Riesgo Operativo: Modelo LDA con Risk Simulator

Investigación Terminada

Laura León Ardila
Ingeniería Financiera
lleon562@unab.edu.co

Diego Hernández Macías
Ingeniería Financiera
dhernandez259@unab.edu.co

Ángela Prada Barajas
Ingeniería Financiera
aprada577@unab.edu.co

Johanna Pérez Quintero
Ingeniería Financiera
kperez626@unab.edu.co

Andrea Delgado Rico
Ingeniería Financiera
Adelgado574@unab.edu.co

Universidad Autónoma de Bucaramanga

RESUMEN

Este artículo aborda un análisis de los fundamentos y simulación del modelo de Distribución de Pérdidas Agregadas como una metodología de medición del Riesgo Operacional, basados en los requerimientos que Basilea II propuso en el 2.004 para controlar, medir y gestionar el riesgo. En el artículo se presenta la conceptualización del modelo LDA y parámetros necesarios para su medición incluyendo métodos numéricos y usando como herramienta para su simulación el Risk Simulator.

ABSTRACT

This article addresses a fundamental analysis and simulation model LDA as a methodology for measuring operational risk, based on Basilea II requirements proposed in 2004 to monitor, measure and manage risk losses. Article conceptualization LDA model and parameters for measurement including numerical methods and using as a tool for simulation the Risk Simulator.

Palabras Clave

Distribución de pérdidas, frecuencia, severidad, riesgo operativo, tipos de evento, líneas de negocio, OPVaR

INTRODUCCIÓN

Para una entidad financiera es de suma importancia minimizar y gestionar los riesgos financieros que se presenten en cada una de sus líneas de negocio y controlar así las actividades y decisiones que puedan afectar la buena imagen de la entidad.

Basilea define al riesgo operacional como riesgo de pérdida resultante de una falta de adecuación o de un fallo de los procesos, el personal o los sistemas internos, o bien como consecuencia de acontecimientos externos [1]. De igual manera el nuevo acuerdo de Basilea II complementa la definición con una clasificación de los siete tipos de evento que se consideran en el riesgo operacional como lo son: fraude interno, fraude externo, relaciones laborales y seguridad en el puesto de trabajo, prácticas con clientes, productos y negocios, daños a activos materiales, incidencia en los negocios y fallos en los sistemas y ejecución, entrega y gestión de procesos.

Uno de los métodos para medir, cuantificar y evaluar el riesgo operacional es el modelo de Distribución de Pérdidas Agregadas (LDA) que tiene como objetivo obtener una función de distribución que conlleve al modelamiento de frecuencia y severidad como dos grandes características y parámetros a tener presentes, proporcionando una estimación de la pérdida por eventos en las líneas de negocio.

El artículo se divide en dos partes: la primera hace alusión a la conceptualización del modelo de Distribución de Pérdidas Agregadas mencionando y describiendo la las dos características y parámetros Frecuencia y Severidad, para obtener así la cuantía de las pérdidas esperadas y no esperadas; y en la segunda parte se mostrará la simulación paso a paso por medio del programa Risk Simulator siendo esto un apoyo para la toma de decisiones de las entidades financieras.

CONTENIDO DEL ARTÍCULO

En el artículo se presenta la metodología del modelo de Distribución de pérdidas Agregadas mencionando sus características, fundamentos de los parámetros Frecuencia y Severidad; el procedimiento para llegar a obtener la pérdida esperada y no esperada y en la última parte presenta la simulación con la herramienta Risk Simulator, paso a paso de tal manera que sea un apoyo para las entidades financieras

Contenido Según Categorías de Participación

Este artículo es una investigación terminada.

MODELO DE DISTRIBUCIÓN DE PÉRDIDAS AGREGADAS

El modelo de Distribución de Pérdidas Agregadas se define como la suma aleatoria de diferentes pérdidas donde Sij, es la pérdida en la celda i que representa las ocho líneas de negocio: finanzas corporativas, negociación y ventas, banca minorista, banca comercial, liquidación y pagos, servicio de agencia, administración de activos e intermediación minorista; y la j determina los siete tipos de evento, dentro de la matriz de pérdidas y se calcula como:

$$S = \sum_{i=1}^{8} \sum_{j=1}^{7} S_{ij}$$

Donde N representa la variable aleatoria de número de eventos de riesgo en la celda i, j y Xn el monto de pérdida en la celda i, j.

$$S_{ij} = \sum_{N=1}^{n} X_n$$

Frecuencia y Severidad

Las variables frecuencia y severidad son dos parámetros necesarios para el modelo de Distribución de pérdidas agregadas. Estos son las que dan vida a la distribución y son modeladas por separado puesto que cada una contempla una serie de características únicas y diferentes.

VARIABLE FRECUENCIA

Variable discreta que representa el número de eventos ocurridos en uno de los siete tipos de evento i y una de las ocho líneas de negocio j, en un determinado intervalo de tiempo.

VARIABLE SEVERIDAD

Variable aleatoria continua que representa la cuantía de la pérdida monetaria en un determinado intervalo de tiempo. Para la modelación de la severidad se tienen presentes las siguientes distribuciones de probabilidad.

PROCESO METODOLÓGICO:

El proceso metodológico que se realiza para el Modelo de Distribución de Pérdidas Agregadas se sostiene del desglose de cada una de las pérdidas operacionales, extraído de la matriz líneas de negocio/tipos de eventos; siguiendo los pasos así:

Con la base de datos de registro de pérdidas por riesgo operativo, se filtran por líneas de negocio y tipos de evento.

Se selecciona el intervalo o periodicidad para la frecuencia, que puede ser mensual, semestral, anual.

Se realiza un ajuste a las variables aleatorias frecuencia y severidad, es decir, se escoge las mejores distribuciones para cada una teniendo presente los supuestos:

-Hace relación a las variables N y Xn, donde el número de eventos es decir la variable aleatoria frecuencia es independiente al monto de la pérdida, en este caso la variable aleatoria severidad.

-Los dos siguientes supuestos toman como base la variable severidad; el segundo toma a las observaciones de tamaño de pérdida Xn (variable severidad), dentro de una misma clase donde se distribuyen idénticamente y el tercer supuesto afirma que las observaciones de tamaño de pérdida Xn (Variable aleatoria de severidad) dentro de una misma clase con independientes.

Para el cálculo de las pérdidas agregadas se realiza la convulución entre la distribución de frecuencias y la distribución de severidad: OPVaR= Frecuencia x Severidad.

Se generan múltiples iteraciones para las dos distribuciones: Frecuencia y Severidad con el software Risk Simulator.

La distribución de pérdidas agregadas se enmarca en la pérdida total ligada a las líneas de negocio y tipos de eventos.

Por último se hace el cálculo de las pérdidas esperadas y no esperadas.

Distribuciones discretas

Se entiende por variables discretas las que toman diferentes valores enteros finitos y se obtiene al azar; sus valores a tomar son positivos.

-Propiedades:

p (xi) <1 Las probabilidades asociadas a cada uno de los valores que toma x deben ser mayores o iguales a cero y menores o iguales a 1.

Ep (xi) = 1 La sumatoria de las probabilidades asociadas a cada uno de los valores que toma x debe ser igual a 1.

Las principales distribuciones discretas para frecuencia: (Binomial negativa)

BINOMIAL

Tiene que ver con el experimento aleatorio que se produce en cada ensayo o prueba uno de dos resultados posibles mutuamente excluyentes: ocurrencia de éxito y no ocurrencia de fracaso.

Sea “p” la probabilidad de éxito cada vez que el experimento se realiza y “q” la probabilidad de fracaso. Sea X la variable aleatoria que representa el número de éxitos en los n ensayos. El interés se centra en conocer la probabilidad de obtener exactamente x éxitos en esos n ensayos.

Resumiendo, podemos definir estos criterios:

1- El experimento aleatorio consiste en n ensayos o pruebas repetidas, e idénticas y fijadas antes del experimento. Son pruebas con reemplazamiento o con reposición.

3- La probabilidad del llamado éxito p, permanece constante para cada ensayo o prueba.

4- Cada prueba o ensayo se repite en idénticas condiciones y es independiente de las demás.

5-Generalmente la distribución binomial es sesgada o asimétrica hacia la derecha, sesgo que se va perdiendo cuanto más grande sea el valor de n y en la medida en que p se acerque a 0.5 y asi con q donde se torna asimétrica.

$$f(x)P(X = x) = \left\{p^{x(1-p)^x} = 0,1,2 \ldots n\right\}$$

POISSON

Representa el número de éxitos independientes que ocurren en un intervalo (amplios) ya sea de tiempo, lugar, espacio, además con una probabilidad de ocurrencia pequeña.

-El número de éxito en el intervalo es limitado

- Las ocurrencias son independientes, no afectan a otros intervalos

-El número promedio de ocurrencias debe permanecer el mismo de intervalo a intervalo.

$P(x) = \frac{e^{-y} * \gamma}{x}$ para x y >0 ; Media = y

Desviación estándar= raíz γ ; Asimetría: $\frac{1}{\sqrt{\gamma}}$

Distribuciones continuas

Se entiende por variable continua las que toman cualquier valor infinito, ya sea fraccionarios o no, encontrado en un intervalo.

-Propiedades:

El área definida bajo la función de densidad de probabilidad deberá ser de 1

$$Fx(x) = prob(X \leq x)$$

Las principales distribuciones continuas para severidad,

LOG NORMAL

Se entiende por distribución log-normal como la probabilidad de una variable aleatoria cuyo logaritmo está normalmente distribuido. Es decir, si X es una variable aleatoria con una distribución normal, entonces exp(X) tiene una distribución log-normal. La distribución además tiende a la función densidad de probabilidad.

Una variable puede ser modelada como log-normal si puede ser considerada como un producto multiplicativo de muchos pequeños factores independientes. Un ejemplo típico es un retorno a largo plazo de una inversión: puede considerarse como un producto de muchos retornos diarios.

Desviación estándar del logaritmo de variable.

El valor esperado es: $E(X) = e^{\mu+\sigma^2/2}$

Varianza es: $var(X) = (e^{\sigma^2} -)\ e^{2\mu+\sigma^2}$

LOGISTICA

Se entiende por distribución logística como la distribución de una probabilidad continua cuya función de distribución es la función logística. Se parece a la distribución normal en su forma, pero tiene colas más pesadas y, por lo tanto, menor curtosis.

SIMULACIÓN

SEVERIDAD

Se realiza la Intro a Excel teniendo en cuenta el programador Risk Simulator. En la primera hoja se realiza la organización de los datos obtenidos en el riesgo operativo - sistemas (se toman en meses por la base de datos), se escoge cual es la severidad (Perdida Asociada) y la Frecuencia (Numero de Fallas en el sistema) para realizar el respectivo análisis, ya organizados los datos si entra a la pestaña de simulación de riesgo que tiene el programa y se crea un perfil de simulación (nombre del perfil Sistemas y números de pruebas 5000 interacciones), continuando con el procedimiento se señala la columna de los datos de severidad, entra en herramienta analíticas y si da la opción 11 (Ajuste de distribución simple), se selecciona las opción perdidas agregadas, opción ajuste de distribución continua (Severidad) – se seleccionan todas las distribuciones para su ajuste. A continuación se abre una pantalla que muestra el orden de distribuciones que mejor se ajustan en un resumen estadístico, mostrando los parámetros de cada distribución.

Distribución	Estudio Estadístico	Valor-P	Rango
Normal	0,03	85,43 %	1
Logística	0,03	69,99 %	2
Gamma	0,04	52,79 %	3
PERT	0,04	45,65 %	4
Parabólico	0,05	35,09 %	5
Logarítmica Normal	0,05	31,35 %	6
Lognormal Desplazada	0,05	31,35 %	7
Laplace	0,05	25,43 %	8
Doble Logaritmo	0,06	7,94 %	9
Triangular	0,13	0,00 %	10
Exponential Desplazada	0,16	0,00 %	11
Coseno	0,16	0,00 %	12
Uniforme	0,25	0,00 %	13
Exponencial	0,30	0,00 %	14
Arcoseno	0,31	0,00 %	15
Pareto	0,47	0,00 %	16

Se toman las dos primeras distribuciones más significativas de severidad y se organiza en hojas diferente para su análisis.

Continuando con el procedimiento se señala la columna de los datos de Frecuencia, entra en herramienta analíticas y si da la opción 11 (Ajuste de distribución simple), se selecciona las opción perdidas agregadas, opción ajuste de distribución separada (Frecuencia) – se seleccionan todas las distribuciones para su ajuste. A continuación se abre una pantalla que muestra el orden de distribuciones que mejor se ajustan en un resumen estadístico, mostrando los parámetros de cada distribución:

Distribución	Estudio Estadístico	MAPE %	Rango
Binomial	5,15	98,36 %	1
Poisson	7,70	99,99 %	2
Discreta Uniforme	34,38	0,00 %	3
Geométrica	58,84	0,26 %	4
Hypergeométrica	769,15	0,00 %	5
Binomial Negativa	1,339740E+005	0,00 %	6
Pascal	Infinito	0,00 %	7
Bemoulli	Infinito	0,00 %	8

Se toman las dos primeras distribuciones más significativas de severidad y se organiza en hojas diferente para su análisis y se realiza una tabla de comparación, de igual forma se hace para las distribuciones más significativas de frecuencia.

DATOS OBTENIDOS DE SEVERIDAD (pérdidas asociadas)	
Distribución Ajustada	Normal
Media	123401,30
Desv.Est	34686,53
Prueba Estadística para P-Value	85,43%
Distribución Ajustada	Logística
Alfa	123763,79
Beta	21295,38
Prueba Estadística para P-Value	69,99%

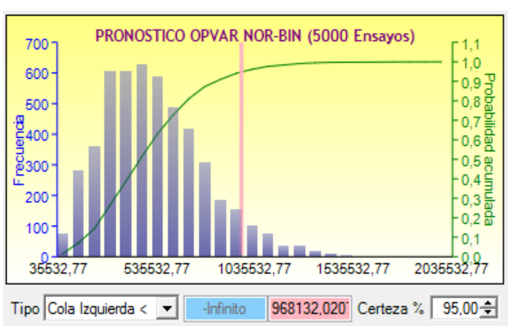
DATOS OBTENIDOS DE FRECUENCIA (Número de fallas en el sistema)	
Distribución Ajustada	Binomial
Pruebas	17,00
Probabilidad	0,23
MAPE % for Test Statistic	98,36%
Distribución Ajustada	Poisson
Lambda	3,94
MAPE % for Test Statistic	99,99%

Después de analizar las distribuciones se realiza una relación entre ellas tomando la frecuencia 1 con la severidad 1, la frecuencia 2 con la severidad 2 y así sucesivamente, para que arrojen los pronósticos OP – VAR.

Teniendo separadas las relaciones entre severidad – frecuencia, armamos una pequeña tabla en la que se encuentran frecuencia, severidad y pronostico OP-VAR sistemas, y para completarla con sus respectivos datos necesitamos para el dato de frecuencia, seleccionamos la celda en la que se encontrara el dato numérico, enseguida nos dirigimos a la pestaña supuesto de entrada, se selecciona la gráfica obtenida en frecuencia y le anexamos los datos más relevantes de esa grafica para su respectivo análisis; se debe realizar el mismo procedimiento para la severidad y para hallar el dato numérico del pronóstico OP-VAR se señala la pestaña pronóstico de salida (se debe tener en cuenta dos parámetros el del 95% y del 99% de confianza; este procedimiento se realiza para cada una de las distribuciones escogidas y organizadas, ya realizado este procedimiento se va para la opción velocidad máxima para que nos arroje el análisis de las 5000 interacciones para un posible escenario, obteniendo los siguientes datos:

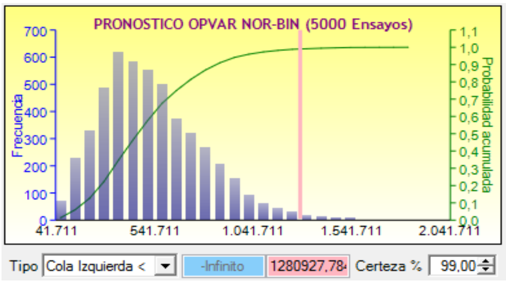
RELACION DE PRONÓSTICO DE LAS DISTRIBUCIONES NORMAL-BINOMIAL

Parámetro del 95% de confianza y un margen de error 5%.



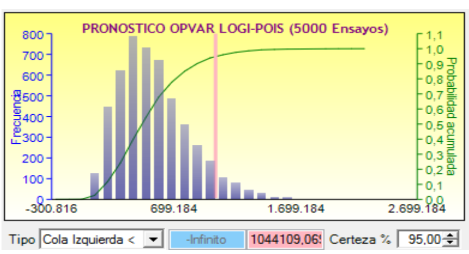
Estadísticas	Resultado
Número de simulaciones	5000
Media	487.229,6522
Mediana	455.175,3647
Desviación Estándar	262.254,9780
Variación	68.777.673.499,4730
Coeficiente de Variación	0,5383
Máximo	1.979.167,0365
Mínimo	-16.350,6934
Rango	1.995.517,7299
Asimetría	0,8141
Curtósis	1,0725
25% Percentil	291.859,4305
75% Percentil	637.702,8558
Precisión de Error al 95% de Confianza	1,4919%

Parámetro del 99% de confianza y un margen de error 1%.



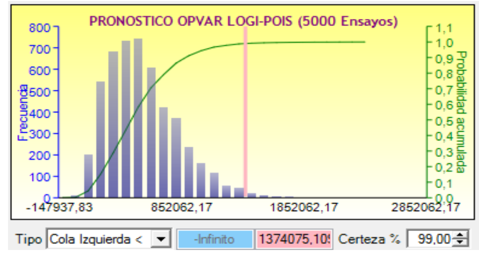
Estadísticas	Resultado
Número de simulaciones	5000
Media	495.157,5130
Mediana	455.252,4632
Desviación Estándar	265.929,1613
Variación	70.718.318.814,6951
Coeficiente de Variación	0,5371
Máximo	1.828.086,6891
Mínimo	-5.974,5112
Rango	1.834.061,2003
Asimetría	0,8120
Curtósis	0,8556
25% Percentil	303.413,7131
75% Percentil	653.373,8914
Precisión de Error al 95% de Confianza	1,4886%

RELACION DE PRONÓSTICO DE LAS DISTRIBUCIONES LOGISTICA-POISSONParámetro del 95% de confianza y un margen de error 5%.



Estadísticas	Resultado
Número de simulaciones	5000
Media	486.870,4219
Mediana	441.650,4914
Desviación Estándar	300.255,1193
Variación	90.153.136.680,7264
Coeficiente de Variación	0,6167
Máximo	2.264.412,4874
Mínimo	-369.291,8719
Rango	2.633.704,3593
Asimetría	1,0067
Curtósis	1,6289
25% Percentil	268.374,1609
75% Percentil	649.488,9201
Precisión de Error al 95% de Confianza	1,7094%

Parámetro del 99% de confianza y un margen de error 1%.



Estadísticas	Resultado
Número de simulaciones	5000
Media	488.208,3892
Mediana	443.341,5539
Desviación Estándar	298.732,2040
Variación	89.240.929.734,9966
Coeficiente de Variación	0,6119
Máximo	2.346.128,9734
Mínimo	-214.514,5655
Rango	2.560.643,5389
Asimetría	1,0167
Curtósis	1,9005
25% Percentil	273.315,9490
75% Percentil	655.583,3340
Precisión de Error al 95% de Confianza	1,6961%

Si se quiere analizar los datos en una forma más completa se puede utilizar la opción de Excel (histograma) que se encuentra en la parte inferior derecha del cuadro obtenido anteriormente, donde nos arroja las diferentes variables de análisis y graficas de pronóstico para cada escenario.

Obtenidos los datos de los dos parámetros escogidos (95% y 99%) podemos realizar el análisis del OP-VAR obtenido en el proceso;

OPVaR	EL + UL	
PÉRDIDA ESPERADA	EL	μ
PÉRDIA NO ESPERADA	UL	OPVaR - μ

En la distribución Normal – Binomial, se obtiene los siguiente datos;

Distribución NORMAL-BINOMIAL	95%	99%
OPVaR	968,132.02	1,280,927.78
Pérdida Esperada μ	487,229.65	495,157.51
Pérdida NO Esperada (OpvaR - μ)	480,902.37	785,770.27

En la Distribución Logística - Poisson , se obtiene los siguientes datos;

Distribución LOGÍSTICA-POISSON	95%	99%
OPVaR	1,044,109.7	1,374,075.1
Pérdida Esperada μ	486,870.4	488,208.4
Pérdida NO Esperada (OpvaR - μ)	557,239.3	885,866.7

Obtenidos los datos del OP-VAR se realiza por último se realiza una simulación con los dos parámetros (95% y 99%) donde une los tres escenarios para su respectivo análisis comparativo.

IDENTIFICACIÓN DEL PROYECTO

Nombre del Semillero	Semillero Investigacion Academica Financiera
Tutor del Proyecto	Luz Helena Carvajal
Grupo de Investigación	GIF
Línea de Investigación	Riesgo Operativo
Fecha de Presentación	

REFERENCIAS

[1] M. Á Nieto Giménez-Montesinos, El tratamiento de Riesgo Operacional en Basilea II, Estabilidad Financiera No 8, Banco de España.

Ding, W. and Marchionini, G. 1997. A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 01 - 06, 2000). CHI '00. ACM, New York, NY, 526-531. DOI= http://doi.acm.org/10.1145/332040.332491.

Tavel, P. 2007. Modeling and Simulation Design. AK Peters Ltd., Natick, MA.

Sannella, M. J. 1994. Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE. In Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology (Vancouver, Canada, November 02 - 05, 2003). UIST '03. ACM, New York, NY, 1-10. DOI= http://doi.acm.org/10.1145/964696.964697.

Yu, Y. T. and Lau, M. F. 2006. A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions. J. Syst. Softw. 79, 5 (May. 2006), 577-590. DOI= http://dx.doi.org/10.1016/j.jss.2005.05.030.

Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender, Ed. ACM Press Frontier Series. ACM, New York, NY, 19-33. DOI= http://doi.acm.org/10.1145/90417.90738.