2).

Este tipo de modelos resultan inapropiados cuando, como en el caso de los estudios de solvencia, la variable de respuesta es cualitativa y más aún, sólo puede producir dos distintos valores (1 = insolvente y 0 = solvente).

Recientes estudios en la teoría estadística, sin embargo, han permitido el uso de métodos análogos a los desarrollados por los modelos lineales en las siguientes condiciones:

a) La distribución de las variables respuestas es diferente a la distribución Normal, además estas variables podrían ser categóricas o continuas;
b) La relación entre las variables respuesta y las variables explicativas no necesariamente es de la forma lineal simple como se muestra en (3.1).

Así, se originan los modelos lineales generalizados que se pueden describir en términos de los siguientes tres componentes.

1. ***Componente Aleatorio***: Las variables respuesta $Y_1,\ldots,Y_n$, son variables aleatorias independientes con una distribución perteneciente a la familia exponencial de distribuciones, que toma la siguiente forma:

$$f(y_i;\theta_i,\varphi) = \left[\frac{y_i\theta_i - b(\theta_i)}{a(\varphi)} + c(y_i,\varphi) \right],$$

Si ϕ (parámetro de dispersión) es conocido, se trata de un modelo con parámetro canónico θ.

Muchas distribuciones conocidas pertenecen a la familia exponencial, como es el caso de la Normal, Binomial, Poisson, Gamma, etc.

2. ***Componente Sistemático***: Las variables explicativas del modelo X_1, ..., X_k, dadas en forma de covariables o factores, producen un predictor lineal $\underline{\eta} = (\eta_1,\ldots,\eta_n)'$ dado por:

$$\underline{\eta} = X\underline{\beta} = \sum_{j=1}^{k} X_j\beta_j \text{ ; es decir, } \eta_i = \sum_{j=1}^{k} x_{ij}\beta_j, \quad \text{con } i = 1, \ldots, n$$

donde,

X es la matriz de variables explicativas,

X_j son las variables explicativas, con j = 1, ..., k,

x_{ij} es el valor de cada variable explicativa j para la observación i de la muestra,

β_j es el parámetro asociado a la variable j, que puede ser estimado a través de los datos.

3. ***Función liga***: Sea $g(\cdot)$ una función monótona diferenciable, tal que relaciona el componente sistemático η_i con el valor esperado de la variable respuesta Y_i; es decir,

$$g(E[Y_i]) = \eta_i = \sum_{j=1}^{k} x_{ij}\beta_j ,$$

Cuando $g(E[Y_i]) = \theta_i = \eta_i$, se dice que la función $g(\cdot)$ es una liga canónica. Para los modelos lineales clásicos la función liga es la función identidad.

Cada una de las distribuciones pertenecientes a la familia exponencial tienen algunas funciones ligas habituales, por ejemplo:

Normal (μ,σ^2), $\eta = \mu$,

Poisson (λ), $\eta = \log(\lambda)$,

Binomial (n,p), $\eta = \log\left[\dfrac{p}{1-p}\right]$,

Gamma $(\alpha, \alpha/\beta)$, $\eta = \beta^{-1}$.

A diferencia de las técnicas estadísticas presentadas anteriormente, los modelos de respuesta cualitativa proporcionan de manera explícita la probabilidad de que una compañía aseguradora sea insolvente.

Un ejemplo de los modelos de respuesta cualitativa son aquellos cuya variable respuesta sigue una distribución Bernoulli(θ) o bien, una distribución Binomial$(1, \theta)$. En particular, si su función de liga es la logit, es decir, $g(\theta) = \log\left[\dfrac{\theta}{1-\theta}\right]$, entonces, éste es conocido como modelo logístico.

Para presentar estos modelos, considere una variable aleatoria cualitativa Y_i que toma los valores 1 y 0, es decir,

$$Y_i = 1 \text{ con } P(Y_i = 1) = \theta_i \text{ y}$$
$$Y_i = 0 \text{ con } P(Y_i = 0) = 1-\theta_i.$$

Por lo tanto, Y_i tiene una función de distribución Bernoulli(θ_i), o bien una función de distribución Binomial ($1,\theta_i$), es decir

$$P(y_i;\theta_i) = \theta_i^{y_i}(1-\theta_i)^{1-y_i},$$

con $E(Y_i) = \theta_i$.

Por otro lado, sea $\underline{x}_i' = (x_{i1}, x_{i2},\ldots, x_{ik})$ un vector con k variables explicativas, observadas en el individuo i, que se suponen relacionadas con la correspondiente probabilidad de insolvencia θ_i. En términos generales, se supone que existe una función $h(\cdot)$ tal que

$$\theta_i = h(x_i).$$

Entonces, para el caso del modelo logístico, el cual presenta una forma más sencilla para trabajar por lo que es utilizado comúnmente, se tiene que

$$\theta_i = F(x_i'\beta) = \frac{e^{x_i'\beta}}{1+e^{x_i'\beta}},$$

o equivalentemente

$$x_i'\beta = \log\left(\frac{\theta_i}{1-\theta_i}\right),$$

en donde $\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)'$ es un vector de coeficientes fijo pero desconocido.

3.1 Estimación puntual de θ_i utilizando el modelo logístico

Cuando se considera simultáneamente la información proveniente de una muestra de tamaño n, el estimador máximo verosímil de θ_i, es decir $\hat{\theta}_i = F(\underline{x}_i'\hat{\beta})$, por la propiedad de invarianza, se obtiene numéricamente al maximizar la función de verosimilitud o equivalentemente la de logverosimilitud que se define como

$$\lambda(\beta; y, X) = \ln\left(\prod_{i=1}^{n} P(y_i;\theta)\right) = \ln\left(\prod_{i=1}^{n} \theta_i^{y_i}(1-\theta_i)^{1-y_i}\right) = \sum_{i=1}^{n}\left[\ln\left(\theta_i^{y_i}(1-\theta_i)^{1-y_i}\right)\right]$$

$$= \sum_{i=1}^{n}[y_i \ln\theta_i + (1-y_i)\ln(1-\theta_i)] = \sum_{i=1}^{n}\left[y_i \ln\left(\frac{e^{x_i'\beta}}{1+e^{x_i'\beta}}\right) + (1-y_i)\ln\left(\frac{1}{1+e^{x_i'\beta}}\right)\right],$$

donde $y = (y_1, y_2, \ldots, y_n)'$ es el vector de la variable de respuesta de los n elementos de la muestra, $X = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k)$ es la matriz de n x k de las variables explicativas observadas, en donde cada renglón de la matriz se refiere a una observación, y cada columna a una diferente variable explicativa, y n es el tamaño de la muestra.

3.2 Intervalos de confianza para θ_i utilizando el modelo logístico

Otros usos de los modelos de respuesta cualitativa es la construcción de intervalos de confianza para la probabilidad de que la variable Y tome el valor de 1 o 0.

Para ello, sea $\hat{\beta}' = (\hat{\beta}_1, \ldots, \hat{\beta}_K)$ el estimador máximo verosímil de $\beta' = (\beta_1, \ldots, \beta_K)$ y $F(x_i'\hat{\beta})$ el correspondiente estimador máximo verosímil de $F(x_i'\beta)$. De las propiedades de los estimadores máximo verosímiles se sabe que si n es suficientemente grande (ver Judge et al, 1988) se cumple que:

$$\sqrt{n}\left(F(x_i'\hat{\beta}) - F(x_i'\beta)\right) \sim N(0, \underline{M}_i' I(\hat{\beta})^{-1} \underline{M}_i) \qquad (3.2)$$

donde, n es tamaño de la muestra,

$I(\hat{\beta})$ es la matriz de información de $\hat{\beta}$, y

$\underline{M}_i$ es el vector de derivadas parciales de $F(x_i'\hat{\beta})$ con respecto a $\hat{\beta}_j$, para $j=1,\ldots,K$.

La matriz $I(\hat{\beta})$, que es una matriz definida positiva, se obtiene a partir de la logverosimilitud ℓ, como sigue:

$$I(\hat{\beta}) = -E\left[\frac{\partial^2 \ell}{\partial\hat{\beta}\partial\hat{\beta}'}\right] = \begin{pmatrix} -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_1^2}\right) & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_1\partial\hat{\beta}_2}\right) & \cdots & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_1\partial\hat{\beta}_k}\right) \\ -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_2\partial\hat{\beta}_1}\right) & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_2^2}\right) & \cdots & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_2\partial\hat{\beta}_k}\right) \\ \vdots & \vdots & \ddots & \vdots \\ -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_k\partial\hat{\beta}_1}\right) & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_k\partial\hat{\beta}_2}\right) & \cdots & -E\left(\frac{\partial^2 \ell}{\partial\hat{\beta}_k^2}\right) \end{pmatrix}$$

Por otro lado, se tiene que:

$$\underline{M}_{ij} = \frac{\partial F(x_i'\hat{\beta})}{\partial\hat{\beta}_j} = \frac{\partial}{\partial\hat{\beta}_j}\left(\frac{e^{x_i'\hat{\beta}}}{1+e^{x_i'\hat{\beta}}}\right) = \frac{\partial}{\partial\hat{\beta}_j}\left(\frac{1}{1+e^{-x_i'\hat{\beta}}}\right) = \frac{e^{-x_i'\hat{\beta}}}{\left(1+e^{-x_i'\hat{\beta}}\right)^2} x_{ij}; \quad j = 1, \ldots, K$$

o equivalentemente

$$M_i = \hat{\theta}_i(1-\hat{\theta}_i)x_i, \quad \text{con } i = 1, \ldots, n$$

Sea $V(\hat{\beta}) = \underline{M}_i' I(\hat{\beta})^{-1}\underline{M}_i = \hat{\theta}(1-\hat{\theta})\underline{x}_i' I(\underline{\hat{\beta}})^{-1}\hat{\theta}(1-\hat{\theta})x_i = \left(\hat{\theta}_i(1-\hat{\theta}_i)\right)^2 x_i' I(\hat{\beta})^{-1} x_i \geq 0$,

debido a que $I(\hat{\beta})^{-1}$ es una matriz definida positiva.

Entonces, a partir de la expresión (3.2), se tiene que

$$\sqrt{n}\left(F(\underline{x}_i'\underline{\hat{\beta}}) - F(\underline{x}_i'\underline{\beta})\right) \sim N(0, V(\hat{\beta})),$$

o bien

$$\sqrt{n}(\hat{\theta}_i - \theta_i) \sim N(0, V(\hat{\beta})),$$

por lo que

$$\frac{\sqrt{n}(\hat{\theta}_i - \theta_i)}{(V(\hat{\beta}))^{1/2}} \sim N(0,1),$$

lo que implica

$$P\left[-Z_{1-\alpha/2} < \frac{\sqrt{n}(\hat{\theta}_i - \theta_i)}{(V(\hat{\beta}))^{1/2}} < Z_{1-\alpha/2}\right] \approx 1-\alpha \qquad (3.3)$$

donde, $Z_{1-\alpha/2}$ es el cuantil $1-\frac{\alpha}{2}$ de una Normal(0,1), y

1-α es un nivel de probabilidad predeterminado.

Entonces, de la expresión (3.3) se deriva que

$$P\left[-Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2} < \sqrt{n}(\hat{\theta}_i - \theta_i) < Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}\right] \approx 1-\alpha \quad \Leftrightarrow$$

$$P\left[\frac{-Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}} < \theta_i - \hat{\theta}_i < \frac{Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}}\right] \approx 1-\alpha \quad \Leftrightarrow$$

$$P\left[\hat{\theta}_i - \frac{Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}} < \theta_i < \hat{\theta}_i + \frac{Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}}\right] \approx 1-\alpha$$

Por lo que el intervalo de confianza de nivel (1-α)x100% para la probabilidad $\theta_i = F(\underline{x}_i'\underline{\beta})$ del individuo i, es un intervalo simétrico de la forma:

$$\left[\hat{\theta}_i - \frac{Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}} , \quad \hat{\theta}_i + \frac{Z_{1-\alpha/2}(V(\hat{\beta}))^{1/2}}{\sqrt{n}} \right]. \qquad (3.4)$$

3.3 Pruebas estadísticas

En los modelos de regresión logística, una medida que permite determinar si el ajuste del modelo resultó adecuado, es la estadística conocida como razón de verosimilitudes generalizada, con la que teóricamente se compara la función de verosimilitud del modelo ajustado con la del modelo llamado saturado. El modelo saturado es el modelo que se ajusta con tantos parámetros como observaciones se tienen, por lo que describiría completamente a los datos. La razón de verosimilitudes para comparar estos dos modelos está definida como

$$\lambda = \frac{L(\hat{\beta}_{max}; y)}{L(\hat{\beta}; y)},$$

donde, $\hat{\beta}_{max}$ es el estimador máximo verosímil bajo el modelo saturado, y

$\hat{\beta}$ es el estimador máximo verosímil del modelo ajustado.

Debido a que el modelo saturado describe completamente a los datos, en caso de que la estadística λ tome valores cercanos a uno, esto significaría que el modelo ajustado describe bien a los datos. Por otra parte, en la medida en que la estadística tome valores grandes, esto significaría que el modelo no se ajusta a los datos.

Equivalentemente, se puede utilizar el logaritmo de λ,

$$Ln(\lambda) = Ln\left(L(\hat{\beta}_{max}; y)\right) - Ln\left(L(\hat{\beta}; y)\right) = \lambda(\hat{\beta}_{max}; y) - \lambda(\hat{\beta}; y),$$

o bien, la devianza que se define como $D = 2^*Ln(\lambda) = 2^*[\lambda(\hat{\beta}_{max}; y) - \lambda(\hat{\beta}; y)]$.

Se sabe que la devianza de un modelo, con k variables explicativas y el término constante, sigue una distribución χ^2 con n-(k+1) grados de libertad (ver Dobson, 1990), si se cumple la hipótesis de que todos los parámetros en el modelo saturado, que no están en el modelo que se pretende ajustar son cero.

En caso de que $D < \chi^2_{(n-(k+1)),\ 1-\alpha}$, se dice que con un nivel de significancia α, el modelo propuesto se ajusta bien a los datos.

La devianza es una estadística útil para elegir el modelo que se ajusta mejor a los datos, al comparar dos modelos que tienen diferente número de parámetros. Para ello, se plantea la prueba de hipótesis especificada en términos del vector de parámetros β, cuya hipótesis nula es H_0: $\beta_i = 0$ para toda $i \in A$, donde $A \subset I = \{0, 1, 2, \ldots, k\}$, esto es, únicamente se considera q parámetros en el modelo ajustado, con $q < k+1$, y la hipótesis alternativa es H_a: $\beta_i \neq 0$ para toda $i \in I$; es decir, se consideran k variables explicativas en el modelo, así como el término independiente o constante, donde $k+1 \leq n$, con n el número de observaciones. La estadística utilizada en esta prueba es la diferencia de la devianza del modelo bajo H_0 menos la devianza del modelo bajo H_a, esta estadística tienen aproximadamente una distribución χ^2 con k-q grados de libertad, esto es,

$$\Delta D = D_o - D_a$$

donde, D_0 es la devianza del modelo bajo la hipótesis nula, por lo que $D_0 \sim \chi^2_{(n-(q+1))}$,

D_a es la devianza del modelo bajo la hipótesis alternativa, por lo que $D_a \sim \chi^2_{(n-(k+1))}$.

Por lo que, la diferencia de devianzas queda de la forma

$$\Delta D = D_0 - D_a = 2[\lambda(\hat{\beta}_{max};y) - \lambda(\hat{\beta}_0;y)] - 2[\lambda(\hat{\beta}_{max};y) - \lambda(\hat{\beta}_a;y)] \sim \chi^2_{[(n-(q+1))-(n-(k+1))]},$$

es decir

$$\Delta D = 2[\lambda(\hat{\beta}_a;y) - \lambda(\hat{\beta}_0;y)] \sim \chi^2_{(k-q)}$$

La regla de decisión de esta prueba consiste en rechazar la hipótesis nula H_0 en favor de la hipótesis alternativa H_a, con un nivel de significancia α, si $\Delta D > \chi^2_{(k-q),1-\alpha}$; es decir, se considerará que el modelo de regresión logística con k variables explicativas y el término constante proporciona una descripción de los datos significativamente mejor, que el modelo con sólo q variables explicativas y el término constante.

Por otro lado, la estadística de Wald, definida como el cuadrado del cociente del parámetro estimado con su error estándar también estimado, es utilizada para verificar la significancia parcial de una variable explicativa; es decir, cuando las demás variables están en el modelo. Esta estadística se distribuye aproximadamente como una χ^2 con un grado de libertad (ver Ryan, 1997). La estadística de Wald, para la variable explicativa i, está dada por

$$W = \frac{(\hat{\beta}_i)^2}{(\hat{S}_{\hat{\beta}_i})^2} \sim \chi^2_{(1)}$$

En caso de que $W > \chi^2_{(1),1-\alpha}$, se dice que con un nivel de significancia α, la variable explicativa i es significativa, al considerar el resto de las variables en el modelo.

3.4 Aplicación al análisis de solvencia

Algunos investigadores comenzaron a usar modelos de regresión con respuesta binaria, también conocidos como modelos lineales de probabilidad, para predecir la probabilidad de insolvencia de compañías del sector asegurador. Algunos de estos estudios fueron publicados por Meyer y Pifer (1970), Collins (1980), Eck (1982), y Harrington y Nelson (1986).

Estos estudios presentaban algunas limitaciones como se mencionó al inicio de este capítulo. Sin embargo, de alguna forma relacionada con esta línea de trabajo, a principios de los años 90, BarNiv comenzó a utilizar los modelos de respuesta cualitativa, en particular el modelo logístico, para identificar probabilidades de insolvencia en compañías aseguradoras que operan el ramo de daños-responsabilidad civil.

En el presente trabajo de tesis se comentan tres artículos recientes en los que se hace uso del modelo logístico para estimar probabilidades de insolvencia de compañías aseguradoras. Dos de estos artículos fueron presentados en la sección 2.7.

El primero artículo fue publicado por BarNiv y Hershbarger en 1990, quienes presentan un estudio para predecir la insolvencia de las compañías que operan el seguro de vida utilizando el modelo lineal generalizado con función de liga logit, y los resultados son comparados con los obtenidos de aplicar las técnicas de análisis discriminante y análisis discriminante no paramétrico, ya mencionadas.

En el segundo estudio, BarNiv McDonald (1992) presentan los modelos de respuesta cualitativa y realizan una comparación con el análisis discriminante y el análisis discriminante no paramétrico utilizando datos de compañías aseguradoras que operan el ramo de daños-responsabilidad civil.

Por último, el tercer artículo publicado por BarNiv et al (1999), es muy reciente y es en donde se introduce otro uso a los modelos de respuesta cualitativa, que consiste en obtener un intervalo de confianza para la probabilidad de insolvencia de las compañías aseguradoras que operan el ramo de daños-responsabilidad civil.

En el primer artículo, en donde se trabajó con compañías que operan el seguro de vida, BarNiv y Hershbarger (1990) obtuvieron como resultado, al utilizar los datos de la muestra compuesta de las compañías solventes obtenidas por un apareamiento con las compañías insolventes, que el modelo logístico fue el que clasificó correctamente un mayor número de compañías, a comparación del análisis discriminante y el análisis discriminante no paramétrico, para uno y dos años previos al estudio o a la declaración de la insolvencia.

Por otra parte, al utilizar BarNiv y Hershbarger la segunda muestra, es decir, aquella compuesta por compañías solventes seleccionadas aleatoriamente, obtuvieron como resultado que al utilizar los datos de un año previo al estudio o a la declaración de la insolvencia, el análisis discriminante no paramétrico fue el que clasificó correctamente un mayor número de compañías insolventes; sin embargo, el modelo logístico fue el que clasificó correctamente un mayor número de compañías del seguro de vida, ya sea solventes o insolventes. Por último, al utilizar datos de dos años previos, se obtuvo que el modelo logístico realizó una mejor clasificación de compañías insolventes, así como del total de las compañías.

En el segundo artículo, en donde se trabajó con compañías que operan el seguro de daños-responsabilidad civil, BarNiv y McDonald (1992) asignaron a cada una de las

compañías estudiadas a una categoría, ya sea solvente (0) o insolvente (1) cuya probabilidad de pertenecer a ella excedió cierto valor, típicamente 0.5, a este valor se le conoce como punto de corte. Como resultado de este estudio, al utilizar un punto de corte de 0.5 para la probabilidad de insolvencia obtenida con los modelos de respuesta cualitativa, y comparar los resultados para todos los modelos, BarNiv y McDonald (1992) observaron que, al utilizar datos de un año previo, el modelo logístico clasificó correctamente un mayor número de compañías, ya sea como solventes o insolventes.

Por otro lado, utilizando datos de dos años previos, el modelo logístico fue el que clasificó correctamente un mayor número de compañías tanto solventes como del total, y junto con el análisis discriminante no paramétrico clasificaron correctamente, a comparación del análisis discriminante, un mayor número de compañías aseguradoras insolventes.

Por último, con los datos de tres años previos, el modelo obtenido mediante el análisis discriminante no paramétrico, fue el que realizó la mejor clasificación de las compañías.

Finalmente, en el tercer artículo, BarNiv et al (1999) comentan que una de las principales dificultades con las que se han topado los investigadores para obtener una estimación puntual confiable de la probabilidad de insolvencia, son los tamaños de muestra. Como se menciona en este artículo, al trabajar con muestras pequeñas, los estimadores sufren de diversas limitaciones. Por otro lado, aunque los estimadores se obtengan de muestras grandes y en principio pueda hacerse uso de algunos resultados como la consistencia, la calidad de la estimación, al final, depende de la estructura de la base de datos, ya que ésta puede tener problemas diversos como errores, observaciones repetidas o incompletas, etc. Es necesario entonces, depurar la base con la consiguiente reducción del tamaño de muestra efectivo.

En estas condiciones y puesto que prácticamente en ningún caso es posible asegurar un tamaño de muestra que permita ignorar el error en la estimación, los autores identifican la necesidad de un estimador de la probabilidad de insolvencia en el que se considere o se mida el error de estimación. Esto se puede lograr a través del cálculo de intervalos de confianza para la probabilidad de insolvencia, ya que éstos proporcionan un rango en el que se encuentra la probabilidad y ese rango depende del error de estimación.

Para el cálculo de los intervalos de confianza para la probabilidad de insolvencia, BarNiv et al (1999) relacionaron la variable de respuesta Y, que toma los valores de 1 y 0 dependiendo de si el asegurador es insolvente o solvente, con las variables que explican la tendencia de un asegurador a ser insolvente de acuerdo con su información contable y financiera. Para describir esta relación utilizaron el modelo logístico, cuya característica principal es que la variable a explicar o variable respuesta no es cuantitativa sino cualitativa.

De esta relación obtuvieron un estimador máximo verosímil de la probabilidad de insolvencia; es decir, de la probabilidad de que la variable aleatoria Y tome el valor 1, que depende de las variables explicativas. A partir de ese estimador construyeron un intervalo de confianza para la probabilidad de insolvencia, de la misma forma que se describió en la sección 3.2, utilizando como cantidad pivotal una expresión derivada de los estimadores máximo verosímiles.

Como resultado de este estudio, BarNiv et al (1999) obtuvieron un intervalo simétrico con un nivel de confianza de $(1-\alpha)x100\%$ para la probabilidad de insolvencia $\theta_i = F(\underline{x}_i'\underline{\beta})$ del asegurador i, como se muestra en la expresión (3.4).

Para complementar el estudio, en el artículo se sugiere examinar ciertas características de los intervalos, como son la mínima cota superior y su longitud

mínima, que resultan de utilidad para proponer estrategias de control de la insolvencia. Al obtener, para un nivel de confianza fijo, la mínima cota superior del intervalo como función de $\underline{X}_i$, se identifican los valores de las variables explicativas en los que se asegura que la probabilidad de insolvencia no sea mayor que la cota superior (mínima) con ese nivel de confianza. Cuando se determinan los valores de $\underline{X}_i$ para los que el intervalo, con un nivel de confianza fijo, alcanza su longitud mínima, se identifican los niveles de las variables explicativas que producen una situación de incertidumbre mínima sobre el valor de la probabilidad de insolvencia. En este artículo se propone tomar decisiones sobre las variables contables que explican la tendencia del asegurador a ser insolvente, con el fin de reducir simultáneamente la longitud del intervalo así como su cota superior.

Para ilustrar este uso del modelo logístico, en este artículo se presentan tres ejemplos en donde se utilizan datos de aseguradores solventes e insolventes del seguro de daños-responsabilidad civil, obtenidos de la NAIC (National Association of Insurance Commissioners) en el periodo de 1984 a 1992. En estos ejemplos, los autores hacen uso de tres variables explicativas para describir la tendencia de las compañías aseguradoras a ser insolventes. Estas variables se describen en el Apéndice C.3.

En su primer ejemplo, BarNiv et al (1999) calculan un intervalo de confianza para la probabilidad de insolvencia, utilizando como variables explicativas NPWSURP, LARAT y LOSRAT. Para efectos de análisis en el modelo ajustado, las últimas dos variables las mantienen fijas en los valores promedio de la industria aseguradora. Al analizar las características del intervalo como función de NPWSURP, se observó que no necesariamente cuando esta variable es muy pequeña el intervalo alcanza su longitud mínima, sino que esto ocurre más bien cuando la variable explicativa se acerca a 1.7 que es un valor mayor al que representa la industria en promedio.

BarNiv et al (1999), en el segundo ejemplo, comparan el intervalo de confianza para la probabilidad de insolvencia obtenido con los valores de $\underline{X}_i$ igual al promedio de la industria aseguradora y los intervalos que, para el mismo nivel de confianza, se obtienen si $\underline{X}_i$ se fija, respectivamente, en los valores que producen la longitud mínima y la mínima cota superior. Con los datos utilizados por los autores, resulta que un mismo valor de $\underline{X}_i$ produce simultáneamente la longitud mínima y la mínima cota superior.

Al utilizar los valores óptimos de las variables explicativas, los autores obtuvieron una considerable reducción en la longitud y en la cota superior del intervalo de confianza, con respecto a los valores que obtuvieron al utilizar los datos promedio de la industria aseguradora.

Este hecho sugiere que, al tomar ciertas decisiones sobre las variables explicativas, es posible reducir de manera considerable los intervalos de confianza para la probabilidad de insolvencia, sin incrementar el riesgo de incertidumbre α.

En el último ejemplo, BarNiv et al (1999) utilizan información de una compañía aseguradora insolvente para calcular el intervalo de confianza de su probabilidad de insolvencia, después, éste es comparado con los intervalos obtenidos al utilizar tanto los valores óptimos de las variables explicativas como con los valores promedio de la industria aseguradora. Como era de esperarse, obtuvieron que la longitud del intervalo de confianza para el asegurador insolvente es excesivamente grande, e incluye valores para la probabilidad de insolvencia desde 0.1 hasta 0.8.

Cabe mencionar que en este artículo, los autores hicieron uso del modelo logístico y de ciertas variables explicativas para calcular los intervalos de confianza para la probabilidad de insolvencia de un asegurador; sin embargo, se pueden utilizar otras variables explicativas derivadas de la situación contable de los aseguradores, así como otros modelos, por ejemplo el probit, para el cual la variable respuesta Y

tiene una distribución Binomial $(1,\theta)$ y su función de liga está basada en la distribución normal estandarizada con media 0 y varianza 1, es decir

$$\theta_i = F(x_i'\beta) = \Phi(x_i'\beta),$$

donde, $\Phi(\cdot)$ la función de densidad normal estandarizada con media 0 y varianza 1.

La propuesta que hace BarNiv el al (1999) en este estudio, es una estrategia para fijar los valores de las variables explicativas que, con un nivel de confianza predeterminado, aseguren que la probabilidad de insolvencia se mantenga relativamente pequeña y/o con poca incertidumbre. En una variante general, esta idea implica tomar decisiones sobre factores que influyan en las variables explicativas como pueden ser reducir la suscripción de pólizas, incrementar las inversiones, o alguna otra situación contable del asegurador. Debido a que la toma de estas decisiones no es sencilla, los autores sugieren considerar el costo y los beneficios que implican los cambios en las variables explicativas.

CAPÍTULO 4
Análisis de Supervivencia

En este capítulo se presentan los elementos estadísticos de la técnica conocida como análisis de supervivencia, así como algunas aplicaciones de ésta en el análisis de solvencia.

En el análisis de supervivencia, el interés principal se centra en el estudio de un grupo de individuos u objetos cada uno de los cuales puede experimentar un evento; este evento puede ser la muerte, la aparición de una enfermedad, un divorcio, etc. El evento sólo ocurre una vez, al final de lo que se conoce como tiempo de vida. Por facilidad en la discusión general de este capítulo se denominará al evento relevante como muerte.

Algunos ejemplos en los que se utiliza el análisis de supervivencia son: el estudio del tiempo de vida útil de las máquinas, el estudio del tiempo que tarda una persona desempleada en encontrar trabajo, o como en nuestro caso, el estudio del tiempo que una compañía aseguradora solvente se mantiene en ese estado.

Para el análisis de supervivencia, el tiempo de vida se trata como una variable aleatoria y a diferencia de las técnicas presentadas anteriormente, el análisis es, por decirlo de alguna manera, de tipo dinámico, ya que se monitorea y estudia la distribución del tiempo de vida. Para describir la distribución de probabilidad de los tiempos de vida se hace uso de cuatro funciones, estas son: la función de supervivencia, la función de distribución acumulada, la función de densidad de probabilidad y la función o tasa de riesgo.

Suponga que T es una variable aleatoria que indica el tiempo de vida de un individuo de una población bajo estudio; por supuesto $T \geq 0$. Suponga, además que T es continua.

La función de supervivencia S(t), se define como la probabilidad de que el individuo se mantenga con vida después de un tiempo t, es decir,

$$S(t) = P(T > t) = 1 - P(T \leq t),$$

Esta función es continua decreciente y en t = 0 vale 1, mientras que tiende a cero si t tiende a infinito.

Por otra parte, la función de distribución acumulada F(t) de la variable aleatoria T, indica la probabilidad de que el individuo muera antes de que transcurra el tiempo t. Entonces, se tiene que

$$S(t) = 1 - F(t) \qquad (4.1)$$

El tercer componte, la función de densidad de probabilidad f(t), se relaciona con F(t) de la siguiente manera

$$F(t) = \int_0^t f(u)du .$$

Por último, la función o tasa de riesgo h(t), representa la tasa de cambio de la probabilidad de que un individuo que ha llegado con vida al tiempo t muera inmediatamente después.

Formalmente, se tiene que

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t},$$

es decir,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(\{t \le T < t + \Delta t\} \cap \{T \ge t\})}{\Delta t} * \frac{1}{P(T \ge t)},$$

$$= \frac{1}{P(T \ge t)} * \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t)}{\Delta t},$$

$$= \frac{1}{P(T \ge t)} * \lim_{\Delta t \to 0} \frac{P(T < t + \Delta t) - P(T < t)}{\Delta t},$$

$$= \frac{1}{P(T \ge t)} * \lim_{\Delta t \to 0} \frac{[1 - P(T > t + \Delta t)] - [1 - P(T > t)]}{\Delta t},$$

$$= -\frac{1}{S(t)} * \lim_{\Delta t \to 0} \frac{S(t + \Delta t) - S(t)}{\Delta t} = -\frac{1}{S(t)} * \frac{d}{dt}(S(t)),$$

$$= -\frac{S'(t)}{S(t)} = -\frac{d}{dt}[\ln(S(t)],$$

o al utilizar (4.1), equivalentemente se tiene que

$$h(t) = \frac{f(t)}{S(t)}.$$

En algunos casos, resulta conveniente utilizar la función acumulada de riesgo H(t), que se define como

$$H(t) = \int_0^t h(u)du,$$

por lo que se tiene, que la función de supervivencia también se puede expresar como

$$S(t) = \exp[-H(t)].$$

Cabe mencionar que la función de riesgo h(t), es utilizada comúnmente en el estudio de distribuciones de tiempos de vida. En muchos casos, la función de riesgo se utiliza para seleccionar un modelo que describa la distribución del tiempo de vida.

4.1 Censura

Una característica distintiva de los datos utilizados en el análisis de supervivencia es la posible existencia de censura. La idea de censura se refiere a que en algunos casos puede suceder que el evento de interés no es observado y únicamente se cuenta con una cota, ya sea superior o inferior, del tiempo de vida. La censura suele deberse a que los tiempos disponibles para los estudios son limitados.

La presencia de censura en los datos objeto de un análisis de supervivencia complica el proceso de inferencia. Dos posibles esquemas de censura son los llamados *censura por la derecha* y *censura por la izquierda*.

La *censura por la derecha* se da cuando el individuo observado se encuentra vivo al momento en que el estudio concluye. Esto es; una observación está censurada por la derecha en A, si el valor exacto del tiempo de vida del individuo es desconocido y únicamente se sabe que es mayor o igual a A. Este tipo de censura es la más común en los datos de tiempos de vida. Existen dos tipos de *censura por la derecha*, censura *tipo I* y censura *tipo II*.

La censura por la derecha *tipo I* se da cuando en una muestra de n individuos el individuo i es observado durante un tiempo A_i (por supuesto, en el caso más simple se tiene que $A_i = A$ para toda $i = 1,2,\ldots,n$). En este caso, se conocerá el tiempo exacto de vida T_i del individuo i, sólo si éste es menor al tiempo de estudio A_i; es decir, si $T_i \leq A_i$. Con este esquema, el número de tiempos de vida exactos (no censurados) que se observan resulta aleatorio.

El dato observado para el individuo i, cuando existe censura por la derecha del *tipo I*, se suele expresar como (t_i, δ_i),donde

$$t_i = \min(T_i, A_i) \quad \text{y} \quad \delta_i = \begin{cases} 1 & \text{si } T_i \leq A_i \\ 0 & \text{si } T_i > A_i \end{cases}$$

con δ_i una variable que indica la presencia de censura.

La función de densidad de probabilidad de t_i y δ_i está dada por

$$P(t_i, \delta_i) = [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i},$$

Debido a que los pares (t_i, δ_i) con $i = 1,2,\ldots,n$ son independientes entre sí, se tiene que la función de verosimilitud resulta

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} . \qquad (4.2)$$

Como contraparte, la censura por la derecha del *tipo II* se presenta cuando el estudio se termina después de haber observado r eventos o muertes ($1 \leq r \leq n$) de entre los n individuos que inician el estudio. En este caso, el número de tiempos de vida no censurados es fijo y está determinado de antemano (es precisamente r).

Con censura por la derecha del *tipo II*, la muestra constará de los r tiempos de vida más pequeños; es decir $T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(r)}$ obtenidos de una muestra aleatoria $T_1, T_2, \ldots, T_n$. Para el caso en que los T_i para $i = 1, \ldots, n$ sean independientes e idénticamente distribuidos con una función de distribución continua se tendrá que la función de densidad de probabilidad conjunta de $T_{(1)}, T_{(2)}, \ldots, T_{(r)}$, es (ver Lawless, 1982):

$$f(t_{(1)}, t_{(2)}, \ldots, t_{(r)}) = \frac{n!}{(n-r)!} f(t_{(1)}) \cdots f(t_{(r)}) \left[S(t_{(r)})\right]^{n-r}.$$

Por otra parte, la *censura por la izquierda* se tiene cuando únicamente se sabe que el individuo estudiado ha experimentado el evento antes de comenzar el estudio. Cabe mencionar que este tipo de censura no se presenta comúnmente en datos utilizados en el análisis de supervivencia.

4.2 Métodos estadísticos

En el análisis de supervivencia se desea realizar inferencias sobre el patrón de supervivencia de la población. Para ello, y por facilidad, se especifican modelos paramétricos para representar la distribución de los tiempos de vida. Sin embargo, en muchas ocasiones los modelos paramétricos no son satisfactorios, por lo que se hace uso de los llamados métodos no paramétricos. Enseguida, se presentaran algunos de los métodos estadísticos básicos para el análisis de datos de tiempos de vida.

4.3 Modelos continuos

4.3.1 Modelos paramétricos continuos

En el análisis de supervivencia se han empleado distintos modelos paramétricos para el análisis de tiempos de vida. Los más comúnmente utilizados son Exponencial, Weibull, Gamma, y Log-normal.

El modelo Weibull es uno de los más populares ya que ha resultado exitoso para describir diferentes tipos de datos de tiempos de vida, además de que presenta expresiones simples para la función de supervivencia, la función de densidad de probabilidad y la función de riesgo.

La función de densidad de probabilidad de una variable aleatoria T con una distribución Weibull (λ,p), donde λ el parámetro de escala y p es el parámetro de forma, es la siguiente

$$f_T(t) = \begin{cases} \lambda p t^{p-1} \exp(-\lambda t^p), & \text{para } t \geq 0, \text{ y } \lambda, p > 0 \\ 0, & \text{en otro caso.} \end{cases}$$

Como puede observarse, la distribución Weibull es una generalización de la distribución exponencial, que se obtiene como caso particular cuando el parámetro de forma es igual a 1, es decir, p = 1. Para la distribución Weibull, se tiene que la función de supervivencia está dada por,

$$S(t) = 1 - F(t) = 1 - \int_0^t f(u)du = 1 - \int_0^t \lambda p u^{p-1} \exp(-\lambda u^p)du = 1 - \left[-\exp(-\lambda u^p)\right]_0^t,$$

es decir,

$$S(t) = 1 - \left(-\exp(-\lambda t^p) + 1\right) = \exp(-\lambda t^p).$$

y la función de riesgo resulta

$$h(t) = -\frac{d}{dt}[\ln(S(t))] = -\frac{d}{dt}\left[\ln\left(\exp(-\lambda t^p)\right)\right] = -\frac{d}{dt}\left(-\lambda t^p\right) = \lambda p t^{p-1}.$$

Esta función es monótona creciente si $p > 1$, es decir, el riesgo de que una persona muera aumenta conforme trascurre el tiempo. Por otra parte, es monótona decreciente si $p < 1$, de manera que el riesgo de que una persona muera va disminuyendo conforme aumenta el tiempo. Por último, la función de riesgo es constante si $p = 1$, esto es, el riesgo de que una persona muera es el mismo para todo tiempo.

4.3.2 Modelos de regresión continuos

Hasta el momento, en el análisis de supervivencia se ha considerado que S(t) y, por tanto, h(t) involucran únicamente los tiempos de vida; sin embargo, en la práctica, es importante considerar la relación existente entre los tiempos de vida y algunas variables explicativas. El análisis correspondiente se lleva a cabo por medio de los modelos de regresión.

Por ejemplo, al analizar el tiempo de vida de pacientes con cáncer en el pulmón, en el estudio pueden ser considerados, además, otros factores como su edad y el tipo de tratamiento al que están sujetos. En el caso que se trata en este trabajo, la probabilidad de que una compañía aseguradora sea insolvente puede cambiar con el

tiempo pero además puede verse afectada por otras variables como, por ejemplo, algunos índices de la situación financiera y contable de la compañía.

De esta manera, se tiene que la distribución del tiempo de vida puede depender de un conjunto de variables explicativas, de manera que la tasa de riesgo cuando se observa el vector de q variables explicativas $\underline{X}' = (X_1, X_2, \ldots, X_q)$ es de la forma

$$h(t \mid x) = W(t, x),$$

que depende tanto de t como de $\underline{x}$.

Ahora, el proceso de modelado requiere la especificación de una estructura para $W(t, \underline{x})$. Una posibilidad es la siguiente:

$$h(t \mid \underline{x}) = h_0(t)G(x, \beta),$$

donde, $h_0(t)$ es una función de riesgo que se denomina función de riesgo basal.

$\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_q)'$ es un vector de coeficientes fijo pero desconocido, y

$G(\cdot)$ es una función de las variables explicativas.

La función $G(\cdot)$ debe tomar valores no negativos debido a que la función de riesgo es no negativa. Una posibilidad, propuesta por Cox (1972), es $G(\underline{x}'\underline{\beta}) = \exp(\underline{x}'\underline{\beta})$ que tiene la ventaja de no imponer restricciones sobre el vector de parámetros $\underline{\beta}$. Así, se tendría

$$h(t \mid \underline{x}) = h_0(t)\exp(\underline{x}'\underline{\beta}),$$

en donde el efecto de las variables explicativas se manifiesta multiplicando el riego basal por una constante (en el sentido de que no depende de t), para obtener un modelo que considere otros factores en la distribución del tiempo de vida.

En el presente trabajo de tesis, se considerará la función exponencial (es decir, $G(\underline{x}'\underline{\beta}) = \exp(\underline{x}'\underline{\beta})$) para incluir las variables explicativas a la función de riesgos.

4.3.3 Modelos de riesgos proporcionales continuos

Una característica de los modelos de regresión en los que las variables explicativas tienen un efecto multiplicativo, es que el cociente de las funciones de riesgo de dos individuos con diferentes vectores de variables explicativas $\underline{x}_1$ y $\underline{x}_2$ (es constante con respecto al tiempo), es decir, diferentes individuos tienen funciones de *riesgos proporcionales* para todo t.

$$\frac{h(t \mid x_1)}{h(t \mid x_2)} = \frac{h_0(t)G(x_1,\beta)}{h_0(t)G(x_2,\beta)} = \frac{G(x_1,\beta)}{G(x_2,\beta)},$$

Por otra parte, debido a que

$$S(t \mid \underline{x}) = \exp\left[-\int_0^t h(u \mid x)du\right],$$

se tiene que

$$S(t \mid \underline{x}) = \exp\left[-\int_0^t h_0(u)\exp(\underline{x}'\underline{\beta})du\right],$$

es decir,

$$S(t \mid \underline{x}) = \left[\exp\left\{-\int_0^t h_0(u)du\right\}\right]^{\exp(\underline{x}'\underline{\beta})} = [S_0(t)]^{\exp(\underline{x}'\underline{\beta})},$$

donde, $S_0(t)$ se conoce como la función de supervivencia basal y corresponde al caso en que $\underline{x}'\underline{\beta} = 0$.

4.3.4 Modelos de riesgos proporcionales paramétricos para datos continuos

Al considerar un conjunto de datos en algunos de los cuales se presenta censura por la derecha del *tipo I*, se tiene la función de verosimilitud de la forma (4.2), es decir

$$L = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} ,$$

si además, se considera un modelo de riesgos proporcionales paramétrico con distribución Weilbull (λ,p) para los tiempos de vida, se tiene que la función de riesgo está dada por

$$h(t \mid x_i) = \left(\lambda p t^{p-1}\right) \exp(x_i' \beta),$$

por lo que la función de supervivencia, dadas las variables explicativas, es de la forma

$$S(t \mid \underline{x}_i) = \exp\left[- \int_0^t \lambda p u^{p-1} du \right]^{\exp(\underline{x}_i' \underline{\beta})} = \exp\left[- \lambda t^p \exp(x_i' \beta) \right] \qquad (4.3)$$

y, debido a que

$$f(t \mid \underline{x}_i) = \frac{d}{dt} F(t \mid \underline{x}_i) = \frac{d}{dt} \left[1 - S(t \mid \underline{x}_i)\right] = - \frac{d}{dt} S(t \mid \underline{x}_i) = h(t \mid \underline{x}_i) S(t \mid \underline{x}_i)$$

se tiene que,

$$f(t \mid \underline{x}_i) = \left[\lambda p t^{p-1} \exp(\underline{x}_i' \underline{\beta})\right] * \left[\exp\left(- \lambda t^p \exp(\underline{x}_i' \underline{\beta})\right)\right] \qquad (4.4)$$

Entonces, al sustituir las ecuaciones (4.3) y (4.4) en la función de verosimilitud se tiene que

$$L = \prod_{i=1}^{n} \left[\left(\lambda p t_i^{p-1} \exp(x_i'\beta)\right)^{\delta_i} \left(\exp\left(-\lambda t_i^p \exp(x_i'\beta)\right)\right)^{\delta_i}\right] \left[\exp\left(-\lambda t_i^p \exp(x_i'\beta)\right)\right]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left[\lambda p t_i^{p-1} \exp(x_i'\beta)\right]^{\delta_i} \left[\exp\left(-\lambda t_i^p \exp(x_i'\beta)\right)\right].$$

Los estimadores de λ, p y $\underline{\beta}$ se obtienen al maximizar la función de verosimilitud o equivalentemente al maximizar la función de logverosimilitud, que resulta:

$$\lambda = \ln L = \sum_{i=1}^{n} \delta_i \ln\left[\lambda p t_i^{p-1} \exp(x_i'\beta)\right] - \sum_{i=1}^{n} \lambda t_i^p \exp(x_i'\beta)$$

$$= \sum_{i=1}^{n} \delta_i \ln(\lambda p t_i^{p-1}) + \sum_{i=1}^{n} \delta_i (x_i'\beta) - \sum_{i=1}^{n} \lambda t_i^p \exp(x_i'\beta).$$

Al obtener los estimadores máximo verosímiles para λ, p y $\underline{\beta}$, se puede obtener el estimador máximo verosímil de cualquier función de estos parámetros, en particular se puede estimar $S(t|\underline{x})$ o $h(t|\underline{x})$.

4.3.5 Modelos de riesgos proporcionales libres de distribución para datos continuos

Hasta el momento, en el análisis de supervivencia, se ha hablado de los modelos de riesgos proporcionales paramétricos; sin embargo, en ocasiones no se conoce con certeza la distribución de los tiempos de vida, por lo que Cox (1972) introdujo los

llamados modelos de riesgos proporcionales libres de distribución, o también llamados semiparamétricos.

Considere la variable aleatoria T continua que representa el tiempo de vida. Entonces, bajo el supuesto de riesgos proporcionales y $G(\cdot)$ exponencial, se tiene que la función de riesgos es

$$h(t \mid x) = h_0(t)\exp(x'\beta).$$

En el caso de los modelos paramétricos la forma de $h_0(t)$ se considera conocida, mientras que en los modelos libres de distribución, no se supone una forma específica para $h_0(t)$.

Para trabajar con este segundo tipo de modelos, Cox propone estimar $\underline{\beta}$ a partir de una función de verosimilitud que no depende de $h_0(t)$. Habiendo estimado $\underline{\beta}$, se procede a estimar $h_0(t)$ ó $S_0(t)$ utilizando métodos no paramétricos.

Suponga se tiene una muestra aleatoria de n individuos, con k distintos tiempos de vida observados, esto es una muestra en la que se observaron k muertes, por lo que n-k individuos muestran tiempos de vida censurados. Sean $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$, los diferentes tiempos de vida y sea $R_i = R(t_{(i)})$ el conjunto de individuos vivos justo antes del tiempo $t_{(i)}$ que no han sido censurados. Entonces dado un conjunto R(t) y dado que sólo ocurre una muerte en el tiempo t, se tiene que la probabilidad $P_i(t \mid \underline{x}_i)$de que el individuo i, con $i \in R(t)$, sea el que muere al tiempo t, dado su vector de variables explicativas, se puede aproximar como sigue

$$P_i(t \mid \underline{x}_i) \approx \Delta t * h(t \mid \underline{x}_i) \approx \frac{1}{\sum_{l \in R(t)} h(t \mid \underline{x}_l)} * h(t \mid \underline{x}_i) = \frac{h_0(t)\exp(\underline{x}_i'\underline{\beta})}{\sum_{l \in R(t)} h_0(t)\exp(\underline{x}_l'\underline{\beta})} = \frac{\exp(\underline{x}_i'\underline{\beta})}{\sum_{l \in R(t)} \exp(\underline{x}_l'\underline{\beta})}$$

donde $\underline{x}_i$ es el vector de variables explicativas asociado al individuo i.

La aproximación se sigue del hecho de que al tiempo t hubo una muerte, por lo que se tiene que $\sum_{l \in R(t)} P_l(t \mid x_l) = 1$, lo que implica que $\sum_{l \in R(t)} \Delta t \cdot h(t \mid x_l) \approx 1$, es decir,

$$\Delta t \approx \frac{1}{\sum_{l \in R(t)} h(t \mid x_l)}.$$

Cox (1972) define entonces una función de verosimilitud para $\underline{\beta}$ aproximada, como:

$$L(\beta) = \prod_{i=1}^{k} \frac{\exp(x_{(i)}'\beta)}{\sum_{l \in R(t)} \exp(x_l'\beta)},$$

donde $x_{(i)}$ es el vector de variables explicativas asociado al individuo que se le observó morir en $t_{(i)}$.

Cox llamó a $L(\underline{\beta})$ función de verosimilitud; sin embargo, ésta no puede ser derivada de la probabilidad de un evento bajo un modelo establecido, como es el caso de los modelos paramétricos; por lo que la validez de esta verosimilitud ha sido analizada por varios autores quienes han llegado a la conclusión de que $L(\underline{\beta})$ puede ser tratada como una función de verosimilitud y que bajo ciertas condiciones (por ejemplo un número reducido de variables explicativas) se tiene que al maximizar esta función, se obtiene un estimador para $\underline{\beta}$ con una distribución asintóticamente normal.

Cabe mencionar que, cuando no existe censura en los datos o se tiene censura por la derecha del *tipo II*, se puede demostrar, que $L(\underline{\beta})$, en efecto, es una función de verosimilitud para $\underline{\beta}$ (ver Lawless, 1976).

Por otro lado, en la práctica frecuentemente se presentan empates entre los tiempos de vida observados, esto puede deberse a que los datos se encuentran agrupados. En caso de que el número de empates sea reducido, es posible trabajar con modelos continuos, pero se debe modificar la función de verosimilitud $L(\underline{\beta})$.

Breslow, en 1974, propone la siguiente función de verosimilitud para $\underline{\beta}$ que modifica la sugerida por Cox,

$$L(\beta) = \prod_{i=1}^{k} \left(\frac{\exp(s_i'\beta)}{\sum_{l \in R(t)} \exp(x_l'\beta)} \right)^{d_i},$$

donde, d_i es el número de muertes que ocurrieron al tiempo $t_{(i)}$, y $\underline{s}_i$ es un vector de q variables explicativas, obtenido como la suma de los valores de las variables explicativas de los individuos que murieron al tiempo $t_{(i)}$.

En la actualidad, este método para los modelos de riesgos proporcionales es utilizado en la mayoría de los estudios con datos de tiempos de vida, ya que no es necesario especificar la distribución de los tiempos. En caso de que los datos realmente provengan de un modelo paramétrico existirá una ligera pérdida de información; sin embargo, se ha visto que se pueden obtener resultados eficientes cuando realmente se trata de un modelo de riesgos proporcionales y el número de variables explicativas es pequeño.

4.3.6 Estimación no paramétrica de $S_0(t)$

Al trabajar con un modelo libre de distribución, Cox propuso una función de verosimilitud aproximada para estimar el vector de coeficientes $\underline{\beta}$; sin embargo,

para estimar $S(t|\underline{x})=[S_0(t)]^{\exp(x'\beta)}$ y realizar inferencias sobre el tiempo de vida de la población bajo estudio, es necesario estimar además $S_0(t)$, contando con el estimador para $\underline{\beta}$. Para tal fin, Kalbfleisch y Prentice, en 1973, propusieron el siguiente método no paramétrico.

Suponga una muestra de k distintos tiempos de vida observados $t_{(1)}, \ldots, t_{(k)}$, tales que $t_{(1)}<\ldots<t_{(k)}$, y que $t_{(0)} = 0$ y $t_{(k+1)} = \infty$. Considere además, que en el tiempo $t_{(i)}$ mueren d_i individuos y que en el intervalo $[t_{(j-1)}, t_{(j)})$ hay λ_i tiempos censurados denotados por $A_j^{(i)}$ ($j = 1,\ldots,\lambda_i$). Entonces, al suponer la posibilidad de censura en algunos de los individuos y debido a que $S(t_i \mid x) - S(t_i + 0 \mid x)$ es la probabilidad de que un individuo muera al tiempo $t_{(i)}$ dado su vector de variables explicativas, donde $S(t_i+0|\underline{x}) = \lim_{y\to 0^+} S(t_i + y \mid x)$, se tiene que la función de verosimilitud es

$$L = \prod_{i=1}^{k}\left[\left(\prod_{j=1}^{\lambda_{i+1}} S\left(A_j^{(i)} \mid x_j^{(i)}\right)\right) * \left(S(t_{(i)} \mid x_{(i)}) - S(t_{(i)} + 0 \mid x_{(i)})\right)^{d_i}\right] \prod_{j=1}^{\lambda_{k+1}} S(A_j^{(k+1)} \mid x_j^{(k+1)}) \quad (4.5)$$

donde, $\underline{x}_j^{(i)}$ es el vector de variables explicativas asociado al individuo censurado en el tiempo $A_j^{(i)}$.

Por consiguiente, el problema se puede resolver al maximizar esta función de verosimilitud con respecto a $S_0(t)$, suponiendo que se conoce $\underline{\beta}$. El resultado es un estimador $\hat{S}_0(t)$, discontinuo en los tiempos $t_{(i)}$ ($i = 1,\ldots,k$) y con valores constantes en los puntos restantes. Por lo que:

$$\hat{S}_0(t_{(1)}) = \hat{S}_0(A_j^{(1)}) = 1, \quad j = 1, \ldots, \lambda_1$$

$$\hat{S}_0(t_{(i)} + 0) = \hat{S}_0(t_{(i+1)}) = \hat{S}_0(A_j^{(i+1)}), \quad i = 1, \ldots, k; \text{ y } j = 1, \ldots, \lambda_{i+1}$$

es decir, se tiene que $\hat{S}_0(t)$ es una función continua por la izquierda.

Con la finalidad de simplificar la expresión para la función de verosimilitud L en la ecuación (4.5), suponga que D_i es el conjunto de los d_i individuos que murieron en $t_{(i)}$ y C_i el conjunto de los λ_i individuos censurados en el intervalo $[t_{(j-1)}, t_{(j)})$. Además, considere $S_0(t_{(i)}+0) = P_i$ $(i = 1,\ldots,k)$ con $P_0 = 1$, entonces la función de verosimilitud se puede escribir como

$$L = \prod_{i=1}^{k}\left[\prod_{l\in D_i}\left(P_{i-1}^{\exp(x_l\beta)} - P_i^{\exp(x_l\beta)}\right)\prod_{l\in C_i}P_{i-1}^{\exp(x_l\beta)}\right] * \prod_{l\in C_{k+1}}P_k^{\exp(x_l\beta)},$$

Si ahora se define $\alpha_i = \dfrac{P_i}{P_{i-1}}$ $(i = 1,\ldots,k)$ y $\alpha_0 = 1$, entonces la función de verosimilitud se puede expresar como

$$L = \prod_{i=1}^{k}\left[\prod_{l\in D_i}\left(1-\alpha_i^{\exp(x_l\beta)}\right)\right] * \prod_{i=1}^{k+1}\left[\prod_{l\in D_i\cup C_i}(\alpha_1\cdots\alpha_i)^{\exp(x_l\beta)}\right] \qquad (4.6)$$

El estimador máximo verosímil de $S_0(t)$ se obtiene al igualar a cero las derivadas parciales con respecto a α_i $(i = 1,\ldots,k)$ del logaritmo de la función de verosimilitud dada por la ecuación (4.6). Procediendo de esta forma se obtiene,

$$\hat{S}_0(t) = \hat{P}_i = \hat{\alpha}_1\cdots\hat{\alpha}_i$$

esto es,

$$\hat{S}_0(t) = \prod_{i:t_{(i)}<t}\hat{\alpha}_i,$$

donde

$$\hat{\alpha}_i = \exp\left(\frac{-d_i}{\sum_{l \in R_i} \exp(x_l \beta)} \right),$$

con R_i el conjunto de individuos vivos no censurados antes de $t_{(i)}$.

4.4 Modelos discretos

En algunos casos se tiene que T es una variable aleatoria discreta, ya sea porque los tiempos de vida se agrupan o porque el proceso sólo se observa en forma discreta.

Entonces, la función de probabilidad de la variable aleatoria T que toma los valores $t_1, t_2, \ldots$, tales que $0 \leq t_1 < t_2 < \ldots$, se define como

$$p(t_i) = P(T = t_i), \text{ para } i = 1,2,\ldots$$

por lo que la función de supervivencia $S(t) = P(T \geq t)$ es monótona decreciente continua por la izquierda, y está dada por

$$S(t) = \sum_{i:t_i \geq t} p(t_i),$$

tal que $S(0) = 1$ y cuando t tiende a infinito $S(t)$ tiende a cero.

Para este caso, en donde se tiene que la variable aleatoria T es discreta, la función de riesgo se define como

$$h(t_i) = P(T = t_i \mid T \geq t_i), \quad \text{con } i = 1,2,\ldots$$

es decir,

$$h(t_i) = \frac{P(T = t_i)}{P(T \geq t_i)} = \frac{p(t_i)}{S(t_i)}, \quad \text{con } i = 1,2,\ldots$$

pero como $p(t_j) = S(t_j) - S(t_{j+1})$, se tiene que

$$h(t_i) = \frac{S(t_i) - S(t_{i+1})}{S(t_i)} = 1 - \frac{S(t_{i+1})}{S(t_i)},$$

por lo que la función de supervivencia S(t) puede expresarse como función de la tasa de riesgo h(t), de la siguiente manera

$$S(t) = \prod_{i:t_i < t} \frac{S(t_{i+1})}{S(t_i)} = \prod_{i:t_i < t} [1 - h(t_i)].$$

4.4.1 Modelos de regresión discretos

Los modelos discretos generalmente se derivan de datos continuos que han sido agrupados, más adelante se presentaran dos modelos libres de distribución para trabajar con datos agrupados o discretos, estos son, el modelo de riesgos proporcionales agrupado, y el modelo logístico.

Para ello, considere una muestra aleatoria de n individuos con tiempos de vida t_1, $t_2,\ldots,t_n$ y vectores de variables explicativas $\underline{x}_i$ con $i = 1,2,\ldots,n$ ($\underline{x}_i = (x_{i1},\ldots,x_{iq})'$). Considere que el eje correspondiente al tiempo está dividido en k+1 intervalos $I_j = [a_{j-1}, a_j)$, para $j = 1,\ldots,k+1$, donde $a_0 = 0$, $a_k = T$ y $a_{k+1} = \infty$. Suponga, además,

que lo único que se conoce es el intervalo en que murió o fue censurado el individuo.

Con el fin de construir la función de verosimilitud para los modelos de regresión libres de distribución basados en datos agrupados, es conveniente considerar las siguientes definiciones: Sea $P_i(\underline{x})$ la probabilidad de que un individuo sobreviva un tiempo mayor que a_i, dado su vector de variables explicativas; es decir

$$P_i(\underline{x}) = P(T>a_i \mid \underline{x}) = S(a_i \mid \underline{x}), \quad \text{para } i = 1,\ldots,k, \quad \text{con } P_0(\underline{x}) = 1$$

y sea $p_i(\underline{x})$ la probabilidad de que un individuo sobreviva un tiempo mayor que a_i, dado que sobrevivió un tiempo mayor que a_{i-1}, (y dado su vector de variables explicativas); es decir

$$p_i(x) = P(T > a_i \mid T > a_{i-1}, x)$$

$$= \frac{P(\{T > a_i \mid x\} \cap \{T > a_{i-1} \mid x\})}{P(T > a_{i-1} \mid \underline{x})}$$

$$= \frac{P(T > a_i \mid \underline{x})}{P(T > a_{i-1} \mid \underline{x})}$$

$$= \frac{P_i(\underline{x})}{P_{i-1}(\underline{x})}.$$

Además, se adopta el supuesto de que todos los casos de censura ocurren al final de cada intervalo; es decir, cuando ya han ocurrido todas las muertes del intervalo correspondiente. Entonces, se puede probar que la función de verosimilitud para modelos de regresión con datos agrupados es de la forma (ver Lawless, 1982).

$$L = \prod_{i=1}^{k}\left(\prod_{l\in D_i}\left[1-p_i(x_l)\right] \prod_{l\in R_i - D_i} p_i(x_l)\right) \qquad (4.7)$$

donde, R_i es el conjunto de individuos vivos y no censurados al tiempo a_{i-1},

D_i es el conjunto de individuos a los que se les observó morir en I_i, y

C_i es el conjunto de individuos censurados en I_i.

4.5 Modelo de riesgos proporcionales para datos agrupados

Cuando los tiempos de vida continuos de un modelo de riesgos proporcionales son agrupados, se obtiene un modelo de regresión para datos agrupados en el cual se tiene que

$$P_i(\underline{x}) = S(a_i|\underline{x}) = [S_0(a_i)]^{\exp(\underline{x}'\underline{\beta})},$$

si se adopta la notación $P_i = P_i(\underline{0})$, entonces

$$P_i(\underline{x}) = P_i^{\exp(\underline{x}'\underline{\beta})}, \qquad \text{para } i = 1,\ldots,k.$$

Por lo tanto,

$$p_i(\underline{x}) = \frac{P_i(\underline{x})}{P_{i-1}(\underline{x})} = \frac{P_i^{\exp(\underline{x}'\underline{\beta})}}{P_{i-1}^{\exp(\underline{x}'\underline{\beta})}} = \left(\frac{P_i}{P_{i-1}}\right)^{\exp(\underline{x}'\underline{\beta})},$$

entonces, si $p_i = p_i(0)$, que se refiere a la probabilidad de que un individuo sobreviva un tiempo mayor que a_i, dado que sobrevivió un tiempo mayor a a_{i-1}, (y dado que su vector de variables explicativas es cero), es decir

$$p_i = \frac{P_i}{P_{i-1}}, \qquad (4.8)$$

se tiene que

$$p_i(x) = p_i^{\exp(x'\beta)}, \qquad \text{para } i = 1,\ldots,k.$$

Por lo que al sustituir $p_i(x) = p_i^{\exp(x'\beta)}$ en la ecuación (4.7), se tiene que la función de verosimilitud para el vector de parámetros $\beta = (\beta_1,\ldots,\beta_q)'$ y el vector de probabilidades condicionales $p = (p_1,\ldots,p_k)'$, está dada por

$$L(\beta,p) = \prod_{i=1}^{k}\left(\prod_{l\in D_i}\left(1-p_i^{\exp(x_l'\beta)}\right)\prod_{l\in R_i-D_i}p_i^{\exp(x_l'\beta)}\right),$$

y la función logverosimilitud resulta

$$\ln\left(L(\underline{\beta},\underline{p})\right) = \sum_{i=1}^{k}\left(\sum_{l\in D_i}\ln\left(\frac{1-p_i^{\exp(\underline{x}_l'\underline{\beta})}}{p_i^{\exp(\underline{x}_l'\underline{\beta})}}\right)+\sum_{l\in R_i}\ln\left(p_i^{\exp(\underline{x}_l'\underline{\beta})}\right)\right)$$

Dado que p_i para $i = 1,\ldots,k$ son probabilidades, se tiene que el estimador máximo verosímil para p_i debe encontrarse dentro del intervalo $[0,1]$. Para considerar esta restricción dentro de la función de logverosimilitud Prentice y Gloeckler (1978) sugirieron utilizar la reparametrización $\gamma_i = \log(-\log(p_i))$, por lo que $p_i = \exp(-\exp(\gamma_i))$, y de esta forma se asegura que las p_i's se encuentran entre cero y uno. Entonces, considerando esta restricción, la función logverosimilitud en términos de γ_i; $i = 1,\ldots,k$ está dada por

$$\ln(L(\beta,\gamma)) = \sum_{i=1}^{k}\left(\sum_{l\in D_i} \ln\left[\exp(\exp(\gamma_i + x_l'\beta)) - 1\right] - \sum_{l\in R_i} \exp(\gamma_i + x_l'\beta)\right).$$

Para obtener ahora los estimadores máximo verosímiles para β y γ, se calculan las derivadas, parciales con respecto a β_r y γ_i, de la función de logverosimilitud, que resultan;

$$\frac{\partial \ln(L(\beta,\gamma))}{\partial \beta_r} = \sum_{i=1}^{k}\left[\sum_{l\in D_i} \frac{x_{lr} z_{il}}{1-\exp(-z_{il})} - \sum_{l\in R_i} x_{lr} z_{il}\right], \qquad \text{para } r = 1,\ldots,q$$

$$\frac{\partial \ln(L(\beta,\gamma))}{\partial \gamma_i} = \sum_{l\in D_i} \frac{z_{il}}{1-\exp(-z_{il})} - \sum_{l\in R_i} z_{il}, \qquad \text{para } i = 1,\ldots,k,$$

donde $z_{il} = \exp(\gamma_i + x_l'\beta)$ con $i = 1,\ldots,k$; y $l = 1,\ldots,n$.

Entonces, al igualar las derivadas parciales a cero y utilizar el método iterativo de Newton-Raphson es posible encontrar los estimadores máximo verosímiles para β y $\underline{\gamma}$. Además, debido a la propiedad de invarianza de los estimadores máximo verosímiles, se tiene que $\hat{p}_i = \exp(-\exp(\hat{\gamma}_i))$.

Finalmente, y dado que de la expresión (4,8) se tiene que $P_i = P_{i-1}p_i$; esto es, $P_i = \prod_{h=1}^{i} p_h$, entonces, el estimador máximo verosímil de la probabilidad de que un individuo sobreviva mas de a_i ($i = 1, \ldots, k$), dado su vector de variables explicativas resulta

$$\hat{P}_i(x) = \hat{P}_i^{\exp(x'\beta)} = \left[\prod_{h=1}^{i} \hat{p}_h\right]^{\exp(x'\beta)} .$$

4.6 Modelo logístico para datos agrupados

Una variante del modelo logístico es otra posibilidad para analizar tiempos de vida agrupados. A diferencia del modelo anterior, no se obtiene de agrupar datos continuos en un modelo de riesgos proporcionales.

En este modelo, la probabilidad condicional de que un individuo sobreviva mas allá de a_i, dado que sobrevivió mas de a_{i-1} y dado su vector de variables explicativas; es decir $p_i(\underline{x})$, está dada por

$$p_i(x) = \frac{1}{(1+\eta_i \exp(x'\beta))}, \qquad \text{para } i = 1,\ldots,k$$

donde $\eta_i = \dfrac{1-p_i}{p_i}$, con p_i definido en la ecuación (4.8)

En este modelo se puede establecer la relación

$$\ln\left(\frac{1-p_i(\underline{x})}{p_i(\underline{x})}\right) = \ln(\eta_i) + \underline{x}'\underline{\beta} .$$

La función de verosimilitud se obtiene al sustituir $p_i(\underline{x})$ en la ecuación (4.6), esto es

$$L(\beta,\eta)=\prod_{i=1}^{k}\left[\prod_{l\in D_i}\left(1-(1+\eta_i\exp(\underline{x}_l'\beta))^{-1}\right)\prod_{l\in R_i-D_i}(1+\eta_i\exp(x_l'\beta))^{-1}\right],$$

o, equivalentemente,

$$L(\beta,\eta)=\prod_{i=1}^{k}\left\{\frac{\left[\prod_{l\in D_i}\left(1-(1+\eta_i\exp(x_l'\beta))^{-1}\right)\prod_{l\in R_i}(1+\eta_i\exp(x_l'\beta))^{-1}\right]}{\prod_{l\in D_i}(1+\eta_i\exp(x_l'\beta))^{-1}}\right\},$$

por lo que la función de verosimilitud del modelo logístico para $\beta=(\beta_1,\ldots,\beta_q)$ y $\eta=(\eta_1,\ldots,\eta_k)'$ es

$$L(\beta,\eta)=\prod_{i=1}^{k}\left[\prod_{l\in D_i}\eta_i\exp(x_l'\beta)\prod_{l\in R_i}(1+\eta_i\exp(x_l'\beta))^{-1}\right].$$

Si se define $\alpha_i=\log(\eta_i)$ para i = 1,...,k, se tiene que la función de log-verosimilitud está dada por

$$\ln\left(L(\underline{\beta},\underline{\alpha})\right)=\sum_{i=1}^{k}\left[\sum_{l\in D_i}\left(\alpha_i+\underline{x}_l'\underline{\beta}\right)-\sum_{l\in R_i}\ln(1+\exp(\alpha_i+\underline{x}_l'\underline{\beta}))\right].$$

Por lo que las correspondiente derivadas parciales con respecto a β_r y α_i son

$$\frac{\partial\ln\left(L(\underline{\beta},\underline{\alpha})\right)}{\partial\beta_r}=\sum_{i=1}^{k}\left[\sum_{l\in D_i}x_{lr}-\sum_{l\in R_i}\frac{x_{lr}\exp(\alpha_i+\underline{x}_l'\underline{\beta})}{1+\exp(\alpha_i+\underline{x}_l'\underline{\beta})}\right],\quad \text{para } r=1,\ldots,q$$

$$\frac{\partial \ln(L(\beta,\alpha))}{\partial \alpha_i} = d_i - \sum_{l \in R_i} \frac{\exp(\alpha_i + x_l'\beta)}{1+\exp(\alpha_i + x_l'\beta)}, \qquad \text{para } i = 1,\ldots,k$$

Entonces, los estimadores máximo verosímiles $\hat{\beta}$ y $\hat{\alpha}$ para β y α, respectivamente, se obtienen al igualar las derivadas parciales a cero y resolver el sistemas de ecuaciones por medio del método iterativo Newton-Raphson. Finalmente, y en virtud de la propiedad de invarianza de los estimadores máximo verosímiles, se tiene que $\hat{\eta}_i = \exp(\hat{\alpha}_i)$ y como consecuencia

$$\hat{p}_i(x) = (1+\hat{\eta}_i \exp(x'\hat{\beta}))^{-1}, \text{ para } i = 1,\ldots,k$$

Entonces, el estimador de la probabilidad de que un individuo sobreviva mas allá de a_i dado su vector de variables explicativas es

$$\hat{P}_i(x) = \prod_{h=1}^{i} \hat{p}_h(x) = \prod_{h=1}^{i} \left(1+\hat{\eta}_h \exp(x'\hat{\beta})\right), \quad \text{para } i = 1,\ldots,k.$$

4.7 Pruebas estadísticas

En el análisis de supervivencia, para comparar la eficiencia de dos modelos de riesgos proporcionales, ajustados al mismo conjunto de datos, los cuales contienen diferente número de parámetros, se usa la prueba llamada razón de verosimilitudes (entre otras), la cual hace uso de la estadística $-2\ln\hat{L}$.

Es decir, considere dos modelos que están siendo contemplados para un particular conjunto de datos, esto es, modelo 1 y modelo 2. En el modelo 1, un grupo de g

variables explicativas fueron ajustadas, estas son $X_1, X_2, \ldots, X_g$. Por lo que la función de riesgos proporcionales bajo este modelo puede escribirse como:

$$h_0(t) * \exp\{\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_g x_g\}.$$

Por otro lado, en el modelo 2 se ajustaron q (q > g) variables explicativas, esto es $X_1, X_2, \ldots, X_g, X_{g+1}, \ldots, X_q$. Así, la función de riesgos proporcionales bajo el modelo 2 es:

$$h_0(t) * \exp\{\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_g x_g + \beta_{g+1} x_{g+1} + \ldots + \beta_q x_q\}$$

Es decir, el modelo 2 contiene q-g variables explicativas adicionales al modelo 1. Como consecuencia de que el modelo 2 tiene un mayor número de variables explicativas éste debe presentar un mejor ajuste al conjunto de datos observados. Sin embargo, el problema estadístico que se desea resolver es el de decidir si al agregar las q-g variables explicativas en el modelo 2, éste mejora significativamente el ajuste obtenido en el modelo 1, en caso contrario, las variables adicionales pueden ser omitidas del modelo final.

Para poder tomar una decisión al comparar ambos modelos ajustados, se hace uso de una estadística que mide el grado en que un modelo se ajusta a los datos. Como consecuencia, y dado que la función de verosimilitud contiene la información acerca de los parámetros desconocidos en un modelo, se utiliza como estadística, el valor de la función de verosimilitud al remplazar los parámetros por su estimador máximo verosímil. O equivalentemente, se puede usar el valor de la estadística $-2\ln\hat{L}$.

Ahora bien, el valor de $\hat{L}$ siempre será menor a uno, esto, debido a que $\hat{L}$ se obtiene como el producto de probabilidades condicionales; y como consecuencia el

valor $-2\ln\hat{L} > 0$. Entonces, para un conjunto de datos, se elegirá aquel modelo que presente el menor valor de $-2\ln\hat{L}$.

Considere que la función de verosimilitud del modelo i es L_i, con i = 1, 2. Así, al comparar ambos modelos, si la diferencia entre $-2\ln\hat{L}_1$ y $-2\ln\hat{L}_2$ es significativa, esto llevaría a la conclusión de que al agregar las q-g variables explicativas en el modelo 2, éste mejora significativamente el ajuste del modelo a los datos observados.

La estadística $-2\ln\hat{L}_1 + 2\ln\hat{L}_2$, puede expresarse como $-2\ln\left(\frac{\hat{L}_1}{\hat{L}_2}\right)$, que se conoce como la estadística de razón de verosimilitudes, utilizada para probar la hipótesis nula H_0 sobre los parámetros de la regresión en un modelo de riesgos proporcionales, ésta es $H_0 : \beta_{g+1} = \beta_{g+2} = \ldots = \beta_q = 0$.

La estadística $-2\ln\left(\frac{\hat{L}_1}{\hat{L}_2}\right)$ sigue una distribución aproximada a una χ^2 con q-g grados de libertad, bajo la hipótesis nula (ver Collett, 1994). Por lo que la regla de decisión de esta prueba, consiste en rechazar la hipótesis nula, a un nivel de significancia α, si el valor de la estadística de prueba es mayor a una $\chi^2_{(q-g),1-\alpha}$.

4.8 Aplicación al análisis de solvencia

En la literatura del análisis de supervivencia, recientemente se publicaron dos artículos en lo que se presenta una aplicación al análisis de solvencia, los autores de estos artículos son Lee y Urrutia (1996) y Kim et al (1995).

En el primer artículo, Lee y Urrutia (1996) comparan dos modelos para estimar la probabilidad de insolvencia de compañías aseguradoras, el modelo logístico, el cual es un modelo de respuesta cualitativa, y el modelo de riesgos proporcionales con función de distribución Weilbull.

Para comparar estos modelos Lee y Urrutia utilizaron algunas variables que explican la situación contable de las compañías aseguradoras. Estas variables se definen de acuerdo a la práctica contable y financiera en los Estados Unidos, los conceptos a que se refieren no necesariamente tienen un equivalente en otros países. Por esta razón, en el Apéndice C.4 se listan las variables con su definición en los términos originales tal como aparece en el artículo y se incluye además una colección aproximadamente equivalente producida a partir de la práctica contable en México.

Lee y Urrutia trabajaron con una muestra seleccionada de 82 aseguradores, que operan el ramo de daños-responsabilidad civil, declarados insolventes durante el periodo de 1980 a 1991; para cada uno de estos aseguradores insolventes se seleccionó un asegurador solvente para el mismo periodo y perteneciente al mismo estado, así como con un tamaño similar a cada una de las compañías declaradas insolventes, en lo que se refiere al portafolio de inversión.

Los datos de compañías insolventes utilizados en el modelo son los correspondientes a uno, dos o tres años anteriores a la insolvencia, dependiendo de la información contable disponible de las compañías.

Debido a como realizaron la selección de la muestra, existe un sobremuestreo de aseguradores insolventes que corrigieron para evitar resultados erróneos o inconsistentes como podría ser la obtención de estimadores sesgados de la probabilidad de insolvencia.

Utilizando un nivel de confiabilidad del 5% o menos, Lee y Urrutia identificaron las variables explicativas más significativas para cada uno de los modelos, resultando para el modelo logístico cuatro variables explicativas y para el modelo de riesgos proporcionales ocho variables explicativas en las que se encuentran contempladas las cuatro variables significativas del modelo logístico.

Después de obtener las variables significativas para cada modelo, los autores realizaron un pronóstico de la probabilidad de insolvencia, para esto utilizaron una validación cruzada, es decir utilizaron los 42 datos disponibles de los aseguradores solventes así como de los declarados insolventes, durante el periodo de tiempo de 1985 a 1988, para pronosticar la insolvencia o solvencia de las 32 compañías que fueron declaradas insolventes, así como de las 32 compañías solventes durante el periodo de 1989 a 1991, con lo cual obtuvieron porcentajes del número incorrecto de clasificaciones para cada modelo con los diferentes niveles de corte.

En general, Lee y Urrutia obtuvieron que al utilizar el modelo logístico existe un menor número de clasificaciones erróneas de compañías solventes que fueron declaradas como insolventes, o viceversa.

Por último, los autores calcularon el costo de clasificar a un asegurador como solvente cuando realmente es insolvente, para cada uno de los modelos, manteniendo igual a uno el costo de clasificar un asegurador como insolvente cuando realmente no lo era. Como resultado, estos costos resultaron ser mayores al estimar la probabilidad de insolvencia con el modelo logístico.

Por lo tanto, se puede concluir que al comparar el modelo logístico y el modelo de riesgos proporcionales, se observa que ninguno de los dos es notablemente mejor que el otro, sino que más bien se complementan; ya que al utilizar el modelo de riesgos proporcionales con función de distribución Weibull se obtiene un mayor número de variables estadísticamente significativas para explicar la tendencia a ser

insolvente de un asegurador, y un menor costo derivado de clasificaciones erróneas, mientras que al utilizar el modelo logístico se obtiene un mejor pronóstico de clasificación de una compañía como solvente o insolvente.

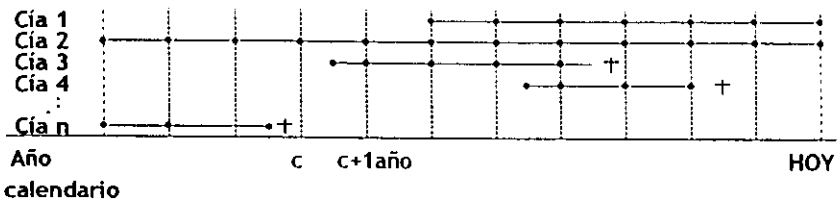
Sin embargo, por la importancia que últimamente representa el regular la situación financiera y contable de las compañías aseguradoras para así proporcionar seguridad a sus asegurados e inversionistas, se deberá analizar el costo - beneficio de utilizar el modelo logístico para obtener un mejor pronóstico de la probabilidad de insolvencia de una compañía aseguradora.

En el segundo artículo, Kim et al (1995) presentaron una aplicación del análisis de supervivencia en el análisis de solvencia al utilizar el modelo exponencial de la función o tasa de riesgo, para estimar la probabilidad de insolvencia de compañías que operan el seguro de daños-responsabilidad civil y el seguro de vida, analizando el tiempo que una compañía solvente permanece en ese estado y su relación con algunos factores que explican la situación contable y financiera de la compañía. Aquí, a diferencia de otros trabajos, se analiza la historia previa al evento, registrando el valor de algunos índices financieros o contables a través del tiempo (durante varios ejercicios). De esta forma, se analiza la influencia que estos pudieran tener sobre la probabilidad de insolvencia en el contexto histórico.

Una contribución interesante de los autores de este artículo, es la metodología utilizada para estimar la probabilidad de insolvencia, que incorpora el análisis de la información contable de las compañías durante todo el periodo observado y no solamente la del ejercicio (año) anterior a la insolvencia. Para ello, el tiempo en el que cada asegurador se mantiene en estado solvente es dividido en intervalos, cada uno de un año, y la información para cada compañía al inicio de cada año es considerada como una observación (Figura 4.1). Es decir, para el análisis se considera la información de una misma compañía para cada uno de sus diferentes

años de forma simultanea. Los autores proponen un modelo exponencial para la función o tasa de riesgo.

Figura 4.1



donde, t_i es el tiempo en el que la compañía i se mantuvo en estado solvente,

• indica la observación de la compañía i en el año calendario c,

+ indica cuando una compañía se vuelve insolvente.

Al estimar el vector de coeficientes β de la función de riesgo por máxima verosimilitud, se puede obtener un estimador de la probabilidad de que una compañía solvente pase a estado insolvente en menos de t años, esto debido a la relación existente entre la función de riesgo λ y la función de distribución acumulada F(t) de la variable aleatoria T que representa el tiempo en que una compañía solvente se mantuvo en ese estado, como se ha mencionado.

Además de estimar la probabilidad de insolvencia de una compañía aseguradora, es posible determinar, de un conjunto de compañías solventes al inicio de un periodo de observación, el número esperado de compañías insolventes al final del periodo, de acuerdo a sus variables explicativas.

Por último, en este artículo, para las compañías de daños-responsabilidad civil, Kim et al (1995) trabajaron con una muestra de 82 aseguradores insolventes, obtenidos

de *Best's Insurance Reports* y de los reportes especiales de NCIGF (National Conference of Insurance Guaranty Funds), de 1984 a 1990 considerando así un periodo de observación de 7 años. Asimismo, para el seguro de vida utilizaron una muestra de 42 aseguradores insolventes para un periodo de 4 años que va de 1987 a 1990. Para esos mismos periodos, existían adicionalmente 132 asegurados insolventes de daños-responsabilidad civil y 33 del seguro de vida para los cuales, sin embargo, la información contable no estaba disponible.

A las muestras compuestas por 82 y 42 compañías insolventes, para el seguro de daños-responsabilidad civil y el seguro de vida, respectivamente, se le agregó el mismo número de compañías solventes. Estas compañías solventes fueron seleccionadas de forma aleatoria, a diferencia de lo que ocurre en otros estudios. En resumen, se trabajó con una muestra compuesta de 878 observaciones compañía-año calendario para el seguro de daños-responsabilidad y 280 observaciones compañía-año calendario para el seguro de vida.

Los estimadores máximo verosímiles son generalmente inconsistentes y sesgados, cuando la muestra se compone de submuestras que se obtienen de estratos distintos de la población. Una forma de corregir este problema es ponderando los estimadores.

Para el estudio de la insolvencia en el seguro de daños-responsabilidad civil y vida, Kim et al (1995) analizaron un conjunto de diez variables explicativas diferentes para cada tipo de seguro. Estas variables se definen de acuerdo a la práctica contable y financiera en los Estados Unidos, los conceptos a que se refieren no necesariamente tienen un equivalente en otros países. Por esta razón, en el Apéndice C.5 se listan las variables con su definición en los términos originales tal como aparece en el artículo y se incluye además una colección aproximadamente equivalente producida a partir de la práctica contable en México.

La significancia de las variables explicativas se determinó a partir de los resultados del modelo. Las variables fueron inicialmente seleccionadas con base en investigaciones previas, los datos disponibles y otras consideraciones. En particular, la 'edad' de cada compañía -en cada año calendario- se incluyó con el propósito de comprobar sus posibles efectos significativos en la probabilidad de insolvencia.

En este artículo, los autores Kim et al (1995) presentan los modelos exponencial de la tasa de riesgo que obtuvieron, estos son:

Para el seguro de daños-responsabilidad civil, el modelo obtenido fue el siguiente:

$$\lambda = \exp[\beta_0 + \beta_1 AGE + \beta_2 CHNPW + \beta_3 YIELD + \beta_4 COMBR + \beta_5 EXP + \\ + \beta_6 RES + \beta_7 REI + \beta_8 RCG + \beta_9 UCG + \beta_{10} EXPANSION],$$

donde β_0 es un término constante (ordenada al origen).

En el seguro de vida, el modelo fue de la forma:

$$\lambda = \exp[\beta_0 + \beta_1 AGE + \beta_2 CHNPW + \beta_3 YIELD + \beta_4 RCG + \beta_5 UCG + \beta_6 NOM \\ + \beta_6 NOM + \beta_7 JUNK + \beta_8 CM + \beta_9 RE + \beta_{10} EXP],$$

donde, β_0 es una constante.

Como se mencionó, los autores estimaron el vector de coeficientes $\underline{\beta} = (\beta_0, \beta_1, \ldots, \beta_{10})'$ a través del método de máxima verosimilitud. Con este método se pueden considerar tanto las observaciones censuradas como las no censuradas entendiendo como observación censurada aquella en la que la compañía se mantiene solvente hasta el final del periodo de observación.

Debido a que se trata de un estudio observacional; es decir, en donde se trabaja con datos que no se producen bajo condiciones controladas, las variables explicativas pueden estar relacionadas entre sí, ocasionando de esta forma problemas de multicolinealidad; por ello, los autores ajustaron varios modelos en los que se omitieron algunas variables de acuerdo a un análisis previo de la matriz de correlación.

En este artículo, para el seguro de daños-responsabilidad civil, Kim et al (1995) ajustaron seis modelos exponenciales diferentes de la tasa de riesgo:

En el primer modelo se consideró la variable *loss reserve*, pero no se estudiaron las variables *realized capital gain, reinsurance* y *expansion to other estates*, debido a que las cuatro variables estaban altamente correlacionadas ya que todas están divididas entre el capital contable. El segundo, tercero y cuarto modelo se obtuvieron de las combinaciones de las variables mencionadas; ninguno de los primero cuatro modelos considera la variable *expense*. El quinto modelo considera todas las variables a excepción de la variable *expense*, para ser comparado con los cuatro modelos anteriores; y por último la sexta versión del modelo considera el efecto de todas las variables a excepción de la variable *underwriting results*.

Como resultado del análisis de los modelos se obtuvo que las variables estadísticamente significativas fueron *age, premium growth, investment yield, underwriting results* y *expense ratio*.

A diferencia de las variables utilizadas en el modelo para aseguradores de daños-responsabilidad civil, las variables utilizadas en el modelo exponencial para aseguradores de vida no están altamente correlacionadas a excepción de la variable *expense*, por ello los autores solamente ajustaron dos diferentes modelos, en el primero consideraron todas las variables incluyendo la variable *expense* y en el segundo además se considero el cuadrado de la variable *expense*.

Como resultado del análisis de los modelos, Kim et al (1995) obtuvieron que de las diez variables analizadas resultaron estadísticamente significativas, *age*, *investment yield*, *realized capital gains*, *unrealized capital gains*, *income performance*, *real estate holdings* y *expense ratio*.

CAPÍTULO 5
Análisis de datos reales

En los capítulos anteriores se han presentado algunas técnicas estadísticas que diferentes autores han aplicado en el análisis de solvencia. En los correspondientes artículos, estas técnicas son ejemplificadas utilizando la información contable y financiera de compañías aseguradoras que operan en los Estados Unidos.

En este capítulo se ilustrará la aplicación del análisis discriminante y del modelo logístico, en un análisis del mismo tipo, pero utilizando información del sector asegurador Mexicano. La técnica de análisis de supervivencia, que también se ha discutido en este trabajo, no se considera en este capítulo en virtud de que la información que finalmente se recolectó no posibilita una aplicación de ese procedimiento. Sin embargo, sería de interés revisar los resultados que obtendrían al utilizar esta técnica estadística junto con una base de datos más completa.

5.1 Consideraciones generales

Para llevar a cabo esta aplicación se recurrió a la información disponible de la Comisión Nacional de Seguros y Fianzas, de las 54 compañías de seguros que para marzo del 2000 se encontraban operando en México, y para las cuales se contara con información de ejercicios anteriores a esta fecha. Reproducir la clasificación de los grupos de compañías solventes e insolventes no es posible, debido a que no se cuenta con información de compañías insolventes. Sin embargo, es posible realizar

un análisis alternativo, en el que se pretende reproducir una clasificación de las compañías aseguradoras en dos grupos: compañías sanas en sentido financiero (Π_0) y compañías con potenciales problemas financieros (Π_1).

La división de las compañías aseguradoras en los dos grupos (Π_0 y Π_1) se realizó con base en información, de marzo del 2000, de tres indicadores regulatorios de la CNSF, que reflejan la situación contable y financiera de cada una de las 54 compañías aseguradoras al primer trimestre del 2000. Estos indicadores son: Cobertura de Reservas Técnicas (CRT), Margen de Capital de Garantía (MCG) y Capital Mínimo Pagado (CMP).

- ***Cobertura de Reservas Técnicas (CRT).***- El resultado de esta prueba permite saber si la Institución de Seguros, Pensiones y Fianzas, mantiene las inversiones suficientes para cubrir sus obligaciones contractuales, se determina con base en las Reglas vigentes, dividiendo el total de inversiones autorizadas entre el monto de la base de inversión y su resultado es de solvencia.

- ***Margen de Capital de Garantía (MCG).***- Este indicador mide la viabilidad financiera de la Institución para hacer frente a eventos extraordinarios no considerados en materia de desviaciones a la siniestralidad, se calcula determinando los requerimientos de capital para cada operación en base a Reglas. El indicador se obtendrá de la división del capital de garantía entre el requerimiento mencionado. Su resultado es de solvencia y muestra si la Institución cuenta con recursos suficientes para cubrir el requerimiento de capital determinado por sus operaciones.

- ***Capital Mínimo Pagado (CMP).***- Es el resultado de restar al requerimiento de capital por cada operación y/o ramo que se practique, el total del capital contable regulatorio. Este último en ningún momento podrá ser inferior al

capital mínimo pagado que fija la Secretaría de Hacienda y Crédito Público y representa el monto mínimo de recursos legales con que debe contar la Institución. Este indicador es de aplicación similar, tanto para el Sector Asegurador, Afianzador como para el de Pensiones, y el no tener cubierto dicho capital mínimo pagado es causal de revocación. Por lo que en caso de presentar un resultado deficitario, se procederá a efectuar el comunicado a la Secretaría de Hacienda y Crédito Público, quien iniciará el proceso conducente.

El procedimiento, arbitrario, que se siguió para determinar el grupo al que pertenece cada compañía aseguradora, fue el siguiente: toda compañía para la cual alguno de estos indicadores regulatorios era menor que uno, fue asignada al grupo Π_1, ya que esto significa que la compañía puede presentar algún problema financiero debido a que no cuenta con inversiones suficientes que respalden sus obligaciones pendientes por cumplir, y/o no cuenta con capital suficiente que respalde las posibles desviaciones en la siniestralidad, y/o su capital mínimo pagado está por debajo del que se requiere por ley. Por otro lado, todas las compañías para las cuales los tres indicadores regulatorios tomaron valores mayores o iguales a uno fueron asignadas al grupo Π_0; esto es, fueron consideradas como sanas en el sentido financiero.

Como resultado de esta división, 43 compañías forman el grupo Π_0 y el resto, esto es 11 compañías, el grupo Π_1. En el Apéndice D, se muestran los indicadores regulatorios publicados por la CNSF para marzo del 2000, tal como aparecen en www.cnsf.gob.mx, así como el grupo al que fueron asignadas las compañías aseguradoras. Es pertinente entonces insistir en que estos dos grupos describen la salud financiera de las empresas en marzo del 2000.

Así con el fin de probar una posible clasificación anticipada de las compañías aseguradoras en los dos grupos, se hizo uso de las técnicas estadísticas análisis

discriminante y modelo de regresión logística. Las variables explicativas utilizadas para estos modelos fueron algunos de los índices financieros y contables, mencionados a lo largo de este trabajo. Debido a que se desea realizar una clasificación de forma anticipada, los índices utilizados fueron calculados con información previa a marzo de 2000.

5.2 Bases de datos

Para la creación de las bases de datos se utilizó información reportada en los estados financieros de las 54 compañías aseguradoras, que para marzo del 2000 se encontraban operando.

Como resultado se cuenta, para cada empresa, con un conjunto de variables que reflejan su situación financiera en cada una de las siguientes fechas:

i) diciembre de 1999
ii) junio de 1999
iii) diciembre de 1998

El propósito es, entonces, utilizar la información de cada una de estas fechas para reconstruir los grupos que se definieron con los indicadores regulatorios en marzo del 2000.

Para aplicar las diferentes técnicas estadísticas que se han mencionado, se crearon tres bases de datos. Las variables explicativas utilizadas fueron 15 de los índices contables y financieros mencionados a lo largo de este trabajo (ver Apéndice E). Cada base de datos está formada por los mismos índices financieros, pero calculados con información (en términos reales a diciembre de 1998) del Balance

General y Estado de Resultados de las tres fechas distintas. Los índices utilizados son,

Variable	Descripción
PR_CC	Prima retenida entre capital contable
I_SI	Índice de costo de siniestralidad
IC	Índice combinado
CRP	Crecimiento real de prima retenida
L_CA	Logaritmo del crecimiento real de activos
R_CC	Reserva de obligaciones contractuales entre capital contable
R_PR	Reserva de obligaciones contractuales entre prima retenida
RE_F	Rentabilidad de financiera
RE_V	Rentabilidad de venta
X_10	Utilidad neta entre activos
RE_C	Rentabilidad de capital
MIC	Margen de ingresos sobre costos
MO	Margen de operación
CC	Capital contable
LARAT	Índice de responsabilidad

Estos índices reflejan la situación contable y financiera que presentaban las compañías aseguradoras en el cierre más cercano a marzo de 2000, a mediados del ejercicio previo al 2000 y en el cierre de ejercicio de dos años previos a marzo de 2000.

En diciembre y junio de 1999, la compañía con el número de observación 35 presentó valores indeterminados para las variables I_SI, IC, R_PR, RE_V, MIC Y MO.

Por otra parte, en diciembre de 1998, la compañía con el número de observación 51 presentó valores indeterminados en las variables CRP y L_CA, mientras que para este mismo conjunto de datos, se obtuvieron valores indeterminados en la variable L_CA para las compañías identificadas como 52, 53 y 54. Esto se debió a falta de información en los estados financieros. Como consecuencia, dichas compañías no fueron consideradas en el estudio del año correspondiente.

La compañía 35 pertenece al grupo Π_0; por ello, finalmente para diciembre de 1999 y junio de 1999 se trabajó con 42 compañías de Π_0 y 11 compañías de Π_1. Por otro lado, las compañías 51 y 54 pertenecen a Π_1 y las compañías 52 y 53 pertenecen a Π_0, de manera que para diciembre de 1998 se trabajó con 41 compañías de Π_0 y 9 compañías de Π_1.

5.3 Análisis preliminar

En principio, se realizó un análisis marginal para cada una de las 15 variables de diciembre de 1999, junio de 1999 y diciembre de 1998, con el fin de observar si se presentan diferencias significativas entre los dos grupos. Para ello, se hizo uso de la prueba no paramétrica de Mann-Whitney, para detectar diferencia entre dos poblaciones.

La prueba Mann-Whitney (ver Apéndice B.2) constituye un procedimiento simple para probar igualdad de poblaciones con muestras independientes y únicamente supone que las observaciones tienen escala al menos ordinal. En este sentido es más robusta que otras pruebas que requieren un mayor número de supuestos.

Al aplicar la prueba Mann-Whitney a cada una de las 15 variables de diciembre de 1999, de junio de 1999 y de diciembre de 1998, para probar la hipótesis H_0 de que

los grupos Π_0 y Π_1 son iguales, y considerando un nivel de significancia α=0.05 se tiene que:

Datos de diciembre de 1999

	T	n.s.d. *	Resultado
CC_D99	153	0.001325	Rechazo
RE_C_D99	155	0.001537	Rechazo
X10_D99	162	0.002550	Rechazo
RE_F_D99	200	0.027714	Rechazo
LARA_D99	391	0.057345	No rechazo
RE_V_D99	212	0.062303	No rechazo
MO_D99	225	0.114327	No rechazo
MIC_D99	228	0.130216	No rechazo
PR_C_D99	369	0.153233	No rechazo
R_PR_D99	255	0.356988	No rechazo
IC_D99	338	0.368553	No rechazo
I_SI_D99	325	0.539161	No rechazo
R_CC_D99	326	0.613762	No rechazo
CRP_D99	324	0.644256	No rechazo
L_CA_D99	289	0.771862	No rechazo

*n.s.d. = nivel de significancia descriptivo

Datos de junio de 1999

	T	n.s.d. *	Resultado
CC_J99	174	0.005786	Rechazo
RE_C_J99	183	0.010277	Rechazo
X10_J99	184	0.010931	Rechazo
MIC_J99	199	0.031617	Rechazo
IC_J99	387	0.048407	Rechazo
RE_V_J99	221	0.095562	No Rechazo
MO_J99	224	0.109383	No Rechazo
RE_F_J99	241	0.186561	No Rechazo
LARA_J99	363	0.193825	No Rechazo
L_CA_J99	361	0.208973	No Rechazo
R_CC_J99	337	0.458720	No Rechazo
CRP_J99	332	0.526360	No Rechazo
I_SI_J99	323	0.568529	No Rechazo
PR_C_J99	327	0.598758	No Rechazo
R_PR_J99	290	0.877988	No Rechazo

*n.s.d. = nivel de significancia descriptivo

Datos de diciembre de 1998

	T	n.s.d. *	Resultado
CC_D98	186	0.012351	Rechazo
X10_D98	221	0.080060	No rechazo
L_CA_D98	326	0.117401	No rechazo
RE_F_D98	370	0.147148	No rechazo
RE_C_D98	238	0.165976	No rechazo
RE_V_D98	248	0.241804	No rechazo
LARA_D98	342	0.396249	No rechazo
I_SI_D98	333	0.512436	No rechazo
CRP_D98	255	0.519631	No rechazo
MO_D98	276	0.569260	No rechazo
PR_C_D98	324	0.644256	No rechazo
IC_D98	324	0.644256	No rechazo
MIC_D98	284	0.691127	No rechazo
R_PR_D98	292	0.821583	No rechazo
R_CC_D98	307	0.923007	No rechazo

*n.s.d. = nivel de significancia descriptivo

De las tablas anteriores se observa que con la información de diciembre de 1999, se obtuvieron 4 variables que presentan diferencias significativas entre ambos grupos, mientras que para junio de 1999 se obtuvieron 5 variables significativas; sin embargo, al utilizar la información de diciembre de 1998, únicamente se obtuvo una variable significativa. Es decir, se tiene un mayor número de variables que

presentan diferencias significativas entre grupos al utilizar información a mediados del ejercicio anterior a marzo de 2000.

Por otro lado, de forma general se observa que conforme se utiliza información de una fecha más lejana a marzo de 2000, las diferencias marginales de las variables entre los grupos disminuyen.

La única variable que presenta diferencias significativas entre ambos grupos sin importar si la información que se utiliza procede de una fecha lejana a marzo de 2000 es CC (capital contable). Por oto lado las variables RE_C (rentabilidad de capital) y X_10 (utilidad neta entre activos) presentan diferencias significativas entre los dos grupos para diciembre de 1999 y junio de 1999.

5.3.1 Análisis de correlaciones

Adicionalmente, se calcularon los coeficientes de correlación que se presentaban entre las 15 variables explicativas de cada una de las tres bases de datos (ver tablas 5.1, 5.2 y 5.3).

Al analizar las correlaciones, se observa que los tres conjuntos de datos presentan la misma estructura en su matriz de correlaciones. Las variables que presentan problemas de correlación entre ellas son IC, RE_V, MIC y MO.

Tabla 5.1 Matriz de correlaciones (Diciembre 1999)

	PR_CC_D99	I_SI_D99	IC_D99	CRP_D99	L(CA)D99	R_PR_D99	R_CC_D99	RE_F_D99	RE_V_D99	X10_D99	RE_C_D99	MIC_D99	MO_D99	CC_D99	LARA_D99
PR_CC_D99	1														
CMS_D99	0.4214	1													
IC_D99	-0.1365	0.0965	1												
CRP_D99	0.0161	0.0036	0.0511	1											
L(CA)D99	0.2521	0.2277	-0.2341	0.3440	1										
R_PR_D99	-0.3084	-0.1046	-0.0801	-0.2562	-0.2436	1									
R_CC_D99	0.5493	0.2538	-0.1611	-0.1718	0.0019	0.0853	1								
RE_F_D99	-0.0032	0.0948	-0.3117	0.0120	0.1384	-0.2936	-0.0443	1							
RE_V_D99	0.1396	-0.0957	-0.9999	-0.0481	0.2352	0.0727	0.1616	0.3133	1						
X10_D99	0.1001	0.0590	-0.4617	-0.4011	0.3156	0.0300	0.1380	0.3424	0.4604	1					
RE_C_D99	-0.2736	-0.0096	-0.1147	-0.1483	0.2199	0.0318	-0.2462	0.4412	0.1126	0.7228	1				
MIC_D99	0.1357	-0.0970	-1.0000	-0.0507	0.2335	0.0790	0.1603	0.3137	0.9999	0.4617	0.1150	1			
MO_D99	0.1399	-0.0953	-0.9999	-0.0482	0.2354	0.0731	0.1620	0.3126	1.0000	0.4603	0.1124	0.9999	1		
CC_D99	0.0413	0.2479	-0.0628	-0.0838	0.0516	-0.1045	0.0824	0.1246	0.0643	0.2777	0.3102	0.0628	0.0642	1	
LARA_D99	0.1168	0.3760	-0.1228	-0.0200	0.1425	0.1740	0.3054	-0.0497	0.1171	0.0882	-0.0447	0.1219	0.1175	-0.0596	1

Tabla 5.2 Matriz de correlaciones (Junio 1999)

	PR_CC_J99	I_SI_J99	IC_J99	CRP_J99	L(CA) J99	R_PR_J99	R_CC_J99	RE_F_J99	RE_V_J99	X10_J99	RE_C_J99	MIC_J99	MO_J99	CC_J99	LARA_J99
PR_CC_J99	1														
CMS_J99	0.3830	1													
IC_J99	-0.1141	-0.3289	1												
CRP_J99	-0.0699	-0.2876	-0.0168	1											
L(CA) J99	0.0633	0.0630	0.4612	0.0195	1										
R_PR_J99	-0.3049	0.0072	-0.1005	-0.3325	-0.2434	1									
R_CC_J99	0.5385	0.3109	-0.1634	-0.1800	-0.1350	0.0292	1								
RE_F_J99	0.2707	0.1712	-0.1821	-0.0081	-0.1329	-0.3315	0.1844	1							
RE_V_J99	0.1165	0.3317	-0.9999	0.0189	-0.4608	0.0943	0.1648	0.1865	1						
X10_J99	0.0676	0.1838	-0.3236	-0.2848	-0.0004	-0.0489	0.1466	0.1176	0.3212	1					
RE_C_J99	-0.3022	-0.0572	-0.0675	-0.0799	-0.1189	-0.0032	-0.0922	0.0632	0.0651	0.7585	1				
MIC_J99	0.1136	0.3279	-1.0000	0.0169	-0.4616	0.0998	0.1629	0.1851	0.9999	0.3235	0.0677	1			
MO_J99	0.1166	0.3319	-0.9999	0.0189	-0.4610	0.0946	0.1649	0.1856	1.0000	0.3212	0.0651	0.9999	1		
CC_J99	0.0088	0.2629	-0.0618	-0.1434	-0.0087	-0.1143	0.1048	0.0809	0.0631	0.2781	0.2949	0.0618	0.0630	1	
LARA_J99	0.3543	0.1093	-0.1874	-0.0235	0.0326	0.1147	0.4336	-0.1582	0.1838	0.1132	-0.1933	0.1865	0.1841	-0.0463	1

Tabla 5.3 Matriz de correlaciones (Diciembre de 1998)

	PR_CC_D98	I_SI_D98	IC_D98	CRP_D98	L(CA)D98	R_PR_D98	R_CC_D98	RE_F_D98	RE_V_D98	X10_D98	RE_C_D98	MIC_D98	MO_D98	CC_D98	LARA_D98
PR_CC_D98	1														
CMS_D98	0.1690	1													
IC_D98	-0.1218	-0.3636	1												
CRP_D98	0.6377	-0.0430	0.0612	1											
L(CA)D98	0.2989	-0.0227	0.4173	0.5339	1										
R_PR_D98	-0.2282	0.0912	-0.0983	0.0160	0.0210	1									
R_CC_D98	0.3527	0.1496	-0.1681	0.0041	-0.0041	0.2962	1								
RE_F_D98	0.5418	0.0116	-0.1240	0.6479	0.2511	-0.1731	-0.0791	1							
RE_V_D98	0.1218	0.3637	-0.9999	-0.0619	-0.4186	0.0940	0.1661	0.1245	1						
X10_D98	-0.0045	0.1379	-0.3386	-0.2214	-0.2176	0.1107	0.1123	0.1111	0.3342	1					
RE_C_D98	-0.4430	0.0310	-0.0883	-0.6207	-0.3801	0.1104	-0.1122	-0.2775	0.0854	0.7458	1				
MIC_D98	0.1202	0.3644	-1.0000	-0.0615	-0.4178	0.1026	0.1671	0.1243	0.9999	0.3387	0.0893	1			
MO_D98	0.1222	0.3636	-0.9999	-0.0618	-0.4184	0.0934	0.1664	0.1244	1.0000	0.3343	0.0853	0.9999	1		
CC_D98	0.0403	0.3309	-0.0685	-0.0871	-0.0718	-0.0638	0.2086	-0.2092	0.0689	0.1840	0.1977	0.0679	0.0690	1	
LARA_D98	0.2865	0.1434	-0.1476	0.3894	0.1493	0.4643	0.3694	0.2960	0.1409	0.1187	-0.0645	0.1481	0.1411	-0.0512	1

Finalmente, al analizar, conjuntamente, los resultados de las pruebas marginales de diferencias entre poblaciones, así como las matrices de correlaciones para las tres bases de datos, las variables IC, RE_V y MO no fueron consideradas para este análisis, debido a que éstas presentaban correlaciones altas con la variable MIC, que fue la única que presentó diferencias significativas entre los dos grupos.

Por ello, únicamente se consideraron 12 variables explicativas, para ejemplificar las tres técnicas estadísticas. Estas variables son:

Variable	Descripción
PR_CC	Prima retenida entre capital contable
I_SI	Índice de costo de siniestralidad
CRP	Crecimiento real de prima retenida
L_CA	Logaritmo del crecimiento real de activos
R_CC	Reserva de obligaciones contractuales entre capital contable
R_PR	Reserva de obligaciones contractuales entre prima retenida
RE_F	Rentabilidad de financiera
X_10	Utilidad neta entre activos
RE_C	Rentabilidad de capital
MIC	Margen de ingresos sobre costos
CC	Capital contable
LARAT	Índice de responsabilidad

5.4 Clasificación utilizando análisis discriminante

La primera técnica utilizada en este trabajo para clasificar a las compañías y recuperar los grupos de empresas Π_0 (sanas) y Π_1 (con posibles problemas financieros) fue el análisis discriminante.

Para ello, y debido a que este análisis se realizó únicamente para efectos de ilustrar la técnica de análisis discriminante, se supuso que los dos grupos, de 12 variables en cada una de las tres fechas, tienen la misma matriz de varianzas y covarianzas.

Por otra parte, y aún cuando se eliminaron las variables que daban origen a los coeficientes de correlación simple más altos, es posible que al considerar el conjunto de 12 variables existan patrones de correlación entre algunos grupos de variables que hagan evidentes problemas de multicolinealidad o casi multicolinealidad. De hecho, así ocurrió y no fue posible construir una regla discriminante que incluyese todas las variables. Por ello, se utilizó un método de selección de variables, para así determinar aquellas variables que deberían entrar en la función discriminante.

En este trabajo, se empleó el método de selección forward stepwise, el cual introduce en la función discriminante una por una aquellas variables que resulten más significativas en lo que se refiere a su contribución a la discriminación entre ambos grupos. El criterio utilizado para incluir una variable en la función discriminante se basa tanto en la estadística λ de Wilks, como el cálculo del nivel de tolerancia (en este trabajo se uso el nivel 0.2).

La λ de Wilks, como se mencionó, es utilizada para determinar si la diferencia entre grupos es estadísticamente significativa, dado un nivel de significancia.

Por otro lado, el nivel de tolerancia se calcula como $1-R^2$, donde R^2 indica la correlación múltiple de la variable que se está analizando con las variables que actualmente están en el modelo; entonces, si alguna variable tiene un nivel de tolerancia menor al especificado (0.2), significará que esta variable es más del 80% redundante con las otras variables ya incluidas en la función discriminante.

Así, se tienen las llamadas funciones de clasificación, dadas por la siguiente expresión para cada una de las poblaciones:

$$L_g = -\frac{1}{2} x_g' S^{-1} x_g + x_g' S^{-1} X, \quad \text{con } g = 0, 1$$

donde, $x_g = (x_{g1}, x_{g2}, \ldots, x_{gk})'$, con k variables explicativas, y

$$S = \frac{(n_0 - 1)S_0 + (n_1 - 1)S_1}{n - 2}.$$

De donde, se tiene que la función discriminante es de la forma

$$L_0 - L_1 = (\bar{x}_0' - \bar{x}_1')S^{-1}X - \frac{1}{2}\left[\bar{x}_0' S^{-1} x_0 - x_1' S^{-1} x_1\right],$$

es decir,

$$L_0 - L_1 = \underline{\beta}' \underline{X} - \mu,$$

donde μ es el punto medio entre la media de ambos grupos (μ_{0Y} y μ_{1Y}), es decir

$$\mu = \frac{1}{2}(\mu_{0Y} + \mu_{1Y}) = \frac{1}{2}\left(\underline{\beta}' \underline{\mu}_0 + \underline{\beta}' \underline{\mu}_1\right) = \frac{1}{2}\underline{\beta}'(\underline{\mu}_0 + \underline{\mu}_1) = \frac{1}{2}\left(\underline{\mu}_0 - \underline{\mu}_1\right)' \Sigma^{-1}(\underline{\mu}_0 + \underline{\mu}_1).$$

Por lo tanto, si se desea clasificar una nueva compañía de seguros en alguno de los dos grupos, entonces se deberá evaluar cada función de clasificación L_g con el vector de variables explicativas de la compañía a clasificar, de tal forma que la nueva compañía aseguradora se clasificará como perteneciente al grupo cuyo valor en la función L_g sea mayor.

5.4.1 Datos de diciembre de 1999

En diciembre de 1999, se obtuvieron dos variables explicativas en la función de discriminante, como se observa en el siguiente cuadro de resultados:

Resumen del análisis discriminante

N=54	λ de Lambda	λ Parcial	F-remover (1,51)	n.s.d. *	Toler.	1-Toler. (R^2)
RE_C_D99	0.833809	0.741301	17.797928	0.000101	0.966656	0.033344
LARA_D99	0.764808	0.808182	12.104630	0.001039	0.966655	0.033345

* n.s.d. = nivel de significancia descriptivo

Y el modelo que incluye a las dos variables presenta un valor de λ de Wilks = 0.618, este valor se puede aproximar utilizando una estadística F de la distribución F de Fisher dada por F (2,51) = 15.755 para la cual el n.s.d es casi cero (Ver Apéndice B.1).

Lo anterior, utilizando la prueba de Wilks, nos lleva a rechazar la hipótesis nula H_0 con un nivel de significancia $\alpha = 0.05$, por lo que se concluye que al considerar las

dos variables en la función discriminante, el poder discriminatorio de ésta es estadísticamente significativo.

Los coeficientes de las dos funciones de clasificación son:

Funciones de Clasificación

	Π_0	Π_1
RE_C_D99	-0.531936	-3.757938
LARA_D99	0.502571	1.405446
Constante	-0.986253	-4.013300

es decir, se tienen las siguientes funciones de clasificación:

L_0 = - 0.986253 - 0.531936 RE_C_D99 + 0.502571 LARA_D99

L_1 = - 4.0133 - 3.757938 RE_C_D99 + 1.405446 LARA_D99

Como resultado, se obtuvo que el porcentaje de compañías de Π_1 (compañías con posibles problemas financieros) correctamente clasificadas es considerable, esto es 72.72%, mientras que el porcentaje de compañías de Π_0 (compañías sanas en el sentido financieros) correctamente clasificadas fue muy alto, es decir 90.69%; esto es:

Matriz de Clasificación*

	Porcentaje Correcto	Π_0	Π_1
Π_0	90.697674	39	4
Π_1	72.727273	3	8
Total	87.037037	42	12

* Renglones: Clasificación Observada
Columnas: Clasificación Pronosticada

Cabe mencionar que fue posible clasificar las 43 compañías sanas y las 11 compañías con posibles problemas financieros, debido a que aunque la compañía 35 (sana) presenta valores indeterminados en algunas variables, ninguna de éstas forman parte de la función discriminante.

Finalmente, las compañías clasificadas incorrectamente son:

	Gpo.	Gpo.	Primer trimestre 2000		
Obs.	Observado	Clasificado	CRT	MCG	CMP
6	0	1	1.21	2.5	3.3
17	0	1	1.19	2.2	4.7
24	1	0	0.99	1.16	15.1
30	0	1	1.18	14.23	2.6
39	0	1	1.38	3.64	1.95
51	1	0	2.6	24.8	0.81
54	1	0	1	0.9	1.11

5.4.2 Datos de junio de 1999

Con información de las 12 variables para junio de 1999, se obtuvo como resultado, dos variables en la función discriminante, tal como se observa en el siguiente cuadro de resultados:

Resumen del análisis discriminante

N=54	λ de Lambda	λ Parcial	F-remover (1,51)	n.s.d.*	Toler.	1-Toler. (R^2)
LARA_J99	0.871770	0.755763	16.481503	0.000169	0.908750	0.091250
X10_J99	0.822495	0.801039	12.667286	0.000815	0.908750	0.091250

* n.s.d. nivel de significancia descriptivo

El modelo, considerando las dos variables explicativas presenta el valor de λ de Wilks = 0.65885, cuya aproximación utilizando la F de Fisher es F(2,51) = 13.204, con un n.s.d. < 0.0001.

Al igual que en el modelo obtenido utilizando la información del cierre de 1999, al aplicar la prueba estadística de Wilks se observa que la función discriminante formada con las dos variables de junio de 1999 tiene poder para discriminar entre los dos grupos (Π_0 y Π_1), esto es, debido a la hipótesis nula H_0 se rechaza con un nivel de significancia $\alpha = 0.05$.

Los coeficientes de las funciones de clasificación son:

Funciones de Clasificación

	Π_0	Π_1
X_10_J99	-3.409080	-12.003447
LARA_J99	1.529519	3.266325
Constante	-1.556765	-5.165544

es decir, las funciones de clasificación son:

L_0 = - 1.556765 + 1.529519 LARA_J99 - 3.409080 X10_J99

L_1 = - 5.165544 + 3.266325 LARA_J99 - 12.003447 X10_J99

Como producto, se observa que el porcentaje de compañías del grupo Π_1 clasificadas correctamente disminuyó a un 63.63% con respecto a la clasificación utilizando la información de diciembre 1999, asimismo el porcentaje de compañías correctamente clasificadas en Π_0 disminuyó a 86.04%, como se muestra en la siguiente matriz de clasificación:

Matriz de Clasificación*

	Porcentaje Correcto	Π_0	Π_1
Π_0	86.046512	37	6
Π_1	63.636364	4	7
Total	81.481481	41	13

* Renglones: Clasificación Observada
Columnas: Clasificación Pronosticada

Al igual que en diciembre de 1999, fue posible clasificar a las 43 compañías sanas y a las 11 compañías que presentaban posibles problemas financieros, debido a que aunque la compañía 35 (sana) presenta valores indeterminados para algunas de sus variables, ninguna de éstas corresponde a las que forman parte de las funciones de clasificación.

Las compañías que fueron clasificadas incorrectamente son:

	Gpo.	Gpo.	Primer trimestre 2000		
Obs.	Observado	Clasificado	CRT	MCG	CMP
6	0	1	1.21	2.5	3.3
17	0	1	1.19	2.2	4.7
19	0	1	1.21	1.73	1.44
23	1	0	0.9	0.5	0.4
24	1	0	0.99	1.16	15.1
30	0	1	1.18	14.23	2.6
38	1	0	0.99	0.57	2.38
39	0	1	1.38	3.64	1.95
48	0	1	1.02	47.7	3.34
54	1	0	1	0.9	1.11

5.4.3 Datos de diciembre de 1998

Por último, con la información del cierre de 1998, se obtuvo únicamente una variable explicativa en la función discriminante, como se observa en el siguiente cuadro de resultados:

Resumen del análisis discriminante

N=54	λ de Lambda	λ Parcial	F-remover (1,52)	n.s.d.*	Toler.	1-Toler. (R^2)
LARA_D98	1.00000	0.87072	7.72069	0.00758	1	0

* n.s.d. nivel de significancia descriptivo

Así, el modelo considerando una sola variable presenta el valor λ de Wilks=0.87072, el cual al aproximarlo utilizando la estadística F de la distribución F de Fisher es F(1,52) = 7.7207 cuyo n.s.d = 0.0076.

Por lo tanto, al aplicar la prueba estadística de Wilks, la discriminación entre los dos grupos (Π_0 y Π_1) es significativa, ya que la hipótesis nula H_0 se rechaza con un nivel de significancia $\alpha = 0.05$.

Los coeficientes de las funciones de clasificación son:

Funciones de Clasificación

	Π_0	Π_1
LARA_D98	0.747511	1.534848
Constante	-1.090404	-2.367965

es decir; se tienen las siguientes funciones de clasificación

$L_0 = -1.090404 + 0.747511$ LARA_D98

$L_1 = -2.367965 + 1.534848$ LARA_D98

Como resultado, se obtuvo que el número de compañías de Π_1 clasificadas correctamente fue menor que al utilizar la información de junio de 1999, es decir 45.45%; por otro lado, el número de compañía de Π_0 correctamente clasificadas es un poco mayor que al utilizar la información de junio de 1999, pero menor que al utilizar la información de diciembre de 1999. Esto se puede observar en la siguiente matriz de clasificación.

Matriz de Clasificación*

	Porcentaje Correcto	Π_0	Π_1
Π_0	88.372093	38	5
Π_1	45.454545	6	5
Total	79.629630	44	10

* Renglones: Clasificación Observada
Columnas: Clasificación Pronosticada

Cabe mencionar, que al igual que los análisis anteriores, fue posible clasificar a las 43 compañías sanas y a las 11 compañías con posibles problemas financieros, debido a que ninguna de estas compañías presenta valores indeterminados en la variable LARA_D98, que es la única que forma parte de la función discriminante.

Las compañías clasificadas incorrectamente son:

	Gpo.	Gpo.	Primer trimestre 2000		
Obs.	Observado	Clasificado	CRT	MCG	CMP
3	0	1	1.49	6.86	13.29
10	0	1	1.06	1.06	2.38
17	0	1	1.19	2.2	4.7
24	1	0	0.99	1.16	15.1
25	0	1	1.13	3.2	2.4
34	0	1	1	1.6	51.3
38	1	0	0.99	0.57	2.38
44	1	0	0.98	-641.04	1.14
47	1	0	0.95	1.94	0.61
51	1	0	2.6	24.8	0.81
54	1	0	1	0.9	1.11

5.5 Clasificación utilizando análisis logístico

La segunda técnica utilizada para clasificar a las compañías y recuperar los grupos de empresas Π_0 (sanas) y Π_1 (con posibles problemas financieros) fue el modelo de regresión logística.

Al igual que con el análisis discriminante, se utilizaron las tres bases de datos correspondientes al cierre de 1999, a junio de 1999 y por último al cierre de 1998.

Al margen de los resultados del análisis discriminante, en un principio, se utilizaron las 12 variables para cada conjunto de datos, junto con el método de selección de variables forward stepwise.

El procedimiento de selección de variables forward stepwise utiliza un modelo inicial formado únicamente por el término constante o independiente. La primera variable que se incluirá en el modelo, será aquella que al considerarla, junto con el término constante, se observa que el modelo presenta un ajuste a los datos significativamente mejor que el modelo dónde solamente se considera el término constante y que los modelos en dónde se consideran de forma marginal el resto de las variables explicativas. La segunda variable que entrará en el modelo, será aquella para la cual el modelo que considera esta variable, además del término constante así como de la primera variable que se incluyó en el modelo, presente un mejor ajuste a los datos que el modelo que se tenía antes de agregar esta segunda variable. De igual forma se irán agregando nuevas variables al modelo, hasta que no se tenga un mejor ajuste a los datos que el que se presenta el modelo sin considerar la nueva variable.

Para la determinación del modelo que se ajusta mejor a los datos, se hizo uso de la estadística formada como la diferencia de devianzas.

El modelo ajustado obtenido al realizar la estimación por máxima verosimilitud, utilizado las q variables obtenidas de la selección de variables será de la forma

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = \underline{x}_i'\hat{\underline{\beta}},$$

donde $\hat{\underline{\beta}} = (\hat{\underline{\beta}}_0, \hat{\underline{\beta}}_1, \ldots, \hat{\underline{\beta}}_q)'$ será el estimado máximo verosímil del vector de parámetros $\underline{\beta} = (\underline{\beta}_0, \underline{\beta}_1, \ldots, \underline{\beta}_q)'$.

Es decir,

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_q X_q$$

Al despejar $\hat{\theta}_i$ de la ecuación anterior, se obtendrá la probabilidad de que la compañía i presente posibles problemas financieros en marzo del 2000, esto es

$$P(Y_i = 1) = \hat{0}_i = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_q X_q}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_q X_q}}, \quad \text{con } i = 1,\ldots,n$$

donde Y_i es una variable aleatoria con función de distribución Bernoulli($\hat{\theta}_i$), de manera que Y_i=1 si la compañía aseguradora i tiene problemas financieros, y $Y_i = 0$ si la compañía aseguradora i es sana en el sentido financiero, con $P(Y_i = 1) = \theta_i$ y $P(Y_i = 0) = 1 - 0_i$.

5.5.1 Datos de diciembre de 1999

Al ajustar el modelo por medio del método de máxima verosimilitud, utilizando las 12 variables explicativas del primer cierre anterior a marzo de 2000; es decir, de diciembre de 1999, así como el método de selección de variables forward stepwise, se obtuvo la siguiente tabla de resultados

Variable	Coeficiente Estimado	Error Estándar	Estadística de Wald	n.s.d.*
Constante	-3.158	0.8383	14.1927	0.0002
RE_C_D99	-2.0869	0.7876	7.0219	0.0081
LARA_D99	0.7171	0.5003	2.0542	0.1518

Devianza del modelo ajustado	35.92091
Grados de libertad	51

* n.s.d. = nivel de significancia descriptivo

Es decir, se obtiene el siguiente modelo ajustado con dos variables explicativas

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = -3.158 - 2.0869\ \text{RE\_C\_D99} + 0.7171\ \text{LARA\_D99}$$

Al realizar la prueba de diferencia de devianzas (ver sección 3.3), la hipótesis nula H_0 se rechaza en favor de la hipótesis alternativa con un nivel de significancia α=0.05, por lo que se concluye que al considerar las variables explicativas RE_C_D99 y LARA_D99, además del término constante en el modelo, éste describe mejor a los datos que el modelo en el que sólo se considera el término constante. Es decir

D_0	Grados de Libertad	D_a	Grados de Libertad	ΔD	Grados de Libertad
54.59337	53	35.92091	51	18.67246	2

Sin embargo, al analizar la estadística de Wald para cada uno de los coeficientes estimados, se observa que el término constante, así como la variable explicativa RE_C_D99 resultan significativos de manera parcial, con un nivel de significancia α=0.05, al considerar el resto de las variables en el modelo ajustado. Sin embargo, la variable LARA_D99 no resultó significativa.

Debido a que tanto la devianza utilizada en la selección de variables, así como la prueba de Wald utilizada para verificar la significancia parcial de las variables explicativas, se basan en una distribución χ^2 aproximada, únicamente se incluirán en el modelo final aquellas variables que resulten significativas con ambas pruebas estadísticas. Por ello, la variable LARA_D99 no se considerará en el modelo.

Al ajustar el modelo por medio del método de máxima verosimilitud, utilizando únicamente el término constante y la variable RE_C_D99 se obtuvo la siguiente tabla de resultados

Variable	Coeficiente Estimado	Error Estándar	Estadística de Wald	n.s.d.*
Constante	-2.011724	0.4665	18.6002	0.0000
RE_C_D99	-2.005646	0.6969	8.2837	0.0040

Devianza del modelo ajustado	43.39807
Grados de libertad	52

* n.s.d. = nivel de significancia descriptivo

Por lo que el modelo ajustado utilizando información de diciembre de 1999, para estimar la probabilidad de que la compañía i tenga posibles problemas financieros, es el siguiente

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = -2.011724 - 2.005646\ RE\_C\_D99$$

La devianza del modelo ajustado es significativa $D = 43.39807 < \chi^2_{(52),\ 0.95} = 67.5$; por lo que se dice que con un nivel de significancia $\alpha = 0.05$ el modelo es aceptable.

Por otro lado, al realizar la prueba de diferencia de devianzas, la hipótesis nula H_0 se rechaza con un nivel de significancia $\alpha = 0.05$, con lo que se concluye que al considerar el término constante además de la variable explicativa RE_C_D99 en el modelo ajustado, éste describe mejor a los datos que el modelo en el que sólo se considera el término constante. Es decir

D_0	Grados de Libertad	D_a	Grados de Libertad	ΔD	Grados de Libertad
54.59337	53	43.39807	52	11.1953	1

De acuerdo con los coeficientes del modelo estimado, se calcularon las probabilidades de que cada una de las compañías presente posibles problemas financieros, es decir $\hat{\theta}_i$ para i = 1,..,54; y al considerar una probabilidad de corte de 0.5, se clasificó a las compañías en uno de los dos grupos. Es decir, todas aquellas compañías que presentaban una probabilidad de presentar problemas financieros mayor o igual a 0.5, es decir, $\hat{\theta}_i \geq 0.5$, fueron clasificadas como pertenecientes al grupo Π_1, mientras que las compañías que presentaban una probabilidad de presentar problemas financieros menor a 0.5, fueron clasificadas en el grupo Π_0. Cabe mencionar, que aunque la compañía identificada con el número

de observación 35 presentaba valores indeterminados para algunas variables, ninguna de éstas corresponde a la incluida en el modelo ajustado, por lo que fue posible clasificarla en alguno de los dos grupos.

El porcentaje de compañías de Π_1 (con posibles problemas financieros) correctamente clasificadas corresponde a un 27.27%, por otro lado, el porcentaje de compañías Π_0 (sanas) correctamente clasificadas fue alto, esto es 95.34%; como se muestra en la siguiente tabla

Matriz de Clasificación*

	Porcentaje Correcto	Π_0	Π_1
Π_0	95.348837	41	2
Π_1	27.272727	8	3
Total	81.481481	49	5

* Renglones: Clasificación Observada
Columna: Clasificación Pronosticada

Finalmente, las compañías que fueron clasificadas incorrectamente son:

Obs.	Gpo. Observado	Gpo. Clasificado	Primer trimestre 2000 CRT	MCG	CMP
1	1	0	0.96	0.92	1.46
23	1	0	0.9	0.5	0.4
24	1	0	0.99	1.16	15.1
30	0	1	1.18	14.23	2.6
33	1	0	0.93	0.02	0.3
38	1	0	0.99	0.57	2.38
39	0	1	1.38	3.64	1.95
44	1	0	0.98	-641.04	1.14
51	1	0	2.6	24.8	0.81
54	1	0	1	0.9	1.11

5.5.2 Datos de junio de 1999

Por otro lado, se ajustó un modelo de regresión logística, utilizando las 12 variables para junio de 1999, obteniendo la siguiente tabla de resultados

Variable	Coeficiente Estimado	Error Estándar	Estadística de Wald	n.s.d.*
Constante	-3.947	1.0606	13.8493	0.0002
X_10_J99	-6.4822	2.2651	8.1895	0.0042
LARA_J99	1.4071	0.6331	4.9405	0.0262

Devianza del modelo ajustado	36.96166
Grados de libertad	51

* n.s.d. = nivel de significancia descriptivo

Por lo que el modelo ajustado con las dos variables explicativas, obtenidas de la selección de variables, es de la forma

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = -3.947 - 6.4822\ X10\_J99 + 1.4071\ LARA\_J99$$

De igual forma que para diciembre de 1999, se analizó la estadística de Wald con un nivel de significancia $\alpha = 0.05$ para cada uno de los coeficientes estimados en el modelo ajustado, de donde se observa la significancia parcial de las dos variables en el modelo (X10_J99 y LARA_J99), así como la del término constante.

Por otro lado, al analizar la devianza del modelo ajustado D = 36.96166 y compararla con una $\chi^2_{(51),\,0.95} = 67.5$, se tiene que $D < \chi^2_{(51),\,0.95}$, por lo que se dice que con un nivel de significancia $\alpha = 0.05$ el modelo ajustado es aceptable.

Al realizar la prueba de diferencia de devianzas, la hipótesis nula H_0 se rechaza a favor de la hipótesis alternativa con un nivel de significancia $\alpha = 0.05$, por lo que se concluye que el modelo ajustado para junio de 1999 con las variables explicativas X10_J99 y LARA_J99 así como el término constante, describe mejor a los datos que el modelo en dónde sólo se considera el término constante. Esto es

D_0	Grados de Libertad	D_a	Grados de Libertad	ΔD	Grados de Libertad
54.59337	53	36.96166	51	17.63171	2

Al utilizar los parámetros estimados del modelo, se obtuvo la probabilidad de presentar posibles problemas financieros para cada una de las 54 compañías

analizadas; es decir $\hat{0}_i$ con i = 1,2,...,54. Al igual que en diciembre de 1999, se consideró una probabilidad de corte de 0.5, y se realizó una clasificación de las compañías aseguradoras en uno de los dos grupos.

Como resultado, se obtuvo que el porcentaje de compañías de Π_1 (con posibles problemas financieros) correctamente clasificadas fue de 45.45%, por otra parte, el porcentaje de compañías Π_0 (sanas) correctamente clasificadas fue el mismo que para diciembre de 1999 es decir, 95.34%; como se muestra en la siguiente tabla

Matriz de Clasificación*

	Porcentaje Correcto	Π_0	Π_1
Π_0	95.348837	41	2
Π_1	45.454545	6	5
Total	85.185185	47	7

* Renglones: Clasificación Observada
Columna: Clasificación Pronosticada

Las compañías que fueron clasificadas incorrectamente, así como los valores de sus indicadores regulatorios a marzo del 2000, se muestran en la siguiente tabla

Obs.	Gpo. Observado	Gpo. Clasificado	Primer trimestre 2000 CRT	MCG	CMP
6	0	1	1.21	2.5	3.3
18	1	0	1.02	1.14	0.3
23	1	0	0.9	0.5	0.4
24	1	0	0.99	1.16	15.1
38	1	0	0.99	0.57	2.38
39	0	1	1.38	3.64	1.95
51	1	0	2.6	24.8	0.81
54	1	0	1	0.9	1.11

5.5.3 Datos de diciembre de 1998

Finalmente, se ajustó un modelo de regresión logística utilizando las 12 variables explicativas de diciembre de 1998, obteniendo como resultados la siguiente tabla

Variable	Coeficiente Estimado	Error Estándar	Estadística de Wald	n.s.d.*
Constante	-2.4378	0.6955	12.2859	0.0005
LARA_D98	0.7846	0.45	3.0397	0.0813

Devianza del modelo ajustado	48.5035
Grados de libertad	52

* n.s.d. = nivel de significancia descriptivo

Esto es, el modelo ajustado obtenido al estimar por máxima verosimilitud y considerar únicamente la variable LARA_D98 así como el término constante, de acuerdo al método de selección de variables forward stepwise, es el siguiente

$$\log\left(\frac{\hat{\theta}_i}{1-\hat{\theta}_i}\right) = -2.4378 + 0.7846\ \text{LARA\_D98}$$

Al realizar la prueba de diferencia de devianzas, de la siguiente tabla se observa que la hipótesis nula H_0 se rechaza con un nivel de significancia $\alpha = 0.05$, a favor de la hipótesis alternativa H_a, con ello, se concluye que el modelo en el que se considera el término constante y la variable explicativa LARA_D98, describe mejor a los datos, que el modelo en dónde únicamente se considera el término constante. Es decir

D_0	Grados de Libertad	D_a	Grados de Libertad	ΔD	Grados de Libertad
54.59337	53	48.5035	52	6.08987	1

Sin embargo al analizar la significancia parcial de los coeficientes estimados por medio de la estadística de Wald, se observa que la variable LARA_D98 no es significativa al considerar un nivel se significancia $\alpha = 0.5$.

Por la anterior, se determinó que ninguna de las variables para diciembre de 1998 resultan significativas para ajustar un modelo que estime la probabilidad de que una compañía presente posibles problemas financieros, ya que se concluyó incluir únicamente aquellas variables que resultaran significativas tanto al utilizar la devianza como la prueba de Wald, debido a que estas se basan en distribuciones aproximadas.

5.6 Resultados generales

Los resultados del análisis de clasificar a las compañías aseguradoras de México como compañías sanas (Π_0), o con posibles problemas financieros (Π_1), que se llevó a cabo en este capítulo, se presentan en la tabla 5.4.

En general, se observa que los resultados obtenidos de reproducir el grupo de compañías consideradas como sanas, pueden ser considerados buenos; sin embargo, no se puede decir lo mismo del número de clasificaciones correctas de compañías con posibles problemas financieros.

Los costos de clasificar erróneamente a una compañía que puede tener problemas financieros, como perteneciente al grupo de compañías sanas, son mayores que los de clasificar a una compañía sana como proveniente del grupo de compañías con posibles problemas financieros, puesto que en el primer caso, al no detectar oportunamente problemas financieros en una compañía aseguradora, está podría presentar insolvencia financiera y como consecuencia no estar en posibilidades de hacer frente a las responsabilidades que adquirió con los asegurados. En el segundo caso, también se ocasionan problemas al requerir la empresa de la adopción de prácticas más conservadoras y restrictivas en su operación, así como la posible aportación de recursos adicionales.

La diferencia, sin embargo, estriba en que el objetivo de la agencia supervisora es proteger el interés de los asegurados antes que el de los accionistas.

Tabla 5.4: Resultados generales

	diciembre 1999		junio 1999		diciembre 1998	
	Análisis Disc.	Modelo Logístico	Análisis Disc.	Modelo Logístico	Análisis Disc.	Modelo Logístico
Núm. de var. en el modelo.	2	1	2	2	1	----
Variables en el modelo	RE_C LARA	RE_C	LARA X_10	LARA X_10	LARA	----
Cías. de Π_0 clasificadas	43	43	43	43	43	----
Cías. de Π_1 clasificadas	11	11	11	11	11	----
Cías. de Π_0 clasificadas correctamente	39 (90.69%)	41 (95.34%)	37 (86.04%)	41 (95.34%)	38 (88.37%)	----
Cías. de Π_1 clasificadas correctamente	8 (72.72%)	3 (27.27%)	7 (63.63%)	5 (45.45%)	5 (45.45%)	----
Total de cías. clasificadas correctamente	47 (87.03%)	44 (81.48%)	44 (81.48%)	46 (85.18%)	43 (79.62%)	----

En los diferentes modelo ajustados, se variaron tanto los datos como el modelo a utilizar. Así, de la tabla se observa que los porcentajes de compañía correctamente clasificadas, obtenidos con la información de diciembre de 1999 y junio de 1999, son superiores a los obtenidos con la información de diciembre de 1998, además de que para esa fecha, ninguna variable resultó significativa al usar el modelo logístico.

Por lo tanto, únicamente se cuenta con dos escenarios a comparar, el primero con información de diciembre de 1999 y el segundo con información de junio de 1999.

Al revisar los resultados de la tabla anterior, se observa que el porcentaje del total de compañías correctamente clasificadas, se encuentra alrededor del 83% para todos los modelos ajustados, considerando los dos escenarios. Si bien es cierto que este porcentaje no muestra una clasificación perfecta, sí puede ser considerado un porcentaje alto. Cabe mencionar, que el porcentaje de compañías mal clasificadas puede deberse a dos situaciones:

1) Ninguna técnica estadística puede reproducir una clasificación perfecta.
2) En esta aplicación, no se trató de reproducir una clasificación de las compañías aseguradoras como solventes o insolventes, sino que se quiso reproducir la clasificación de las compañías en dos grupos, que fueron obtenidos a partir de un criterio alternativo arbitrario, es decir, sanas o con posibles problemas financieros.

Siguiendo el criterio de los costos de una mala clasificación, y con el fin de proteger a los asegurados, es preferible utilizar un modelo estadístico en el que se clasifique correctamente un mayor número de compañías que en un futuro podrían presentar problemas financieros, sin incurrir en extremos inaceptables. Por lo tanto, con los datos utilizados en este capítulo y sin verificar ningún supuesto, de la tabla anterior se observa que el análisis discriminante es superior al modelo logístico, para ambos escenarios, ya que con este modelo se obtienen los mayores porcentajes de compañías con posibles problemas financieros correctamente clasificadas, además de que el número de compañías sanas correctamente clasificadas no es malo.

Por último, debido a que con el análisis discriminante, utilizando información de diciembre de 1999 se obtiene, además de un mayor número de compañías de Π_1

clasificadas correctamente, un mayor número de compañías de Π_0 correctamente clasificadas, que al utilizar el análisis discriminante con datos de junio de 1999, se concluye que el modelo candidato para detectar oportunamente compañías con posibles problemas financieros, es el que se obtiene con el análisis discriminante al utilizar datos de tres meses previos a la fecha de clasificación de las variables RE_C y LARA (a diciembre de dos ejercicios previos a la fecha de clasificación). Es pertinente aclarar que se llega a esta conclusión bajo estos datos y sin verificar ninguno de los supuestos

Así mismo, es importante señalar que si se desea realizar una clasificación en cualquier otro trimestre de un ejercicio, es necesario realizar un nuevo análisis, para poder decidir cual es la información previa a este trimestre, que se requiere para poder llevar a cabo una clasificación óptima. Esto, debido a que las condiciones del sector asegurador mexicano varían de un trimestre a otro, ya que por ejemplo, al inicio del ejercicio las compañías emiten un menor número de pólizas que al final del mismo ejercicio, como consecuencia de la cuesta de enero.

CONCLUSIONES

En este trabajo se presentó, discutió e ilustró el papel que algunas técnicas estadísticas pueden jugar en el desarrollo de sistemas para la detección oportuna de insolvencia en empresas de seguros. El problema de detección está muy relacionado con el de clasificación estadística y también puede hacer uso del análisis de supervivencia. Sin embargo, no necesariamente estas herramientas son las únicas que podrían emplearse.

Las técnicas consideradas en esta tesis fueron el análisis discriminante, los modelos de respuesta cualitativa, en particular el modelo logístico, y por último el análisis de supervivencia.

Al revisar los elementos teóricos de estas técnicas, se apreció que el análisis discriminante es una herramienta simple con la que se obtienen resultados satisfactorios al realizar una clasificación estadística, mientras que en los modelos de respuesta cualitativa es necesario incorporar algunos supuestos adicionales. A cambio, se obtiene una expresión explícita para la probabilidad de pertenecer a cada una de las poblaciones involucradas en la clasificación. Por último, el análisis de supervivencia considera modelos más complejos que hacen uso de más información, pero que proporcionan resultados más detallados ya que estiman las probabilidades de pertenecer a cada una de las poblaciones, tanto en función de un conjunto de covariables asociadas a cada unidad observada como en función de tiempo.

El análisis discriminante y el modelo logístico fueron ilustrados con una base de datos reducida que procede del sector asegurador mexicano, y que sólo en forma aproximada refleja la insolvencia de este tipo de compañías.

En este trabajo no fue posible utilizar el análisis de supervivencia para llevar a cabo una clasificación financiera de las compañías aseguradoras de México, debido a que la información con la que se trabajó no posibilitaba este tipo de estudio; sin embargo, sería interesante revisar los resultados que se obtendrían al contar con una base de datos más completa y utilizar los elementos teóricos presentados en esta tesis para realizar un análisis de este tipo; por lo que es de destacar la importancia de recabar información financiera detallada de las compañías aseguradoras a lo largo del tiempo.

A pesar de ello, los resultados son sugerentes y prometedores. De cualquier manera, un estudio posterior con una colección de datos más completa, que incluya otros índices, actualizados de acuerdo a la situación del mercado y que se relacionen más de cerca con la potencial insolvencia de las compañías aseguradoras de México, además del empleo de otras técnicas estadísticas, por ejemplo Bayesianas que consideran un énfasis especial en los aspectos predictivos de todo proceso de inferencia, podría complementar el primer esfuerzo reportado en esta tesis.

En cualquier caso, el objetivo de plantear la conveniencia del análisis estadístico en la versión mexicana de este problema ha sido establecido.

APÉNDICE A

A.1 Desigualdad de Cauchy - Schwarz

Sea x, y dos vectores de $p \times 1$, entonces $(x'y) \leq (x'x)(y'y)$.

Demostración:

Considere el vector $x - ay$ con a un escalar cualquiera, entonces,

$$0 \leq (x-ay)'(x-ay) = x'x - ax'y - ay'x + a^2 y'y = x'x - 2ax'y + a^2 y'y,$$

complementando un cuadrado, se tiene que:

$$0 \leq x'x - \frac{(x'y)^2}{y'y} + (y'y)\left[a - \frac{x'y}{y'y}\right]^2,$$

si $\quad a = \dfrac{x'y}{y'y}, \quad \Rightarrow \; a - \dfrac{x'y}{y'y} = 0,$

por lo tanto:

$$0 \leq \underset{\sim}{x}'\underset{\sim}{x} - \frac{(\underset{\sim}{x}'\underset{\sim}{y})^2}{\underset{\sim}{y}'\underset{\sim}{y}}, \quad \Leftrightarrow \quad \frac{(\underset{\sim}{x}'\underset{\sim}{y})^2}{\underset{\sim}{y}'\underset{\sim}{y}} \leq \underset{\sim}{x}'\underset{\sim}{x}, \quad \Leftrightarrow \quad (\underset{\sim}{x}'\underset{\sim}{y})^2 \leq (\underset{\sim}{x}'\underset{\sim}{x})(\underset{\sim}{y}'\underset{\sim}{y}).$$

■

A.2 Extensión de la desigualdad Cauchy - Schwarz

Sea B una matriz de p x p, definida positiva, y x, y dos vectores de p x 1,

Entonces $(\underline{x}'\underline{y})^2 \leq (\underline{x}'B\underline{x})(\underline{y}'B^{-1}\underline{y})$.

Demostración:

$x'y = x'Iy = x'B^{1/2}B^{-1/2}y = (B^{1/2}x)'(B^{-1/2}y)$,

Por la desigualdad de Cauchy - Schwarz, se tiene que:

$$(x'y)^2 = \left[(B^{1/2}x)'(B^{-1/2}y)\right]^2 \leq \left[(B^{1/2}x)'(B^{1/2}x)\right] * \left[(B^{-1/2}y)'(B^{-1/2}y)\right] \quad \Rightarrow$$

$$(x'y)^2 \leq (x'B^{1/2}B^{1/2}x) * (y'B^{-1/2}B^{-1/2}y), \quad \Rightarrow$$

$$(x'y)^2 \leq (x'Bx) * (y'B^{-1}y) \quad \blacksquare$$

A.3 Lema:

Dado que B es una matriz de p x p definida positiva y x un vector de p x 1, tal que $\underline{x} \neq \underline{0}$, y $\underline{d}$ un vector de p x 1, se tiene que

$$\underset{\underline{x}}{\text{Max}} \frac{(\underline{x}'\underline{d})^2}{\underline{x}'B\underline{x}} = \underline{d}'B^{-1}\underline{d}, \quad \text{si } \underline{x} = cB^{-1}\underline{d}, \text{ con } c \neq 0.$$

Demostración:

Debido a que B es una matriz definida positiva de p x p y $\underline{x}$, $\underline{d}$ dos vectores de p x 1; por la extensión de la desigualdad de Cauchy-Schwarz (ver Apéndice A.2), se tiene que:

$$(x'd)^2 \leq (x'Bx)(d'B^{-1}d),$$

debido a que $\underline{x} \neq \underline{0}$ y B es una matriz definida positiva, se tiene que $\underline{x}'B\underline{x} \geq 0$, entonces

$$\frac{(x'd)^2}{x'Bx} \leq \frac{(x'Bx)(d'B^{-1}d)}{x'Bx}, \Rightarrow$$

$$\frac{(x'd)^2}{x'Bx} \leq (d'B^{-1}d).$$

Si se considera $\underline{x} = cB^{-1}\underline{d}$, con $c \neq 0$, se tiene que:

$$\frac{(x'd)^2}{x'Bx} = \frac{((cB^{-1}d)'d)^2}{(cB^{-1}d)'B(cB^{-1}d)} = \frac{(cd'B^{-1}d)((cB^{-1}d)'d)}{c^2(d'B^{-1}BB^{-1}d)}, \Rightarrow$$

$$\frac{(\underline{x}'\underline{d})^2}{\underline{x}'B\underline{x}} = \frac{(\underline{d}'B^{-1}\underline{d})(\underline{d}'B^{-1}\underline{d})}{\underline{d}'IB^{-1}\underline{d}} = (\underline{d}'B^{-1}\underline{d})$$

Por lo tanto:

$$\text{Max}_{\underline{x}\neq\underline{0}} \frac{(\underline{x}'\underline{d})2}{\underline{x}'B\underline{x}} \leq \frac{(\underline{x}'B\underline{x})(\underline{d}'B^{-1}\underline{d})}{\underline{x}'B\underline{x}}, \text{ si } \underline{x} = cB^{-1}\underline{d}, \text{ con } c \neq 0. \blacksquare$$

APÉNDICE B

B.1 λ de Wilks

La prueba de la λ de Wilks para p grupos se define como:

$$H_0: \underline{\mu}_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_p \quad \text{Vs} \quad H_a: \underline{\mu}_i \neq \underline{\mu}_j, \text{ para algún } i \neq j$$

donde $\underline{\mu}_g$ es el vector de medias del grupo g con g = 1,...,p.

La estadística de prueba λ de Wilks se calcula como el cociente del determinante de la matriz de varianzas y covarianzas dentro de grupos entre el determinante de la matriz de varianzas y covarianzas total; esto es

$$\lambda = \frac{|\Sigma|}{|\Sigma + B|} = \frac{|\Sigma|}{|T|},$$

donde, Σ es la matriz de varianzas y covarianzas dentro de grupos
B es la matriz de varianzas y covarianzas entre grupos
T es la matriz de varianzas y covarianzas total

Esta estadística toma valores entre cero y uno, mientras más cercano sea su valor a cero, esto indicará que se tiene un mejor poder discriminante. Por lo que, se rechaza H_0 para valores pequeños de λ; es decir, cuando se tiene que la matriz de varianzas y covarianzas entre grupos, B, es grande.

Cuando se tienen dos grupos (Π_0 y Π_1) y cualquier número de variables explicativas, la transformación utilizada para aproximar la estadística λ a una estadística F, cuya distribución es una F de Fisher con k y n-k-1 grados de libertad, donde n es el número de observaciones considerando los dos grupos, y k es el número de variables en la función discriminante, es

$$F = \frac{(n-k-1)}{k} \frac{1-\lambda}{\lambda} \sim F_{(k,n-k-1)}$$

La regla de decisión de esta prueba consiste en rechazar la hipótesis nula H_0: $\underline{\mu_0} = \underline{\mu_1}$ al nivel de significancia α si $F < F^{1-\alpha}_{(k,n-k-1)}$, es decir con un nivel de $(1-\alpha)*100\%$ de confianza el vector de medias de ambos grupos es diferente, por lo que se considera que existe una diferencia significativa entre los dos grupos.

B.2 Prueba no paramétrica Mann-Whitney

La prueba no paramétrica Mann-Whitney para probar igualdad de poblaciones con muestras independientes, presenta las siguientes hipótesis:

$$H_0: F(x) = G(x) \text{ para toda } x \quad Vs \quad H_a: F(x) \neq G(x), \text{ para alguna } x,$$

donde, F(x) y G(x) son las funciones de distribución de cada una de las dos poblaciones.

Esta prueba se basa en los siguientes supuestos:

- Las muestras de ambas poblaciones son aleatorias.
- Además de la independencia dentro de la muestra debe existir independencia entre las dos muestras.
- Los valores deben de estar dados en escala ordinal.

Comúnmente esta prueba se plantea utilizando las siguientes hipótesis, en términos de las medias de ambas poblaciones

$$H_0: E(X) = E(Y) \quad Vs \quad H_a: E(X) \neq E(Y),$$

donde, X y Y son las variables aleatorias de cada una de las dos poblaciones.

Para ello, se hace uso adicionalmente del siguiente supuesto:

- Si existe alguna diferencia entre la función de distribución de ambas poblaciones, esta diferencia se dará en la localización de la distribución, es decir, si F(x) no es idéntica a G(x), entonces F(x) es idéntica a G(x+c), donde c es una constante.

La prueba Mann-Whitney consiste en ordenar de menor a mayor los datos de ambos grupos y asignar rangos a los valores del más pequeño al más grande, sin importar a que grupo pertenecen. Entonces, la estadística de prueba para cuando no existen demasiados empates al asignar los rangos es

$$T = \sum_{i=1}^{n_1} R(x_i),$$

donde $R(x_i)$ es el rango asignado a la observación i, y

n_1 es el número de elementos de la población 1.

Si se hace uso de los cuantiles superiores, T' se utiliza como alternativa de T, y se define como

$$T' = n_1 (n+1) - T,$$

donde n es e número de elementos considerando todas las poblaciones.

Una aproximación del nivel de significancia descriptivo, puede darse como:

$$n.s.d. = 2 * P\left(Z < \frac{t + \frac{1}{2} - n_1 \frac{n+1}{2}}{\sqrt{\frac{n_0 n_1 (n+1)}{12}}} \right),$$

donde $t = \min(T, T')$

La regla de decisión consiste en rechazar la hipótesis nula H_0, al nivel de significancia α, si $T < w_{\alpha/2}$ ó $T > w_{1-\alpha/2}$, donde w es un cuantil de la distribución de la estadística de prueba bajo H_0. Es decir, se rechaza la hipótesis nula H_0 si el nivel de significancia descriptivo (n.s.d.) es menor al nivel significancia utilizado α.

Los cuantiles superiores, es decir w_{1-p}, para T pueden ser obtenidos de la siguiente relación:

$$W_{1-p} = n_1 (n+1) - w_p$$

Para cuando no existen empates, o estos son muy pocos, la aproximación de los cuantiles para n_0 o n_1 mayores a 20, es la siguiente

$$w_p \cong \frac{n_1(n+1)}{2} + z_p \sqrt{\frac{n_0 n_1 (n+1)}{12}} .$$

APÉNDICE C

C.1 Variables explicativa utilizada por BarNiv y Hershbarger en 1990

- I_2 - Net gain to total income
- I_3 - Commissions and other expense to premiums
- I_5 - Nonadmitted to admitted assets
- I_{10} - Change in Product Mix
- I_{11} - Change in Assets Mix
- Medidas de descomposición para activos y pasivos

$$DM_s = \sum_i Q_i \left(\ln \frac{Q_i}{P_i} \right)$$

$$NDM_s = \sum_i Q_i \left| \ln \frac{Q_i}{P_i} \right|$$

donde, i es cada uno de componentes de s para s = pasivos, activos, i=1,..,n con n el número de componentes,

Q_i es la proporción relativa de del componente i entre el total de la hoja de balance del año actual, $0 \leq Q_i$

P_i es la proporción relativa de del componente i entre el total de la hoja de balance de un año previo, $P_i \leq 1$

- Size Surplus
- Size Premiums
- Size Assets
- Ln(GRA)= Ln of Growth in Assets = Ln (assets $_t$ / assets $_{t-1}$)
- GP - Gains to Premium
- I_4 a measure of gain

- P/S - Premium to surplus

La adecuación de estas variables explicativas a la práctica contable en México es la siguiente:

- I_2 es un indicador financiero del IRIS definido por

$$I_2 = \frac{\text{Utilidad Neta}}{\text{Prima Retenida}},$$

este indicador, conocido como rentabilidad de venta, se refiere al rendimiento obtenido por cada peso de prima retenida y sugiere problemas de insolvencia en caso de que este tome valores pequeños o negativos, debido a que esto significaría una falta de eficiencia administrativa.

- I_3 es un indicador financiero del IRIS definido por

$$I_3 = \text{Índice Combinado},$$

es decir,

$$I_3 = \frac{\text{Costo Neto Siniestralidad}}{\text{Prima Retenida Devengada}} + \frac{\text{Gasto Neto Operación}}{\text{Prima Directa}} + \frac{\text{Gasto Neto Adqusición}}{\text{Prima Retenida}},$$

este indicador sugiere problemas de insolvencia cuando toma valores cercanos a uno o mayores, debido a que estaría diciendo que gran parte del ingreso por primas es destinado a cubrir los gastos en que incurre la compañía aseguradora.

- I_5 es un indicador financiero del IRIS definido por

$$I_5 = \frac{\text{Activos Afectos a Cobertura de Reservas Técnicas}}{\text{Reservas Técnicas}},$$

éste muestra que la compañía tiene problemas de insolvencia si tomo valores menores a uno ya que podría suceder que ésta no pueda pagar sus compromisos porque las inversiones destinadas para ello no son suficientes.

- I_{10} Cambio en mezcla de productos, definido por

$$I_{10} = \frac{\text{Prima Producto } i_t - \text{Prima Producto } i_{t-1}}{\text{Prima Producto } i_{t-1}},$$

este indicador sugiere que la compañía podría tener problemas de insolvencia cuando toma valores grandes, ya que a mayor cambio la compañía tiene una menor estabilidad.

- I_{11} Cambio en mezcla de activos, definido por

$$I_{11} = \frac{\text{Activos}_t - \text{Activos}_{t-1}}{\text{Activos}_{t-1}},$$

esta variable indica que a mayor cambio se tiene una mayor probabilidad de insolvencia, esto debido a que la empresa no es estable.

- Medidas de descomposición para activos y pasivos

$$DM_s = \sum_i Q_i \left(\ln \frac{Q_i}{P_i} \right),$$

$$NDM_s = \sum_i Q_i \left| \ln \frac{Q_i}{P_i} \right|,$$

donde i es cada uno de componentes de s, para s igual a pasivos o activos, i = 1,..,n con n el número de componentes, Q_i es la proporción relativa de del componente i entre el total de la hoja de balance del año actual, $0 \leq Q_i$, y P_i es la proporción relativa de del componente i entre el total de la hoja de balance de un año previo, $P_i \leq 1$.

En caso de que estas medidas tomen valores grandes, implica que hay peligro financiero, debido a que es un indicador de que no existe estabilidad en la empresa.

- Volumen de variables

  Volumen del Capital Contable,
  Volumen de Primas,
  Volumen de activos,

  estas variables sugieren que existen problemas de insolvencia cuando toman valores pequeños, ya que la experiencia ha mostrado que las compañías pequeñas son las más propensas a ser insolventes.

- Ln(GRA) este índice permite observar la velocidad de crecimiento de las compañías, se define como

$$\text{Ln(GRA)} = \ln\left(\frac{\text{Activos}_t}{\text{Activos}_{t-1}}\right),$$

  por lo tanto, esta variable sugiere problemas de insolvencia cuando toma valores grandes, lo que estaría indicando un rápido crecimiento de la empresa, por lo que está es más vulnerable a tener problemas financieros.

- GP se define como el porcentaje de ganancias del total de las primas

$$GP = \frac{\text{Utilidad de Operación}}{\text{Prima Retenida}},$$

este índice sugiere problemas de insolvencia cuando toma valores pequeños o negativos, ya que significa que los ingresos por primas disminuyen considerablemente después de destinar parte de estos al pago de costos incluyendo los de operación, es decir la compañía aseguradora tiene costos altos por lo que la rentabilidad por cada peso de prima retenida es pequeña. Por otra parte, también es útil comparar este indicador con el promedio del mercado.

- I_4 Comportamiento de inversiones, se define como

$$I_4 = \frac{\text{Productos Financieros}}{\left(\frac{\text{Inversiones}_t + \text{Inversiones}_{t+1}}{2}\right)},$$

este índice, conocido como rentabilidad financiera, representa el rendimiento obtenido por el total de recursos productivos o resultado integral del financiamiento, por lo tanto si toma valores pequeños significa que los recursos de la compañía no se están siendo invertidos adecuadamente, lo que podría ocasionar la insolvencia de la compañía aseguradora. Este indicador puede ser comparado con los CETES.

- P/S se define como el crecimiento de la prima con respecto al capital contable, es decir

$$P/S = \frac{\text{Prima Retenida}}{\text{Capital Contable}},$$

Este índice, también conocido como Índice de Riesgo Neto, mide la exposición del capital contable en términos de la emisión de prima retenida, por lo que sugiere problemas de insolvencia si toma valores grandes, es decir, si el monto de las primas retenidas no se equilibra con el capital contable de la empresa aseguradora.

C.2 Variables explicativas utilizadas por BarNiv y McDonald en 1992

- X_{10} Net income/total assets, where net income comprises underwriting profits, net investment income, and other investment gains(losses)
- X_{20} Surplus
- X_{29} Net income/surplus
- X_{35} Mean/standar desviation of an operating ratio. The operating ratio is identified as:1 - CTR+(NII+OIG)/NPW, where CTR=combined ratio (loss ratio plus expense ratio), NII=net investment income, OIG=other investment gains(losses) and NPW=net premium written
- X_{37} Mean/standar desviation of an operating ratio. The operating ratio is identified as: 1-[(UE+LE-NII-OIG)/surplus], where UE=Underwriting expenses, and LE=loss expenses
- X_{42} Liability decomposition = $\sum_{i}^{k} Q_i \left(\ln \frac{Q_i}{P_i} \right)$
- X_{43} Liability decomposition = $\sum_{i} Q_i \left| \ln \frac{Q_i}{P_i} \right|$

La adecuación de estas variables explicativas a la práctica contable en México es la siguiente:

- X_{10}

$$X_{10} = \frac{\text{Utilidad Neta}}{\text{Activo}},$$

en caso de que este índice tome valores grandes, significa que con relación a los activos los ingresos de la compañía son cada vez mayores, por lo tanto, si esta variable toma valores grandes, podría ser un indicador de inestabilidad.

- X_{20}

$$X_{20} = \text{Volumen de capital contable},$$

esta variable sugieren que existen problemas de insolvencia cuando toman valores pequeños, ya que la experiencia ha mostrado que las compañías pequeñas son las más propensas a ser insolventes.

- X_{29}

$$X_{29} = \frac{\text{Utilidad Neta}}{\text{Capital Contable}},$$

este índice, conocido en como rentabilidad de capital, mide el porcentaje de rendimiento sobre los recursos invertidos y generados por las operación de cada institución, en caso de que tome valores negativos o pequeños significa que pueden existir problemas de insolvencia.

- X_{35}

$$X_{35} = \frac{\text{Media}\left[\left(1+\frac{\text{Poductos Financieros}}{\text{Prima Retenida}}\right)-\text{Índice Combinado}\right]}{\text{Desviación Estandar}\left[\left(1+\frac{\text{Poductos Financieros}}{\text{Prima Retenida}}\right)-\text{Índice Combinado}\right]}$$

Esta variable es una medida estandarizada con el objeto de quitar el efecto del tiempo. Representa la proporción en que los ingresos totales cubren los costos. Por lo tanto, en caso de que tome valores negativos significa que pueden existir problemas de insolvencia debidos a los altos costos en los que incurre la compañía.

- X_{37}

$$X_{37} = \frac{\text{Media}\left[\left(1+\frac{\text{Poductos Financieros}}{\text{Capital Contable}}\right)-\frac{\text{Gastos Siniestralidad}}{\text{Capital Contable}}\right]}{\text{Desviación Estandar}\left[\left(1+\frac{\text{Poductos Financieros}}{\text{Capital Contable}}\right)-\frac{\text{Gastos Siniestralidad}}{\text{Capital Contable}}\right]}$$

- X_{42} Medida de descomposición de activos o pasivos, definida como

$$DM_s = \sum_i Q_i \left(\ln \frac{Q_i}{P_i}\right).$$

- X_{43} Medida de descomposición de activos o pasivos, definida como

$$NDM_s = \sum_i Q_i \left|\ln \frac{Q_i}{P_i}\right|,$$

donde i es cada uno de componentes de s, para s igual a pasivos o activos, i = 1,..,n con n el número de componentes, Q_i es la proporción relativa de del componente i entre el total de la hoja de balance del año actual, $0 \leq Q_i$, y P_i es la proporción relativa de del componente i entre el total de la hoja de balance de un año previo, $P_i \leq 1$.

En caso de que estas medidas de descomposición tomen valores grandes, implica que hay peligro financiero, debido a que es un indicador de que no existe estabilidad en la empresa.

Cabe mencionar que

Prima Emitida = Prima Tomada + Prima Directa.

Al ceder parte de la Prima Emitida en reaseguro obtenemos la Prima Retenida, es decir

Prima Retenida = Prima Emitida - Prima Cedida.

La Prima Retenida se utiliza parcialmente para incrementar las reservas técnicas (pasivos) tales como:

- Reserva de Riesgos en Curso,
- Reserva de Siniestros Ocurridos,
- Reserva de Previsión, (utilizada para cubrir las desviaciones en la siniestralidad),

el resto constituye la Prima Retenida Devengada.

Por otro lado,

Capital Contable = Activos - Pasivos,

donde los activos pueden ser bienes inmuebles, disponibilidad en efectivo, deudores e inversiones a corto y largo plazo, y los pasivos pueden ser las reservas técnicas y los acreedores.

C.3 Variables Explicativas Utilizada por BarNiv et al en 1999

- NPWSURP se define como el crecimiento de la prima con respecto al capital contable, es decir

$$\text{NPWSURP} = \frac{\text{Prima Retenida}}{\text{Capital Contable}},$$

este índice, también conocido como Índice de Riesgo Neto, mide la exposición del capital contable en términos de la emisión de prima retenida, por lo que sugiere problemas de insolvencia si toma valores grandes, es decir, si el monto de las primas retenidas no se equilibra con el capital contable de la empresa aseguradora.

- LARAT es un índice de responsabilidad, y se define como sigue

$$\text{LARAT} = \frac{\text{Pasivos}}{\text{Disponibilidad} + \text{Inversiones}},$$

este índice sugiere problemas de insolvencia cuando toma valores grandes, esto es en el caso de que los pasivos sean mayores que las inversiones y activos disponibles, ya que se podría llegar a un punto en el que no se contasen con recursos suficientes para hacer frente a las responsabilidades del asegurador.

- LOSRAT es un índice de siniestralidad, que se construye como:

$$\text{LOSRAT} = \frac{\text{Costo Neto de Siniestralidad}}{\text{Prima Retenida Devengada}},$$

este índice representa la proporción de ingresos por primas que se destinan a cubrir los costos de siniestralidad de la compañía, cuando este índice toma valores grandes, es decir, cuando los gastos en que se está incurriendo son mayores a la prima devengada retenida, esto indica que los ingresos del asegurador son menores a sus gastos, lo que puede ocasionar problemas de insolvencia, ya que no es equiparable la emisión con los costos de la compañía.

C.4. Variables explicativas utilizada por Lee y Urrutia en 1996

- Leverage Ratio

  NWP/S = ratio of net premiums written to surplus.

- Profitability Ratios

  OM = Operating Margin, defined as the ratio of net operating income to premiums earned.

  RPS = Return to policyholders' surplus.

- Liquidity Premium

  CL= Current liquidity ratio.

- Product Mix Variables

  W_A = Ratio of all auto lines net premiums written to total net premiums written.

  W_W = Ratio of net premiums written for workers´ compensation, medical malpractice, and other liabilities to total net premiums written.

  LP = Ratio of total loss reserve to total net premiums written.

- Asset Mix Variable

  I_B= ratio of market value of invested bonds to total admitted assets.

- Growth Ratios

GRS= Rate of growth of statutory surplus.

GRP= Rate of growth of net premiums written.

- Other Insurers' Characteristics

MUT= A dummy variable equal to one if the insurer is a mutual company an zero otherwise.

DWR= A dummy variable equal to one if the insurer is a direct writer an zero otherwise.

La adecuación de estas variables explicativas a la práctica contable en México es la siguiente:

- Índice de Apalancamiento

NWP/S se define como el crecimiento de la prima con respecto al capital contable, es decir

$$NPW / S = \frac{\text{Prima Retenida}}{\text{Capital Contable}},$$

esta variable, también conocido como Índice de Riesgo Neto, mide la exposición del capital contable en términos de la emisión de prima retenida, por lo que sugiere problemas de insolvencia si toma valores grandes es decir, si el monto de las primas retenidas no se equilibra con el capital contable de la empresa aseguradora.

- Índices de Rentabilidad

OM es el margen de operación, y se define como:

$$OM = \frac{\text{Utilidad de Operación}}{\text{Prima Retenida}},$$

este índice mide que tan rentable es la empresa en relación a su emisión y productividad, por lo que, este índice sugiere problemas de insolvencia cuando toma valores pequeños o negativos, ya que significa que los ingresos por primas disminuyen considerablemente después de destinar parte de estos al pago de costos incluyendo los de operación, es decir, la compañía aseguradora tiene costos altos por lo que la rentabilidad por cada peso de prima retenida es pequeña.

RPS, este índice se define como

$$RPS = \frac{\text{Productos Financieros}}{\text{Capital Contable}},$$

este índice mide que tan productivo es la empresa en relación al manejo de sus inversiones, sugiere que la compañía puede tener problemas de insolvencia cuando toma valores pequeños, debido a que sus ingresos por inversiones son pequeños con respecto al capital contable de la compañía al inicio del año; por ello, la compañía puede no ser rentable y no contar con recursos suficientes para responder a todas sus obligaciones.

- Liquidez

CL es el índice de liquidez, y mide la proporción que guardan los activos a corto plazo con respecto a las reservas técnicas, es decir

$$CL = \frac{\text{Activos a Corto Plazo}}{\text{Reserva Técnicas a Corto Plazo}},$$

este índice sugiere problemas de insolvencia cuando toma valores menores a uno, ya que señala que no se cuenta con recursos disponibles para hacer frente a las responsabilidades a corto plazo de una compañía.

- Variables de Mezcla de Productos

W_A es una variable que indica el porcentaje de participación del ramo de automóviles, de una compañía aseguradora, es decir,

$$W_A = \frac{\text{Primas Retenidas del Ramo de Autos}}{\text{Total de Primas Retenidas}},$$

W_W es una variable que muestra el porcentaje de participación del ramo de responsabilidad civil, de una compañía aseguradora, es decir,

$$W_W = \frac{\text{Primas Retenidas de Responsabilidad Civil}}{\text{Total de Primas Retenidas}},$$

LP variable definida como:

$$LP = \frac{\text{Reservas Obligaciones Cotractuales}}{\text{Prima Retenida}},$$

En México, las reservas técnicas únicamente se comparan con las inversiones que las cubren, sin embargo, este indicador se aplica en Estados Unidos para detectar insolvencia de compañías aseguradoras, por lo que sería interesante analizar su aplicación en México.

- Variable de Mezcla de Activos

I_B este índice se construye como

$$I_B = \frac{\text{Bonos de Inversión}}{\text{Activos Afectos}},$$

los activos afectos son aquellos activos que respaldan las reservas, ejemplos de estos son: bonos, inversiones a corto plazo, efectivo y sus equivalentes, activos invertidos, deudores por primas, entre otros; es decir, dentro de los activos afectos no son considerados los bienes que no dan rendimientos como son los bienes inmobiliarios. Por ello, este índice es el porcentaje de activos afectos que corresponden a bonos de inversión, con los que cuenta una compañía para hacer frente a obligaciones a corto plazo.

- Índices de Crecimiento

GRS tasa de crecimiento del capital estatutario, definida como

$$GRS = \frac{\text{Capital Estatutario}_t - \text{Capital Estatutario}_{t-1}}{\text{Capital Estatutario}_{t-1}},$$

el capital estatutario es aquel conformado por las aportaciones a que se obligan los socios de una compañía. Una vez fijado el capital estatutario, para aumentarlo o disminuirlo es necesario convocar a una asamblea de socios para acordar la reforma de los estatutos. La disminución en el capital estatutario se podría llevar a cabo cuando el capital social con el que se está trabajando sea superior al requerido por las necesidades de la empresa, lo que implicaría que parte de ese capital sea improductivo, esta disminución se podrá realizar siempre y cuando no se reduzca en menos del mínimo legal. Un aumento en el

capital estatutario podría deberse a que la compañía necesita recursos, por consiguiente un mayor crecimiento del capital estatutario indica insuficiencia de recursos, lo que podría mostrar problemas de insolvencia.

GRP tasa de crecimiento de las primas retenidas, definida como:

$$GRP = \frac{\text{Primas Retenidas}_t - \text{Primas Retenidas}_{t-1}}{\text{Primas Retenidas}_{t-1}},$$

un mayor crecimiento o decremento de las primas retenidas indican falta de estabilidad en las operaciones de la compañía aseguradora, debido a que un incremento en las primas retenidas pueden indicar que la compañía aseguradora tuvo una entrada de capital repentina que puede deberse a que está intentando incrementar su ingreso en efectivo para satisfacer el pago de los siniestros. Por esto, a mayor tasa de cambio se tiene mayor tendencia a ser insolvente.

- Otras Características de los Aseguradores

MUT = 1 si el asegurador es una compañía mutualista,
MUT = 0 en otro caso.

DWR = 1 si el asegurador realiza sus ventas de forma directa,
DWR = 0 en otro caso.

C.5 Variables explicativas utilizadas por Kim et al en 1995

Las variables analizadas para el estudio de la insolvencia en el seguro de daños-responsabilidad civil, definidas de acuerdo a la práctica contable en los Estados Unidos son

- AGE = Age Effect
- CHNPW = Premium Growth
- YIELD = Investment performance
- COMBR = Underwriting results
- EXP = Expense
- RES = Loss Reserve
- REI = Reinsurance
- RCG = Realized Capital Gains
- UCG = Unrealized Capital Gains
- EXPANSION = Expansion to Other States

Las variables analizadas para el estudio de la insolvencia en el seguro de vida, definidas de acuerdo a la práctica contable en los Estados Unidos.

- AGE = Age Effect
- CHNPW = Premium Growth
- YIELD = Investment performance
- RCG = Realized Capital Gains
- UCG = Unrealized Capital Gains
- NOM = Net Operating Margin
- JUNK = Junk Bond
- CM =Commercial Mortgages
- RE = Real Estate

- EXP = Expense

La adecuación de estas variables explicativas a la práctica contable en México es la siguiente:

- Edad

$$\text{Edad} = \text{Año de Observación} - \text{Año que Comenzó a Operar},$$

a mayor edad, se espera que la probabilidad de insolvencia disminuya, ya que la compañía se ha podido mantener estable durante muchos años.

- Crecimiento de primas, definido como

$$\text{Crecimiento de Primas} = \frac{(\text{Prima Retenida}_t - \text{Prima Retenida}_{t-1})}{\text{Prima Retenida}_{t-1}},$$

un mayor crecimiento o decremento de las primas retenidas indican falta de estabilidad en las operaciones de la compañía aseguradora, debido a que un incremento en las primas retenidas pueden indicar que la compañía aseguradora tuvo una entrada de capital repentina que puede deberse a que la compañía está intentando incrementar su ingreso en efectivo para satisfacer el pago de los siniestros. Por esto, a mayor tasa de cambio se tiene mayor tendencia a ser insolvente.

- Comportamiento de inversiones, se define como

$$\text{Comportamiento de Inversiones} = \frac{\text{Productos Financieros}}{\left(\frac{\text{Inversiones}_t + \text{Inversiones}_{t+1}}{2}\right)},$$

este índice, conocido como rentabilidad financiera, representa el rendimiento obtenido por el total de recursos productivos, por lo tanto si toma valores pequeños significa que los recursos de la compañía no se están siendo invertidos adecuadamente, lo que podría ocasionar la insolvencia de la compañía aseguradora.

- Índice Combinado

$$IC = \left[\frac{\text{Gastos Siniestralidad}_t}{\text{Prima Devengada}_t} + \frac{\text{Gatos Operación}_t}{\text{Prima Directa}_t} + \frac{\text{Gatos Adquisición}_t}{\text{Prima Retenida}_t} \right],$$

este indicador sugiere problemas de insolvencia cuando toma valores cercanos a uno o mayores, debido a que estaría diciendo que gran parte del ingreso por primas es destinado a cubrir los gastos en que incurre la compañía aseguradora, por lo que el resto puede llegar a ser no suficiente para cubrir las obligaciones de la compañía.

- Gastos

$$\text{Gatos} = \frac{\text{Gatos Adquisición}_t}{\text{Priama Retenida}_t},$$

en caso de que este índice se grande implica que existen problemas administrativos, lo que a lo largo del tiempo podrían provocar la insolvencia de compañías, debido a que los gatos administrativos son altos en relación con la prima retenida

- Reserva Obligaciones Contractuales

$$\text{Reserva Obligacioens Contractuales} = \frac{\text{Reserva Obligacioenes Contractuales}_t}{\text{Capital Contable}_t},$$

este índice mide el grado en que los recursos de los accionistas se encuentran comprometidos con las obligaciones pendiente por cumplir, por lo que en caso de que este índice tomo valores grandes podrían tenerse problemas de insolvencia.

- Reaseguro

$$\text{Re aseguro} = \frac{\text{Siniestro Retenidos por Reaseguro Tomado}_t}{\text{Capital Contable}_t}.$$

Este indicador mide el impacto de los compromisos de otras compañías en relación con el capital contable de la empresa.

- Expansión a Otros Estados

$$\text{Expansión Otros Estados} = \frac{\text{Número de Estados donde el Asegurador tiene Licencia}}{\text{Capita Contable}_t},$$

cuando un asegurador se expande a un nuevo territorio, sin contar con recursos adecuados para soportar su expansión, implica que se pueden desarrollar problemas debido a la pérdida de control sobre las operaciones de seguros.

- Margen de Operación

$$\text{Margen Operación} = \frac{\text{Utilidad de Operación}_t}{\text{Prima Retenida}_t},$$

este índice sugiere problemas de insolvencia cuando toma valores pequeños o negativos, ya que significa que los ingresos por primas disminuyen considerablemente después de destinar parte de estos al pago de costos incluyendo los de operación, es decir la compañía aseguradora tiene costos altos por lo que la rentabilidad por cada peso de prima retenida es pequeña.

- Junk Bond.- No se aplican en México y en EU ya no existen.

- Inmuebles

$$\text{Inmobiliario} = \frac{\text{Inmuebles}_t}{\text{Capital Contable}_t}.$$

APÉNDICE D

En esta parte se muestran las compañías de cada una de las dos poblaciones, junto con el valor que presentan los indicadores financieros en marzo de 2000.

Compañías de Π_0 (sanas)

Obs	Gpo	Primer trimestre 2000		
		CRT	MCG	CMP
2	0	1	1.95	14.41
3	0	1.49	6.86	13.29
4	0	1.4	15.6	2.9
5	0	1.61	24.64	1.95
6	0	1.21	2.5	3.3
7	0	1.66	4.78	11.04
8	0	1.24	8.19	1.89
9	0	1.2	6	4.9
10	0	1.06	1.06	2.38
11	0	1	1.4	3
12	0	1.2	3.2	1.2
13	0	1.4	1.26	1.53
14	0	1.48	28.86	5.87
15	0	1.04	N/D	31
16	0	1.52	11.68	12.31
17	0	1.19	2.2	4.7
19	0	1.21	1.73	1.44

Obs	Gpo	Primer trimestre 2000		
		CRT	MCG	CMP
20	0	1.09	4.9	1.13
21	0	1.04	1.9	6
26	0	1.5	7.47	2.6
22	0	1.01	2.32	2.4
27	0	1	1.5	8.7
28	0	1.07	2.88	3.11
29	0	1.03	1.55	36.14
30	0	1.18	14.23	2.6
31	0	1.1	32.3	5.2
32	0	1	2	8.4
34	0	1	1.6	51.3
35	0	4.04	N/D	1.67
36	0	1.13	2.2	4.24
37	0	1.16	6.53	98.57
39	0	1.38	3.64	1.95
40	0	1.5	10.3	2.7
41	0	2.87	10.96	20.61
42	0	1.19	3	15.6
43	0	1.48	5.57	1.91
45	0	1.32	2	2.9
46	0	1.34	5.36	7.04
48	0	1.02	47.7	3.34
49	0	1.06	67.4	4.2
52	0	1.9	45.5	2.5
53	0	1.7	N/D	2.43

Compañías de Π_1 (con problemas financieros)

Obs	Gpo	Primer trimestre 2000		
		CRT	MCG	CMP
1	1	0.96	0.92	1.46
18	1	1.02	1.14	0.3
23	1	0.9	0.5	0.4
24	1	0.99	1.16	15.1
33	1	0.93	0.02	0.3
38	1	0.99	0.57	2.38
44	1	0.98	-641.04	1.14
47	1	0.95	1.94	0.61
50	1	0.7	0.3	0.7
51	1	2.6	24.8	0.81
54	1	1	0.9	1.11

Nota: N/D se refiere a que el indicador no está disponible, de acuerdo a la CNSF.

APÉNDICE E

En esta parte, se definen las variables que fueron utilizadas dentro del análisis de datos de compañías aseguradoras que operan en México.

$$PR\_CC_t = \frac{\text{Prima Retenida}_t}{\text{Capital Contable}_t}$$

$$I\_SI_t = \frac{\text{Costo Neto de Siniestralidad}_t}{\text{Prima Retenida Devengada}_t}$$

$$IC_t = \frac{\text{Costo Neto Siniestralidad}_t}{\text{Prima Retenida Devengada}_t} + \frac{\text{Costo Neto Adquisición}_t}{\text{Prima Retenida}_t} + \frac{\text{Costo Neto Operación}_t}{\text{Prima Directa}_t}$$

$$CRP_t = \frac{\text{Prima Retenida}_t - \text{Prima Retenida}_{t-1}}{\text{Prima Retenida}_{t-1}}$$

$$L\_CA_t = \ln\left(\frac{\text{Activos}_t}{\text{Activos}_{t-1}}\right)$$

$$R\_CC_t = \frac{\text{Reserva de Obligaciones Contractuales}}{\text{Capital Contable}}$$

$$R\_PR_t = \frac{\text{Reserva de Obligaciones Contractuales}}{\text{Prima Retenida}}$$

$$RE\_F_t = \frac{\text{Productos Financieros}_t}{\left(\frac{\text{Inversiones}_{t-1} + \text{Inversiones}_t}{2}\right)}$$

$$RE\_V_t = \frac{\text{Utilida Neta}_t}{\text{Prima Retenida}_t}$$

$$X\_10_t = \frac{\text{Utilida Neta}_t}{\text{Activos}_t}$$

$$RE\_C_t = \frac{\text{Utilida Neta}_t}{\text{Capital Contable}_t}$$

$$MIC_t = 1 + \frac{\text{Producto Financieros}_t}{\text{Prima Retenida}_t} - IC_t$$

$$MO_t = \frac{\text{Utilidad de Operación}_t}{\text{Prima Retenida}_t}$$

$$CC_t = \text{Capital Contable}_t$$

$$LARA_t = \frac{\text{Pasivos}_t}{\text{Disponibilidad}_t + \text{Inversiones}_t}$$

BIBLIOGRAFÍA Y REFERENCIAS

Actualidad en Seguros y Fianzas. No.30, Comisión Nacional de Seguros y Fianzas, México, Octubre - Diciembre 1998.

Actualidad en Seguros y Fianzas. No.32, Comisión Nacional de Seguros y Fianzas, México, Junio de 1999.

Actualidad en Seguros y Fianzas. No.34, Comisión Nacional de Seguros y Fianzas, México, Diciembre de 1999.

AM Best Company. *Best's Key Rating Guide, Property-Casualty*, USA, 1990.

Ambrose, J.M. y Carroll A.M. Using Best´s Ratings in Life Insurer Insolvency Prediction. . *The Journal of Risk and Insurance*, Vol.61, No.2, 317-327, 1994.

Ambrose, J.M. y Seward, J.A. Bets´s Rating, Financial Ratios and Prior Probabilities in Insolvency Predictions. *The Journal of Risk and Insurance*, Vol.55, 229-244, 1988.

BarNiv, R., Hathorn, J., Mehrez, A. y Kline, D. Confidence Intervals for the Probability of Insolvency in the Insurance Industry. *The Journal of Risk and Insurance*, Vol.66, No.1, 125-137, 1999.

BarNiv, R. y Hershbarger, R.A. Classifying Financial Distress in the Life Insurance Industry. *The Journal of Risk and Insurance*, Vol.57, No.1, 110-136, 1990.

BarNiv, R. y McDonald, J.B. Identifying Financial Distress in the Insurance Industry: A Synthesis of Methodological and Empirical Issues. *The Journal of Risk and Insurance*, Vol. LIX, No.4, 543-574, 1992.

Baz, G.G. *Curso de Contabilidad de Sociedades*. México, 1962.

Breslow, N.E. Covariance Analysis of Censored Survival Data. *Biometrics*, Vol.30, 89-99, 1974

Collett, D. *Modeling Survival Data in Medical Research*. Chapman and Hall, London, 1994.

Conover, W.J. *Practical Nonparametric Statistics*, John Willey and Sons, Canada, 1980.

Cooley, P.L. Bayesian and Cost Considerations for Optimal Classification with Discriminant Analysis. *The Journal of Risk and Insurance*, Vol.42, 277-287, 1975.

Cox, D.R. *Regression Models and Life Time Tobles*. J. R. Stat. Soc. B, 34, 1972.

Cox, D.R. y Oakes, D. *Analysis of Survival Data*. Chapman and Hall, London, 1984.

Dillon, W.R. y Goldstein, M. *Multivariate Analysis, Methods and Applications*. John Wiley and Sons, Estados Unidos de América, 1984.

Dobson, A.J. *An Introduction to Generalized Linear Models.* Chapman and Hall, London, 1990.

Emmet, J. *Fundamentals of Risk and Insurance.* John Wiley and Sons, 1986.

García, D.O. *Razanes Financieras de Sector Asegurador,* Documento de Trabajo No.24, Comisión Nacional de Seguros y Fianzas, México, 1993.

Hand, D.J. *Discrimination and Classification.* John Willey and Sons, Estados Unidos de América, 1981.

Hershbarger, R.A. y Miller, R.K. The NAIC Information System and the Use of Economic Indicators in Predicting Insolvencies. *Journal of Insurance Issues and Practice,* Vol. 9, 21-43, 1986.

Hind, A. *Survival Models, Demographic Methods.*

Judge, G.G., Hill, R.C., Griffith, W.E., Lutkepohl, H. y Lee, T.C. *Introduction to the Theory and Proctice of Econometrics.* John Wiley and Son, New York, 1988.

Johnson, R. A. y Wichern, D. W. *Applied Multivariote Statistical Analysis.* Prentice-Hall, New Jersey, 1992.

Kalbfleisch, J.D. y Prentice, R.L. Marginal Likelihoods Based on Cox's Regression and Life Model. *Biometrika,* Vol 60, 267-279, 1973

Kiefer, N.M. Economic Durations Data and Hazard Functions. *Journal of Economic Literature,* Vol. XXVI, 646-679, 1988.

Kim, Y.D., Anderson, D.R., Amburgey, T.L. y Hickman, J.C. The Use of Event History Analysis to Examine Insurer Insolvencies. *The Journal of Risk and Insurance*, Vol.62, No.1, 94-110, 1995.

Klein, J.P. y Moeschberger, M.L. *Survival Analysis: Techniques for Censored and Truncated Data*, Springer-Verlag, New York, 1997.

Lawless, J.F. Confidence Interval Estimation in the Inverse Power Law Model. *Applic.Stat*, 25, 128-138, 1976.

Lawless, J.F. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons, Canada, 1982.

Lee, E.T. *Statistical Methods for Survival Data Anolysis*. John Wiley and Sons, Canada, 1992.

Lee, S.H. y Urrutia, J.L. Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models. *The Journal of Risk and Insurance*, Vol.63, No.1, 121-130, 1996.

Maddala, G.D. *Limited Dependence and Qualitative Variables in Econometrics*, Cambridge University Press, 1983.

Madrigal, A.M.G, *Tesis de Licenciatura: Análisis Bayesiano del Problema de Clasificación Estadística*. México, 1994.

Mardia, K.V., kent, J.T. y Bibby, J. M. *Multivariate Analysis*. Academic Press, London, 1995.

McCullagh, P. y Nelder, J.A. *Generalized Linear Models*. Chapman and Hall, Estados Unidos de América, 1989.

Mood, A.M., Graybill, F.A. y Boes, D.C. *Introduction to the Theory of Statistics*. McGraw-Hill, New York, 1989.

NAIC. *Insurance Regulatory Information System Ratio Results for 1997-1998*. Estados Unidos de América.

OECD, *Insurance Solvency Supervisión*. OECD Publications, France, 1995

Prentice, R.L. y Gloeckler, L.A. Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data. *Biometrics*, Vol.34, 57-67, 1978.

Pinches, G.E. y Trieschmann, J.S. The Efficiency of Alternative Models for Solvency Surveillance in the Insurances Industry. *The Journal of Risk and Insurance*, Vol.41, 563-577, 1974.

Pinches, G.E. y Trieschmann, J.S. Discriminant Analysis, Classification Results and Financially Distressed P-L Insurers. *The Journal af Risk and Insurance*, Vol.44, 289-298, 1977.

Seber, G.A.F. *Multivariate Observations*. John Wiley and Sons, New York, 1984.

Trieschmann, J.S. y Pinches, G.E. A Multivariate Model for Predicting Financially Distressed P-L Insurers. *The Journal of Risk and Insurance*, Vol.40, 327-338, 1973.

Zangwill, W. Minimizing a Function Without Calculating Derivatives. *Computer Journal*, 293-296, 1939.

Direcciones de Internet:

http://www.cnsf.gob.mx

http://www.naic.org

http://www.iaisweb.org/

http://www.ambest.com