

Synergy Vision

Riesgo de Crédito

Ciencia de los Datos Financieros



Índice general

Índice de cuadros	VII
Índice de figuras	IX
Prefacio	XI
Acerca del Autor	XV
1. Introducción	1
2. Introducción a la teoría de probabilidad	1
2.1. Espacio de probabilidad	1
2.2. Variables aleatorias y función de distribución. . . .	3
2.2.1. Variable aleatoria	3
2.2.2. Tipos de variables aleatorias	4
2.2.3. Distribución de probabilidad:	4
2.2.4. Función de distribución:	5
2.2.5. Algunos tipos de variables aleatorias	6
2.2.6. Esperanza de una variable aleatoria.	8
2.2.7. Varianza de una variable aleatoria.	9
2.2.8. Covarianza de variables aleatorias.	9
3. Conceptos de la gestión del riesgo de crédito	1
3.1. Pérdida esperada	1
3.2. Probabilidad de incumplimiento (DP)	2
3.3. Exposición (<i>EAD</i>)	3
3.4. Pérdida dado el incumplimiento (<i>LGD</i>)	4
3.5. Pérdida inesperada (<i>UL</i>)	4
3.6. Capital económico o valor en riesgo	5
3.7. Matriz de transición	6

4. Scoring de credito	1
4.1. Estructura del Scoring de Crédito	2
4.1.1. Los modelos probit y logit.	4
4.1.2. Interpretación del modelo	6
4.1.3. Relación entre el score y el riesgo	7
4.1.4. Uso de modelos lineales para la selección de variables (Métodos de búsqueda secuencial).	8
5. CreditRisk+	1
5.1. La distribución del número de incumplimientos con probabilidades de impago fijas.	2
5.2. La agrupación por bandas de exposición las pérdidas	4
5.3. La determinación de la <i>distribución de pérdidas</i> de la cartera.	5
5.4. Obteniendo la distribución de las pérdidas	6
5.5. Obtención del VaR	8
6. CreditMetrics	1
6.1. Elementos que usa Credimetrics	1
6.2. El caso de un solo crédito	4
6.2.1. Ejemplo 1	5
6.2.2. Ejemplo 2	7
6.3. Cartera de dos de créditos con probabilidades de transición independientes.	8
6.3.1. Ejemplo 3	8
6.4. Aproximación de la distribución de pérdidas de una cartera con n créditos.	10
6.4.1. Simulación de MonteCarlo	10
7. Manual técnico de la aplicación Vision CreditRisk	1
7.1. Presentación de la aplicación.	1
7.2. Primera sección: CreditRisk+	1
7.2.1. Datos	3
7.2.2. Estadísticos:	3
7.2.3. Selección de variables	5
7.3. Pérdida por incumplimiento	6
7.4. Score de Crédito.	6
7.4.1. Selección y resultados del modelo	6

7.4.2. Score de la cartera de crédito.	7
7.4.3. Proyección a nuevos clientes.	7
7.5. Parametros y resultados.	7
7.5.1. Parámetros iniciales	8
7.5.2. Resultados	9
7.6. Stress Testing	9
7.7. Segunda sección: Credimetrics	10
7.8. Créditos	11
7.9. Matriz de transición	12
7.9.1. Cálculo de la matriz de transición	12
7.9.2. Selección de la matriz de transición	12
7.10. Pérdida por clase.	14
7.10.1. Cálculo de pérdida esperada	14
7.10.2. Pérdida esperada	14
7.11. Simulación y resultados.	15
7.12. Stress Testing.	15
7.13. Tercera sección: Indicadores contables	17
7.14. Cálculo del RAROC	17



Índice de cuadros



Índice de figuras

3.1.	6
3.2.	7
6.1.	12
7.1.	2
7.2.	2
7.3.	3
7.4.	4
7.5.	4
7.6.	5
7.7.	5
7.8.	6
7.9.	7
7.10.	8
7.11.	8
7.12.	9
7.13.	10
7.14.	10
7.15.	11
7.16.	11
7.17.	13
7.18.	13
7.19.	14
7.20.	15
7.21.	16
7.22.	16
7.23.	17
7.24.	18



Prefacio



La versión en línea de este libro se comparte bajo la licencia Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License¹.

¿Por qué leer este libro?

En este libro presentamos dos metodologías para la cuantificación del riesgo de crédito: *CreditRisk+* y *CreditMetrics*, las cuales tienen como principal objetivo crear previsiones objetivas que las instituciones financieras logren tener para afrontar momentos de dificultad que sean consecuencia de eventos de impagos de sus clientes.

Además, este libro nos ofrece herramientas adicionales tales como el *Scoring* la cual es una herramienta que nos permite dar una calificación crediticia a las carteras de clientes de las instituciones, lo cual es fundamental en metodologías como la de *CreditRisk+*.

Este libro tiene la virtud de que el lector no debe poseer un *Back-ground* e conocimientos matemáticos tan profundos para abordarlo, pero se incluye una sección dedicada a los conceptos provenientes de las probabilidades que es de gran valor si se pretende profundizar en el área.

La finalidad final de ese libro es orientar a los especialistas a

¹<http://creativecommons.org/licenses/by-nc-sa/4.0/>

usar la aplicación web *Vision CreditRisk*, la cual nos proporciona un ambiente facil y comodo para poder implementar todos los conceptos que veremos en este libro, es por esto que se incluye un capítulo final dedicado a esto.

Estructura del libro

Este libro esta estructurado en dos capitulos teóricas, tres capítulos que contemplan las metodologias a usar y un capítulo orientado al uso de la aplicación *Vision CreditRisk*.

Los capitulos teóricos estan dispuesto para que el lector tengo los conomocimientos mínimos para que pueda obtener el mayor provecho de la informacion que en este texto encontrara, el primer capítulo posee los conceptos teóricos de la teoría de probabilidad los cuales son las bases de la mayoría de los métodos existentes para la cuantificación de riesgo de crédito y un segundo capítulo que contiene los conceptos relacionados con el riesgo de crédito los cuales son fundamentales para cualquier analista o especialista que se encuentre dentro de un departamento de riesgo en una institución financiera.

Los capítulos metodologicos nos orientan a las herramientas que usaremos para los propositos practicos que nos interesan, el primero hace referencia a medicion de las probabilidades de que un cliente incumpla sus compromisos de deuda, este tipo de herramientas es conocida como *Scoring*, el segundo capítulo nos presente la metodología *CreditRisk+* la cual tiene una visión mas actuarial y el tercer capítulo nos presenta la metodología credimetrics la cual tiene una visión mas estadística y posee un mayor nivel de dificultad con respecto al computo necesario para su aplicación.

El último capítulo como ya se ha mencionado esta orientado al uso correcto de la aplicación *Vision CreditRisk* la cual aplica las metodologias que se explican en este libro, este capítulo tiene una

descripción detallada la cual esta acompañada de imagenes del paso por paso para la obtención de nuestros resultados.

Información sobre los programas y convenciones

Este libro es posible gracias a una gran cantidad de desarrolladores que contribuyen en la construcción de herramientas para generar documentos enriquecidos e interactivos. En particular al autor de los paquetes Yihui Xie xie2015.

Agradecimientos

A todo el equipo de Synergy Vision que no deja de soñar. Hay que hacer lo que pocos hacen, insistir, insistir hasta alcanzar. Lo más importante es concretar las ideas. La idea es sólo el inicio y solo vale cuando se concreta.

Synergy Vision, Caracas, Venezuela



Acerca del Autor

Este material es un esfuerzo de equipo en Synergy Vision, (<http://synergy.vision/nosotros/>). En especial al equipo conformado por Danny Morales y Arturo Carreño. Los cuales en conjunto y con su experiencia en el sector financiero crearon este sistema para la facilitar a las instituciones cumplir con las regulaciones obligatorias exigidas por el ente regulatorio.

Danny Morales: Licenciado en Ciencias Actuariales, cuenta con varios años de experiencia en el sector financiero y en la actualidad es Gerente General de Synergy Vision y profesor universitario (UCV)

Arturo Carreño: Liceciado en Matemáticas, cuenta con experiencia en el sector financiero y en la actualidad es Cientifico de Datos Financieros de Synergy Vision y profesor universitario (UCV)

Nuestra cuenta de Twitter es (<https://twitter.com/bysynergyvision>) y nuestros repositorios están en GitHub (<https://github.com/synergyvision>).

Somos Científicos de Datos Financieros



1

Introducción

En este texto estudiaremos las metodologías mas usados para medir el riesgo crediticio. A lo largo de la historia la necesidad de las instituciones financieras de tener una medida que permita tomar previsiones en momentos de crisis ha ido evolucionando a la mano con el desarrollo de las teorías y modelos estadísticos. Esta necesidad es natural, debido a que cuando se realiza un prestamo, el banco esta corriendo el riesgo de que en cualquier instante de tiempo el cliente deje de cumplir sus compromisos, este incumplimiento puede ocurrir desde el primer momento en que deba empezar a pagarse el prestamo o al final del mismo, por lo que el banco esta expuesto independientemente de la confianza que tenga en el cliente a perder el 100 % del monto dado (a menos que hayan garantias). Un banco puede tener una cartera de miles inclusive millones de clientes, por lo tanto esta expuesto constantemente a sufrir pequeñas pérdidas que ocasionen los incumplimientos, por lo que la institución debe contar con fondos para poder cubrir dichos riesgos, es ridiculo pensar que la institución deba guardar la misma cantidad de fondos de lo que presta, por lo tanto debe crear mecanismos estadisticos o contables que permitan medir de manera objetiva los fondos que deben poseerse para el resguardo del banco y permitir superar posibles crisis económicas. En este texto hacedmon una descripción de las principales metodologías que se de este problema, mas precisamente las metodologías CreditRisk+ y Credimetrics.

Como hemos mencionado, las herramientas provenientes de la estadística y matemáticas son fundamentales para la construcción de las metodologías, por lo que en el primer capítulo realizaremos un recuento de las principales definiciones de la probabilidad y estadística que serán usadas a lo largo de todo el texto.

Mencionamos que los bancos están propensos a que los clientes incumplan, por lo cual es natural asignarle a cada cliente una probabilidad de que caiga en mora, aunque a primera instancia pareciera fácil dicha asignación, esta puede ser tan complicada como queramos, pues una población puede estar compuesta por varios sectores, los cuales pueden ser tan heterogéneos como queramos. Los bancos contruyen herramientas para asignar un valor o score a los clientes y a través del puntaje hallar la probabilidad de que los clientes incumplan sus compromisos. Este tipo de metodologías suelen recibir el nombre de CreditScore o Scoring de crédito. En el capítulo 4 damos una descripción de uno de los CreditScore más populares usando modelos lineales generalizados.

En el capítulo 5 y 6 daremos una descripción de las metodologías CreditRisk+ y Creditmetrics, la primera mucho más estadística y propensa a crear errores de aproximación grandes si no se cuenta con una buena herramienta de Score, además tiene supuestos fuertes relacionados con la independencia entre los préstamos y pide que las probabilidades de incumplimiento sean bajas, en caso contrario no será ideal su uso, la segunda Creditmetrics es una herramienta más precisa y más fácil de implementar pero con la desventaja de que necesita para sus cálculos, información que solo está disponible en mercados desarrollados, por lo que usarla en mercados emergentes, suele ser complicado por la falta de datos históricos y de registros.

Una vez presentado este texto se espera que el personal encargado de las áreas de crédito pueda adaptar la información disponible y poder realizar una buena aproximación de las métricas de riesgo y así lograr que la institución financiera sea capaz de tomar previsiones para los momentos de crisis, que suelen ser imprevistos y algunas veces aleatorios.

2

Introducción a la teoría de probabilidad

El mundo actual esta regido por metodologías que en su mayoría están estructuradas por hipótesis matemáticas, estas en la mayoría de los casos son fundamentales y de no cumplirse, los modelos pueden dejar de ser prácticos, por ejemplo, imagínese que usted quiera introducir un motor de un automóvil de la década de los sesenta en su automóvil moderno, con esto no queremos decir que no pueda funcionar, solo que estaría ignorando los principios mecánicos de su carro y con mucha seguridad provocaría fallas futuras. Igual pasa con los modelos de riesgo de crédito, por esto, dedicaremos este capítulo a dar un breve repaso, a la teoría de probabilidad y algún otro concepto matemático que sea necesario conocer.

2.1. Espacio de probabilidad

Los espacios de probabilidad son las estructuras que se utilizan en matemáticas para estudiar los fenómenos aleatorios, es decir, aquellos fenómenos que tienen un alto grado de incertidumbre en sus resultados.

Los modelos de probabilidad trabajan con las “probabilidades” de los diferentes resultados o sucesos que nos puedan interesar en un fenómeno aleatorio, y nos dan respuestas más que satisfactorias en el estudio de dichos fenómenos.

El enorme interés de la Teoría de la Probabilidad reside en que, aparte de tener unos preciosos desarrollos matemáticos, permite

estudiar, a través de la *Inferencia Estadística*, los problemas que surgen en los estudios científicos que intentan obtener conclusiones a partir de datos reales.

Definición 2.1. El *espacio muestral*, Ω , es el conjunto de todos los posibles resultados del fenómeno aleatorio que deseamos estudiar.

Definición 2.2. Una σ -álgebra de sucesos, S , es un subconjunto de partes de Ω que verifica:

Si $A \in S$, entonces $A^c \in S$ (c denota el complemento de un conjunto).

Si $A_1, A_2, \dots, A_n, \dots \in S$, entonces $\cup_n A_n \in S$

La idea de una σ -álgebra es que contenga todos los sucesos en los que podemos estar interesados en un fenómeno aleatorio.

Ejemplo: Cuando el espacio muestral, Ω , es finito o numerable, la σ -álgebra habitualmente utilizada es partes de Ω . Es decir, consideramos que cualquier subconjunto de Ω puede ser un suceso de interés. Esta elección es la más sencilla y no ofrece ningún problema.

Con este par de definiciones podemos definir lo que es una medida de probabilidad.

Definición 2.3. Consideramos un espacio muestral Ω y una σ -álgebra S . Una *función de probabilidad* o *medida de probabilidad* es una función $P : S \rightarrow [0, 1]$, que verifica:

$$P(\Omega) = 1$$

Si $A_1, A_2, \dots, A_n, \dots \in S$ son disjuntos, entonces:

$$P\left(\cup_n A_n\right) = \sum_n P(A_n)$$

La manera más sencilla de entender el motivo de esta definición es pensar que el concepto de probabilidad trata de formalizar lo que ocurre con las frecuencias relativas de los distintos sucesos que podemos encontrarnos al lanzar un dado con 6 caras numeradas del 1 al 6:

- La frecuencia relativa de cualquier suceso siempre es un número entre 0 y 1.
- La frecuencia relativa de que salga algún número de los posibles (es decir, entre 1 y 6) es, evidentemente, 1.
- La frecuencia relativa de la unión de sucesos disjuntos es la suma de sus frecuencias relativas.

Esta forma de interpretar la medida de probabilidad no significa que esta sea la única posible, solo que a manera de ejemplo es la más fácil de explicar y visualizar.

Definición 2.4. Definimos como *espacio de probabilidad* a la terna (Ω, S, P) .

2.2. Variables aleatorias y función de distribución.

En este tema se tratará de formalizar numéricamente los resultados de un fenómeno aleatorio. Por tanto, una variable aleatoria es un valor numérico que corresponde a un resultado de un experimento aleatorio. Algunos ejemplos son: número de caras obtenidas al lanzar seis veces una moneda, número de llamadas que recibe un teléfono durante una hora, tiempo de fallo de una componente eléctrica, etc.

2.2.1. Variable aleatoria

Definición 2.5. Dado un experimento aleatorio y asociado al mismo, un espacio de probabilidad Ω, S, P , una variable aleatoria es una aplicación $X : \Omega \rightarrow \mathbb{R}$, que a cada valor de X , del espacio muestral, le hace corresponder un número real.

Ejemplo: Consideramos un experimento aleatorio de lanzar una moneda al aire tres veces y anotamos el resultado. Se define la

variable aleatoria X como número de caras aparecidas en los tres lanzamientos.

2.2.2. Tipos de variables aleatorias

Las variables aleatorias se clasifican en discretas y continuas:

- *Discreta*: La variable aleatoria X se dice que es discreta si los números asignados a los sucesos elementales de Ω son puntos aislados. Sus posibles valores constituyen un conjunto finito o infinito numerable. Por ejemplo, supongamos el experimento consistente en lanzar tres veces una moneda no trunca; si consideramos la variable aleatoria X ="número de caras obtenidas en los tres lanzamientos", los valores que puede tomar esta variable aleatoria son finitos"(0,1,2,3)".
- *Continua*: La variable aleatoria X será continua si los valores asignados pueden ser cualesquiera, dentro de ciertos intervalos, es decir, puede tomar cualquier valor real. Por ejemplo, si consideramos el experimento aleatoria consistente en medir el nivel de agua en un embalse y tomamos la variable aleatoria X ="nivel de agua", esta puede tomar cualquier valor positivo.

2.2.3. Distribución de probabilidad:

Es un modelo teórico que describe la forma en que varían los resultados de un experimento aleatorio, es decir, nos da todas las probabilidades de todos los posibles resultados que podrían obtenerse cuando se realiza un experimento aleatorio. Se clasifican como discretas o continuas. En la distribución de probabilidad discreta está permitido tomar sólo un número limitado de valores. En la continua, llamada función de densidad, la variable que se está considerando puede tomar cualquier valor dentro de un intervalo dado.

2.2.3.1. Distribución de probabilidad discreta

Sea X una variable aleatoria discreta que toma como posibles valores x_1, x_2, \dots, x_n , se define la distribución de probabilidad de X como el conjunto de pares (x_i, p_i) que a cada valor de la variable le asocia una probabilidad, donde $p_i = P(X = x_i)$, tal que la suma de todas las probabilidades es igual a uno.

Por ejemplo, en el experimento del dado la distribución de probabilidad viene dado por: $(1;1/6), (2;1/6), (3;1/6), (4;1/6), (5;1/6), (6;1/6)$

2.2.3.2. Distribución de probabilidad continua

Si la variable aleatoria es continua, hay infinitos valores posibles de la variable y entre cada dos de ellos se podrían definir infinitos valores. En estas condiciones no es posible deducir la probabilidad de un valor puntual de la variable como se puede hacer en el caso de las variables discretas. Pero sí es posible calcular la probabilidad acumulada hasta un cierto valor (función de distribución) y cómo cambia esa probabilidad acumulada en cada punto (densidad de probabilidad). Por tanto, cuando la variable aleatoria sea continua hablaremos de función de densidad.

Definición 2.6. Sea X una variable aleatoria continua, se llama función de densidad y se representa como $f(x)$ a una función no negativa definida sobre la recta real, tal que para cualquier intervalo (a, b) que estudiemos se verifica:

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

2.2.4. Función de distribución:

La función de distribución describe el comportamiento probabilístico de una variable aleatoria X asociada a un experimento aleatorio y se representa como F_x . De igual forma estudiaremos el caso discreto y continuo.

2.2.4.1. Caso discreto

Sea X una variable aleatoria discreta asociada a un espacio de probabilidad, se define la función de distribución:

$$F_x : R \rightarrow [0, 1] \quad \text{que verifica} \quad F_x(x) = P[X \leq x] = \sum_{x_i \leq x} P_i$$

2.2.4.2. Caso continuo

Sea X una variable aleatoria continua con función de densidad $f(x)$, se define la función de distribución, F_x , como:

$$F_x(x) = P[X \leq b] = \int_{-\infty}^b f(x) dx$$

2.2.5. Algunos tipos de variables aleatorias**2.2.5.1. Bernoulli**

Es un experimento en el cual solo existen dos eventos, en general éxito(1) y fracaso(0), y además tiene asociada un parámetro p el cual es la probabilidad de éxito, es decir,

$$P(X = 1) = p \quad \text{y} \quad P(X = 0) = 1 - p$$

2.2.5.2. Binomial

La distribución de probabilidad discreta más usada es la distribución binomial y consiste en :

- Repetir un experimento de Bernoulli n veces de forma idéntica
- Cada experimento es independiente de los anteriores

La variable aleatoria de Bernoulli es la que expresa el número de éxitos, por lo tanto esta variable aleatoria tiene asociada dos parámetros: un parámetro p que nos indica la probabilidad de éxito de cada evento, y un número natural n que nos indica la cantidad de veces que se realizó el experimento.

2.2.5.3. Poisson

Esta distribución es una de las más importantes distribuciones de variable discreta. Sus principales aplicaciones hacen referencia a la modelización de situaciones en las que nos interesa determinar el número de hechos de cierto tipo que se pueden producir en un intervalo de tiempo o de espacio, bajo supuestos de aleatoriedad y ciertas circunstancias restrictivas. Otro de sus usos frecuentes es la consideración límite de procesos dicotómicos reiterados un gran número de veces si la probabilidad de obtener un éxito es muy pequeña .

La distribución de Poisson tiene ajustado un parámetro positivo λ que representa el número de veces que se espera que ocurra el fenómeno durante un intervalo dado. La probabilidad de que un experimento de Poisson tenga k ocurrencias es

$$\frac{e^{-\lambda} \lambda^k}{k!}$$

2.2.5.4. Distribución uniforme

Es la distribución que sigue una variable aleatoria X que toma valores en un intervalo $[a, b]$ con la misma probabilidad, su función de densidad es:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{en otro caso} \end{cases}$$

2.2.5.5. Distribución exponencial

Es usada muchas veces para modelizar el comportamiento de variables aleatorias del tipo “tiempo transcurrido hasta el fallo de un componente industrial” o “el tiempo que se tarda en completarse un proceso determinado”. Depende de un parámetro positivo λ , que esta relacionado con la media de ocurrencia de un evento. La

función de densidad es:

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \lambda e^{-\lambda x} & \text{si } x > 0 \end{cases}$$

2.2.5.6. Distribución normal

La distribución normal o gaussiana es muy importante puesto que se utiliza para modelar muchísimos fenómenos aleatorios; además incluso para aproximar otras distribuciones. La distribución normal aproxima lo observado en muchos procesos de medición sin errores sistemáticos, por ejemplo medidas físicas del cuerpo humano, medidas en la calidad de ciertos procesos industriales, entre otros.

Esta distribución depende de dos parámetros: un parámetro μ que es un indicador de la promedio del experimento que se está estudiando y un parámetro positivo σ que es un indicador de la volatilidad del mismo. Su función de densidad es

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

2.2.6. Esperanza de una variable aleatoria.

La esperanza matemática de una variable aleatoria X es el número que expresa el valor medio del fenómeno que representa dicha variable.

2.2.6.1. Esperanza de una variable aleatoria discreta.

Dada una variable aleatoria X que toma valores $x_1, x_2, x_3, \dots, x_n$ con distribución de probabilidad $P(X = x_i) = P_i$, se define la esperanza matemática de una variable aleatoria como:

$$E[X] = \sum_{i=1}^n x_i P(X = x_i) = \sum_{i=1}^n x_i P_i$$

2.2.6.2. Esperanza de una variable aleatoria continua.

Sea X una variable aleatoria continua con función de densidad $f(x)$, se define la esperanza matemática de esa variable aleatoria como:

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

2.2.7. Varianza de una variable aleatoria.

La varianza de una variable aleatoria es una característica numérica que proporciona una idea de la dispersión de la variable aleatoria respecto de su esperanza. Decimos que es un parámetro de dispersión y se define por:

$$Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$$

2.2.8. Covarianza de variables aleatorias.

La covarianza es un valor que indica el grado de variación conjunta de dos variables aleatorias. Es el dato básico para determinar si existe una dependencia entre ambas variables y además es el dato necesario para estimar otros parámetros básicos y se define como: Sean X, Y variables aleatorias se define la covarianza entre X y Y como:

$$Cov(X, Y) = E((X - E(X))(Y - E(Y)))$$



3

Conceptos de la gestión del riesgo de crédito

A la hora de aplicar o construir cualquier modelo matemático debemos conocer los conceptos que están involucrados en la integración de este, la falta de dicho conocimiento casi seguramente nos conducirá a tesis erradas y modelos mal elaborados. El manejo del riesgo de crédito no se escapa a esto, en el presente capítulo definiremos las nociones y conceptos básicos que estarán presente prácticamente en todos los capítulos del libro.

3.1. Pérdida esperada

En general las instituciones financieras se encuentran bajo distintos escenarios en los cuales se ven comprometidas, esto se debe a que dependen de que sus clientes cumplan sus compromisos adquiridos. Debido a la necesidad de tener que controlada la incertidumbre e intentar cuantificar los posibles riesgos de la insolvencia de las contrapartes, las instituciones usan herramientas estadísticas para lograr lo anterior. La idea es que la institución asigne una probabilidad de incumplimiento (DP), una fracción de pérdida llamada pérdida en caso de incumplimiento (LGD), la cantidad que esta en riesgo, llamada exposición en caso de impago (EAD). Así, podemos definir la pérdida de un deudor, como la variable de pérdida:

$$\tilde{L} = EAD * LGD * L$$

donde L una variable aleatoria Bernoulli, que es 1 si el deudor incumple 0 si no, el parametro de esta variable Bernoulli es $P(D) = DP$ donde D denota el evento de incumplir.

Ahora, estamos interesados no tanto en la pérdida, si no en el valor medio que se espera perder, es decir, $\mathbb{E}(\tilde{L})$ que denotaremos por EL y es la pérdida esperada. Para calcular, veamos que:

$$\mathbb{E}(\tilde{L}) = \mathbb{E}(EAD * LGD * L) = EAD * LGD * \mathbb{E}(L) = EAD * LGD * DP$$

Este modelo, tienes algunos detalles, por ejemplo, estamos suponiendo que las variables EAD y LGD son constantes, pero puede haber situaciones en que estas variables sean aleatorias, esto impide el cálculo anterior, pero en general, EAD se considera constante, y si se toma LGD como variable aleatoria, definimos por severidad su esperanza, es decir: $SEV = \mathbb{E}(LGD)$

3.2. Probabilidad de incumplimiento (DP)

La tarea de asignar una probabilidad de incumplimiento (default probability) a cada uno de los clientes de la cartera de crédito del banco dista mucho de ser fácil. Hay esencialmente dos enfoques para las probabilidades de incumplimiento:

- Aproximación a partir de datos de mercado:

El más famoso representante de este tipo de aproximación es el concepto de de frecuencias por defecto esperadas (expected default frequencies [EDF]) de KVM Corporation.

Otro método para calibrar las probabilidades de incumplimiento a partir de los datos de mercado se basa en los márgenes de crédito de los productos negociados que conllevan riesgo de crédito.

- Aproximación a partir de ratings:

En esta aproximación, la probabilidad de incumplimiento esta asociada con ratings, y estos ratings son asociado a los clientes por agencias externas de ratings o por metodologías internas de la institución financiera.

3.3. Exposición (*EAD*)

El *EAD* es la cantidad que representa la exposición que tiene el deudor sobre la institución. En general, la exposición consta de dos partes principales, los valores en circulación y los compromisos. Los valores en circulación se refieren a la parte de la exposición que ya ha sido tomada por el deudor. En caso de incumplimiento del prestatario, el banco está expuesto a la importe total de los valores en circulación. Los compromisos se pueden dividir en dos porciones, retirados y no retirados, en el tiempo anterior al incumplimiento. El total el importe de los compromisos es la exposición que el banco ha prometido prestar al deudor a petición de éste. La experiencia histórica predeterminada muestra que los deudores tienden a recurrir a líneas de crédito comprometidas en épocas de crisis financiera. La angustia. Por lo tanto, el compromiso también está sujeto a pérdidas en en el caso de incumplimiento del deudor, pero sólo el importe retirado (anterior al incumplimiento).

La fracción que describe la descomposición de los compromisos en partes retiradas y no retiradas es una variable aleatoria debido al carácter opcional que tienen los compromisos (el deudor tiene el derecho, pero no la obligación, de utilizar líneas de crédito comprometidas). Entonces es natural definir la *EAD* por :

$$EAD = OUTST + (\gamma COMM)$$

donde *OUTST* denota los valores en circulación y *COMM* los compromisos y γ es la proporción que se espera el deudor use, γ toma valores entre 0 y 1, si asumimos que *EAD* es determinística, entonces γ también, en caso contrario, γ es una variable aleatoria.

3.4. Pérdida dado el incumplimiento (LGD)

El LGD de una operación está más o menos determinada por “1 menos la tasa de recuperación”, es decir, el LGD cuantifica la parte de la pérdida que el banco realmente sufrirá en caso de incumplimiento. La estimación de estas cotizaciones de pérdidas dista mucho de ser sencilla, ya que los tipos de recuperación dependen de muchos factores determinantes, por ejemplo, de la calidad de las garantías (valores, hipotecas, etc.) y de la antigüedad del crédito del banco sobre los activos del prestatario. Esta es la razón detrás de nuestra convención para considerar la pérdida por incumplimiento como una variable aleatoria que describe la gravedad de la pérdida. La noción LGD se refiere entonces a la expectativa de la severidad. Una fuente externa del banco para determinar la tasa de recuperación proviene de las agencias de calificación. Por ejemplo, Moody's proporciona los valores de recuperación de los bonos impagados, distinguiendo así entre distintas antigüedades.

Desafortunadamente, muchos bancos no tienen buenos datos internos para estimar las tasas de recuperación. De hecho, aunque la severidad es un factor clave de la pérdida esperada, en comparación con otros factores de riesgo como la DP , se ha avanzado poco en el camino hacia una calibración sofisticada.

3.5. Pérdida inesperada (UL)

Al principio de este capítulo presentamos el EL de una transacción como un seguro o reserva para pérdidas con el fin de cubrir las pérdidas que el banco espera de la experiencia histórica de los incumplimientos. Pero no basta con tener capital como colchón frente a las pérdidas esperadas. De hecho, el banco deberá, además de la reserva para pérdidas esperadas, ahorrar dinero para cubrir

pérdidas inesperadas que excedan el promedio de pérdidas experimentadas en el pasado. Una medida de magnitud de la desviación de las pérdidas esperadas, es la desviación estándar de la variable \tilde{L} , por esta razón, se define la pérdida inesperada por:

$$UL = \sqrt{\mathbb{V}(\tilde{L})} = \sqrt{\mathbb{V}(EAD \times SEV \times L)}$$

3.6. Capital económico o valor en riesgo

Hasta ahora hemos aprendido que los bancos deberían tener algún tipo de colchón de capital contra pérdidas inesperadas. Sin embargo, definir la UL de una cartera como el capital de riesgo ahorrado para casos de dificultades financieras no es la mejor opción, porque podría haber una probabilidad significativa de que las pérdidas superen la pérdida esperada de la cartera en más de una desviación estándar de la pérdida de la cartera. Por lo tanto, se buscan otras formas de cuantificar el capital riesgo, teniendo en cuenta un nivel objetivo de confianza estadística. La forma más común de cuantificar el capital riesgo es el concepto de capital económico o valor en riesgo (VaR). Para un nivel de confianza específico α , donde α representa el nivel de confianza de la pérdida esperada, es decir:

$$VaR_{\alpha} = q_{\alpha} - EL$$

donde q_{α} es el cuantil α de la distribución anterior mencionada:

$$q_{\alpha} = \inf\{q > 0 | P[\tilde{L} \leq q] \geq \alpha\}$$

Un ejemplo gráfico de lo que está ocurriendo es presentado a continuación:

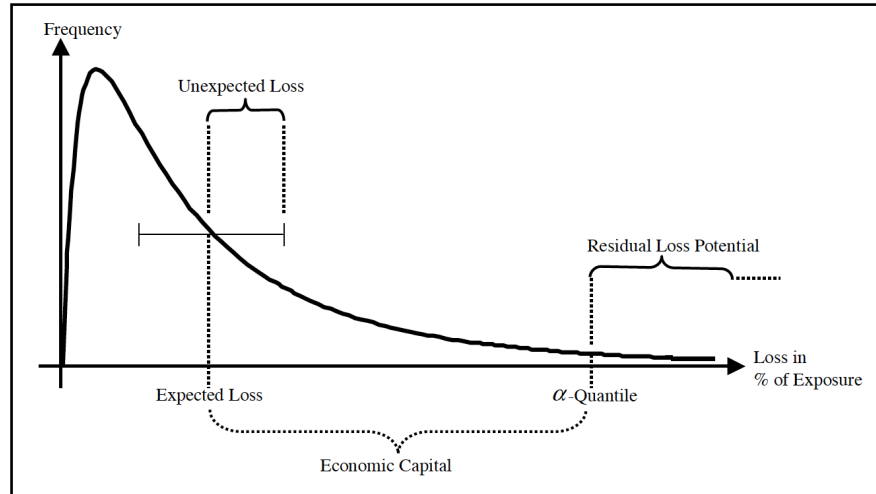


Figura 3.1

3.7. Matriz de transición

La matriz de transición es un instrumento que se usa de manera fundamental en la metodología Credimetrics, que será desarrollada mas adelante. Como su nombre lo indica es una matriz, que además es cuadrada, el orden de la matriz depende de la cantidad de calificaciones que use la institución financiera para calificar sus créditos, por ejemplo, si la institución usa 5 calificaciones, la matriz será una matriz cuadrada de orden 5

Ahora, la matriz de calificación indica proporciona la probabilidad de que un crédito con una calificación específica migre a otra calificación determinada, por ejemplo, por la imagen anterior, la probabilidad de que un crédito con calificación BB migre a una calificación B es de 0.188, es decir, se espera que el 18,8 % de los créditos con calificación BB migren a la calificación B.

La matriz de transición también es conocida con el nombre de matriz de transición de probabilidad. Para que una matriz sea

Tabla 4. Matriz de transición con el método discreto

	AA	A	BB	B	CC	D
AA	0,855	0,034	0,032	0,030	0,027	0,026
A	0,116	0,384	0,174	0,135	0,109	0,082
BB	0,009	0,079	0,394	0,188	0,172	0,158
B	0,008	0,039	0,087	0,396	0,264	0,237
CC	0,000	0,000	0,024	0,122	0,473	0,359

Fuente: cálculos de los autores.

Figura 3.2

considerada como una matriz de transición crediticia, debe cumplir las siguientes propiedades:

- Todos sus elementos son no negativos y menores o iguales a 1
- La suma de los elementos de cualquier columna es exactamente 1.

En general las matrices de transición esta intimamente relacionadas con las políticas internas de las instituciones financieras, y son generadas por metologías internas, y por lo general son calculadas con simples medias exponenciales de migraciones.



4

Scoring de credito

Los métodos o modelos de credit scoring, a veces denominados score-cards o classifiers, son algoritmos que de manera automática evalúan el riesgo de crédito de un solicitante de financiamiento o de alguien que ya es cliente de la entidad. Tienen una dimensión individual, ya que se enfocan en el riesgo de incumplimiento del individuo o empresa, independientemente de lo que ocurra con el resto de la cartera de préstamos. Este es uno de los aspectos en los que se diferencian de otras herramientas de medición del riesgo de crédito, como son los modelos de cartera y los VaR marginales, que tienen en cuenta la correlación de la calidad crediticia de los deudores de una cartera de préstamos.

Los modelos de scoring fueron introducidos a partir del año 1970 en el análisis del otorgamiento del crédito, pero generalizados después de 1990 gracias al desarrollo estadístico y tecnológico. Entre los métodos para la construcción de modelos de scoring se pueden tener en cuenta como los más comunes, el modelo LOGIT, PROBIT, las Redes Neuronales y el Análisis Discriminante.

El resultado de la evaluación se refleja en la asignación de alguna medida que permita comparar y ordenar a los evaluados en función de su riesgo, a la vez que cuantificarlo. Por lo general, los modelos de credit scoring le asignan al evaluado un puntaje o score, o una calificación, clasificación o rating. Algunos métodos los asignan a grupos, en donde cada grupo tiene un perfil de riesgo distinto; sin embargo, en la práctica esto equivale a una calificación. A su vez, estos ordenamientos de los deudores permiten obtener estimaciones más concretas del riesgo; en general se busca obtener alguna estimación de la probabilidad de incumplimiento del deudor

(PD, por probabilidad de default) asociada a su score, rating o calificación

4.1. Estructura del Scoring de Crédito

Cuando al plantear un modelo la variable dependiente o a explicar toma valores discretos, se emplean modelos de regresión discreta. El caso más simple se da cuando ella es binaria y toma los valores 0 o 1, y se puede estimar con distintos enfoques como el modelo de probabilidad lineal, análisis discriminante, los modelos de tipo probit y logit o con una regresión logística.

Sea Y una variable aleatoria binaria que toma el valor 1 si ocurre el evento (el deudor cumple con los pagos normalmente) y 0 si entra en mora, se cuenta con una muestra aleatoria de n observaciones, Y_i , $i : 1, \dots, n$, y se define como Ω_i al conjunto de información relevante asociado con el individuo i , que se utilizará para explicar a Y_i .

Un modelo de elección binaria es un modelo de la probabilidad de ocurrencia del evento Y_i condicional en el conjunto de información Ω_i .

$$P_i = Pr(Y_i = 1|\Omega_i) \quad (4.1)$$

Dado que Y_i toma los valores 0 y 1, la esperanza de Y_i condicional en Ω_i es:

$$E(Y_i|\Omega_i) = (1 * P_i) + (0 * (1 - P_i)) = P_i = Pr(Y_i = 1|\Omega_i) \quad (4.2)$$

En consecuencia, cuando la variable a explicar es binaria, su esperanza condicional es la probabilidad condicional de ocurrencia del evento.

##El modelo de probabilidad lineal

Supóngase que Ω_i está constituido por un vector columna X_i compuesto por k variables explicativas, incluyendo a la ordenada al origen, β es un vector columna que contiene los parámetros correspondientes a las variables explicativas, y que se intenta modelar a la variable Y_i a través del modelo de probabilidad lineal, postulando la siguiente relación:

$$Y_i = \beta^T X_i + \epsilon_i, \quad \text{donde} \quad E(\epsilon_i|X_i) = 0 \quad \text{y} \quad E(\epsilon_i) = 0$$

y usando (4.2)

$$E(Y_i|X_i) = P_i = \beta^T X_i$$

El modelo de probabilidad lineal, como se observa en la ecuación (4.1), implica estimar un modelo lineal en los parámetros para Y_i . Los valores predichos deberían en su mayoría ubicarse en el intervalo $[0, 1]$, pudiendo ser interpretados como la probabilidad de que la variable a explicar tome alguno de estos valores.

Mientras que su estimación e interpretación es simple, su utilización se ha visto desalentada por dos problemas en la metodología. En primer lugar, como la esperanza condicionada de Y_i es igual a la probabilidad condicionada de ocurrencia del evento (de $Y_i = 1$), ella debería estar restringida al intervalo $[0, 1]$. Sin embargo, el modelo lineal no impone ninguna restricción sobre $\beta^T X_i$, asumiendo implícitamente que la variable dependiente puede tomar cualquier valor. Es por esto que el modelo puede estimar probabilidades negativas o mayores que uno, lo cual carece de significado económico. A su vez, el término de error de este modelo no es homocedástico, ya que la varianza condicional varía según las observaciones, por lo que las estimaciones de β no son eficientes.

Para resolver estos inconvenientes hay modelos econométricos, generalmente estimados por máxima verosimilitud, que tienen en cuenta la naturaleza discreta de la variable dependiente: se trata de los modelos de respuesta o elección binaria. Ellos utilizan ciertas funciones de distribución para la innovación, con el objetivo de

limitar las probabilidades estimadas al intervalo $[0,1]$: las más usadas son la función de probabilidad acumulada normal estándar y la función logística. Cuando se usa la normal estándar se trata de un modelo probit, y logit cuando se usa la función logística.

4.1.1. Los modelos probit y logit.

El modelo probit es una de varias alternativas para estimar modelos de respuesta binaria. La idea consiste en utilizar una función de transformación $F(x)$ que tenga las siguientes propiedades:

$$F(-\infty) = 0, \quad F(\infty) = 1 \quad \text{y} \quad f(x) \equiv \frac{dF(x)}{dx} > 0$$

$F(x)$ es una función monótona creciente que mapea de la línea real al intervalo $[0,1]$. Varias funciones de distribución acumulada tienen estas propiedades: la normal, la logística, la de Cauchy y la de Burr, entre otras. Estas distintas alternativas para los modelos de respuesta binaria consisten en una función de transformación $F(x)$ aplicada a una función índice que depende de las variables explicativas del modelo y que tiene las propiedades de una función de regresión, pudiendo ser lineal o no lineal.

La siguiente es una especificación general para cualquiera de los modelos de elección binaria:

$$E(Y_i|\Omega_i) = F(h(X_i))$$

Donde h es la función índice.

Si bien h puede ser cualquier función, en general se usa una función lineal, es decir:

$$E(Y_i|\Omega_i) = F(\beta^T X_i)$$

Por lo cual el modelo de elección binaria es simplemente una transformación no lineal de una regresión lineal, y si bien $\beta^T X_i$ puede

tomar cualquier valor sobre la línea real, $F(\beta^T X_i)$ está limitado al intervalo $[0,1]$. En el modelo probit, la función de transformación $F(x)$ es la función de distribución acumulada normal estándar, y por definición satisface las condiciones impuestas en (4.1.1). En este caso, el modelo de elección binaria puede escribirse de la siguiente manera:

$$P_i = E(Y_i|\Omega_i) = F(\beta^T X_i) = \Phi(\beta^T X_i) = \int_{-\infty}^{\frac{\beta^T X_i}{\sigma}} \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} ds$$

Cuando se trata del modelo logit, $F(x)$ es la función logística y el modelo de respuesta binaria se escribe como:

$$P_i = E(Y_i|\Omega_i) = F(\beta^T X_i) = \Phi(\beta^T X_i) = \frac{e^{\beta^T X_i}}{1 + e^{\beta^T X_i}}$$

Los modelos probit y logit pueden ser derivados de otro modelo que introduce una variable no observada o latente y^* , de la siguiente manera. Sea:

$$y_i^* = \beta^T X_i + \epsilon_i, \quad \text{con } \epsilon_i \sim iid(0, 1)$$

Si bien y^* no se observa, decimos que:

$$Y_i = 1 \quad \text{si } y^* > 0 \quad \text{y} \quad Y_i = 0 \quad \text{si } y^* < 0$$

Luego, la probabilidad que $Y_i = 1$ viene dada por:

$$\begin{aligned} P(Y_i = 1) &= P(y_i^* > 0) = P(\beta^T X_i + \epsilon_i > 0) = 1 - P(\beta^T X_i + \epsilon_i \leq 0) = 1 - P(\epsilon_i \leq -\beta^T X_i) \\ &= 1 - F(-\beta^T X_i) = F(\beta^T X_i) \end{aligned} \tag{4.3}$$

ya que se supone que ϵ_i tiene una distribución simétrica. Cuando $\epsilon_i \sim N(0, 1)$ F es Φ , la función de distribución de probabilidades

acumuladas normal estándar y se trata del modelo probit, mientras que si F es la función logística se trata de un modelo logit y su densidad también es simétrica alrededor de cero.

Estimación de los modelos logit y probit

La estimación de estos modelos se hace por el método de máxima verosimilitud. Con métodos numéricos se buscan los valores de β que maximizan la siguiente función logarítmica de verosimilitud:

$$l(\beta) = \sum_{i=1}^n \left(Y_i \log(F(\beta^T X_i)) + (1 - Y_i) \log(1 - F(\beta^T X_i)) \right)$$

Las condiciones de primer orden para un máximo en (4.1.1) son:

$$\sum_{i=1}^n \frac{(Y_i - \hat{F}_i) \hat{f}_i X_{ij}}{\hat{F}_i(1 - \hat{F}_i)} = 0, \quad j = 1, \dots, k$$

donde $\hat{F}_i \equiv F_i(b^T X_i)$ y $\hat{f}_i \equiv f_i(b^T X_i)$

Los modelos probit, logit y otros tipos de modelos de respuesta binaria satisfacen las condiciones de regularidad necesarias para que las estimaciones de los parámetros sean consistentes y asintóticamente normales, con la matriz de covarianzas asintótica dada por la inversa de la matriz de información.

4.1.2. Interpretación del modelo

En el contexto de los modelos de credit scoring se puede asociar $\beta^T X_i$ a la calidad crediticia del individuo (variable latente o no observada). Cambiando su denominación por Z_i , esta variable representa la calidad crediticia del individuo, que se puede suponer el resultado de una función lineal en sus parámetros, como por ejemplo:

$$Z_i = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \epsilon_i$$

Las estimaciones de los parámetros β_j se obtienen por máxima verosimilitud como se explicó en los párrafos anteriores, y las variables X_j contienen la información de los deudores. Habiendo obtenido las estimaciones b_j , el modelo empírico con el que trabajará el analista de riesgo es,

$$Z_i = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \beta_3 X_{3j}$$

cuando se trata de un modelo que emplea cinco variables ($j = 5$). La variable z_i es el score estimado del deudor, una medida de su calidad crediticia obtenida a partir de los parámetros estimados y de su propia información. Este score, aplicado a las funciones de distribución de probabilidades acumuladas normal o logística, permite conocer la probabilidad de incumplimiento y en consecuencia el riesgo del deudor.

4.1.3. Relación entre el score y el riesgo

Habiendo definido al score, como

$$P(Y_i = 1) = F(Z_i)$$

y

$$P(Y_i = 0) = 1 - F(Z_i)$$

donde queda claro que cambios en Z_i implican cambios en la probabilidad de incumplimiento (PD) del individuo. La relación entre score y riesgo (la PD) no es lineal, por lo que el cambio en el riesgo derivado de un cambio en el score depende de los valores que este último tome. Para valores del score muy bajos, un aumento en el mismo produce una rápida subida en la probabilidad de cumplimiento y una rápida disminución de la PD, mientras que para valores del score altos, una mejora en el mismo hace que la probabilidad de incumplimiento aumente poco y genera una leve caída en el riesgo. Es decir, cuanto mayor es el score, menor es la caída en el riesgo derivada de un aumento en el primero.

4.1.4. Uso de modelos lineales para la selección de variables (Métodos de búsqueda secuencial).

Los métodos de búsqueda secuencial tienen en común la aproximación general de estimación de las ecuaciones de regresión con un conjunto de variables y a continuación añadir o eliminar selectivamente variables hasta que se consiga alguna medida criterio conjunta. Esta aproximación proporciona un método objetivo de selección de variables que maximizan la predicción con el número más pequeño de variables empleadas. Existen dos tipos de aproximaciones de búsqueda secuencial: (1) la estimación por etapas, y (2) la eliminación progresiva y regresiva. En cada aproximación, se valoran las variables individualmente en función de su contribución a la predicción de la variable dependiente y se añaden o eliminan según su contribución relativa.

4.1.4.1. Estimación por etapas (paso a paso o stepwise)

La estimación por etapas es quizá la aproximación más popular para seleccionar variables. Esta aproximación permite examinar la contribución de cada variable predictor al modelo de regresión. Se considera la inclusión de cada variable antes de desarrollar la ecuación. Se añade primero la variable independiente con la contribución más grande. Las variables independientes se seleccionan entonces para la inclusión basada en su contribución incremental sobre la(s) variable(s) ya existente(s) en la ecuación. Las cuestiones específicas en cada etapa son:

1. Empezar con el modelo de regresión simple en el cual sólo se utiliza la única variable predictor que es la que se muestra más altamente correlacionada con la variable criterio. La ecuación sería $Y = \beta_0Q + \beta_1X_1$.
2. Examinar los coeficientes de correlación parcial para encontrar una variable predictor adicional que explique además de una parte significativa, la mayor parte del error que queda de la primera ecuación de regresión.

3. Recalcular la ecuación de regresión utilizando las dos variables predictor, y examinar el valor parcial F de la variable original del modelo para ver si todavía realiza una contribución significativa, dada la presencia de la nueva variable predictor. Si no lo hace, eliminamos la variable. Esta capacidad de eliminar variables presentes en el modelo distingue el modelo por etapas de los modelos de adición progresiva/eliminación regresiva. Si las variables originales todavía representan una contribución significativa, la ecuación sería $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$.
4. Continúa este procedimiento examinando todas las variables independientes no presentes en el modelo para determinar si deberían incluirse en la ecuación. Si se incluye una nueva variable independiente, hay que examinar todos los predictores previamente incluidos en el modelo para juzgar si se deben mantener. Existe un sesgo potencial en el procedimiento por etapas que resulta de considerar sólo una variable a seleccionar cada vez. Supongamos que las variables X_3 y X_4 explicaran conjuntamente una parte significativa de la varianza (cada una considerando la presencia de la otra), pero no son significativas por sí solas. En esta situación, ninguna debería ser considerada para el modelo final.

4.1.4.2. La adición progresiva (forward) y la eliminación regresiva (backward)

La adición progresiva y la eliminación regresiva son fundamentalmente procesos de ensayo y error para buscar los mejores estimadores de la regresión. El modelo de adición progresiva es similar al procedimiento por etapas arriba explicada, mientras que el procedimiento de eliminación regresiva implica calcular una ecuación de regresión con todas las variables independientes, para a continuación ir eliminando las variables independientes que no contribuyan significativamente. La distinción principal de la aproximación por etapas respecto de los procedimientos de adición progresiva y eliminación regresiva es su capacidad de añadir o eliminar las variables

en cada etapa. Una vez que se añade o elimina una variable en las esquemas de adición progresiva o eliminación regresiva, no existe posibilidad de revertir la acción posteriormente.

4.1.4.3. Advertencias sobre los métodos de búsqueda secuencial

El investigador debe ser consciente de dos advertencias cuando se usa cualquier procedimiento de búsqueda secuencial. En primer lugar, la multicolinealidad entre variables independientes puede tener un impacto sustancial sobre la especificación final del modelo. Examinemos esta situación con dos variables altamente correlacionadas que tienen similares correlaciones con la variable independiente. El criterio de inclusión o eliminación en estas aproximaciones es maximizar el incremento de potencia predictiva de la variable adicional. Si una de estas variables entra en el modelo de regresión, es muy probable que la otra variable también entre, dado que estas variables están altamente correlacionadas y existe poca varianza singular para cada variable por separado.

Por esta razón, se deben evaluar los efectos de la multicolinealidad en la interpretación del modelo y examinar las correlaciones directas de todas las variables independientes potenciales. Esto ayudará a evitar concluir que las variables independientes que no entren en el modelo no sean trascendentes cuando en realidad están altamente relacionadas con la variable dependiente, pero también correlacionadas con las variables ya existentes en el modelo. Aunque las aproximaciones de búsqueda secuencial maximizarán la capacidad predictiva del modelo de regresión, el investigador debe ser cuidadoso en la interpretación del modelo. Una segunda advertencia pertenece principalmente al procedimiento por etapas. En esta aproximación, los test de significación múltiple se realizan en el proceso de estimación del modelo. Para asegurar que la tasa de error conjunto a lo largo de todos los test de significación es razonable, el investigador debería emplear umbrales muy conservadores (por ejemplo, 0,01) al añadir o sacar las variables.

5

CreditRisk+

CreditRisk+ es un modelo de impago en el que se parte de la idea de que los eventos de incumplimiento de los diferentes deudores son eventos *Bernoulli independientes*. Sin embargo, es importante tener en cuenta que los eventos de incumplimiento de deudores ocurren de manera causística en el tiempo y es imposible pronosticar el momento preciso en el que ocurrirán dichos eventos y el número de eventos que sucederán en un cierto intervalo de tiempo.

El desarrollo de una teoría estadística que explique estos procesos, dentro del contexto de riesgo crediticio, comienza considerando el caso simple de una cartera de créditos que incluye a N deudores, a cada uno de los cuales se les puede asociar una *probabilidad fija de incumplimiento*; es decir, se conoce:

$$p_i = \text{Probabilidad de incumplimiento del deudor } i$$

La distribución de pérdida que puede resultar del incumplimiento de los deudores de todo el portafolio la obtiene **CreditRisk** de manera indirecta, a través de las funciones de probabilidad, y procede en dos pasos. Primero, obtenemos la función generadora de probabilidad (FGP) del número de incumplimientos y después, haciendo un supuesto sobre el nivel de pérdidas asociadas al incumplimiento de cada deudor, obtiene la FGP las pérdidas que puede ceder la cartera. A continuación se obtiene la primera para el caso simple bajo consideración.

5.1. La distribución del número de incumplimientos con probabilidades de impago fijas.

La FGP de números de incumplimientos se define como:

$$F(s) = \sum_{n=0}^{\infty} \text{Prob}(n \text{ incumplimientos}) s^n$$

Si la cartera tuviera un sólo deudor, lo único que puede suceder es que éste cumpla o incumpla. Por lo tanto, la FGP de incumplimientos para un sólo deudor es simplemente:

$$F_i(s) = (1 - p_i)s^0 + p_i(s - 1) = 1 + p_i(s - 1)$$

Como se considera que los eventos de incumplimiento son independientes, la FGP de toda la cartera es el producto de las FGP individuales, es decir:

$$F(s) = \prod_{i=1}^N F_i(s) = \prod_{i=1}^N [1 + p_i(s - 1)]$$

Aplicando logaritmo a ambos lados de la expresión, se obtiene:

$$\text{Ln}(F(s)) = \sum_{i=1}^N \text{Ln}[1 + p_i(s - 1)]$$

Generalmente, las probabilidades de incumplimiento de obligaciones crediticias a nivel individual son pequeñas y, por lo tanto, las potencias de estas son más pequeñas y pueden ser ignoradas.

Esto implica que para los valores pequeños de p_i :

$$\text{Ln}[1 + p_i(s - 1)] \approx p_i(s - 1)$$

Lo anterior es el gran supuesto de **CreditRisk+**, que intuitivamente dice que mientras la probabilidad de incumplimiento se p_i sea “pequeña”, se puede ignorar el hecho de que un deudor no

5.1 La distribución del número de incumplimientos con probabilidades de impago fijas. 3

puede incumplir más que una sola vez. Haciendo esta sustitución, se obtiene:

$$\ln(F(s)) = \sum_{i=1}^N p_i(s-1)$$

Lo anterior conduce directamente a:

$$F(s) = e^{\sum_{i=1}^N p_i(s-1)} = e^{\mu(s-1)}; \text{ donde } \mu = \sum_{i=1}^N p_i$$

La FGP corresponde a la FGP de la distribución de *Poisson* con parámetro μ , a continuación se comprueba esto, usando el desarrollo de Taylor, es decir:

$$F(s) = \sum_{n=0}^{\infty} \mu^n e^{-\mu} s^n$$

Por lo tanto, bajo el supuesto que las probabilidades de incumplimientos son bajas, obtenemos que:

$$Prob(\text{número incumplimientos} = n) = \frac{1}{n!} \mu^n e^{-\mu}$$

En la expresión anterior, ' μ ' es la *tasa promedio de incumplimiento* y es el único parámetro de la distribución. Como " N ", la distribución es independiente del número de deudores y de las probabilidades individuales de impago, siempre y cuando éstas sean "uniformemente pequeñas". Sin embargo queda claro que los deudores tengan la misma probabilidad de incumplir. Es más, éstas pueden ser diferentes y el cálculo de la tasa esperado de incumplimiento, sólo depende de la existencia de esta información.

Para concluir con este tema, es importante hacer hincapié en que el supuesto clave que permite llegar a la distribución Poisson es cuando un número no-negativo " y " es muy pequeño ($y \approx 0$), se puede utilizar la aproximación " $\ln(1+y) \approx y$ ". Esta forma exponencial de la FGP de la distribución de Poisson, es la clave de la facilidad de cálculo que tiene **CreditRisk+**.

Además, es necesario señalar que en todo el análisis anterior hay implícito un período de tiempo que está relacionado con el parámetro

de la distribución. Así, la distribución obtenida debe interpretarse de la misma manera, es decir:

$$\frac{1}{n!} \mu^n e^{-\mu} = \text{Probabilidad de que ocurran } n \text{ incumplimientos}$$

5.2. La agrupación por bandas de exposición las pérdidas

Dado que pueden existir varios deudores que implican niveles de pérdidas semejantes, **CreditRisk+** empieza por agrupar a los deudores por bandas de exposición iguales. Aunque esta agrupación introduce errores de redondeo en la estimación de la distribución, facilita mucho el desarrollo del modelo y reduce significativamente el número de datos requeridos para realizar los cálculos. Además, si el número de niveles de exposición es grande y el ancho de las bandas es pequeño, en relación con el tamaño de exposición promedio de la cartera, se puede demostrar que el error introducido es despreciable. De hecho, en la práctica es muy difícil determinar *a priori* el nivel de exposición y, por lo tanto, éste no debe representar un elemento crítico en la estimación del riesgo total de la cartera.

Así, supóngase que la pérdida esperada por deudor que cae en impago es una proporción fija “ λ_i ” del monto total “ D_i ” que debe el deudor “ i ”. A su vez, para mantener la elegancia del modelo, supóngase que el nivel de exposición que representa el deudor para el acreedor “ L_i ” es un múltiplo entero de una unidad fija de pérdida “ L ”. Estos múltiplos enteros de L reciben el nombre de “*niveles de exposición estandar*”. Así la pérdida que puede representar el incumplimiento del deudor “ i ” para el acreedor, se mide en términos de múltiplos de “ v_i ”, de la unidad fija de pérdida L , y es simplemente la cantidad siguiente:

$$v_i = \text{Redondeo} \left(\frac{\lambda D_i}{L} \right)$$

De esto se obtiene el nivel estándar de exposición que representa

cada deudor para el acreedor, mediante la identidad siguiente:

$$L_i = L_{v_i}$$

5.3. La determinación de la *distribución de pérdidas de la cartera*.

El último paso es encontrar la *la distribución de pérdidas de la cartera*, para lo cual se procede de la misma manera, es decir, encontrar la FGP de dicha distribución. Nótese que la distribución de pérdidas es necesariamente diferente a la del número de impagos, ya que puede resultar un cierto nivel de pérdidas con diferentes combinaciones de incumplimientos de deudores. Por ejemplo una pérdida de 100,000 dólares puede ser la consecuencia del incumplimiento que debe esta cantidad o del incumplimiento de 10 de deudores que deben 10,000 dólares cada uno. Además, el conocimiento de la forma en que están distribuidos los diferentes niveles de exposición entre los diferentes deudores que componen la cartera es indispensable para poder obtener la distribución de pérdidas. Contrario a lo que sucede con las diferencias entre probabilidades de incumplimiento individuales, que son pequeñas aun en términos relativos y por lo tanto no afectan la distribución del número de incumplimiento, las diferencias en los niveles de exposición a los deudores representados en la cartera si son significativos y, por lo tanto, la distribución de pérdidas resultante no tiene por que ser *Poisson*. Sin embargo, aunque no se va a poder obtener una fórmula explícita de la distribución de pérdidas, se verá que es posible obtener una expresión cerrada sencilla de la FGP, que se presta fácilmente al cálculo de las probabilidades correspondientes.

Así, recordemos que la FGP de las pérdidas es:

$$G(s) = \sum_n Prob(\text{Pérdida acumulada} = nxL) S^n$$

Sea “ G_j ” la FGP de las pérdidas de la banda “ j ”, la probabilidad

de que se pierda $n \times v_j$ unidades en la banda “ j ” es igual a la probabilidad de que “ n ” deudores de esta banda incumplan, y ya se sabe que el incumplimiento de estos deudores sigue una distribución de *Poisson*, es decir:

$$\text{Prob. de } n \text{ incumplimientos en la banda } j = \frac{1}{n!} \mu_j^n e^{-\mu_j}$$

Por lo tanto:

$$G_j(s) = \sum_{n=0}^{\infty} \frac{1}{n!} \mu_j^n e^{-\mu_j} s^{nv_j} = e^{-\mu_j} \sum_{n=0}^{\infty} \frac{1}{n!} (\mu_j s^{v_j})^n = e^{-\mu_j} e^{\mu_j s^{v_j}} = e^{\mu_j(s^{v_j}-1)}$$

Suponiendo independencia de eventos de incumplimientos, las pérdidas asociadas a cada banda también serán independientes. En este caso, la FGP de las pérdidas de la cartera de crédito es la suma de las pérdidas ocurridas en cada banda, que por propiedades de las funciones generadoras de probabilidad es el producto de las FGP de las bandas:

$$G(s) = \prod_j G_j(s) = \prod_{j=1}^m e^{\mu_j(s^{v_j}-1)} = e^{\sum_{j=1}^m \mu_j(s^{v_j}-1)}$$

es decir:

$$G(s) = e^{\sum_{j=1}^m \mu_j(s^{v_j}-1)}$$

5.4. Obteniendo la distribución de las pérdidas

Aunque a diferencia de la distribución del número de incumplimientos, en el caso de las pérdidas no se sabe la forma de la distribución, la probabilidad de que se pierdan n unidades de L en la cartera total de créditos se puede obtener mediante la expansión de Taylor de la expresión anterior; es decir: la probabilidad de que se pierdan n unidades L es el coeficiente de s^n en la expansión de Taylor. El desarrollo se facilita definiendo el polinomio siguiente:

$$P(s) = \frac{1}{\mu} \sum_{j=1}^m \mu_j s^{v_j} = \frac{\sum_{j=1}^m \frac{\epsilon_j}{v_j} s^{v_j}}{\sum_{j=1}^m \frac{\epsilon_j}{v_j}}$$

donde ϵ_j es la pérdida esperada de la banda j y v_j es la exposición común que comparten los deudores de la banda j .

Con esto, la FGP de pérdidas se puede expresar de la forma siguiente:

$$G(s) = e^{\sum_{j=1}^m \mu_j (s^{v_j} - 1)} = e^{\mu[P(s) - 1]} = F[P(s)]$$

La función generadora de probabilidades de las pérdidas comprende dos fuentes de incertidumbre, a saber: el comportamiento *Poisson* del número de incumplimientos y la aleatoriedad de las pérdidas dado incumplimiento, que está asociado a los distintos niveles de exposición del acreedor a los diferentes deudores. Notese que la única información requerida para obtener la distribución de pérdidas es ϵ_j y v_j , estas parejas representan un número de datos significativamente menor, que si tuviera que manejar información semejante por cada deudor, aun para carteras con un número muy grandes de créditos. Para encontrar la distribución de pérdidas, se parte de la expresión:

$$G(s) = \sum_n \text{Prob}(\text{pérdidas acumuladas} = n \times L) S^n$$

.

Tomando la expansión de Taylor para $G(s)$, la probabilidad de pérdida de nL unidades es

$$P_n(L) = \frac{1}{n!} \left. \frac{d^n G(s)}{ds^n} \right|_{s=0}$$

. En el apéndice A 4.1 del manual de CreditRisk+ se deduce una fórmula recursiva sencilla para obtener estas probabilidades, las cuales es:

$$P_n(L) = \sum_{j|v_j \leq n} \frac{\epsilon_j}{n} P_{n-v_j}(L) = \frac{1}{n} \sum_{j|v_j \leq n} \epsilon_j P_{n-v_j}(L)$$

donde $P_0(L) = e^{-\mu}$ donde $\mu = \sum_{j=1}^m \mu_j$

5.5. Obtención del VaR

Para obtener el VaR de una cartera de créditos bajo el esquema de CreditRisk+, simplemente se obtiene la distribución de probabilidad acumulada de las pérdidas y, una vez decidido el nivel de confianza α con el que se quiere trabajar se lee de la tabla el número de unidades estándar de pérdidas que corresponden a una probabilidad acumulada de pérdida $1 - \alpha$. El paso final es convertir las unidades estándar de pérdida a unidades monetarias, multiplicando las unidades de pérdidas correspondientes al nivel de confianza escogido, por la unidad de pérdida L .

6

CreditMetrics

Mientras que en el caso de **CreditRisk+** se usa un enfoque actuarial mediante el cual se llega a obtener una forma explícita de la distribución de pérdidas a través de la función generadora de probabilidades, en el caso de **CreditMetrics**, aunque en teoría existe una representación explícita de la distribución de pérdidas, su obtención supera la capacidad de cálculo disponible en un computador convencional, y por lo tanto, está se estima a través de un proceso de simulación de *MonteCarlo*, otra característica importante es que **CreditMetric** hace referencia, y utiliza extensamente, información que solo esta disponible en los mercados en desarrollo, aunque esto no impide que se aplique en los mercados emergentes si se hacen las edecuaciones correspondientes.

Para la explicación de esta metodología usamos como guía el documento técnico de **CreditMetrics**. Iniciaremos con el caso más sencillo en el cual nuestra cartera de créditos esta integrada por un único credito para luego generalizar el procedimiento a una cartera que puede incluir cualquier número N de créditos. Pero primero establezcamos los elementos básicos que usaremos en esta metodología.

6.1. Elementos que usa Credimetrics

- *Conceptos relevantes al riesgo de crédito*: probabilidades de incumplimiento, patrones de migración entre los distintas clasificaciones que posean los clientes, umbrales que separen las calificaciones,

el valor de los créditos, la tasa de recuperación en caso de incumplimiento de las calificaciones crediticia.

- Información del mercado en general: Tales como datos sobre las calificaciones e indicadores sobre su calidad.
- Información generada internamente por las instituciones que se pueda mapear a los estándares del mercado.

Historicamente Creditmetrics fue desarrollado para abordar el problema de medir el riesgo de crédito de que los emisores que colocan papel que se negocia en el mercado de los Estados Unidos incumplan sus compromisos. En general, los elementos que son necesarios en esta metodología están disponibles en mercados desarrollados, pero las instituciones financieras pueden crear procedimientos y metodologías internas para lograr suplir la falta de información. Los elementos clave del sistema son:

- Las calificaciones que otorgan las instituciones financieras a los distintos elementos que conforman su portafolio.
- Las diferencias de riesgo que expresan las distintas calificaciones.
- Información sobre la tasa de recuperación sobre créditos que caen en la cartera vencida.

Una característica importante de las calificaciones emitidas es que ellas aportan de manera implícita piezas clave para la evaluación del riesgo de crédito, como por ejemplo:

- Probabilidades de incumplimiento históricas asociadas a cada calificación, así como las probabilidades de transición entre las categorías.
- Las distintas tasas de interés asociadas a cada categoría.

Otra parte fundamental del análisis es la cantidad que se espera recuperar en caso de que el crédito caiga en la cartera de vencidos. Aunque si el sistema de calificación es robusto raramente se espera que un crédito que haya pasado a la cartera de vencido este en una calificación alta. Pero al final, la calificación no solo depende del

sistema interno de la institución, sino también de las garantías que ofrezca el crédito.

En general, todos los requerimientos están entrelazados, por ejemplo las calificaciones que se otorgan dependen de cosas como: historial de incumplimiento, garantías, utilidades de los créditos en los últimos años, el país de origen, entre otros.

A manera de ilustración se presenta una matriz de transición publicada por Standard & Poor's en Abril de 1966. En ella se aprecian las probabilidades de transición de que un bono que se encuentra en una calificación dada migre a otra calificación, por ejemplo en la tabla se aprecia que la probabilidad de que un bono que está en la calificación BB migre a una calificación C es de un 1 %.

Calificación inicial	AAA	AA	A	BBB	BB	B	C
AAA	90.81	8.33	0.68	0.06	0.12	0.00	0.00
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00
B	0.00	0.11	0.24	0.43	6.48	83.46	4.07
C	0.22	0.00	0.22	1.30	2.38	11.24	64.86

Las tasas de recuperación de los créditos como ya se menciono son fundamentales, y no son únicamente proporcionadas por las instituciones financieras, estas pueden recurrir a terceros por las calificaciones de los créditos. A continuación se presenta una tabla que muestra las tasas de recuperación de créditos elaboradas por Moody's y otros académicos.

Prelación	Media (%)	Desv. Std (%)
Senior Garantizada	53.80	26.86
Senior sin Garantía	51.13	25.45
Senior Subordinado	38.52	23.81
Subordinada	32.74	20.18

Prelación	Media (%)	Desv. Std (%)
Junior Subordinada	17.09	10.90

Una acotación interesante, es que originalmente credimetrics esta direccionada a medir el riesgo de crédito de una cartera de bonos. Sin embargo puede ser aplicado a distintos tipos de posiciones que involucren riesgo de crédito, como por ejemplo: posiciones cortas en divisas, futuros, opciones o como es el caso de intereses, créditos tradicionales. Ahora, daremos el ejemplo de credimetrics usando un único crédito.

6.2. El caso de un solo crédito

Una vez descrito los elementos necesarios para esta metodología vamos a empezar por realizar los pasos que son necesarios para la medición del riesgo de crédito de una cartera de un solo activo. Utilizaremos un ejemplo parecido al que procede del documento de Credimetrics con la diferencia que ellos usan un bono y nosotros un crédito personal. Los tres pasos a realizar son:

- Dependiendo de la calificación del bono, se obtiene la probabilidad de que migre hacia un estado de migración diferente.
- Se calcula el saldo o la exposición de los créditos.
- Se aplican las probabilidades de migración a la exposición que posee el crédito para obtener la distribución de probabilidad de pérdida o perfil de riesgo del instrumento y las estadísticas que corresponden, es decir, la pérdida esperada, su varianza o desviación estandar, su VaR con un nivel de confianza deseado, entre otros.

6.2.1. Ejemplo 1

Supongamos que tenemos un crédito con calificación BB con exposición de 250, la institución financiera posee la siguiente información sobre las probabilidades calificación de los créditos que estan en una calificación BB:

Calificación	Probabilidad de transición (%)
AAA	4
AA	5
A	9
BBB	15
BB	44
B	18
CCC	5

Tambien la institución cuenta con las perdidad esperadas por calificación (las cuales representan la inversa aritmética de las tasas de recuperación en caso de que el crédito caiga en default o mora)

Calificación	Pérdida esperada (%)
AAA	0.1
AA	3
A	5
BBB	9
BB	18
B	25
CCC	40

Ya con esta información podemos establecer una tabla con la pérdida esperada del crédito en las distintas calificaciones, la cual es:

Calificación	Pérdida esperada del Crédito	Proabilidad
AAA	0.25	4

Calificación	Pérdida esperada del Crédito	Proababilidad
AA	7.5	5
A	12.5	9
BBB	22.5	15
BB	45	44
B	62.5	18
CCC	100	5

De esta forma para obtener la pérdida esperada se procede a multiplicar la pérdida esperada de una calificación por la probabilidad de que el crédito caiga en la respectiva calificación y luego se realiza la suma global, de la siguiente forma:

$$\begin{aligned}
 PE = & (0,25 * 0,04) + (7,5 * 0,05) \\
 & + (12,5 * 0,09) + (22,5 * 0,15) + (45 * 0,44) \\
 & + (62,5 * 0,18) + (100 * 0,05) = 40,93
 \end{aligned}$$

Así la pérdida esperada de una cartera de un solo crédito de calificación BB de una exposición de 250 es de 40.935. Ahora para calcular el VaR a un nivel de confianza en específico por ejemplo al 5 % simplemente sumamos las probabilidades de manera descendentes de las calificaciones hasta llegar 95 en el caso de que se pase se toma la calificación en la cual ocurre el exceso, en nuestro ejemplo el VaR ocurre en la calificación B, así el $VaR_{5\%} = 62,5$.

En general, se suele pensar que una aproximación normal a la distribución de pérdidas puede ser idónea, pero esto es incorrecto pues suele ocurrir que las probabilidades de transición se concentren en los extremos de las calificaciones, esto se verá en el siguiente ejemplo.

6.2.2. Ejemplo 2

Consideremos un bono con una calificación AA con un valor de 300 donde las probabilidades de transición a las otras categorías esta representada en la siguiente tabla

Calificación	Probabilidad de transición (%)
AAA	15
AA	48
A	15
BBB	12
BB	5
B	4
CCC	1

Además la tasas de pérdida por calificación son las mismas que el ejemplo anterior

Ya con esta información podemos establecer una tabla con la pérdida esperada del crédito en las distintas calificaciones, la cual es:

Calificación	Pérdida esperada del Crédito	Proababilidad
AAA	0.3	15
AA	9	48
A	15	15
BBB	27	12
BB	54	5
B	75	4
CCC	120	1

Ahora de igual forma para obtener la paérdida esperada procedemos con el siguiente cálculo.

$$PE = (0,3 * 0,15) + (9 * 0,48) + (15 * 0,15) + (27 * 0,12) + (54 * 0,05) + (75 * 0,04) + (120 * 0,01) = 16,75$$

Para calcular el valor en riesgo al 95 % es suficiente tomar la pérdida esperada de la calificación BB, así el $Var_{95\%} = 54$.

Con este ejemplo podemos notar que la distribución normal no se ajusta de manera adecuada a la distribución de pérdidas, pues esta está concentrada en un extremo de las pérdidas por calificación.

6.3. Cartera de dos de créditos con probabilidades de transición independientes.

Como en realidad el caso de un solo instrumento no es de interés, pues una cartera de créditos de una institución financiera puede estar conformada por millones de créditos daremos un ejemplo ilustrativo con una cartera de dos créditos, con esto se pretende mostrar como aumenta el nivel de dificultad, mas adelante veremos que en general, no calcularemos la distribución de pérdidas explícitamente, sino realizaremos una aproximación por el método de simulación de Montecarlo.

6.3.1. Ejemplo 3

Para facilitar los cálculos supondremos que nuestra cartera esta conformada por los créditos de los primeros dos ejemplos, la idea es crear una tabla con todas las combinaciones posibles de las sumas de pérdidas posibles entre los créditos, la cual es:

Crédito AA	AAA	AA	A	BBB	BB	B	CCC
Crédito BB							

6.4 Cartera de dos de créditos con probabilidades de transición independientes. 9

Crédito	AA	AAA	AA	A	BBB	BB	B	CCC
AAA	0.55	9.25	15.25	27.25	54.25	75.25	120.25	
AA	7.8	16.5	22.5	34.5	61.5	82.5	127.5	
A	12.8	21.5	27.5	39.5	66.5	87.5	132.5	
BBB	22.8	31.5	37.5	49.5	76.5	97.5	142.5	
BB	45.3	54	60	72	99	120	165	
B	62.8	71.5	77.5	99.5	106.5	137.5	182.5	
CCC	100.3	109	115	127	154	175	220	

Esta tabla se puede entender como sigue, la pérdida esperada si el crédito BB migra a la calificación BBB y el crédito AA migra a la calificación CCC es 142.5, ahora para hallar la probabilidad debemos multiplicar las probabilidades correspondientes de que estas migraciones ocurran, es decir, $0,15 * 0,01 = 0,0015$. Ahora para hallar la pérdida esperada de esta cartera debemos encontrar todas las probabilidades de transición, lo cual se muestra en la siguiente tabla:

Crédito	AA	AAA	AA	A	BBB	BB	B	CCC
Crédito BB								
AAA	0.006	0.0192	0.0060	0.0048	0.0020	0.0016	0.0004	
AA	0.0075	0.0240	0.0075	0.0060	0.0025	0.0020	0.0005	
A	0.0135	0.0432	0.0135	0.0108	0.0045	0.0036	0.0009	
BBB	0.0225	0.0720	0.0225	0.0180	0.0075	0.0060	0.0015	
BB	0.0660	0.2112	0.0660	0.0528	0.0220	0.0176	0.0044	
B	0.0270	0.0864	0.0270	0.0216	0.0090	0.0072	0.0018	
CCC	0.0075	0.0240	0.0075	0.0060	0.0025	0.0020	0.0005	

Ahora realizando todas las correspondientes multiplicaciones de las pérdidas por sus probabilidades y sumandolas, obtenemos que la pérdida esperada de una cartera con estos dos créditos en particular es de 57,86. Ahora para obtener el VaR, ordenamos de menor a mayor las pérdidas y sumamos las respectivas probabilidades hasta llegar al nivel de confianza requerido, por ejemplo el VaR al 95 % de esta cartera es de 115.

6.4. Aproximación de la distribución de pérdidas de una cartera con n créditos.

En el ejemplo anterior notamos como una cartera con 2 créditos con 7 posibles calificaciones existen 7^2 combinaciones, si tuvieramos 50 créditos en nuestra cartera hubiera 7^{50} combinaciones posibles, es decir, el coste computacional aumenta exponencialmente, por lo tanto, una cartera realista con miles e incluso millones de créditos se vuelve inviable si queremos realizar el cálculo de la misma forma que el ejemplo anterior, por lo tanto nos vemos obligados a realizar una aproximación a la distribución de pérdidas.

6.4.1. Simulación de MonteCarlo

Ahora veremos como aproximar la distribución de pérdida con el método más común, conocido como método de aproximación por simulación de MonteCarlo. Antes de describir el método debemos acotar que en general los eventos de migración crediticia pueden ser supuestos independientes o no entre las migraciones de otros créditos. Sin embargo en mercados poco desarrollados o donde la información es limitada se tiende a suponer la independencia entre los créditos de una cartera.

6.4.1.1. ¿Que es un Método de Montecarlo?

El término Monte Carlo se aplica a un conjunto de métodos matemáticos que se empezaron a usar en los 1940s para el desarrollo de armas nucleares en Los Alamos, favorecidos por la aparición de los ordenadores digitales modernos. Consisten en resolver un problema mediante la invención de juegos de azar cuyo comportamiento simula algún fenómeno real gobernado por una distribución de probabilidad (e.g. un proceso físico) o sirve para realizar un cálculo (e.g. evaluar una integral).

Más técnicamente, un método de MonteCarlo es un proceso esto-

cástico numérico, es decir, una secuencia de estados cuya evolución viene determinada por sucesos aleatorios. Recordemos que un suceso aleatorio es un conjunto de resultados que se producen con cierta probabilidad. Veamos un ejemplo ahora:

6.4.1.2. Ejemplo: Gotas de lluvia para estimar π

Consideremos un círculo de radio unidad circunscrito por un cuadrado. Suponiendo una lluvia uniforme sobre el cuadrado, podemos hallar el valor de π a partir de la probabilidad de que las gotas caigan dentro del círculo (próxima figura)

Sea P dicha probabilidad, por calculo:

$$P = \frac{\text{área del círculo}}{\text{área del cuadrado}} = \frac{\int_{-1}^1 dx \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy}{\int_{-1}^1 dx \int_{-1}^1 dy} = \frac{2 \int_{-1}^1 \sqrt{1-x^2} dx}{2,2} = \frac{\pi}{4}$$

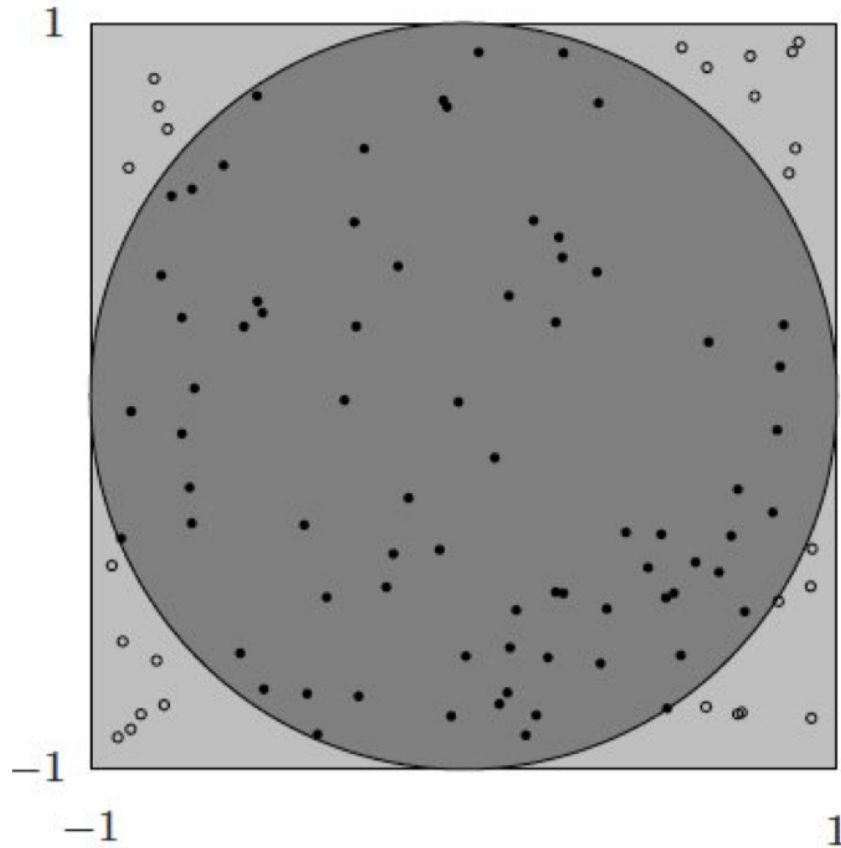
Es decir $\pi = 4P$. Ahora notemos lo siguiente:

- Gracias a los métodos de integración de cálculo podemos relacionar rápidamente la probabilidad con π
- Con un ordenador podemos generar números o pares de números pseudo aleatorios (x, y) entre 0 y 1.
- Entre mayor número de simulaciones realicemos mayor sera nuestra aproximación a π

Para finalizar el ejemplo debemos encontrar la frecuencia de puntos que caen dentro de la circunferencia para aproximar P de una manera correcta, así despues de 1000 simulaciones $\pi \approx 3,158733$.

6.4.1.3. Números pseudoaleatorios

Hemos visto que el método de MonteCarlo es un proceso estocástico numérico que nos permite resolver problemas de aproximación. Para ello se requiere mostrar variables aleatorias según una ley de distribución de probabilidad, por lo tanto es vital que nuestro

**Figura 6.1**

método de muestreo sea efectivo, por lo tanto presentamos el siguiente test de calidad de números pseudoaleatorios.

- Equidistribución: los números pseudoaleatorios deben repartirse por igual, como correspondería a una verdadera distribución uniforme
- Largo periodo: todos los métodos de generación de números pseudoaleatorios tienen un periodo en el cual la secuencia de números se vuelve a repetir, por lo tanto, el periodo debe ser largo para no agotar la secuencia en un cálculo concreto.

- Repetibilidad: a veces se necesita repetir un cálculo con exactamente los mismos números pseudoaleatorios. Así que conviene que el generador permita almacenar su estado
- Largas subsecuencias disjuntas: Si la simulación es muy extensa resulta conveniente subdividirla en otras más pequeñas, para lo que es importante que sean estadísticamente independientes y así se puedan recombinar sin introducir correlaciones.

Portabilidad: La rutina debe generar exactamente la misma secuencia de números no solamente por distintos lenguajes de programación sino también en distintas máquinas.

Eficiencia: La generación de cada número debe consumir muy poco tiempo.

Ahora para comprobar la eficiencia del método de generación de números pseudo aleatorios existen test estadísticos para comprobar su bondad, los mas usados son:

- Test de frecuencia: Sirve para comprobar la equidistribución de N valores generados. Se divide el intervalo $[0, 1]$ en k sub intervalos. Uno esperaría encontrar N/k valores en cada sub intervalo. Sea N_j la cantidad de valores encontrados en el intervalo j -ésimo. Entonces la siguiente χ^2 permite hallar la probabilidad de que la distribución generada sea compatible con una verdadera distribución aleatoria uniforme,

$$\chi^2 = \frac{k}{N} \sum_{j=1}^k \left(N_j - \frac{N}{k} \right)^2$$

que se comportara asintóticamente como una $\chi^2(k-1)$

- Test de la serie: Sirve para comprobar la independencia entre sucesivos numeros en una secuencia. Es una generalización del test anterior. Dividimos el intervalo $[0, 1]$ en r sub intervalos y miramos ternas de $s \geq 2$ puntos $x_n = (x_n, x_{n+1}, \dots, x_{n+s-1})$ consecutivamente generados. Cada una de las N ternas x_n generadas caerá dentro de uno de los r^s bins en los que se divide este espacio s dimensional. Si llamamos N_{j_1, j_2, \dots, j_s} con $j_i \in \{1, 2, 3, \dots, r\}$, el número de valores que caen en el bin (j_1, j_2, \dots, j_s) , la siguiente χ^2

nos permita hallar la probabilidad de que la distribución generada sea uniforme,

$$\chi^2 = \frac{r^s}{N} \sum_{\{j_1, j_2, \dots, j_s\}} \left(N_{j_1, j_2, \dots, j_s} - \frac{N}{r^s} \right)$$

Se comportara asintóticamente como una $\chi^2(r^2 - 1)$

6.4.1.4. Aplicando simulación de MonteCarlo para hallar la distribución de pérdidas de una cartera.

Ahora una vez se pretenda usar un método algorítmico de generación de números aleatorios, se usa este método para generar eventos probabilísticos en los cuales se consideren todos los posibles eventos de transición entre las distintas calificaciones. Cada evento se considera como una simulación de un evento particular, por cada simulación los créditos se encuentran en una calificación determinada por la probabilidad de transición respectiva, esta calificación trae con si una pérdida esperada por crédito, así por cada simulación se obtiene una pérdida.

Ahora si nuestra cartera de crédito posee 1000 créditos, y existen 5 calificaciones, en la sección vimos que existen 5^{1000} posibles eventos, para imaginarnos esta cantidad basta saber que este número es mayor que un billón, por lo tanto debemos realizar un gran número de simulaciones para poder captar información verídica de la pérdida esperada. Una de las bondades de estimar la pérdida esperada de esta forma es que la distribución tendera asintóticamente a una normal donde su media tendera la pérdida esperada de la cartera.

6.4.1.5. Obtención de las métricas de riesgo.

Una vez realizadas las simulaciones y obtenida la aproximación de la distribución de pérdidas usando la aproximación normal es facil obtener las métricas de riesgo, pues estas ya estan definidas para la distribución normal. Si la distribución de pérdidas es como una normal de media μ y desviación estandar σ las métricas de riesgo son:

$$\text{Pérdida esperada} = \mu$$

6.4 Aproximación de la distribución de pérdidas de una cartera con n créditos. 15

,

$$VaR_p = \mu + \sigma \Phi^{-1}(p)$$

,

$$TVaR_p = \mu + \sigma \frac{\phi[\Phi^{-1}(p)]}{1-p}$$

Donde ϕ y Φ representan la función de densidad de probabilidad y la función de probabilidad acumulada de una normal estandar.



7

Manual técnico de la aplicación Vision CreditRisk

En el presente capítulo, presentaremos el manual técnico de la aplicación Vision CreditRisk, el cual utiliza las nociones teóricas vistas anteriormente para lograr de una forma agradable y de fácil entendimiento permitir que el usuario encargado del área de crédito de su institución maneje de forma precisa y adecuada las métricas de riesgo.

7.1. Presentación de la aplicación.

La aplicación esta compuesta de tres secciones, la primera referente a la metodología CreditRisk+, la segunda correspondiente a la metodología Credimetrics y la tercera sección referente a indicadores técnicos.

7.2. Primera sección: CreditRisk+

En la primera sección que se observa en la columna izquierda, consta de las siguientes secciones: Datos, Estadísticos, Pérdida por incumplimiento, Score de crédito, parámetros y resultados, stressting.



Figura 7.1

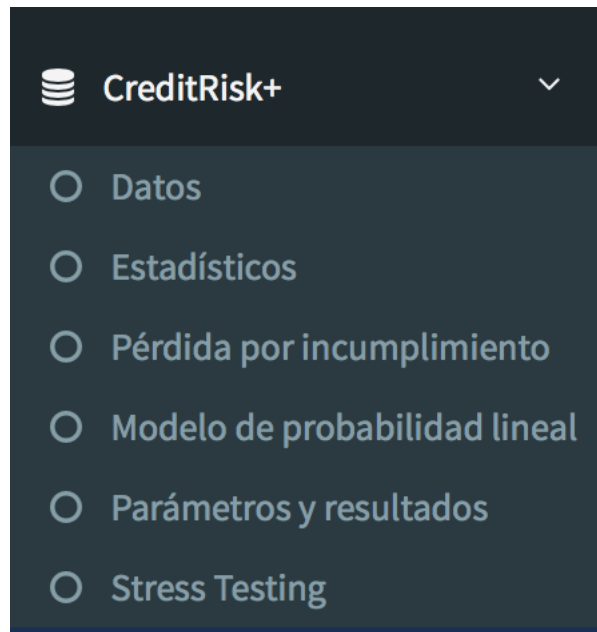


Figura 7.2

7.2.1. Datos

En esta sección se cargan los datos con los que se alimentara la construcción del score de crédito y parte de la metodología, al iniciar veremos la siguiente pantalla:



Figura 7.3

Tendremos dos opciones, una a manera de ejemplo para entender como debe ser el formato de los datos, al seleccionarla se nos desplegara la siguiente pantalla.

Ahora del lado que debemos seleccionar, se nos desplagara:

Debemos hacer click en buscar los datos a seleccionar, estos datos deberan estar conformados por la lista de los clientes, similar a los datos de ejemplo, una de la columnas debe tener el nombre de “Creditability” (la cual indica si el cliente esta en mora, valor de 1 o no valor de 0)

7.2.2. Estadísticos:

En esta seccion tendremos dos pestañas:

☒ Selecciona para iniciar Datos de Ejemplo
 ☐ Cargar Datos Propios

Selecciona variable de estudio
 Creditability

Show 10 entries Search:

Creditability	Account.Balance	Duration.of.Credit..month.	Payment.Status.of.Previous.Credit	Purpose	Credit.Amount	Value.Savi
0	1	18	4	2	1049	1
0	1	9	4	0	2799	1
0	2	12	2	9	841	2
0	1	12	4	0	2122	1
0	1	12	4	0	2171	1

Figura 7.4

☐ Selecciona para iniciar Datos de Ejemplo
 ☒ Cargar Datos Propios

Selecciona variable de estudio
 Seleccione primero los datos

Cargar el archivo con los datos

Selecciona el archivo
 Buscar Aun no seleccionas el archivo...

Encabezado de los datos
☒ Con encabezado

Separador
☐ Coma
☒ Punto y coma
☐ Tab

Comillas
☒ Ninguna
☐ Comilla doble
☐ Comilla simple

Figura 7.5

7.2.2.1. Relación de las variables independientes

En esta pestaña, veremos el gráfico estadístico de boxplot para ver el comportamiento de la variable al segmentarla en deudores o no, además en la parte inferior tendremos un resumen estadístico de la variable seleccionada.

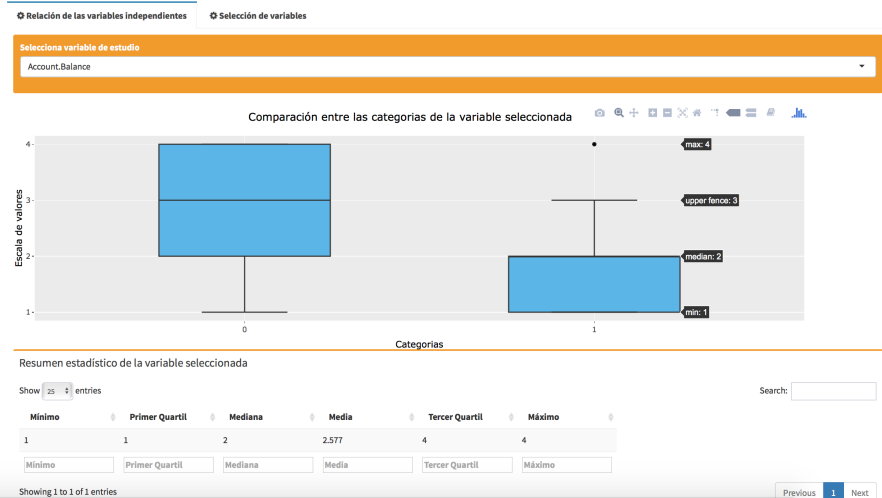


Figura 7.6

7.2.3. Selección de variables

En esta sección el usuario podrá colocar el significancia para escoger las variables cualitativas y cuantitativas que se usaran en el modelo, la pestaña se divide en 2, una para cada tipo de variable.

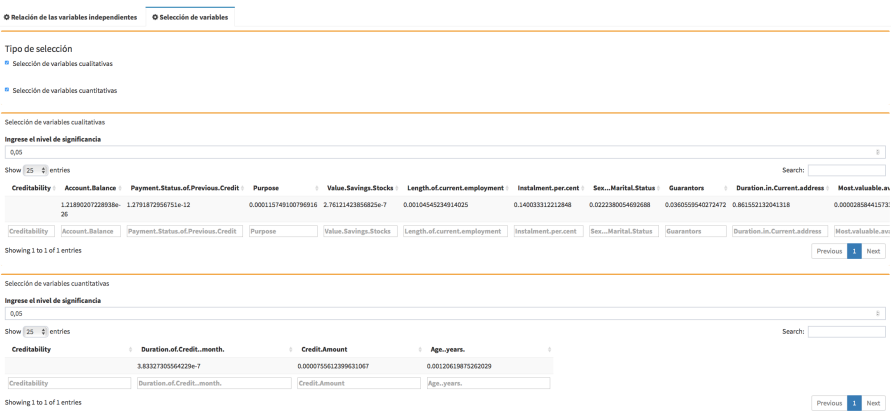


Figura 7.7

7.3. Pérdida por incumplimiento

En esta sección podrá calcular la pérdida histórica por incumplimiento. El usuario podrá ver graficamente lo que espera recuperar en caso de incumplimiento, mientras transcurre el tiempo.

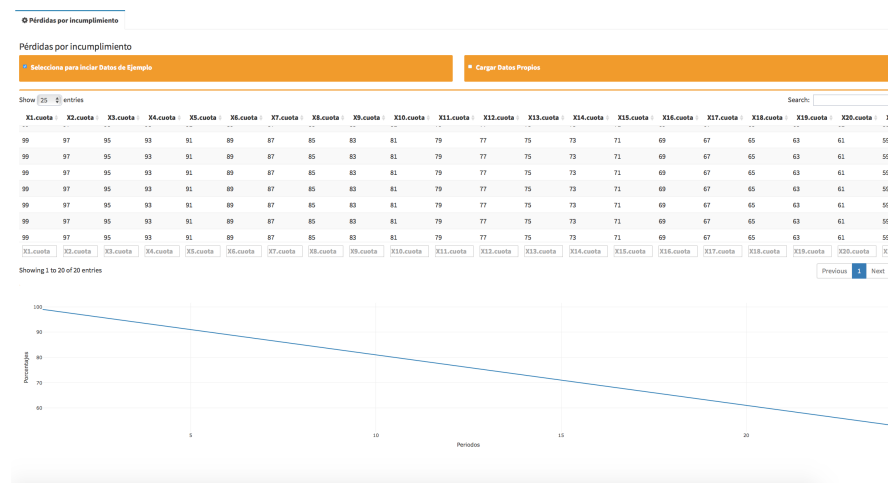


Figura 7.8

7.4. Score de Crédito.

En esta sección se calculara la probabilidad de incumplimiento, score o puntuja crediticio por cliente, y se haran proyecciones a nuevos clientes.

7.4.1. Selección y resultados del modelo

En esta sección el usuario seleccionara el modelo que mejor se ajuste a los datos, una vez seleccionado se desplegara la matriz con

los resultados del modelo y el grafico ROC para mostrar el nivel de acierto del modelo.

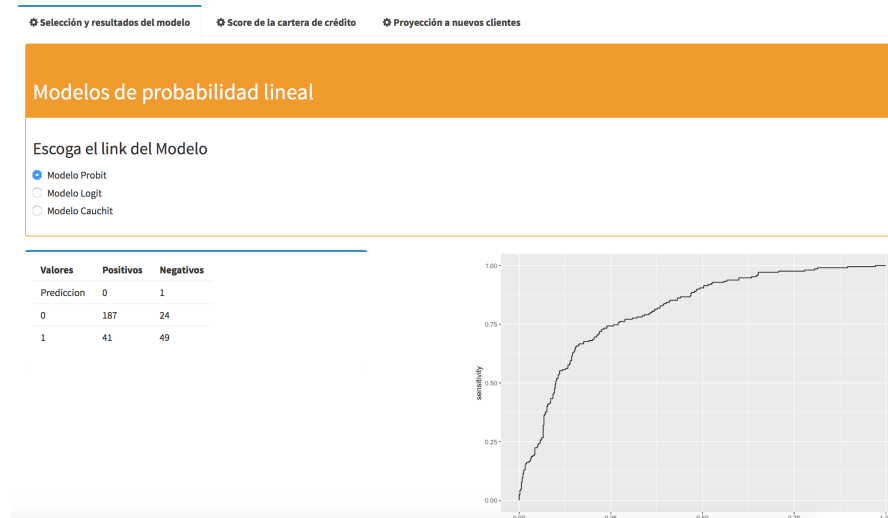


Figura 7.9

7.4.2. Score de la cartera de crédito.

Una vez seleccionado el modelo, en esta pestaña el usuario podra ver la probabilidad de incumplimiento y el score o puntaje crediticio de los clientes.

7.4.3. Proyección a nuevos clientes.

7.5. Parametros y resultados.

En esta sección se pedirán los valores de los parámetros básicos y se mostraran los resultados de las principales métricas de riesgo para la metodología CreditRisk+.

Selección y resultados del modelo

Score de la cartera de crédito

Proyección a nuevos clientes

Score y probabilidad de incumplimiento de los clientes

Show 25 entries

Search:

Posición	Score	Probabilidad de incumplimiento
12	-0.189121603344142	0.424998754996813
13	-0.519009850319188	0.301876936087892
14	0.535815149641517	0.70395684456471
15	-0.41654549163871	0.338505443106238
16	-0.411296768332866	0.340427469478762
17	-0.634721570330218	0.262805011455004
18	-0.504360443700255	0.30700405482693
19	-0.737795823464082	0.230319265834138

Posición

Score

Probabilidad de incumplimiento

Showing 1 to 25 of 1,000 entries

Previous

1

2

3

4

5

...

40

Next

Figura 7.10

Parametros iniciales

Resultados

Modelo CreditRisk+

Ingrese unidad de pérdida

1000

Ingrese porcentaje de recuperación luego del default

80

Probabilidades de incumplimiento provenientes del Score

Probabilidades de incumplimiento propias

Figura 7.11

7.5.1. Parámetros iniciales

En esta sección, deberemos establecer la unidad de pérdida, que no es mas que la unidad de medida que inteta cuantificar la pérdida de una forma mas compacta. Deberemos ingresar el porcentaje de recuperación que se espera recuperar luego que un crédito esta en mora. Debemos cargar las probabilidades de incumplimiento de la cartera de clientes, podemos cargar una data propia generada por mecanismos internos o podemos cargar la proveniente de la sección

de Score de crédito, en ambos casos la data debe tener la siguiente estructura:

Parametros iniciales Resultados

Modelo CreditRisk+

Ingreso unidad de pérdida: 1000

Ingreso porcentaje de recuperación luego del default: 80

☒ Probabilidades de incumplimiento provenientes del Score

☐ Probabilidades de incumplimiento propias

Show 25 entries Search:

PD
0.537908770372401
0.308774686109323
0.29358863894729
0.161006965882968
0.218276911952371
0.186457643125225

Figura 7.12

7.5.2. Resultados

En la sección resultados, podremos escoger el nivel de significancia deseado, para mostrar las métricas usuales de riesgo, tendremos también la opción para descargar el reporte proveniente de esta metodología

7.6. Stress Testing

En esta sección de una manera controlada porcentualmente simularemos resultados extremos de pérdida, para luego mostrar la pérdida esperada tras esto.

Parametros Iniciales Resultados

Escoga nivel de confianza para el VaR

☐ 90%
☒ 95%
☐ 99%

Pérdida esperada	Valor en riesgo	TVaR
578870.4	632000	687884.2

Reporte

Descargar

Figura 7.13

StressTesting

Nivel de estrés de la prueba

☒ 1 %
☐ 5 %
☐ 10 %

Resultado

El valor de la prueba

642000

Figura 7.14

7.7. Segunda sección: Credimetrics

En la segunda sección que se observa en la columna izquierda, consta de las siguientes secciones: créditos, matriz de transición, Pérdida por clase, simulación y resultados, stressting.



Figura 7.15

7.8. Créditos

En esta sección, deberemos cargar las exposiciones de los créditos y sus respectivas calificaciones:

Créditos

☒ Selecciona para Iniciar Datos de Ejemplo

☐ Cargar Datos Propios

Show 25 entries
Search:

creditos	calif
8000	A
4440	A
3270	A
1900	A
4720	A
2000	A
5500	A
3000	A

Showing 1 to 25 of 132 entries

Previous
1
2
3
4
5
6
Next

Figura 7.16

En esta sección cargaremos los datos para el cálculo de la matriz de transición crediticia

En esta sección debemos cargar los datos históricos de migraciones crediticias en periodos de tiempo especificos, es decir, con que calificación inicio el crédito y con cual finalizo.

Una vez cargados los datos en la parte inferior se mostrara la matriz de transición calculada:

En esta sección se deberá seleccionar la matriz de transición, se puede escoger la ya calculada o en caso de que el banco no posea los datos históricos necesarios, podrá cargar su matriz de transición propia, pero debe tener el siguiente formato.

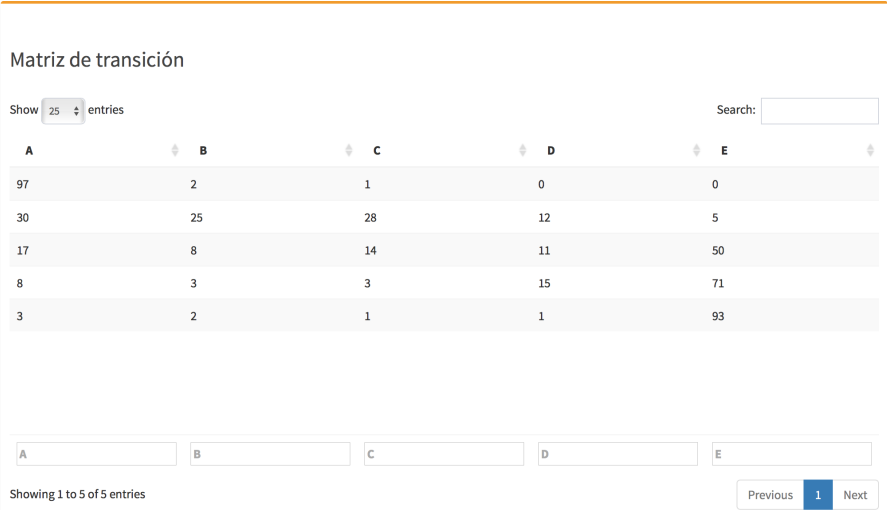


Figura 7.17



Figura 7.18

7.10. Pérdida por clase.

En esta sección se calcula o solicita las pérdidas esperadas por calificación crediticia.

7.10.1. Cálculo de pérdida esperada

En esta pestaña el usuario debe cargar los datos históricos de las pérdida ocasionadas por clientes dado su calificación.

✱ Cálculo de las pérdidas esperadas ✱ Pérdida esperada

Cálculo de las Pérdidas esperada por clases

☒ Selecciona para Iniciar Datos de Ejemplo ☐ Cargar Datos Propios

Show 25 entries Search:

clases	perdida
A	0
A	0
A	0
A	0
A	2
A	0
A	0
A	1

Figura 7.19

7.10.2. Pérdida esperada

Al igual que en la sección de la matriz de transición, el usuario podrá seleccionar la matriz calculada o introducir una propia:



Figura 7.20

7.11. Simulación y resultados.

En esta sección el usuario indicara el número de simulaciones para la simulación de MonteCarlo, la aplicación realizara los cálculos y nos presentara las métricas de riesgo y la opción de descargar el reporte.

7.12. Stress Testing.

Una vez llevado a cabo todos los pasos requeridos para la metodología, el usuario podrá realizar una prueba de estres sobre los parámetros del modelo.

Simulación y Resultados

Número de simulaciones

2

Escoga nivel de confianza para el VaR

- ☐ 90%
- ☒ 95%
- ☐ 99%

Pérdida esperada:

32837.77

Valor en riesgo:

33891.87

TVaR:

34159.65

 Descargar**Figura 7.21**

StressTesting

Nivel de estrés de la prueba

- ☐ 1 %
- ☐ 5 %
- ☒ 10 %

Resultado

El valor de la prueba

37281.06

Figura 7.22

7.13. Tercera sección: Indicadores contables

En esta sección se calculan indicadores contables, poniendo atención al RAROC de crédito, y otros que dependen exclusivamente del balance de la entidad financiera.

7.14. Cálculo del RAROC

Dedicamos una subsección completa debido a su gran uso en las entidades financieras. Lo primero que veremos es una sección informativa:

✳ Cálculo del RAROC ✳ RAROC

Calculo del RAROC

La ecuacion estandar para el calculo del RAROC es:

$$\text{RAROC} = \frac{\text{RAR}}{C}$$

Donde:

RAR = Risk Adjusted Return

C = Capital necesario para cubrir el riesgo

Figura 7.23

7.14.1. RAROC

En esta pestaña, requerimos que el usuario cargue la información del balance necesaria para el cálculo del RAROC y la metodología de riesgo de crédito que uso

The screenshot displays the 'RAROC' tab of the 'Cálculo del RAROC' application. The interface is divided into three main sections:

- Metodología:** A section with two radio buttons. 'CreditRisk+' is unselected, and 'Credimetrics' is selected.
- Ingreso Balance:** A section containing a search bar with the text 'Buscar' and a placeholder message 'Aun no seleccionas el archivo...'.
- RAROC Result:** A large orange box on the right displaying the calculated RAROC value of 2.90216.

Figura 7.24

A

Software Tools

For those who are not familiar with software packages required for using R Markdown, we give a brief introduction to the installation and maintenance of these packages.

A.1. R and R packages

R can be downloaded and installed from any CRAN (the Comprehensive R Archive Network) mirrors, e.g., <https://cran.rstudio.com>. Please note that there will be a few new releases of R every year, and you may want to upgrade R occasionally.

To install the **bookdown** package, you can type this in R:

```
install.packages("bookdown")
```

This installs all required R packages. You can also choose to install all optional packages as well, if you do not care too much about whether these packages will actually be used to compile your book (such as **htmlwidgets**):

```
install.packages("bookdown", dependencies = TRUE)
```

If you want to test the development version of **bookdown** on GitHub, you need to install **devtools** first:

```
if (!requireNamespace('devtools')) install.packages('devtools')
devtools::install_github('rstudio/bookdown')
```

R packages are also often constantly updated on CRAN or GitHub, so you may want to update them once in a while:

```
update.packages(ask = FALSE)
```

Although it is not required, the RStudio IDE can make a lot of things much easier when you work on R-related projects. The RStudio IDE can be downloaded from <https://www.rstudio.com>.

A.2. Pandoc

An R Markdown document (*.Rmd) is first compiled to Markdown (*.md) through the **knitr** package, and then Markdown is compiled to other output formats (such as LaTeX or HTML) through Pandoc. This process is automated by the **rmarkdown** package. You do not need to install **knitr** or **rmarkdown** separately, because they are the required packages of **bookdown** and will be automatically installed when you install **bookdown**. However, Pandoc is not an R package, so it will not be automatically installed when you install **bookdown**. You can follow the installation instructions on the Pandoc homepage (<http://pandoc.org>) to install Pandoc, but if you use the RStudio IDE, you do not really need to install Pandoc separately, because RStudio includes a copy of Pandoc. The Pandoc version number can be obtained via:

```
rmarkdown::pandoc_version()
## [1] '2.7.3'
```

If you find this version too low and there are Pandoc features

only in a later version, you can install the later version of Pandoc, and **rmarkdown** will call the newer version instead of its built-in version.

A.3. LaTeX

LaTeX is required only if you want to convert your book to PDF. The typical choice of the LaTeX distribution depends on your operating system. Windows users may consider MiKTeX (<http://miktex.org>), Mac OS X users can install MacTeX (<http://www.tug.org/mactex/>), and Linux users can install TeXLive (<http://www.tug.org/texlive>). See <https://www.latex-project.org/get/> for more information about LaTeX and its installation.

Most LaTeX distributions provide a minimal/basic package and a full package. You can install the basic package if you have limited disk space and know how to install LaTeX packages later. The full package is often significantly larger in size, since it contains all LaTeX packages, and you are unlikely to run into the problem of missing packages in LaTeX.

LaTeX error messages may be obscure to beginners, but you may find solutions by searching for the error message online (you have good chances of ending up on StackExchange¹). In fact, the LaTeX code converted from R Markdown should be safe enough and you should not frequently run into LaTeX problems unless you introduced raw LaTeX content in your Rmd documents. The most common LaTeX problem should be missing LaTeX packages, and the error may look like this:

```
! LaTeX Error: File `titling.sty' not found.
```

```
Type X to quit or <RETURN> to proceed,
```

¹<http://tex.stackexchange.com>

```
or enter new name. (Default extension: sty)

Enter file name:
! Emergency stop.
<read *>

1.107 ^~M

pandoc: Error producing PDF
Error: pandoc document conversion failed with error 43
Execution halted
```

This means you used a package that contains `titling.sty`, but it was not installed. LaTeX package names are often the same as the `*.sty` filenames, so in this case, you can try to install the `titling` package. Both MiKTeX and MacTeX provide a graphical user interface to manage packages. You can find the MiKTeX package manager from the start menu, and MacTeX’s package manager from the application “TeX Live Utility”. Type the name of the package, or the filename to search for the package and install it. TeXLive may be a little trickier: if you use the pre-built TeXLive packages of your Linux distribution, you need to search in the package repository and your keywords may match other non-LaTeX packages. Personally, I find it frustrating to use the pre-built collections of packages on Linux, and much easier to install TeXLive from source, in which case you can manage packages using the `tlmgr` command. For example, you can search for `titling.sty` from the TeXLive package repository:

```
tlmgr search --global --file titling.sty
# titling:
# texmf-dist/tex/latex/titling/titling.sty
```

Once you have figured out the package name, you can install it by:

```
tlmgr install titling # may require sudo
```

LaTeX distributions and packages are also updated from time to time, and you may consider updating them especially when you run into LaTeX problems. You can find out the version of your LaTeX distribution by:

```
system('pdflatex --version')  
## pdfTeX 3.14159265-2.6-1.40.19 (TeX Live 2018)  
## kpathsea version 6.3.0  
## Copyright 2018 Han The Thanh (pdfTeX) et al.  
## There is NO warranty. Redistribution of this software is  
## covered by the terms of both the pdfTeX copyright and  
## the Lesser GNU General Public License.  
## For more information about these matters, see the file  
## named COPYING and the pdfTeX source.  
## Primary author of pdfTeX: Han The Thanh (pdfTeX) et al.  
## Compiled with libpng 1.6.34; using libpng 1.6.34  
## Compiled with zlib 1.2.11; using zlib 1.2.11  
## Compiled with xpdf version 4.00
```