



## **MAESTRÍA EN FINANZAS CORPORATIVAS**

**Trabajo de grado**

**Modelo de predicción de punto de quiebra de las empresas manufactureras  
Pymes en Colombia**

**Presentado por**

**Heymar Augusto Ruiz Cotrino**

**Director**

**Enrique ter Horst**

**Bogotá D.C., 29 de octubre de 2015**

## TABLA DE CONTENIDO

Introducción .....	7
<b>1. Estado del Arte.....</b>	<b>11</b>
<b>1.1. Modelos de predicción de la insolvencia empresarial de Alejandro Astorga Hilbert.....</b>	<b>11</b>
<b>1.2. Modelo de predicción de quiebras de las empresas colombianas – Jorge Rosillo .....</b>	<b>15</b>
<b>1.3. Determinantes de fragilidad en las empresas colombianas – Óscar Martínez</b>	<b>19</b>
<b>1.4. Modelo Z de Altman.....</b>	<b>20</b>
<b>1.5. Modelo Z1 de Altman.....</b>	<b>22</b>
<b>1.6. Modelo Z2 de Altman.....</b>	<b>23</b>
<b>1.7. Modelo de Gordon Springate.....</b>	<b>24</b>
<b>1.8. Modelo Fulmer .....</b>	<b>25</b>
<b>1.9. Modelo CA-Score.....</b>	<b>26</b>
<b>1.10. Modelo Ricardo Pascale.....</b>	<b>27</b>
<b>2. Marco Teórico .....</b>	<b>29</b>
<b>2.1. Punto de quiebra o insolvencia financiera empresarial .....</b>	<b>29</b>
<b>2.1.1. Régimen de insolvencia empresarial en Colombia – Ley 1116 de 2006</b>	<b>30</b>
<b>2.2. Razones financieras .....</b>	<b>33</b>
<b>2.2.1. Razones de liquidez.....</b>	<b>34</b>
<b>2.2.1.1. Capital de trabajo.....</b>	<b>34</b>
<b>2.2.1.2. Razón Corriente.....</b>	<b>35</b>
<b>2.2.1.3. Prueba ácida.....</b>	<b>35</b>
<b>2.2.2. Razones operacionales o de actividad.....</b>	<b>36</b>
<b>2.2.2.1. Rotación de cuentas por cobrar o período de cobro.....</b>	<b>36</b>
<b>2.2.2.2. Rotación de inventarios o período de reposición .....</b>	<b>37</b>
<b>2.2.2.3. Rotación de cuentas por pagar o período de pago: .....</b>	<b>38</b>

2.2.2.4.	Ciclo operacional.....	38
2.2.3.	Razones de endeudamiento o apalancamiento .....	39
2.2.4.	Razones de rentabilidad o rendimiento .....	39
2.2.4.1.	ROA (Return On Asset).....	39
2.2.4.2.	ROE (Return On Equity).....	40
2.2.4.3.	ROI (Return On Investment) .....	40
3.	Metodología .....	41
3.1.	Elaboración del modelo Logit .....	42
3.2.	Minería de datos .....	58
3.3.	Modelo LOGIT final .....	62
3.4.	Análisis de resultados Modelo LOGIT vs. Z1 Altman y CA-Store .....	64
	CONCLUSIONES.....	67
	Lista de referencias .....	71

## LISTA DE TABLAS

Tabla 1. Empresas seleccionadas por Alejandro Astorga.....	12
Tabla 2. Resultados Estudio Astorga.....	13
Tabla 3. Factores y comportamiento de default.....	42
Tabla 4. Datos de salida de la función LOGIT.....	50
Tabla 5. Resultados individuales de los factores.....	51
Tabla 6. Aplicación del modelo LOGIT a la muestra de empresas, con 5 indicadores financieros.....	52
Tabla 7. Prueba de restricciones conjuntas con el LR test.....	55
Tabla 8. Medidas estadísticas de las variables en el modelo Logit.....	59
Tabla 9. Winsorización de los valores de las variables.....	61
Tabla 10. Modelo Logit final después de winsorizar los datos de las variables.....	63
Tabla 11. Análisis modelo Logit para empresas insolventes.....	65
Tabla 12. Análisis modelo Logit para empresas solventes.....	66

## LISTA DE GRÁFICOS

Gráfico 1. Scores y probabilidades de default en el modelo Logit.....	44
Gráfico 2. Empresas que entran en liquidación judicial a partir de 2007 por sector.....	47

## LISTA DE ANEXOS

Anexo 1. Codificación del modelo en Visual Basic.....	75
---	----

## Introducción

Uno de los hechos más significativos en la práctica del desarrollo económico ha sido el descubrimiento del gran potencial de las pequeñas y medianas empresas, junto con su gran contribución en factores macroeconómicos como la generación de empleo y el Producto Interno Bruto (PIB) (Rodríguez, 2003).

La Encuesta Anual Manufacturera, elaborada por el DANE, permite valorar la incidencia de las MIPYME en el panorama empresarial colombiano, en donde éstas representan el 96,4% de los establecimientos, aproximadamente el 63% del empleo; el 45% de la producción manufacturera, el 40% de los salarios y el 37% del valor agregado. Son más de 650.000 empresarios cotizando en el sistema de seguridad social (Portal Businesscol.com, 2008).

Sin embargo, un estudio realizado por la Confederación Colombiana de Cámaras de Comercio (Confecámaras), basado en la información que recopila en el Registro Único empresarial y Social (RUES), señala que las Pymes no superan la edad madura y entran en estado de insolvencia o punto de quiebra tan sólo a los 12 años.

En este sentido, el estado de insolvencia empresarial se convierte en una de las principales amenazas para toda empresa. Este tipo de amenaza a la estabilidad financiera de las empresas puede llegar a afectar no sólo a

inversionistas y empleados, sino también a prestamistas del sector financiero, auditores y autoridades reguladoras, entre otros.

Es así como el interés de muchos financieros se ha volcado en poder contar con algún tipo de modelo matemático, con el que logren predecir el estado de insolvencia o punto de quiebra de las empresas con cierto nivel de anticipación. Esto ha llevado a diferentes expertos en la materia de todo el mundo, a trabajar desde diferentes perspectivas en la estimación de dichos modelos, a partir de diferentes razones financieras, consideradas como las más relevantes de acuerdo con el concepto, experiencia e investigación de cada uno de ellos.

En la medida en que se pueda encontrar un modelo financiero de predicción de estado de insolvencia que se ajuste a las pequeñas y medianas empresas en Colombia, las cuales son el gran motor, no sólo de nuestra economía sino de las economías mundiales, se les estaría brindando a estos empresarios una herramienta adicional para la toma de decisiones oportunas que le permitan, a través de la identificación de las razones financieras críticas, tomar las medidas tanto preventivas como correctivas, según sea el caso, con el fin de permitirles mantenerse en el tiempo.

Basados en lo anterior, el objetivo general de este trabajo de investigación es encontrar un modelo financiero multivariable que se ajuste a las características de las pequeñas y medianas empresas colombianas, de tal manera que le permita a los socios, dueños o accionistas de las mismas, predecir con al menos un año



de anticipación, la posibilidad de que su empresa entre en estado de insolvencia o punto de quiebra, basados en probabilidades de default.

Es importante mencionar que existen tanto modelos estructurales como no estructurales; en este último se encontraría el modelo Logit construido.

Para aclarar este término, Enrique ter Horst, en su paper publicado en 2014, *Inferencia Bayesiana para un modelo estructural de riesgo de crédito con volatilidad y tasas de interés estocásticas*, afirma que “los modelos estructurales de riesgo de crédito tienden a evaluar la solvencia de una empresa a través de la modelación del balance general de la firma como una cartera que consiste en el activo, las opciones implícitas y la deuda libre de riesgo, en la cual el valor de mercado de los activos debe sumarse al valor de mercado de los pasivos. Los modelos estructurales de riesgo de crédito se remontan a Merton quien aplicó el conocido modelo de valoración de opciones Black and Scholes”.

Para poder cumplir con el objetivo planteado, se han propuesto los siguientes objetivos específicos:

- Definir la muestra de las empresas en proceso de liquidación y de empresas consideradas como solventes, de acuerdo con la información reportada a la Superintendencia de Sociedades de Colombia.
- Consolidar los estados financieros de las empresas muestra, con el fin de calcular los indicadores financieros relevantes.

- Crear un método de regresión multivariable o Logit que permita determinar con mayor exactitud el punto de quiebra de las empresas manufactureras en Colombia.
- Correr los modelos conocidos, y aplicables a las empresas manufactureras, de predicción de insolvencia de las empresas colombianas, tales como el Z1 de Altman y el CA-Score, con el fin de estimar su precisión.
- Realizar una comparación entre el modelo creado y los mencionados en el punto anterior, de tal manera que permita determinar el de mayor precisión, identificar posibles fallas en el mismo y las opciones de mejora o recomendaciones para futuros estudios de investigación.

El presente trabajo de investigación está conformado por cuatro grandes secciones: la primera, Estado del Arte, realiza una descripción de los principales trabajos realizados en este campo; la segunda, Marco Teórico, contiene todo el soporte teórico y científico utilizado como insumo para el desarrollo y cumplimiento de los objetivos de este trabajo; una tercera parte, Metodología, que describe el procedimiento paso a paso utilizado para poder construir el modelo Logit, a partir de la muestra de empresas seleccionada y, por último, las conclusiones a las que se llega después de cumplir con el objetivo de construir el modelo Logit que sirva para la predicción del punto de quiebra, o la probabilidad de default, de las empresas manufactureras en Colombia. Para cerrar, se lista las diferentes referencias consultadas a lo largo del desarrollo del trabajo de investigación.

## **1. Estado del Arte**

Si bien es cierto, son diversas las investigaciones y los trabajos que se han realizado en torno a modelos financieros para la determinación del punto de quiebra o el estado de insolvencia financiera empresarial en el mundo, son pocos los que se han desarrollado para empresas colombianas.

Como ejemplo de estudios realizados en otros países, Grant Samkin (2012), estudió el uso del Z-Score de Altman para establecer el momento en que se debió predecir el colapso de las compañías financieras en Nueva Zelanda.

### **1.1. Modelos de predicción de la insolvencia empresarial de Alejandro Astorga Hilbert**

Dentro de los estudios internacionales realizados, se destaca el del ingeniero Alejandro Astorga Hilbert, secretario técnico del Instituto Federal de Especialistas de Concursos Mercantiles de México (IFECOM), quien seleccionó los modelos de predicción de insolvencia considerados como los más conocidos y confiables de ese momento, y con el fin de probar la aplicabilidad de cada uno de los modelos, obtuvo información financiera de 124 empresas mexicanas y se seleccionaron 44 de ellas para conformar una muestra en la que 22 de ellas han sido insolventes y

las otras 22 han sido solventes. La distribución por sector de dicha muestra se presenta en la tabla 1.

Tabla 1. Empresas seleccionadas por Alejandro Astorga

Sector	No. Empresas	%
Manufactura	26	59%
Comercio	9	20%
Servicios	5	11%
Construcción	3	7%
Comunicación	1	2%
Total	44	100%

Nota: Adaptado de "Modelos de predicción de la insolvencia empresarial", por Astorga, Alejandro, 2000. P. 4

Como empresas insolventes se tomaron aquellas que hubiesen incumplido en el pago de sus créditos y cayeron en cartera vencida con instituciones financieras. El resultado de esta muestra fue de 17 manufactureras, 1 de servicios, 1 comercial y 3 constructoras. En este caso, los modelos se aplicaron tomando estados financieros con fecha de un año previo al incumplimiento (Astorga, 2000).

Como empresas solventes se consideraron empresas que no hubiesen presentado problemas financieros graves, es decir, que no mostraran incumplimiento en el pago de sus obligaciones. La composición de esta muestra fue de 9 manufactureras, 4 de servicios, 8 comerciales y 1 de comunicación. Para esta muestra, se tomaron estados financieros con por lo menos un año de antigüedad, con el fin de tener la certeza que durante ese periodo no habían sido insolventes.

Posterior a esto, se aplicó cada uno de los modelos de predicción a cada una de las empresas y se comparó el resultado obtenido por el modelo contra la realidad. Los resultados de este estudio se resumen en la tabla 2.

Tabla 2. Resultados Estudio Astorga

Modelo	Insolventes		Solventes		Total	
	Aciertos	%	Aciertos	%	Aciertos	%
Z1 de Altman	18	82%	22	100%	40	91%
Z2 de Altman	17	77%	21	95%	38	86%
Springate	21	95%	16	73%	37	84%
CA-Score	21	95%	18	82%	39	89%
Fulmer	12	55%	19	86%	31	70%

Nota: Adaptado de "Modelos de predicción de la insolvencia empresarial", por Astorga, Alejandro, 2000. P. 5

De esta manera Astorga concluyó (Astorga, 2000):

“El modelo Z1 de Altman tuvo sólo el 82% de certeza en el caso de empresas insolventes, sin embargo, debe tomarse en cuenta que este modelo genera tres resultados posibles: alta probabilidad de insolvencia, baja probabilidad de insolvencia y un valor intermedio llamado "zona gris". Si el resultado es "zona gris", se debe interpretar como una empresa con posibilidades de caer en insolvencia aunque no necesariamente son altas. Si los resultados de "zona gris" se hubieran tomado como posibles casos de insolvencia, la precisión de este modelo sería del 95%. El modelo está diseñado para aplicarse a empresas manufactureras”.

El modelo Z2 de Altman al igual que el Z1 puede dar tres resultados diferentes y si se consideran los "zona gris" como casos de insolvencia, la precisión del modelo sería del 88%. Puede aplicarse a empresas de cualquier tipo.

El modelo Springate (Astorga 2000) resultó altamente confiable (95%) para pronosticar la insolvencia de las empresas, aunque para el caso de empresas solventes no fue tan preciso (73%).

El comportamiento del modelo CA-Score fue muy similar al Springate. Se debe tomar en cuenta que fue diseñado para empresas manufactureras.

El modelo Fulmer (Astorga, 2000) no resultó confiable para predecir la insolvencia.

A excepción de los modelos Fulmer y Z2 de Altman, los modelos estudiados tuvieron una alta precisión para predecir la insolvencia.

Por su parte, aunque en Colombia no se han replicado estos modelos, se han realizado algunos estudios, en su gran mayoría derivados de los modelos de Altman, que han permitido determinar diferentes criterios para la predicción del punto de quiebra de las empresas. De éstos se destaca el elaborado por el Ingeniero Industrial y Administrador de Empresas de la Universidad Javeriana,

Jorge Rosillo en 2002 y el desarrollado por Óscar Martínez en 2003, profesional especializado en seguimiento financiero del Departamento de Estabilidad Financiera del Banco de la República.

## **1.2. Modelo de predicción de quiebras de las empresas colombianas – Jorge Rosillo**

El objetivo de este estudio se basó en definir qué indicadores financieros permitirían predecir con mayor certeza una situación financiera difícil o detectar con anticipación una quiebra empresarial. Para el logro de este objetivo, utilizó el modelo propuesto por el profesor Altman aplicado al caso colombiano de manera que se generara una alerta que indicara de antemano cuándo una empresa presenta problemas y cómo puede el empresario tomar los correctivos del caso.

De acuerdo con Jorge Rosillo, según publicación realizada en la revista Innovar en 2.002, el discriminante tiene su origen en la regresión multivariable, pero a diferencia de esta, la variable dependiente es categórica u ordinal, es decir, que sirve para clasificar una población en diferentes categorías, utilizando varias variables independientes cuantitativas. El punto de partida es definir las variables que permitan discriminar la muestra. En el caso de Altman utilizó indicadores financieros, en donde el total de la muestra debió contener empresas fuertes y débiles diagnosticadas con base en estos indicadores.

Rosillo tomó una muestra de 106 empresas de diferentes sectores, 80 para diseñar el modelo y las 26 restantes para probarlo, a las cuales se les asignó una posición según indicadores financieros promedio, tales como: razón corriente, prueba ácida, razón de liquidez, endeudamiento, días periodo de cobro, días de inventario, rotación de activos, margen neto, rentabilidad del patrimonio, rentabilidad del activo, el leverage a corto plazo y el leverage a largo plazo.

Los resultados de dichos indicadores son la base para calcular la función discriminante; inicialmente se había asignado 12 ratios para calcular la función discriminante, el sistema paso a paso lo redujo a tres (endeudamiento, rentabilidad del patrimonio y leverage), y determinó la constante de 1,563 por medio del cálculo estadístico SPSS, utilizado por Jorge Rosillo en su modelo de predicción de quiebra, los cuales son los que realmente establecen la diferencia entre las empresas fuertes y débiles. Los datos obtenidos de las 106 empresas de acuerdo a su posición promedio en los estados financieros son los siguientes:

Endeudamiento:	-7,165
Rentabilidad del Patrimonio:	9,852
Laverage Largo Plazo:	1,097
Constante:	1,563

*Z = Función discriminante*

$$Z = -7,165X + 9,852X + 1,097X + 1,563$$



Esta función ayudaría a definir cuándo una empresa es fuerte o débil.

Inicialmente se analizaron los resultados de los indicadores financieros de las 106 empresas y de acuerdo con los parámetros considerados acerca de cuándo un indicador es bueno o malo, se definió que de las 106 empresas objeto de análisis, las primeras 65 pueden considerarse financieramente fuertes y el resto débiles. De las 106 empresas se tomó una muestra de 80 (50 fuertes y 30 débiles) para calcular los promedios de las razones discriminantes (Innovar, 2002).

De las primeras 65 empresas clasificadas como fuertes se tomaron 50 y se procedió a promediar sus indicadores obteniendo los siguientes resultados:

Endeudamiento:	0,26017006
Rentabilidad del Patrimonio:	0,06113952
Leverage largo plazo:	0,50273351
Constante:	1,563

Con base en estos resultados se calculó la función discriminante de las empresas financieramente fuertes:

$$Z_{fuertes} = -7,165(0,26017006) + 9,852(0,06113952) + 1,097(0,50273351) + 1,563$$

$$Z_{fuertes} = 0,85272677$$

De las 41 empresas que se consideraron débiles dentro de las 106 consideradas, se tomaron 30 y se calcularon los indicadores financieros promedio de la función discriminante:

Endeudamiento:	0,46342418
Rentabilidad del patrimonio:	-0,07814256
Leverage largo plazo:	1,00787766
Constante:	1,563

$$Z_{débiles} = -7,165(0,46342418) + 9,852(-0,07814256) + 1,097(1,00787766) + 1,563$$

$$Z_{débiles} = -1,42165289$$

Con base en los resultados del Z para las empresas fuertes y débiles obtenidos anteriormente se procede a determinar el  $Z_{oc}$  que es el punto de corte o Z en el que una empresa fuerte pasa a ser débil.

$$Z_{oc} = ((Z_{fuertes})(\text{número de empresas fuertes}) + Z_{débiles}(\text{número de empresas débiles})) / \text{total de empresas}$$

$$Z_{oc} = (50(0,85272677) + 30(-1,42165289)) / 80 = -0,0001656$$

Todas las empresas que tengan función discriminante mayor al  $Z_{oc}$  son fuertes, de lo contrario son débiles.

De acuerdo con lo anterior se probó la clasificación de las empresas de la muestra (80 empresas) y se obtuvo un 94% de aciertos para las fuertes y un 87% para las débiles.

Finalmente, se probó el modelo con 26 empresas (15 fuertes y 11 débiles) que no hicieron parte de la muestra para el cálculo del  $Z_{oc}$  discriminante y se obtuvo un 100% de aciertos para las fuertes y 82% para las débiles. Por esta razón se concluye que el modelo sí es efectivo para predecir si una empresa es fuerte o débil con los indicadores financieros obtenidos a través del análisis discriminante (Innovar, 2002).

### **1.3. Determinantes de fragilidad en las empresas colombianas – Óscar Martínez**

Una de las mayores amenazas para toda empresa es caer en un estado de insolvencia. Este tipo de amenaza a la estabilidad financiera de las empresas es relevante no sólo para inversionistas y empleados, sino también para prestamistas del sector financiero, auditores y autoridades reguladoras entre otros. Usando técnicas de regresión probit este estudio desarrolla un modelo en el que se identifican las variables relevantes para pronosticar el estrés o fragilidad financiera de las empresas en Colombia en el año 2001.

A partir de los estados financieros que 9000 empresas reportaron en el año 2000, el modelo identificó correctamente al 82% de las empresas frágiles y no

frágiles. Los resultados confirman la importancia de los indicadores de rentabilidad, endeudamiento y liquidez en la solvencia presentada por las empresas, puntualmente la utilidad antes de impuestos sobre activo, obligaciones financieras sobre activo y disponible sobre activo respectivamente (Martínez, 2003).

Dentro de sus conclusiones incluyó:

“Este estudio permite identificar las razones financieras relevantes a la hora de pronosticar un deterioro en el estado legal de las empresas. Sin embargo, el modelo utilizado es de corte transversal y no es conveniente hacer análisis de carácter intertemporal a partir de estos resultados. Variables relevantes en estudios anteriores como edad y valor de mercado de la firma no se incluyeron por su falta de disponibilidad. Futuros trabajos en esta área permitirán saber si las razones financieras aquí presentadas siguen siendo determinantes a pesar de los cambios macroeconómicos de la economía.” (Martínez, 2003).

#### **1.4. Modelo Z de Altman**

Creado por Edward Altman con base en un análisis estadístico iterativo de discriminación múltiple en el que se ponderan y suman cinco razones de medición para clasificar las empresas en solventes e insolventes.

Para el desarrollo de su modelo en 1966, Altman tomó una muestra de 66 empresas de las cuales 33 habían quebrado durante los 20 años anteriores y 33

seguían operando a esa fecha. A la muestra le calculó 22 razones financieras que clasificó en 5 categorías estándar: liquidez, rentabilidad, apalancamiento, solvencia y actividad. Después de numerosas corridas, se seleccionaron las cinco variables que juntas dieron el mejor resultado en la predicción de la insolvencia.

Esto se hizo mediante:

- 1) Observación estadística de varias funciones, incluyendo la contribución relativa de cada variable independiente,
- 2) Evaluación de intercorrelaciones entre las variables relevantes,
- 3) Observación de la precisión en la predicción de los modelos y,
- 4) Criterio del analista.

La función discriminante final fue:

$$Z = 1,2 X1 + 1,4 X2 + 3,3 X3 + 0,6 X4 + 0,99 X5$$

Donde:

X1 = Capital de trabajo / Activo total

X2 = Utilidades retenidas / Activo total

X3 = Utilidades antes de intereses e impuestos / Activo total

X4 = Valor de mercado del capital / Pasivo total

X5 = Ventas / Activo total

El resultado indica que, si  $Z \geq 2,99$ , la empresa no tendrá problemas de insolvencia en el futuro; si  $Z \leq 1,81$ , entonces es una empresa que, de seguir así,

en el futuro tendrá altas posibilidades de caer en insolvencia. El modelo considera que las empresas se encuentran en una "zona gris" o no bien definida si el resultado de Z se encuentra entre 1,82 y 2,98.

Debido a que este modelo aplicaba sólo a empresas manufactureras que cotizaban en bolsa, Altman hizo una revisión del mismo y obtuvo dos nuevas versiones, el Z1 y el Z2.

### **1.5. Modelo Z1 de Altman**

Este modelo es una variación del modelo Z original, en el que se sustituye, por un lado, el numerador en X4 por el valor del capital contable en lugar del valor de mercado del capital y en el que la ponderación de cada índice también se modifica. Las adaptaciones se hicieron con el fin de aplicarlo a todo tipo de empresas y no solamente a las que cotizaran en bolsa (Altman, 1977).

Esta versión se desarrolló con empresas manufactureras y pondera de manera importante el activo total de la empresa y su rotación.

La función quedó de la siguiente manera:

$$Z1 = 0,717 X1 + 0,847 X2 + 3,107 X3 + 0,420 X4 + 0,998 X5$$

Si  $Z1 \geq 2,90$ , la empresa no tendrá problemas de insolvencia en el futuro; si  $Z1 \leq 1,23$ , entonces es una empresa que, de seguir así, en el futuro tendrá altas

posibilidades de caer en insolvencia, Si el resultado de Z1 es entre 1,24 y 2,89, se considera que la empresa se encuentra en una "zona gris" o no bien definida.

### **1.6. Modelo Z2 de Altman**

Ante el señalamiento por parte de analistas y críticos del modelo, de que también existían empresas que no eran manufactureras, cotizadas o no, para las cuales las dos versiones ya publicadas no eran aplicables, es decir que existían las empresas comerciales y de servicio, Altman realizó modificaciones en su estudio y llegó a obtener el valor Z2, para ser usado por cualquier tipo de empresa diferente a las contempladas en Z y Z1 (Altman, 1978).

Esta versión es un ajuste del modelo anterior Z1 en la que se elimina la razón de rotación de activos X5, para aplicarlo a todo tipo de empresas y no sólo a las manufactureras. Este nuevo modelo pondera de manera importante la generación de utilidades en relación al activo, así como su reinversión.

La función final es:

$$Z2 = 6,56 X1 + 3,26 X2 + 6,72 X3 + 1,05 X4$$

Si  $Z2 \geq 2,60$ , la empresa no tendrá problemas de insolvencia en el futuro; si  $Z2 \leq 1,10$ , entonces es una empresa que, de seguir así, en el futuro tendrá altas

posibilidades de caer en insolvencia. Las empresas se ubicarán en una zona no bien definida si el resultado de Z2 se encuentra entre 1,11 y 2,59.

### **1.7. Modelo de Gordon Springate**

Este modelo fue desarrollado en 1978 por Gordon L.V. Springate de la Universidad Simon Fraser de Canadá, siguiendo los procedimientos desarrollados por Altman (Springate, 1978). Springate usó el análisis estadístico iterativo de discriminación múltiple para seleccionar cuatro de 19 razones financieras de uso frecuente que mejor distinguieron entre los buenos negocios y los candidatos a insolvencia. El modelo Springate tiene la siguiente forma:

$$Z = 1,03A + 3,07B + 0,66C + 0,40D$$

Donde:

A = Capital de trabajo / Activo total,

B = Utilidad neta antes de intereses e impuestos / Activo total,

C = Utilidad neta antes de impuestos / Pasivo circulante,

D = Ventas / Activo total.

Cuando  $Z < 0.862$ , la firma podría considerarse como "insolvente".



Este modelo logró una precisión del 92,5% en 50 empresas que examinó Springate. Botheras (1979) probó el modelo de Springate en 50 empresas con un activo promedio de 2,5 millones de dólares canadienses y encontró el 88,0% de exactitud. Sands (1980) tomó el modelo de Springate para 24 empresas con un activo promedio de 63,4 millones de dólares canadienses y encontró una precisión de 83,3%.

### **1.8. Modelo Fulmer**

Desarrollado en 1984 por Fulmer (Fulmer, Moon, Gavin, Erwin, 1984), también utiliza el análisis iterativo de discriminación múltiple. El autor evaluó 40 razones financieras aplicadas a una muestra de 60 empresas, 30 solventes y 30 insolventes.

El modelo final toma 9 razones financieras ponderadas, de acuerdo con la siguiente ecuación:

$$H = 5.528 X1 + 0.212 X2 + 0.073 X3 + 1.270 X4 - 0.120 X5 + 2.335 X6 + 0.575 X7 + 1.083 X8 + 0.894 X9 - 6.075$$

Donde:

X1 = Utilidades retenidas / Activo total,

X2 = Ventas / Activo total,

X3 = Utilidades antes de impuestos / Capital contable,

X4 = Flujo de caja / Pasivo total,

X5 = Deuda / Activo total,

X6 = Pasivo circulante / Activo total,

X7 = Activo total tangible,

X8 = Capital de trabajo / Pasivo total,

X9 = log Utilidad operativa / Gastos financieros.

Cuando  $H < 0$ , la empresa puede calificarse como "insolvente".

Fulmer obtuvo el 98 % de precisión aplicando su modelo con un año de anticipación a la insolvencia y 81% con más de un año.

### **1.9. Modelo CA-Score**

La Orden de Contadores Certificados de Quebec (Quebec CA's) recomienda este modelo y, según dice quien lo desarrolló, lo usan cerca de mil analistas financieros en Canadá (Bilanas, 1987).

Este modelo fue desarrollado por Jean Legault de la Universidad de Quebec en Montreal, usando el análisis estadístico iterativo de discriminación múltiple. Fueron analizadas 30 razones financieras de una muestra de 173 empresas manufactureras con ventas anuales del rango de entre 1 a 20 millones de dólares canadienses.

El modelo tiene la siguiente forma:

$$CA\text{-SCORE} = 4.5913 X1 + 4.5080 X2 + 0.3936 X3 - 2.7616$$

Donde:

X1 = Capital contable / Activo total,

X2 = (Utilidades antes de impuestos y Rubros extraordinarios + Gastos financieros) / Activo total,

X3 = Ventas / Activo total.

Cuando CA-SCORE < -0.3, la empresa puede considerarse como "insolvente".

El autor del modelo informó (1987) que éste tiene un promedio de confianza del 83% y está restringido a evaluar empresas manufactureras.

### **1.10. Modelo Ricardo Pascale**

En 1988, RICARDO PASCALE desarrolló un modelo que también busca predecir la bancarrota de las empresas en el contexto latinoamericano, con alta inestabilidad. Trabajando para la industria uruguaya, usando el análisis discriminante múltiple, su modelo es el siguiente:

$$Z = -3,70992 + 0,99418 X1 + 6,55340 X2 + 5,51253 X3$$

Donde:

X1 = Ventas / Deudas totales

X2 = Ganancias Ajustadas por Inflación / Activos Totales

X3= Deuda a Largo Plazo / Deuda Total

Valor Crítico:  $Z = 0$  y la Zona de Ignorancia:  $-1,05 < Z < 0,4$

En las distintas pruebas de significación, la clasificación resultó correcta en un 92% para un año anterior a la quiebra y de un 81/82 % para dos o tres años antes a la quiebra. Utilizando un modelo como el de Pascale para países de América Latina (Uruguay), los resultados deben interpretarse de la siguiente manera: para el modelo de Pascale, si reporta un valor de Z superior a cero (0), la empresa será clasificada como con características similares a firmas que no han presentado serios problemas financieros y aquellas que tienen un Z menor que cero (0) como aquellas con características similares a las empresas que han tenido serios problemas financieros. (Pascale, 2000).

## **2. Marco Teórico**

En el marco del cumplimiento del objetivo general de este trabajo de investigación, se hace necesario revisar conceptos básicos como punto de quiebra o insolvencia financiera empresarial, así como analizar cuáles son sus principales causas en Colombia y, sin ser lo menos importante, realizar un repaso de los principales indicadores financieros.

### **2.1. Punto de quiebra o insolvencia financiera empresarial**

Dirigiéndonos a la definición más básica, la Real Academia Española (RAE), define a la quiebra, en la primera acepción del término, como la rotura de una superficie de cierta rigidez o dureza.

El concepto, de todas formas, es mucho más frecuente a nivel comercial para nombrar a la acción y efecto de quebrar un comerciante. La quiebra, también conocida como bancarrota, se produce cuando una persona u organización no está en condiciones de afrontar pagos que debe realizar.

Esto quiere decir que si una empresa tiene que pagar una deuda y no tiene dinero para ello, puede declararse en estado de quiebra. Esta es una situación jurídica que implica diversas obligaciones y responsabilidades. En otras palabras,

la quiebra se produce cuando el pasivo exigible a una entidad supera los activos (recursos económicos disponibles) de ésta.

La declaratoria de quiebra tiene varios efectos. En principio, la persona (física o jurídica) queda inhabilitada para administrar sus bienes. Dicha administración queda en manos de un tercero conocido como síndico. La quiebra, por otra parte, fija los derechos de los acreedores (que no pueden exigir mejoras en la situación tras la declaración de quiebra) y reúne todos los juicios pendientes contra el deudor ante un mismo juez.

La Superintendencia de Sociedades (2004) presentó un libro llamado "Causas de liquidación obligatoria. Estrategias para prevenir la crisis", en donde se hizo una encuesta a 409 empresas que se encontraban en liquidación obligatoria. Dentro de la encuesta se le consultó a las sociedades cuáles habían sido las posibles causas de la liquidación, en donde las respuestas señalan que las principales cinco causas de las liquidaciones fueron el elevado endeudamiento, la reducción en las ventas, los malos manejos administrativos, la alta competencia y la falta de personal competente. Y, con menor incidencia, la inseguridad, la corrupción, la terminación de incentivos tributarios y el contrabando.

### **2.1.1. Régimen de insolvencia empresarial en Colombia – Ley 1116 de 2006**

De acuerdo con la Superintendencia de Sociedades cuando un deudor se ve en la imposibilidad de pagar sus deudas y cumplir sus obligaciones cuando vencen los plazos, la mayoría de los ordenamientos jurídicos prevén mecanismos legales para satisfacer colectivamente las reclamaciones pendientes afectando a su pago todos los bienes del deudor.

Son muchos y muy diversos los intereses que tienen que atender esos mecanismos; en primer lugar, los de las partes afectadas por el procedimiento, entre ellas el deudor, los propietarios y los administradores de la empresa de éste, los acreedores que estén respaldados por garantías de diverso grado (incluidas las administraciones tributarias y otros acreedores públicos), los empleados, los garantes de la deuda y los proveedores de bienes y servicios, así como las instituciones jurídicas, comerciales y sociales que tienen interés en la implementación del régimen de la insolvencia. En general, estos mecanismos no sólo deben compaginar los distintos intereses de las partes directamente interesadas, sino también conjugar esos intereses con las consideraciones sociales, políticas y formativas pertinentes que repercuten en los objetivos económicos y jurídicos del procedimiento de insolvencia.

La Ley 1116 de 2006 tiene por objeto la protección del crédito y la recuperación y conservación de la empresa viable como unidad de explotación económica y fuente generadora de empleo, a través de procesos de reorganización y de liquidación judicial.

El proceso de reorganización de empresas está destinado a salvar a un deudor, que puede tratarse de una empresa, una persona natural comerciante o un patrimonio autónomo afecto a la realización de actividades empresariales. El salvamento se realiza a través del acuerdo que celebre entre acreedores internos y externos, con las mayorías estipuladas en la Ley, para pagar las acreencias vigentes al momento de la apertura del proceso.

En este orden de ideas, se utiliza el término “reorganización” en sentido amplio, para referirse a los procedimientos cuya finalidad básica sea la de permitir al deudor superar sus dificultades financieras y reanudar o continuar el funcionamiento de sus operaciones comerciales normales, aun cuando en algunos casos pueda incluir la reducción de la capacidad de la empresa, su venta como negocio en marcha a otra empresa y de no lograrlo extinguirse a través de un procedimiento de adjudicación o en caso de incumplimiento del acuerdo celebrado.

El régimen de la insolvencia regula el tipo de procedimiento denominado de Liquidación judicial y prevé en general que ante el juez del concurso se disponga de los bienes del deudor con miras a poner fin a la actividad comercial de la empresa, transformando en dinero los bienes a través de la venta directa o subasta privada y distribuyendo después el producto de la venta o, en caso de no ser posible la venta en todo o en parte, celebrando un acuerdo de adjudicación entre los acreedores aplicando la prelación legal de créditos o en su defecto adjudicándolos a través de providencia judicial (Supersociedades, 2004).



En nuestro régimen se permite que se proceda a la venta de unidades productivas de la empresa o a la venta de esta como unidad de explotación económica.

La liquidación suele concluir con la extinción o desaparición del deudor que sea una entidad jurídica mercantil y la exoneración de todo deudor que sea persona física, comerciante, a menos que dentro del proceso de liquidación judicial se negocie un acuerdo de reorganización que permita que el deudor reanude operaciones.

## **2.2. Razones financieras**

Los indicadores financieros, determinados mediante la aplicación de razones, son, posiblemente, la herramienta más utilizada para el análisis de los estados financieros (Carrillo de Rojas, 2011).

Hay algunas razones más comúnmente usadas porque la mayoría de los casos, ofrecen información útil para el análisis.

La forma más clara de agrupar las razones financieras es la siguiente:

- Razones de liquidez
- Razones operacionales o de actividad
- Razones de endeudamiento o apalancamiento
- Razones de rentabilidad o rendimiento
- Razones de eficiencia

Es importante anotar que existen otras razones que pueden ser empleadas, pero no incluidas dentro de la agrupación anterior.

### **2.2.1. Razones de liquidez**

Se entiende por liquidez, la capacidad de tesorería que tiene una empresa, para pagar sus deudas de corto plazo y llevar a cabo sus operaciones normales (León, 2009).

Debido a que los activos corrientes son aquellos que la empresa espera convertir en efectivo a corto plazo, entendemos que su realización proveerá los recursos de tesorería, con los cuales la empresa contará en el futuro inmediato, tanto para el pago de sus obligaciones a corto plazo, como para realizar las compras y gastos propios de sus operaciones normales.

#### **2.2.1.1. Capital de trabajo**

$$\textit{Capital de Trabajo} = \textit{Activo Corriente} - \textit{Pasivo Corriente}$$

Se puede interpretar como si todos los activos corrientes se pudieran realizar en un momento dado y todos los pasivos corrientes hubiera que pagarlos en ese mismo momento, a la empresa le sobraría, en efectivo, un monto equivalente a su capital de trabajo.

Una parte de los activos corrientes está financiada por pasivos corrientes y la otra parte por fuentes de financiamiento del largo plazo como lo son el pasivo no corriente y el capital contable. Por ende, el capital de trabajo, debe corresponder a la parte lenta de los activos corrientes como lo son el inventario mínimo y las cuentas por cobrar mínimas que la empresa mantiene en el transcurso del período contable (Brigham, Houston, 2001).

#### **2.2.1.2. Razón Corriente**

$$\text{Razón corriente} = \frac{\text{Activos corrientes}}{\text{Pasivos corrientes}}$$

Suele interpretarse en veces y significa que por cada peso que la empresa debe pagar de su pasivo a corto plazo, tiene en activos realizables a corto plazo, tantos pesos cuantas veces haya dado la razón corriente (León, 2009).

#### **2.2.1.3. Prueba ácida**

$$\text{Prueba ácida} = \frac{\text{Activos corrientes} - \text{Inventarios}}{\text{Pasivos corrientes}}$$

Se resta a los activos corrientes los inventarios, por suponerse que estos son los más lentos del grupo. Sin embargo, en el grupo de los activos corrientes se

clasifican activos menos líquidos como son los pagos anticipados, por esta razón Gladys Carrillo, 2011, propone el siguiente ajuste:

$$\text{Prueba ácida} = \frac{\text{Efectivo} + \text{Inversiones temporales} + \text{Cartera comercial}}{\text{Pasivos corrientes}}$$

Sin embargo, la anterior fórmula pueda variar en función del objeto social y el desempeño de cada empresa (Brigham, Houston, 2001).

## **2.2.2. Razones operacionales o de actividad**

Este grupo de razones sirve para medir la eficiencia de la administración, en el manejo de los recursos de la empresa, y algunas de ellas pueden determinar el grado de confiabilidad de las razones de liquidez (Brigham, Houston, 2001).

### **2.2.2.1. Rotación de cuentas por cobrar o período de cobro**

Se refiere al análisis que se hace de las cuentas por cobrar que se generan en las ventas a crédito que realiza la empresa, ya sea de mercancías, productos o de servicios.

$$\text{Días venta CxC} = \frac{\text{Cuentas por cobrar}}{\text{Venta promedio diaria}}$$

Entendiendo por rotación de cartera, las veces que la cartera se recauda y se vuelve a otorgar en el período, se puede determinar de la siguiente manera:

$$\text{Rotación de cartera (veces)} = \frac{\text{Ventas}}{\text{Cuentas por cobrar}}$$

El recaudo o período de cobro se puede determinar así:

$$\text{Período de cobro (días)} = \frac{\text{Cuentas por cobrar} \times 360}{\text{Ventas}}$$

#### **2.2.2.2. Rotación de inventarios o período de reposición**

De la misma manera que las cuentas por cobrar rotan, porque se vende a crédito y se recauda, los inventarios rotan, porque se venden y se reponen, ya sea por compra, como en el caso de mercancías o por fabricación, en el caso de productos manufacturados.

$$\text{Rotación de inventarios (veces)} = \frac{\text{Costo de ventas}}{\text{Inventario}}$$

$$\text{Período de reposición (días)} = \frac{\text{Inventario} \times 360}{\text{Costo de ventas}}$$

### 2.2.2.3. Rotación de cuentas por pagar o período de pago:

Con el fin de analizar las cuentas por pagar, se toman únicamente las que se adeudan a proveedores por concepto de compras de mercancías, en el caso de la empresa comercial o de materias primas, en el caso de la empresa manufacturera y se comparan con las compras totales o a crédito.

$$\text{Rotación CxP (días)} = \frac{\text{Cuentas por pagar} * 360}{\text{Compras}}$$

### 2.2.2.4. Ciclo operacional

La suma del período de recaudo de la cartera más el período de reposición de los inventarios se llama el período de conversión en efectivo o ciclo operacional. En este sentido, si los inventarios tienen un período de reposición de 90 días y la cartera un período de recaudo.

Si se compara el ciclo operacional con el período de pago a proveedores, se puede determinar las necesidades de capital de trabajo de la empresa y si la razón corriente es adecuada o no.

### 2.2.3. Razones de endeudamiento o apalancamiento

Este grupo mide el volumen de endeudamiento de la empresa comparando el valor de los pasivos con otros grupos de cuentas.

$$\text{Endeudamiento total (\%)} = \frac{\text{Total pasivo}}{\text{Total activo}}$$

$$\text{Leverage (\%)} = \frac{\text{Total pasivo}}{\text{Capital}}$$

$$\text{Concentración (\%)} = \frac{\text{Pasivo corriente}}{\text{Pasivo total}}$$

$$\text{Financiación a largo plazo (\%)} = \frac{\text{Capital} + \text{Pasivo a largo plazo}}{\text{Activo total}}$$

### 2.2.4. Razones de rentabilidad o rendimiento

#### 2.2.4.1. ROA (Return On Asset)

$$ROA = \frac{\text{Utilidad Neta}}{\text{Activo total}}$$

**2.2.4.2. ROE (Return On Equity)**

$$ROE = \frac{\textit{Utilidad Neta}}{\textit{Patrimonio}}$$

**2.2.4.3. ROI (Return On Investment)**

$$ROI = \frac{\textit{Utilidad Neta}}{\textit{Inversión}}$$



### 3. Metodología

A partir de la literatura consultada, revisada y analizada, se mostrará a continuación cómo determinar un modelo de scoring usando una técnica estadística denominada regresión logística o simplemente Logit. Esto equivale a codificar información en un valor específico y encontrar la combinación o los factores que explican de la mejor manera el comportamiento histórico de incumplimiento.

La herramienta básica que se utilizará para la estimación del modelo, anteriormente mencionado, es Excel con una codificación en Visual Basic.

En este sentido, es importante tener presente que un score resume la información contenida en factores que afectan la probabilidad de default. Los modelos estándar de scoring toman el enfoque más directo, combinando linealmente esos factores (Löffler y Posch, 2007). Se denotará  $\mathbf{x}$  como los factores (el número de factores será  $k$ ) y  $\mathbf{b}$  los pesos (o coeficientes) ligados a ellos; representando el score que se obtiene en un momento de scoring  $i$ , de la siguiente manera:

$$Score_i = b_1x_{i1} + b_2x_{i2} + \dots + b_kx_{ik}$$

### 3.1. Elaboración del modelo Logit

La Tabla 3 presenta un resumen del insumo necesario para la elaboración del modelo Logit, con una muestra de empresas con estados financieros de diferentes años y en donde  $y_i$  toma los valores de “0”, en el caso que la compañía no haya caído en default y “1”, en caso contrario. De la misma manera reúne un número de indicadores financieros (factores). Tras el default, las empresas pueden permanecer en este estado por varios años, caso en el cual sólo se tomarán en cuenta los estados financieros hasta el año en que entró en default.

Tabla 3. Factores y comportamiento de default

Empresa	Año Scoring	Indicador de Default	Factores			
		$y_i$	$x_{i1}$	$x_{i2}$	...	$x_{ik}$
800036662	2010	0	0,18	0,02	...	-1,53
800036662	2011	1	0,18	0,02	...	-1,53
800042706	2008	0	0,27	0,60	...	0,24
800042706	2009	0	0,25	0,22	...	0,38
800042706	2010	0	0,13	0,09	...	-0,08
800042706	2011	1	-0,14	0,07	...	-0,76
800108557	2008	0	0,44	1,93	...	0,33
800108557	2009	0	0,49	2,64	...	0,32
800108557	2010	0	0,53	1,77	...	0,18
800108557	2011	0	0,39	1,20	...	0,13
800108557	2012	1	0,36	0,38	...	0,19
800123441	2007	0	-0,04	0,57	...	-0,07
800123441	2008	0	-0,21	1,10	...	-1,26
800123441	2009	0	0,23	0,00	...	0,00
800123441	2010	1	0,24	0,00	...	0,00
...	...	...	...	...	...	...
N	2013	0	0,49	0,06	0,32	4,04

Nota: Adaptado de “Credit risk modeling using Excel and VBA”, por Löffler, G. y Posch

El modelo de scoring deberá predecir altas probabilidades para las observaciones con registro de default 1 y bajas en caso contrario. Con el fin de poder encontrar los pesos adecuados ( $b$ ), se necesita vincular los registros a probabilidades de incumplimiento:

$$Prob(Default_i) = F(Score_i)$$

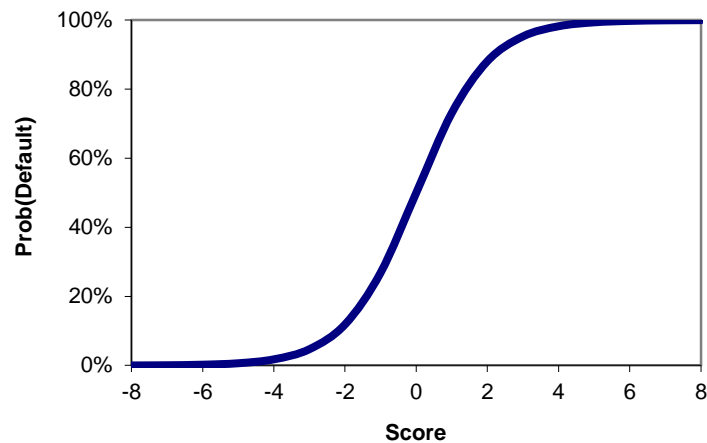
Así como en las probabilidades de default, la función  $F$  debe tomar valores entre 0 y 1. Una distribución frecuentemente utilizada para este propósito es la distribución logística  $\Lambda(z)$ :

$$Prob(Default_i) = \Lambda(Score_i) = \frac{\exp(b'x_i)}{1 + \exp(b'x_i)} = \frac{1}{1 + \exp(-b'x_i)}$$

Los modelos que vinculan información a probabilidades usando la función de distribución logística son denominados “modelos Logit” (Löffler y Posch, 2007).

La gráfica 1 muestra la relación entre las probabilidades de default asociadas con algunos valores de score. Se puede evidenciar claramente que los más altos scores se relacionan con las más altas probabilidades.

Gráfica 1: Scores y probabilidades de default en el modelo Logit



Nota: Tomado de "Credit risk modeling using Excel and VBA", por Löffler, G. y Posch

Habiendo reunido los factores  $\mathbf{x}$  y escogiendo la distribución de la función  $\mathbf{F}$ , una forma clara de determinar los pesos  $\mathbf{b}$  es el método de máxima probabilidad (ML por sus siglas en inglés). Para una empresa que entró en default ( $Y_i = 1$ ), la probabilidad sería:

$$Prob(Default_i) = \Lambda(b'x_i)$$

Para una empresa que no entró en default ( $Y_i = 0$ ), la probabilidad sería:

$$Prob(No Default_i) = 1 - \Lambda(b'x_i)$$

Teniendo presente que cualquier número elevado a la potencia 0 es igual a 1, la probabilidad para la observación  $i$  puede ser:

$$(L_i) = (\Lambda(b'x_i))^{y_i} (1 - \Lambda(b'x_i))^{1-y_i}$$

Asumiendo que los defaults son independientes, la probabilidad de un grupo de observaciones es el producto de las probabilidades individuales:

$$L = \prod_{i=1}^N L_i = \prod_{i=1}^N (\Lambda(b'x_i))^{y_i} (1 - \Lambda(b'x_i))^{1-y_i}$$

Para la maximización, es mejor examinar el logaritmo de la probabilidad (ln L):

$$\ln L = \sum_{i=1}^N y_i \ln(\Lambda(b'x_i)) + (1 - y_i) \ln(1 - \Lambda(b'x_i))$$

Se puede maximizar calculando su primera derivada con respecto de **b** a 0:

$$\frac{\partial \ln L}{\partial \mathbf{b}} = \sum_{i=1}^N (y_i - \Lambda(b'x_i)) x_i$$

El método Newton se puede utilizar para resolver la ecuación con respecto a **b**.

para aplicar este método, es necesario aplicar la segunda derivada:

$$\frac{\partial^2 \ln L}{\partial \mathbf{b} \partial \mathbf{b}'} = - \sum_{i=1}^N \Lambda(b'x_i) (1 - \Lambda(b'x_i)) x_i x_i'$$

Dado que Excel no posee una función que permita estimar modelos Logit, es necesario construir una que desarrolle esta tarea (para este fin, se desarrolló un código, el cual está descrito paso a paso en Anexo 1). La función completa se denomina LOGIT, cuya sintaxis se puede resumir de la siguiente manera:

**LOGIT (y; x; [const], [statistics])**

**y:** especifica el rango de la variable dependiente que, para este caso, es el indicador de default.

**x:** rango de las variables explicativas o significantes.

**[const]:** valor lógico para la inclusión de una constante, 1 u omitido si la constante es incluida y 0 en caso contrario.

**[statistics]:** cálculo de la regresión estadística, 1 si va a ser calculada y 0, u omitido, en caso contrario.

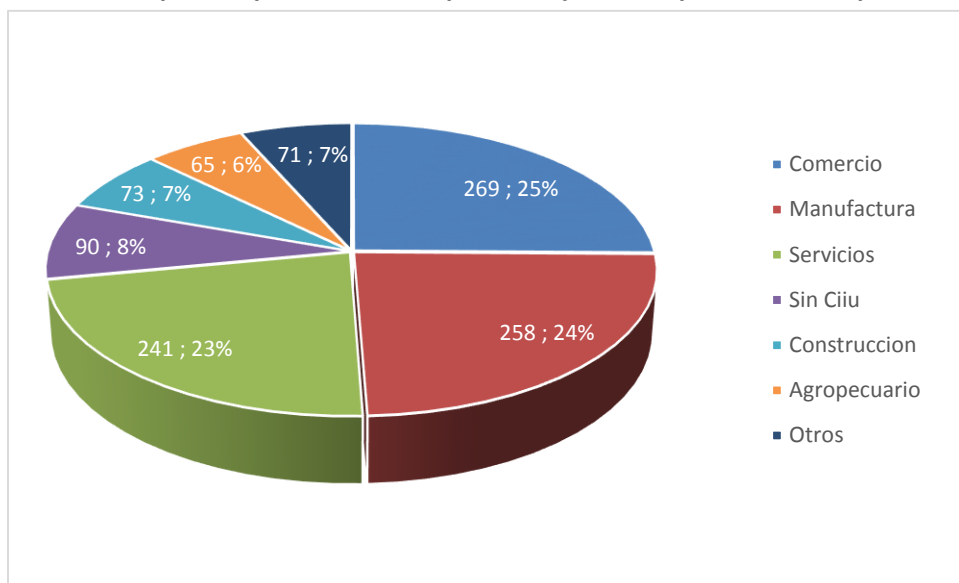
Los corchetes significan que es un argumento opcional.

La función devuelve una matriz, por lo tanto debe ser ejecutada en un rango de celdas y terminar con [Ctrl] + [Shift] + [Enter].

Para la elaboración del modelo Logit, objeto estudio de la presente investigación, se reunieron los estados financieros históricos desde el Sistema de Información y Reporte Empresarial (SIREM), de la Superintendencia de Sociedades, de una muestra de 838 empresas Mipymes, del sector manufacturero colombiano. Adicional a esto se tomó una muestra de 57 empresas que hubieran caído en default (liquidación judicial) en los últimos cinco años, con corte al 30 de junio de 2015.

Desde 2007, hasta el 30 de junio de 2015, un total de 1.067 empresas han entrado en liquidación judicial, las cuales se encuentran distribuidas por sector de la siguiente manera:

**Gráfica 2: Empresas que entran en liquidación judicial a partir de 2007 por sector**



**Nota: Construcción propia - SIREM**

Sin embargo, se tienen en cuenta sólo las manufactureras (258) y que reúnan la condición de ser tamaño mediana, micro y pequeña, pero sólo se cuenta con información de Estados Financieros de 57 de ellas.

Habiendo reunido los diferentes estados financieros, se procede a hacer el cálculo de diferentes indicadores financieros tales como: capital de trabajo, utilidades retenidas, utilidades antes de intereses e impuestos (EBIT), ventas; todos estos utilizados en los modelos reconocidos de Altman; adicionalmente se calcularon: razón corriente, prueba ácida, endeudamiento, EBITDA, utilidad neta, utilidad operacional, margen bruto, margen operacional, margen neto, ROA, ROE, entre otros.

Por su parte, retomando algunos conceptos de Econometría, para evaluar si una variable ayuda a explicar el evento de default o no, se puede examinar un ratio  $t$ , para la hipótesis que el coeficiente de la variable es cero. Para el coeficiente de orden  $j$ , como un ratio  $t$  es construido como:

$$t_j = \frac{b_j}{SE(b_j)}$$

Donde SE es el error estándar estimado del coeficiente. Tomamos  $b$  de la última iteración del esquema Newton y los errores estándar de los parámetros estimados son derivados de la matriz Hessiana (Gray and Malone, 2008).

En el modelo Logit, el ratio  $t$  no sigue una distribución  $t$  como en la clásica regresión lineal. Más bien, es comparado con una distribución estándar normal. Para obtener el **p-value** de una prueba en doble vía, se explota la simetría de la distribución normal:

$$p - value = 2 * (1 - distr. norm. n(ABS)(t))$$

En el modelo, la función LOGIT entregará errores estándar, ratios  $t$  y p-values, en las filas 2 a la 4 si el valor estadístico lógico se determina en 1.

Adicional a esto, en una regresión lineal, se puede reportar un  $R^2$  como una medida de la bondad del ajuste global. En modelos no lineales estimados con la máxima probabilidad, usualmente se utiliza el Pseudo- $R^2$  sugerido por McFadden.



Este se calcula como 1 menos la razón entre el log-probabilidad del modelo estimado ( $\ln L$ ) y el de un modelo restringido que tiene sólo una constante ( $\ln L_0$ ):

$$Pseudo - R^2 = 1 - \frac{\ln L}{\ln L_0}$$

Como el  $R^2$  estándar, esta medida está limitada por 0 y 1. Altos valores indica un mejor ajuste. El log-probabilidad  $\ln L$  está dado por la función log-probabilidad de la última iteración del procedimiento Newton, y está por lo tanto disponible. Las dos probabilidades empleadas para el Pseudo- $R^2$  pueden ser utilizadas para conducir una prueba estadística del modelo completo. La prueba es estructurada como un ratio de probabilidad (Likelihood Ratio por sus siglas en inglés):

$$LR = 2(\ln L - \ln L_0)$$

El test estadístico es distribuido asintóticamente al chi cuadrado con los grados de libertad iguales al número de restricciones impuestas. Cuando se prueba la significancia de la regresión completa, el número de restricciones iguala el número de variables  $k - 1$ . La función CHIDIST (prueba estadística, restricciones) da el p-value del test LR. El comando LOGIT devuelve los dos el LR y su p-value.

Las probabilidades  $\ln L$  y  $\ln L_0$  están también reportadas, como es el número de iteraciones que fueron necesarias para lograr la convergencia. A manera de resumen, los datos de salida de la función LOGIT se organizan en la tabla 4:

**Tabla 4. Datos de salida de la función LOGIT**

$b_1$	$b_2$	...	$b_K$
SE( $b_1$ )	SE( $b_2$ )	...	SE( $b_K$ )
$t_1 = b_1/SE(b_1)$	$t_2 = b_2/SE(b_2)$	...	$t_K = b_K/SE(b_K)$
$p$ -value( $t_1$ )	$p$ -value( $t_2$ )	...	$p$ -value( $t_K$ )
Pseudo- $R^2$	# iterations	#N/A	#N/A
LR test	$p$ -value (LR)	#N/A	#N/A
log-likelihood (model)	log-likelihood (restricted)	#N/A	#N/A

**Nota:** Tomado de “Credit risk modeling using Excel and VBA”, por Löffler, G. y Posch

Con el fin de poder encontrar el modelo adecuado, para la predicción del punto de quiebra de las empresas manufactureras en Colombia, y teniendo la data necesaria, se empieza a realizar la evaluación de cada uno de los indicadores financieros dentro del modelo, de manera individual, buscando hallar los que presentaran un  $p$ -value, en lo posible, menor al 1%, y que nos diera un LR-test alto con el fin de poder determinar que este factor (indicador financiero) es altamente significativo dentro del modelo. Es así como se llega a los principales cinco factores (indicadores financieros): capital de trabajo / total de activos, ventas / total activos, margen bruto (utilidad bruta / ingresos operacionales netos), razón corriente y prueba ácida, resultados que pueden ser observados en la tabla 5:

**Tabla 5. Resultados individuales de los factores**

<b>Factor</b>	<b>Modelo 1</b>	<b>CONST</b>	<b>CT/TA</b>
Capital Trabajo / Total Activos	<b>b</b>	-3,723	-3,434
	<b>SE(b)</b>	0,141	0,440
	<b>t</b>	-26,40	-7,80
	<b>p-value</b>	0,000	0,000
	<b>Pseudo R<sup>2</sup> / # iter</b>	0,141	8
	<b>LR-test / p-value</b>	81,0	0,000
	<b>InL / InL0</b>	-247,2	-287,7
Ventas / Total Activos	<b>b</b>	-2,822	-1,243
	<b>SE(b)</b>	0,250	0,267
	<b>t</b>	-11,30	-4,65
	<b>p-value</b>	0,000	0,000
	<b>Pseudo R<sup>2</sup> / # iter</b>	0,048	7
	<b>LR-test / p-value</b>	27,5	0,000
	<b>InL / InL0</b>	-273,9	-287,7
Razón Corriente	<b>b</b>	-2,676	-0,861
	<b>SE(b)</b>	0,307	0,210
	<b>t</b>	-8,72	-4,10
	<b>p-value</b>	0,000	0,000
	<b>Pseudo R<sup>2</sup> / # iter</b>	0,046	9
	<b>LR-test / p-value</b>	26,6	0,000
	<b>InL / InL0</b>	-274,4	-287,7
Prueba Ácida	<b>b</b>	-2,817	-1,185
	<b>SE(b)</b>	0,269	0,283
	<b>t</b>	-10,47	-4,18
	<b>p-value</b>	0,000	0,000
	<b>Pseudo R<sup>2</sup> / # iter</b>	0,049	9
	<b>LR-test / p-value</b>	28,1	0,000
	<b>InL / InL0</b>	-273,6	-287,7
Margen Bruto	<b>b</b>	-3,723	-1,332
	<b>SE(b)</b>	0,157	0,370
	<b>t</b>	-23,76	-3,60
	<b>p-value</b>	0,000	0,000
	<b>Pseudo R<sup>2</sup> / # iter</b>	0,032	12
	<b>LR-test / p-value</b>	18,3	0,000
	<b>InL / InL0</b>	-278,5	-287,7

**Nota: Construcción propia**

Posterior al análisis individual, se integran los cinco factores en un solo modelo con el fin de poder determinar cuál es su comportamiento: si se mantiene un bajo p-value y un alto LR test, o si cambia de manera tal que dejen de ser significantes para el modelo (tabla 6).

Tabla 6. Aplicación del modelo LOGIT a la muestra de empresas, con 5 indicadores financieros

NIT	Año	Default	CT/TA	V/TA	MB	RC	PA	Model 1	CONST	CT/TA	V/TA	MB	RC	PA
800036662	2010	0	0,18	0,02	-1,53	2,35	1,75	<b>b</b>	-2,476	-2,512	-0,937	-0,688	0,000	-0,239
800036662	2011	1	0,18	0,02	-1,53	2,35	1,75	<b>SE(b)</b>	0,330	0,484	0,275	0,260	0,048	0,225
800042706	2008	0	0,27	0,60	0,24	2,85	1,76	<b>t</b>	-7,51	-5,19	-3,41	-2,65	0,00	-1,06
800042706	2009	0	0,25	0,22	0,38	2,62	1,66	<b>p-value</b>	0,000	0,000	0,001	0,008	0,999	0,288
800042706	2010	0	0,13	0,09	-0,08	1,70	1,03	<b>Pseudo R<sup>2</sup> / # iter</b>	0,183	13				
800042706	2011	1	-0,14	0,07	-0,76	0,49	0,13	<b>LR-test / p-value</b>	105,5	0,000				
800108557	2008	0	0,44	1,93	0,33	2,21	2,21	<b>InL / InL0</b>	-234,9	-287,7				
800108557	2009	0	0,49	2,64	0,32	2,71	2,71		={logit(C4:C3293;D4:H3293;1;1)}					
800108557	2010	0	0,53	1,77	0,18	2,85	2,85							
800108557	2011	0	0,39	1,20	0,13	1,91	1,91							
800108557	2012	1	0,36	0,38	0,19	1,75	1,75							
800123441	2007	0	-0,04	0,57	-0,07	0,96	0,54							
800123441	2008	0	-0,21	1,10	-1,26	0,70	0,09							
800123441	2009	0	0,23	0,00	0,00	1,39	1,39							
800123441	2010	1	0,24	0,00	0,00	1,39	1,39							
800131627	2008	0	0,52	0,08	-0,99	2,06	2,06							
800131627	2009	0	0,67	0,00	0,00	3,02	3,02							
800131627	2010	1	0,02	0,00	0,00	1,02	1,02							
800210495	2008	0	0,18	2,16	0,24	2,18	1,89							
800210495	2011	1	0,34	5,12	0,28	2,61	2,46							
800213025	2007	0	0,21	0,56	0,07	2,77	2,40							
800213025	2008	0	0,12	0,55	0,09	2,01	1,65							

Nota: Construcción propia

La tabla 6 está conformada por el listado de empresas, con el año de estados financieros (los cuales no son empleados en el modelo); la columna default, que contiene 1 en el caso en que la empresa haya entrado en default en ese año específico, o cero en caso contrario; y cinco indicadores financieros que fueron los que arrojaron los p-value más bajos de manera individual.

Como se puede observar en la misma tabla, por los coeficientes de regresión obtenidos, se puede determinar que tres de los cinco indicadores

financieros tienen coeficientes **b** que son significantes en un nivel del 1% o mejor (por debajo del 1%). Si rechazáramos la hipótesis que uno de estos coeficientes es cero, podríamos errar con una probabilidad menor al 1%. Cuatro de las cinco variables tienen coeficientes negativos, lo cual indica que incrementando los valores de las variables, se reduce la probabilidad de default. Si observamos el p-value de la constante, se puede concluir que también es altamente significativa.

Por su parte, al observar los coeficientes de la razón corriente y de la prueba ácida, muestran una significancia del 99,9% y 28,8%, respectivamente, por lo que se consideran no significantes o marginalmente significantes y podrían excluirse del modelo.

Al interpretar el Pseudo-R<sup>2</sup>, es útil tener en cuenta que éste no mide si el modelo predijo correctamente las probabilidades de default - esto no es factible, porque no sabemos las verdaderas probabilidades de incumplimiento. En cambio, el Pseudo-R<sup>2</sup> (hasta cierto punto) mide si predijimos correctamente los defaults (Gray, Malone; 2008).

Estos dos aspectos están relacionados, pero no son idénticos. Por ejemplo, en el caso en que una empresa entró en default a pesar de haber registrado una baja probabilidad de default: si el modelo era correcto con respecto a esta baja probabilidad de default, se ha cumplido con su objetivo, pero el resultado pasó a estar desalineado, reduciendo así el Pseudo-R<sup>2</sup>. En un portafolio de crédito, la mayoría de las probabilidades de default están en el rango de 0,05% a 5%. Incluso si obtenemos cada probabilidad de default de manera individual, habrá muchos

casos en los que los datos observados (default) no están en línea con la predicción (probabilidad de default baja) y que por lo tanto no puede esperar para conseguir un Pseudo-R<sup>2</sup> cercano a 1. Una situación en la que el Pseudo-R<sup>2</sup> sería cercano a 1 se vería de la siguiente manera: la empresas se clasifican en uno de dos grupos: el primero se caracteriza por probabilidades de default muy bajas (0,1% y menos), el segundo grupo por probabilidades muy altas (99,9% o más). Esto es claramente poco realista para los típicos portafolios de crédito (Löffler y Posch, 2007).

Volviendo al modelo, si se llegaran a descartar las variables razón corriente y prueba ácida, basados en sus *t* ratios, se podría omitir la posibilidad de que unidas puedan explicar defaults, a pesar de que sean no significantes individualmente. Por esta razón, es mejor correr una segunda regresión en la cual se excluyan estas variables y luego conducir al LR test (tabla 7).

En el modelo 2 de la tabla 7, se han descartado las variables no significantes, teniendo la restricción que los coeficientes de estas dos variables es cero. De esta manera el LR test se basa en una comparación del log LR  $\ln L$  de los dos modelos:

$$LR = 2[\ln L(\text{modelo 1}) - \ln L(\text{modelo 2})]$$

Además, se refiere a una distribución chi-cuadrado con dos grados de libertad porque se incluyen dos restricciones. El LR test da un valor de 2,3 con un

p-value de 31,62%, lo que significa que si se adicionan las variables razón corriente y prueba ácida al modelo 2, existe una probabilidad del 31,62% que no agreguen significancia al modelo. De esta manera, el LR test reafirma el resultado del análisis tanto individual como en conjunto de las variables: estas dos variables son apenas marginalmente significantes.

Tabla 7. Prueba de restricciones conjuntas con un LR test

Aplicación del modelo LOGIT a la muestra de empresas								Model 1	CONST	CT/TA	V/TA	MB	RC	PA	
								b	-2,476	-2,512	-0,937	-0,688	0,000	-0,239	
								SE(b)	0,330	0,484	0,275	0,260	0,048	0,225	Default probab
800036662	2010	0	0,18	0,02	-1,53	2,35	1,75	t	-7,51	-5,19	-3,41	-2,65	0,00	-1,06	8,96%
800036662	2011	1	0,18	0,02	-1,53	2,35	1,75	p-value	0,000	0,000	0,001	0,008	0,999	0,288	8,96%
800042706	2008	0	0,27	0,60	0,24	2,85	1,76	Pseudo R <sup>2</sup> / # iter	0,183	13					1,32%
800042706	2009	0	0,25	0,22	0,38	2,62	1,66	LR-test / p-value	105,5	0,000					1,85%
800042706	2010	0	0,13	0,09	-0,08	1,70	1,03	InL / InL0	-234,9	-287,7					4,35%
800042706	2011	1	-0,14	0,07	-0,76	0,49	0,13		{=logit(C4:C3293;D4:H3293;1;1)}						15,39%
800108557	2008	0	0,44	1,93	0,33	2,21	2,21								0,21%
800108557	2009	0	0,49	2,64	0,32	2,71	2,71	Model 2	CONST	CT/TA	V/TA	MB			0,09%
800108557	2010	0	0,53	1,77	0,18	2,85	2,85	b	-2,765	-2,874	-0,887	-0,614			0,19%
800108557	2011	0	0,39	1,20	0,13	1,91	1,91	SE(b)	0,255	0,429	0,271	0,253			0,59%
800108557	2012	1	0,36	0,38	0,19	1,75	1,75	t	-10,83	-6,71	-3,28	-2,43			1,36%
800123441	2007	0	-0,04	0,57	-0,07	0,96	0,54	p-value	0,000	0,000	0,001	0,015			4,73%
800123441	2008	0	-0,21	1,10	-1,26	0,70	0,09	Pseudo R <sup>2</sup> / # iter	0,179	12					10,63%
800123441	2009	0	0,23	0,00	0,00	1,39	1,39	LR-test / p-value	103,2	0,000					3,24%
800123441	2010	1	0,24	0,00	0,00	1,39	1,39	InL / InL0	-236,1	-287,7					3,16%
800131627	2008	0	0,52	0,08	-0,99	2,06	2,06		{=logit(C4:C3293;D4:F3293;1;1)}						2,51%
800131627	2009	0	0,67	0,00	0,00	3,02	3,02								0,76%
800131627	2010	1	0,02	0,00	0,00	1,02	1,02								5,86%
800210495	2008	0	0,18	2,16	0,24	2,18	1,89	LR- Test for b(WC/TA)=b(S/TA)=0 in mod							0,38%
800210495	2011	1	0,34	5,12	0,28	2,61	2,46	LR	2,30	=2*(J8-J18)					0,01%
800213025	2007	0	0,21	0,56	0,07	2,77	2,40	DF	2						1,53%
800213025	2008	0	0,12	0,55	0,09	2,01	1,65	p-value	31,62%	CHDIST(J23,J24)					2,30%
800213025	2009	0	0,12	0,51	0,17	1,66	1,40								2,39%
800213025	2010	0	0,03	0,56	0,09	1,10	0,92								3,38%

Nota: Construcción propia

Con base en estos resultados, se podría proceder bajo las reglas estándar de restricción de significancia estadística y se eliminarían las variables que no son significantes en un nivel del 5% o mejor que esto, conduciendo a inclinarse por el

modelo 2. Sin embargo, es aconsejable complementar estas reglas con pruebas adicionales.

Ahora bien, después de haber determinado el modelo de scoring, lo que se quiere es determinar probabilidades de default, de tal manera que se calcula el score y se traduce en términos de probabilidades de default:

$$Prob(Default_i) = \Lambda(Score_i) = \Lambda(b'x_i) = \frac{1}{1 + \exp(-b'x_i)}$$

Para este caso se necesita evaluar el score  $b'x_i$ , cuyo procedimiento puede ser fácilmente implementado a través de la utilización de la función de Excel sumaproducto de los cinco coeficientes de las variables arrojados por el modelo, por los cinco indicadores financieros para cada  $i$  (o input), sumándole el valor de la constante.

Revisando el resultado de la probabilidad de default calculada (tabla 7), se puede apreciar casos específicos de empresas con un valor muy bajo, a pesar de que su indicador de default es uno:

- Nit 800210495: Prob. Default = 0,01%
- Nit: 815002520: Prob. Default = 0,55%
- Nit: 890919912: Prob. Default = 0,33%



Esto puede presentarse porque una empresa puede caer en default a pesar de tener una probabilidad de default muy baja y, además, aunque el modelo se ajuste bastante bien en predicción de defaults, como lo indicó el LR test del modelo completo, puede presentar fallas en el momento de predecir probabilidades de default individuales. Otro aspecto importante a tener en cuenta es que, por lo general, los modelos de scoring se estiman con datos pasados y aplicados a datos actuales.

Con el fin de poder cuantificar el impacto de una variable, se puede evaluar de manera individual a una empresa y calcular posibles cambios en la probabilidad de default relacionados con cambios en el valor de la variable. Una forma simple de hacerlo es realizar cambios en el valor de una de las variables y mantener las demás constantes y determinar el cambio presentado en la probabilidad de default (Löffler y Posch, 2007).

Realizando el ejercicio es coherente el resultado obtenido, dado que al mejorar el valor del indicador financiero, claramente la probabilidad de default disminuye y, en caso de empeorarlo, ésta aumenta. Otra posibilidad de modelarlo es tratar de encontrar qué valor de la variable se requeriría para llegar a una probabilidad de default deseada, por ejemplo menor al 5%, a través de herramientas en Excel como buscar objetivo o el mismo solver.

### 3.2. Minería de datos

Löffler y Posch en 2007, emplearon un procedimiento para suavizar los valores atípicos en las variables de entrada, debido a que en la creación de los modelos, las variables significantes contienen algunos valores extremos, ya sea por situaciones particulares para cada empresa o por errores en los datos; pudiendo tener un amplio efecto en el cálculo de los coeficientes y, por ende, afectar el modelo.

El primer paso que dieron fue el de analizar la distribución de las variables, empleando algunas herramientas de estadística descriptiva para cada una de las variables e incluidas en Excel. Es bastante útil analizar la distribución normal dado que en ésta los valores extremos son prácticamente promediados.

La tabla 8 muestra algunas medidas estadísticas para los cinco indicadores financieros implementados en el modelo, tales como: la media aritmética que es, quizás, la medida de tendencia central más utilizada, conocida y sencilla de calcular; la mediana que es el valor central de la muestra de datos siempre y cuando éstos se encuentren ordenados; la desviación estándar que es una medida bastante precisa de la dispersión de los datos en torno a la media aritmética de la distribución; la asimetría que es una medida que permite identificar si los datos se distribuyen de manera uniforme alrededor de la media aritmética, caso en el cual se presenta una distribución simétrica en forma de campana, cuando los valores altos de la distribución se presentan con mayor frecuencia

(estando por encima de la media aritmética) se presenta una distribución asimétrica, con sesgo negativo con cola alargada a la izquierda y, en caso contrario, se presenta una distribución asimétrica, con sesgo positivo y con cola alargada a la derecha; la curtosis es la medida que determina el grado de aplastamiento o afilamiento de la curva de frecuencias de los datos y se refleja en la mayor o menor propensión de la curva a tomar valores próximos a la moda, pudiendo denominarse planocúrtica, mesocúrtica o leptocúrtica (Fernández y Sarmiento, 2014).

**Tabla 8. Medidas estadísticas de las variables en el modelo logit**

	CT/TA	V/TA	MB	RC	PA
Media Aritmética	0,21	1,27	0,28	2,50	1,68
Mediana	0,20	1,10	0,28	1,61	1,08
Desviación Estándar	0,25	1,24	0,23	7,51	5,84
Asimetría	-1,85	12,83	-10,29	25,05	39,07
Curtosis	18,91	283,41	264,84	808,39	1870,32
<b>Cuantiles / Valores extremos</b>					
Min	-3,09	0,00	-6,38	0,00	0,00
0,50%	-0,69	0,00	-0,37	0,14	0,08
1%	-0,49	0,00	-0,11	0,24	0,14
5%	-0,14	0,26	0,06	0,69	0,33
95%	0,61	2,58	0,57	5,36	3,81
99%	0,76	4,29	0,74	12,77	9,36
99,50%	0,79	5,44	0,84	25,53	18,05
Max	0,94	35,66	1,00	291,81	291,81

**Nota: Construcción propia**

Una buena medida para determinar la existencia de valores atípicos es la curtosis, partiendo del hecho que la distribución normal tiene una curtosis igual a cero. Como se puede observar en la muestra de datos de las variables las curtosis van desde 18,91 hasta 1.870,3 y además positivas indicando que pueden existir

muchos datos alejados de la media. Esta afirmación se puede afianzar al observar y analizar las desviaciones estándar de cada una de las variables y compararlas con las típicas de la distribución normal.

Teniendo entonces presente que los datos utilizados para el modelo Logit, poseen valores atípicos, se podría proceder a encontrar errores corregibles en los datos, aunque esto no asegura que se eliminen por completo dichos valores, por lo que es importante utilizar un procedimiento como la winsorización, que ayude a controlar la influencia de los valores atípicos de una manera automática y de una forma objetiva (Löffler y Posch, 2007).

En otras palabras, lo que hace la winsorización es contraer los datos extremos o atípicos a observaciones más centrales. Para tal fin, se determina un nivel de winsorización  $\alpha$ , cuyo valor puede estar entre 0,5% y 5%, y para los valores por debajo del percentil  $\alpha$  de la distribución de las variables, se establecen en el percentil  $\alpha$ , y para los valores por encima del percentil  $1 - \alpha$  se establecen en este nivel  $1 - \alpha$ . Este nivel  $\alpha$  puede ser el mismo para todas las variables, o puede determinarse de manera independiente para cada una de ellas dependiendo del comportamiento de sus valores (Winsor, 1947).

Tabla 9. Winsorización de los valores de las variables

	Percentil más bajo					Percentil más alto					Nivel de Winsorización														
	-0,31					0,06					0,00					0,69					0,33				
	0,70					3,53					0,69					5,36					3,81				
	2,0%					2,0%					2,0%					5,0%					5,0%				

	Asimetría					Curtosis																			
	(1,85)					12,83					(10,29)					25,05					39,07				
	18,91					283,41					264,84					808,39					1.870,32				

Datos Iniciales						Datos Winsorizados					
NIT	CT/TA	V/TA	MB	RC	PA	CT/TA	V/TA	MB	RC	PA	
800036662	0,18	0,02	-1,53	2,35	1,75	0,18	0,06	0,00	2,35	1,75	
800036662	0,18	0,02	-1,53	2,35	1,75	0,18	0,06	0,00	2,35	1,75	
800042706	0,27	0,60	0,24	2,85	1,76	0,27	0,60	0,24	2,85	1,76	
800042706	0,25	0,22	0,38	2,62	1,66	0,25	0,22	0,38	2,62	1,66	
800042706	0,13	0,09	-0,08	1,70	1,03	0,13	0,09	0,00	1,70	1,03	
800042706	-0,14	0,07	-0,76	0,49	0,13	-0,14	0,07	0,00	0,69	0,33	
800108557	0,44	1,93	0,33	2,21	2,21	0,44	1,93	0,33	2,21	2,21	
800108557	0,49	2,64	0,32	2,71	2,71	0,49	2,64	0,32	2,71	2,71	
800108557	0,53	1,77	0,18	2,85	2,85	0,53	1,77	0,18	2,85	2,85	
800108557	0,39	1,20	0,13	1,91	1,91	0,39	1,20	0,13	1,91	1,91	
800108557	0,36	0,38	0,19	1,75	1,75	0,36	0,38	0,19	1,75	1,75	
800123441	-0,04	0,57	-0,07	0,96	0,54	-0,04	0,57	0,00	0,96	0,54	
800123441	-0,21	1,10	-1,26	0,70	0,09	-0,21	1,10	0,00	0,70	0,33	
800123441	0,23	0,00	0,00	1,39	1,39	0,23	0,06	0,00	1,39	1,39	
800123441	0,24	0,00	0,00	1,39	1,39	0,24	0,06	0,00	1,39	1,39	
800131627	0,52	0,08	-0,99	2,06	2,06	0,52	0,08	0,00	2,06	2,06	
800131627	0,67	0,00	0,00	3,02	3,02	0,67	0,06	0,00	3,02	3,02	
800131627	0,02	0,00	0,00	1,02	1,02	0,02	0,06	0,00	1,02	1,02	
800210495	0,18	2,16	0,24	2,18	1,89	0,18	2,16	0,24	2,18	1,89	
800210495	0,34	5,12	0,28	2,61	2,46	0,34	3,53	0,28	2,61	2,46	
800213025	0,21	0,56	0,07	2,77	2,40	0,21	0,56	0,07	2,77	2,40	
800213025	0,12	0,55	0,09	2,01	1,65	0,12	0,55	0,09	2,01	1,65	
800213025	0,12	0,51	0,17	1,66	1,40	0,12	0,51	0,17	1,66	1,40	
800213025	0,03	0,56	0,09	1,10	0,92	0,03	0,56	0,09	1,10	0,92	
800213025	0,03	0,56	0,08	1,11	0,91	0,03	0,56	0,08	1,11	0,91	

La tabla 9 muestra los datos iniciales de las variables, así como sus medidas estadísticas de asimetría y curtosis, esta última con datos muy elevados; posterior a la winsorización en un nivel del 2% uniforme para todas las variables, es claro cómo las medidas estadísticas se reducen y empiezan a aproximarse a cero. De hecho, si se quisieran ajustar mucho más, podríamos aplicar un nivel de winsorización del 5% para las variables RC y PA, y la curtosis resultante sería del

1,60 y 1,30, respectivamente. Otra posibilidad es aplicar el logaritmo a las variables que mantienen asimetría y curtosis altas o, inclusive, a todas las variables dado que nos generaría un mayor Pseudo-R<sup>2</sup> en el momento de realizar la regresión de nuevo. Para el caso en que los valores de las variables sean negativos se requeriría realizar una transformación matemática para poder aplicar el logaritmo:  $-\ln(1 - \text{variable})$ , debido a que si la asimetría es negativa, aplicando el logaritmo lo que haría es aumentar la cola a la izquierda. Sin embargo, es importante tener cuidado con la minería de datos y definir restricciones para la utilización de los métodos comunes para la winsorización de los datos.

### 3.3. Modelo LOGIT final

Después de realizar todo el procedimiento para llegar al modelo Logit definitivo, entendiendo que las variables razón corriente y prueba ácida no dan valor al mismo, no son significantes, el modelo resultante es el siguiente:

$$Z = -2,765 - 2,874x_1 - 0,887x_2 - 0,614x_3$$

Teniendo presente que las variables poseen datos atípicos o extremos y después de haber realizado la winsorización de los mismos (procedimiento revisado en el numeral anterior), el modelo resultante se puede apreciar en la

tabla 10, el cual posee coeficientes diferentes y mejora su Pseudo-R<sup>2</sup>, pasando de 17,9% a 18,6%, indicando que la minería de datos ha sido acertada:

$$Z = -1,881 - 4,545x_1 - 0,837x_2 - 3,694x_3$$

En donde:

$x_1$  = Capital de trabajo / Total activos

$x_2$  = Ventas / Total activos

$x_3$  = Margen bruto (utilidad bruta / ingresos operacionales netos)

Tabla 10. Modelo LOGIT final después de winsorizar los datos de las variables

Aplicación del modelo LOGIT a la muestra de empre						Model 1	CONST	CT/TA	V/TA	MB	Probabilidad de default
						<b>b</b>	-1,881	-4,545	-0,837	-3,694	
						<b>SE(b)</b>	0,286	0,697	0,246	1,002	
						<b>t</b>	-6,56	-6,52	-3,40	-3,69	5,97%
						<b>p-value</b>	0,000	0,000	0,001	0,000	5,97%
						<b>Pseudo R<sup>2</sup> / # iter</b>	0,186	7			1,10%
						<b>LR-test / p-value</b>	107,2	0,000			0,98%
						<b>lnL / lnL<sub>0</sub></b>	-234,1	-287,7			7,19%
						{=logit(C6:C3295;D6:F3295;1;1)}					21,19%
											0,12%
											0,05%
											0,16%
											0,58%
											1,07%
											10,03%
											13,72%
											4,75%
											4,55%
											1,35%
											0,69%
											11,59%
											0,45%
											0,06%
											2,71%
											3,81%
											3,00%
											5,63%
											5,92%

Nota: Construcción propia

El valor resultante es una probabilidad de default, calculada para cada una de las empresas de la muestra, tanto para las que entraron en default, como para las que se consideran sólidas y nunca han entrado en default.

Al revisar el comportamiento de la probabilidad de default en la muestra de empresas, se puede evidenciar aspectos relevantes en los que algunas de ellas muestran una baja probabilidad de incumplimiento en los años anteriores al que entraron en default y una probabilidad mucho más alta en el año del default. Un claro ejemplo de este tipo es la empresa cuyo Nit es 800042706 cuyas probabilidades de default son 1,10% en el año 2008, 0,98% en el año 2009, 7,19% en el año 2010 y 21,19% en el año 2011 (año del default). Para este caso, cada una de sus variables se fue deteriorando con el paso de los años, hasta llegar a tener un valor de  $x_1$  negativo. Caso contrario sucede con la empresa con Nit 815002520 cuyas probabilidades de default son 9,38% en el año 2009, 1,32% en el año 2010, 3,65% en el año 2011 y 0,47% en el año 2012 (año del default). Para este caso, la empresa mejoró su  $x_1$  y su  $x_2$  en el transcurso de los años, pero fue deteriorando su  $x_3$ .

### **3.4. Análisis de resultados Modelo LOGIT vs. Z1 Altman y CA-Store**

Teniendo el modelo definido, es importante determinar el nivel de confiabilidad que éste puede tener en el momento de estimar la probabilidad de que una empresa caiga en default en un momento específico, bajo el esquema planteado.



Por esta razón, vale la pena recordar que dentro de la muestra de empresas utilizada existen dos tipos: 57 empresas insolventes (que se pueden denominar de esta manera debido a que entraron en default) y 838 empresas solventes (que hasta el momento del estudio no habían entrado en default). Según esta clasificación se enfocará el análisis del modelo.

Para poder hacer un análisis es necesario comparar el modelo Logit con los principales modelos ya conocidos, y que se han enfocado en las empresas manufactureras, tales como el Z1 de Altaman y el CA-Score. De las empresas contempladas dentro de la muestra se toman los estados Financieros del año más reciente, que para las empresas consideradas como insolventes coincide con el año en el que entraron en default. Paso siguiente, se corren los diferentes modelos, resaltando que en el modelo Logit se toma como punto de referencia la probabilidad de default del 3%, tanto para las empresas solventes como para las insolventes: si su probabilidad de default es menor o igual que el 3%, se considera que la empresa es solvente y, si ésta es mayor que el 3% la empresa es considerada insolvente y puede entrar en default en los años próximos, si no llegara a tomar las medidas necesarias para evitarlo.

**Tabla 11. Análisis modelo LOGIT para empresas insolventes**

Calificación	Z1	%	CA-Score	%	LOGIT	%
Insolvente	47	82,5%	32	56,1%	38	66,7%
Zona Gris	9	15,8%	0	0,0%	0	0,0%
Solvente	1	1,8%	25	43,9%	19	33,3%
<b>Total</b>	<b>57</b>	<b>100,0%</b>	<b>57</b>	<b>100,0%</b>	<b>57</b>	<b>100,0%</b>

Nota: Construcción propia

La tabla 11 contiene la comparación de los diferentes modelos para la muestra de empresas que entraron en default. Para este caso el modelo que mejor se ajusta para predecir la insolvencia de las empresas de la muestra es el Z1 Altman y le seguiría el modelo Logit.

**Tabla 12. Análisis modelo LOGIT para empresas solventes**

Calificación	Z1	%	CA-Score	%	LOGIT	%
Insolvente	224	26,7%	108	12,9%	98	11,7%
Zona Gris	467	55,7%	0	0,0%	0	0,0%
Solvente	147	17,5%	730	87,1%	740	88,3%
<b>Total</b>	<b>838</b>	<b>100,0%</b>	<b>838</b>	<b>100,0%</b>	<b>838</b>	<b>100,0%</b>

Nota: Construcción propia

La tabla 12 contiene la comparación de los diferentes modelos para la muestra de empresas que no entraron en default (consideradas solventes). Para este caso el modelo que mejor se ajusta para predecir la solvencia de las empresas de la muestra es el Logit, dado que el 88,3% de las empresas se considerarían solventes, y le seguiría el modelo CA-Score con el 87,1%.

Después de revisar y comparar los resultados obtenidos, se puede determinar que el modelo Logit es útil para predecir la probabilidad de insolvencia de las empresas manufactureras en Colombia.

## CONCLUSIONES

La lógica nos obligaría a asumir que las empresas que entraron en default deberían presentar probabilidades muy altas y, por su parte, las empresas sólidas deberían tener probabilidades de default muy bajas. Sin embargo, esta hipótesis no se puede generalizar debido a varios factores:

- El LR test del modelo puede presentar fallas en el momento de determinar probabilidades de default individuales, por problemas en la data, antes que por el modelo per se.
- Este modelo posee las limitaciones de todo instrumento predictor que utiliza información histórica para pronosticar el futuro y se asume que el comportamiento del pasado se a mantendrá en el futuro.
- Una empresa que entró en default, teniendo una probabilidad baja, se puede ver afectada por un cambio drástico en los ratios financieros de mayor peso dentro del modelo, que para nuestro caso es las ventas / total activos.
- La muestra puede presentar errores en los datos, debido a que en la Superintendencia de Sociedades existen empresas con años faltantes en su reporte de estados financieros y otras que aunque tienen la secuencia de éstos, su información no coincide con la realidad económica de la misma.

Los ratios financieros empleados en este tipo de modelos, puede variar dependiendo del tipo de empresas analizadas, del país donde se desarrolla la actividad económica, del momento histórico en que se analizan afectado por variables exógenas como cambios importantes en las variables macroeconómicas del país, entre otros.

El  $R^2$  es muy útil para determinar la precisión del ajuste en las regresiones del tipo lineal. Para el caso de las regresiones no lineal, como la utilizada en el modelo construido durante el objeto de la presente investigación, el Pseudo- $R^2$  es el más acertado, teniendo presente que altos valores mostrarán un mejor ajuste del modelo.

Si bien es cierto, el signo del coeficiente revela el efecto directo de una variable, en donde si éste es positivo la probabilidad de default aumenta cuando el valor de la variable (o indicador financiero utilizado) aumenta, y viceversa; no es fácil cuantificar el efecto por cuanto la probabilidad de default no es una función lineal de todas las variables y los coeficientes.

Los resultados arrojados por el modelo Logit, para el caso de empresas insolventes, se pueden considerar como confiables, basados en que llegó a superar por 10,6 pp el resultado obtenido por el modelo CA-Score, a pesar de quedar por debajo en 15,8 pp del Z1 de Altman, que es de los más conocidos y

estudiados durante los últimos años. Sin embargo, este resultado puede estar influenciado por varios factores:

- A pesar de haber utilizado las principales fuentes de información para la recolección de los datos, como la Superintendencia de Sociedades, este dato depende del reporte que realizan las empresas, los cuales no siempre corresponden con la situación económica real de ellas o, en algunos casos, no los reportan.
- Los tres modelos utilizan variables diferentes (ratios financieros) que, basados en el país de las empresas estudiadas, pueden conducir a resultados diferentes.
- El punto de referencia tomado para determinar el punto de quiebra de las empresas en el modelo Logit fue del 3% y puede mejorar en la medida en que éste disminuya. Sin embargo, este nivel funciona no sólo para las empresas insolventes sino para las solventes.

Para el caso de las empresas solventes, el modelo logit arrojó mejores resultados, superando inclusive al modelo Z1 de Altman por 70,8 pp y al CA-Score por 1,2%, pudiendo estar influenciado por los factores enunciados anteriormente. De la misma manera, es importante destacar que el punto de referencia que se tomó fue el 3%: probabilidades por encima se considerarían insolventes, o tendrían una alta posibilidad de entrar en default y, probabilidades de default por

debajo, se considerarían solventes. Este punto de referencia se puede disminuir con el fin de sensibilizarlo, empezando así a ajustarse mejor a empresas insolventes y alejándose del ajuste para empresas solventes.

### Lista de referencias

Adams, T.; Low, M; Samkin, Grant. (2012). The use of Z-Scores to predict finance company collapses: A Research Note. *New Zealand Journal of applied business research*. 10, 2, p. 69.

Altman, Edward. Financial Ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of finance*. Septiembre 23 de 1968.

Altman, Edward. Examining Moyer's reexamination of forecasting financial failfure. *Financial management* 7, 1978.

Altman, Edward. Zeta Analisis: a new model to identify bankruptcy risk. *Journal of banking and finance*, 1977.

Astorga Hilbert, A. (2002). Modelos de predicción de la insolvencia empresarial. *Instituto Federal de Especialistas de Concursos Mercantiles*.

Carrillo de Rojas, Gladys. Análisis y administración financiera. *Corcas Editores Ltda*. Séptima Edición, 2011.

CA-Score, a warning system for small business failures. Bilanas, 1987.

Fernández, Felipe; Sarmiento, Benjamín. Estadística descriptiva, introducción al análisis de datos. *Ediciones de la U*. Primer Edición, 2014.

Fulmer, John G. Jr.; Moon, James E.; Gavin, Thomas A.; Erwin, Michael J. A Bankruptcy Classification Model for Small Firms. *Journal of Commercial Bank Lending*, 1984.

Gil Zuluaga, J. (2010). Detección de alertas tempranas que indiquen la necesidad de acuerdos de reestructuración financiera en compañías colombianas. Tesis de maestría no publicada. *Universidad Nacional de Colombia*. Bogotá, Colombia.

Gran parte de las Pymes colombianas son informales, (2010). Recuperado de <http://actualicese.com/actualidad/2010/09/20/gran-parte-de-las-pymes-colombianas-son-informales/>

Gray, D.; S. Malone. Macrofinancial Risk Analysis. *The Wiley Finance Series*. London, Wiley, 2008.

León García, O. Administración financiera – Fundamentos y aplicaciones. *Prensa*



*moderna impresores*. Cuarta edición. 2009.

Las 1.000 pymes más grandes del país. En *Latinpyme*, (2010). *Revista especializada en formación empresarial y negocios*. Edición No. 73. P. 15-22.

Löffler, G. y Posch, P. Credit risk modeling using Excel and VBA. *John Wiley & Sons, Ltd.*, 2007.

Malone, Samuel; Rodríguez, Abel; Ter Horst, Enrique. Bayesian Inference for a Structural Credit Risk Model with Stochastic Volatility and Stochastic Interest Rates. *Journal of Financial Econometrics*. 2014.

Martínez, O. (2003). Determinantes de fragilidad en las empresas colombianas. Banco de la República. *Reporte de estabilidad financiera*.

Oceda, C. y Apari, H. Excel for Masters – Macros y Aplicaciones VBA. *Empresa Editora Macro*. Primera Edición, 2011.

Ortega, J; Martínez, J. y Valencia J. El modelo de calificación crediticia Z-Score – Aplicación en la evaluación del riesgo crediticio de HB Fuller Colombia Ltda. *Revista MBA, EAFIT*. P. 102.

Rodríguez, Astrid Génova. La realidad de la PYME Colombiana, desafío para el desarrollo. *Fundes Internacional*, 2003.

Rosillo, J. (2002). Modelo de predicción de quiebras de las empresas colombianas. *INNOVAR, revista de ciencias administrativas y sociales*. No. 19.

Springate, Gordon L.V. J. Predicting the Possibility of Failure in a Canadian Firm. *Unpublished M.B.A. Research Project, Simon Fraser University*. January, 1978.

Superintendencia de Sociedades (2.015). Estados financieros empresas manufactureras en Colombia y empresas en liquidación judicial acumulada al 30 de junio de 2015, bases SIREM (Sistema de Información y Reporte Empresarial), recuperado el 31 de julio de 2015.  
<http://sirem.supersociedades.gov.co/Sirem2/>

Winsor, Charles P. Low moments for small samples: a comparative study of order statistics. *Annals of Mathematical Statistics*. 1947.

## Anexo 1. Codificación del modelo en Visual Basic

Option Explicit

Function Logit(y As Range, xraw As Range, Optional constant, Optional stats)

If IsMissing(constant) Then constant = 1

If IsMissing(stats) Then stats = 0

'Count variables

Dim i As Integer, j As Integer, jj As Integer

'Read data dimensions

Dim K As Integer, N As Integer

N = y.Rows.Count

K = xraw.Columns.Count + constant

'Some error checking

If xraw.Rows.Count <> N Then MsgBox "error"

'Adding a vector of ones to the x matrix if constant=1, name xraw=x from now on

Dim x() As Double

```
ReDim x(1 To N, 1 To K)
```

```
For i = 1 To N
```

```
    x(i, 1) = 1
```

```
    For j = 1 + constant To K
```

```
        x(i, j) = xraw(i, j - constant)
```

```
    Next j
```

```
Next i
```

```
'Initializing the coefficient vector (b) and the score (bx)
```

```
Dim b() As Double, bx() As Double, ybar As Double
```

```
ReDim b(1 To K)
```

```
ReDim bx(1 To N)
```

```
ybar = Application.WorksheetFunction.Average(y)
```

```
If constant = 1 Then b(1) = Log(ybar / (1 - ybar))
```

```
For i = 1 To N
```

```
    bx(i) = b(1)
```

```
Next i
```

```
'Defining the variables used in the Newton procedure
```

```
Dim sens As Double, maxiter As Integer, iter As Integer, change As Double
```

```
Dim lambda() As Double, lnL() As Double, dlnL() As Double, hesse() As
Double, hinv(), hinvg()
```

```
ReDim lambda(1 To N)
```

```
sens = 1 * 10 ^ (-11): maxiter = 50
```

```
ReDim lnL(1 To maxiter)
```

```
change = sens + 1: iter = 1: lnL(1) = 0
```

```
'Loop for Newton iteration
```

```
Do While Abs(change) > sens And iter < maxiter
```

```
    iter = iter + 1
```

```
    'reset derivative of log likelihood and Hessian
```

```
    Erase dlnL, hesse
```

```
    ReDim dlnL(1 To K): ReDim hesse(1 To K, 1 To K)
```

```
'Compute prediction Lambda, gradient dlnL, Hessian hesse, and log
likelihood lnL
```

```
For i = 1 To N
```

```
    lambda(i) = 1 / (1 + Exp(-bx(i)))
```

```
For j = 1 To K
```

```
    dlnL(j) = dlnL(j) + (y(i) - lambda(i)) * x(i, j)
```

```

For jj = 1 To K
    hesse(jj, j) = hesse(jj, j) - lambda(i) * (1 - lambda(i)) * x(i, jj) * x(i, j)
Next jj
Next j
lnL(iter) = lnL(iter) + y(i) * Log(1 / (1 + Exp(-bx(i)))) + (1 - y(i)) * Log(1 -
1 / (1 + Exp(-bx(i))))
Next i

'Compute inverse Hessian (=hinv) and multiply hinv with gradient dlnL
hinv = Application.WorksheetFunction.MInverse(hesse)
hinvg = Application.WorksheetFunction.MMult(dlnL, hinv)

change = lnL(iter) - lnL(iter - 1)

'If convergence achieved, exit now and keep the b corresponding with the
estimated hessian
If Abs(change) <= sens Then Exit Do

' Apply Newton's scheme for updating coefficients b
For j = 1 To K
    b(j) = b(j) - hinvg(j)
Next j

```

```
'Compute new score (bx)
For i = 1 To N
    bx(i) = 0
    For j = 1 To K
        bx(i) = bx(i) + b(j) * x(i, j)
    Next j
Next i
Loop

'some error handling
If iter > maxiter Then
    MsgBox "Maximum Number of Iteration exceeded. No convergence
achieved. Exiting. Sorry."
GoTo myend
End If

'output
Dim reLogit()
ReDim reLogit(1 To 1, 1 To K)
If stats = 1 Then ReDim reLogit(1 To 7, 1 To K)

'Coefficients
```

```
For j = 1 To K
```

```
  reLogit(1, j) = b(j)
```

```
Next j
```

```
'Additional statistics if requested
```

```
If stats = 1 Then
```

```
  For j = 1 To K
```

```
    reLogit(2, j) = Sqr(-hinv(j, j))
```

```
    reLogit(3, j) = reLogit(1, j) / reLogit(2, j)
```

```
    reLogit(4, j) = (1 -
```

```
Application.WorksheetFunction.NormSDist(Abs(reLogit(3, j)))) * 2
```

```
    reLogit(5, j) = "#N/A"
```

```
    reLogit(6, j) = "#N/A"
```

```
    reLogit(7, j) = "#N/A"
```

```
  Next j
```

```
'ln Likelihood of model with just a constant(lnL0)
```

```
Dim lnL0 As Double
```

```
lnL0 = N * (ybar * Log(ybar) + (1 - ybar) * Log(1 - ybar))
```



reLogit(5, 1) = 1 - lnL(iter) / lnL0    'McFadden R2

reLogit(5, 2) = iter - 1        'Number of iterations

reLogit(6, 1) = 2 \* (lnL(iter) - lnL0)    'LR test

reLogit(6, 2) = Application.WorksheetFunction.ChiDist(reLogit(6, 1), K - 1)

'p-value for LR

reLogit(7, 1) = lnL(iter)

reLogit(7, 2) = lnL0

End If

Logit = reLogit

GoTo myend

'Error Handler

error:

MsgBox ("Fatal Error. Reasons might be: y not {0,1}, not the same number  
of N for y and x's...or anything else")

myend:

End Function

Function XTRANS(defaultdata As Range, x As Range, numranges As  
Integer)

```
Dim bound, numdefaults, obs, defrate, N, j, defsum, obssum, i
ReDim bound(1 To numranges), numdefaults(1 To numranges)
ReDim obs(1 To numranges), defrate(1 To numranges)

N = x.Rows.Count

End Function
```