Simulación

# Modelos de pérdida agregada en riesgo de crédito

Mario Humberto Becerra Contreras 124362 Juan Bernardo Martínez Parente Castañeda 124458

10 de diciembre de 2015

**Resumen**

En el presente trabajo se desea estimar la pérdida agregada de una cartera crediticia utilizando los modelos individual y colectivo, los cuales agrupan la información de forma distinta. Primero se presentan las metodologías que cada uno utiliza; después se presentan los resultados de las distintas simulaciones realizadas. Se obtuvo un valor esperado de la pérdida agregada de 1,109M de pesos bajo el modelo individual y de 787M pesos bajo el modelo colectivo.

**Keywords:** Simulación, Monte Carlo, riesgo de crédito.

## Introducción

La cartera crediticia de un crédito es el saldo de montos efectivamente entregados a los acreditados más los intereses devengados no cobrados correspondientes. Los modelos de pérdida agregada permiten estimar la pérdida total que el conjunto de individuos acreditados genera por impago a partir de una probabilidad de riesgo de incumplimiento de pago dada.

En el presente reporte se presentan dos metodologías para estimar el valor esperado de la pérdida agregada de una cartera cuyos acreditados pertenecen a cuatro grupos homogéneos de riesgo: industrial, construcción, comercio y servicios. En la tabla de abajo se muestran los componentes de riesgo de cada grupo, en donde $p_j$ es la probabilidad de impago del grupo $j$ y $\rho_j$ es la correlación intragrupo del grupo $j$. Además, se sabe que la correlación de que cualquier pareja de variables pertenezca a grupos distintos (correlación extragrupo) es $\rho = 0.5\,\%$.

| | Grupo | | | |
|---|---|---|---|---|
| | **Industrial** | **Construcción** | **Comercio** | **Servicios** |
| $p_j$ | 0.7 % | 0.9 % | 0.65 % | 0.6 % |
| $\rho_j$ | 0.09 % | 0.04 % | 0.05 % | 0.07 % |

Cuadro 1: Componentes de Riesgo de cada grupo.

Se cuenta con datos sobre los montos de diferentes créditos otorgados a personas de los cuatro sectores. Se realizó un análisis exploratorio de los mismos que se muestra en la siguiente tabla:

| | Grupo | Media | Varianza | Mínimo | Cuantil 25 | Mediana | Cuantil 75 | Máximo |
|---|---|---|---|---|---|---|---|---|
| 1 | Industrial | 548.21 | 293327.48 | 2.00 | 138.00 | 369.50 | 821.25 | 2429.00 |
| 2 | Construcción | 549.38 | 256209.43 | 1.00 | 159.50 | 374.00 | 797.00 | 1949.00 |
| 3 | Comercio | 438.51 | 182619.29 | 7.00 | 148.00 | 302.50 | 644.50 | 2292.00 |
| 4 | Servicios | 492.93 | 282577.22 | 3.00 | 144.50 | 357.50 | 674.50 | 3190.00 |

Cuadro 2: Estadísticos descriptivos.

## Marco teórico

Un *modelo de pérdida agregada*, permite obtener la pérdida total de un conjunto de individuos en una cartera crediticia, para ello existen dos maneras de agregar las pérdidas individuales, a saber: *modelo individual* y *modelo colectivo*.

Se puede hacer el supuesto de que los acreditados están agrupados en $m$ grupos homogéneos de riesgo, entonces se tiene que todos los acreditados pertenecientes a un mismo grupo comparten las mismas probabilidades de incumplimiento y las mismas correlaciones lineales de incumplimientos para cualquier pareja de acreditados tomada dentro del mismo grupo.

### Modelo individual

En el modelo individual la pérdida agregada se define por la siguiente variable aleatoria:

$$L = \sum_{j=1}^{m} L_j \quad \text{con:} \quad L_j = \sum_{k=1}^{n_j} I_k^{(j)} \, f_k^{(j)}, \quad I_k^{(j)} \sim \text{Bernoulli}(p_j) \quad \text{y} \quad Corr(I_k^{(j)}, I_i^{(j)}) = \rho_j \quad \forall i \neq j$$

donde:
$j$: Subíndice que denota al número de grupo ($j \in \{1, \ldots, m\}$).
$n_j$: Número de créditos del grupo $j$.
$I_k^{(j)}$: Variable indicadora del incumplimiento del acreditado $k$ del grupo $j$.
$f_k^{(j)}$: Monto del crédito del acreditado $k$ del grupo $j$.
$p_j$: Probabilidad de incumplimiento de cualquier acreditado del grupo $j$.
$\rho_j$: Correlación lineal de incumplimiento para cualquier pareja de acreditados del grupo $j$ (llamada también *correlación intragrupo*).

Además de las *correlaciones intragrupo* de incumplimiento, existen correlaciones lineales para parejas de acreditados tomadas de grupos diferentes (llamadas también *correlaciones extragrupo*), que también deben de ser consideradas para modelar todas las posibles correlaciones que pueden existir. Esto es:

$$Corr(I_k^{(j)}, I_r^{(i)}) = \rho_{j,i} \quad \forall i \neq j \ \text{ y } \ \forall k \neq r$$

Para que la variable aleatoria de las pérdidas de la cartera quede bien definida, se debe de considerar una cierta estructura de dependencia entre los incumplimientos de la cartera (i.e. la dependencia entre las variables indicadoras "$I_k$"), la cual se puede modelar con un modelo en el que se calibre la estructura de dependencia con las correlaciones lineales entre las variables indicadoras. Se tiene que una forma eficiente de simular la v.a. $L$ es a través de *variables latentes* definidas con un *modelo de factor* utilizando dos *factores*, utilizando una única correlación ($\tilde{\rho}$) entre cualquier pareja de *variables latentes* que no pertenezcan al mismo grupo para simplificar el modelo, quedando definidas las variables de la siguiente forma:

$$I_i^{(j)} = 1_{X_i^{(j)} \leq u} \quad \text{con:} \quad X_i^{(j)} = \sqrt{\tilde{\rho}} \cdot Z_0 + \sqrt{\tilde{\rho}_j - \tilde{\rho}} \cdot Z_j + \sqrt{1 - \tilde{\rho}_j} \cdot \epsilon_i \quad \text{y con:} \ \{Z_j\}_{i=0}^{m} \text{ y } \{\epsilon_i\}_{i=1}^{n} \overset{iid}{\sim} N(0,1) \quad (1)$$

Con la siguiente calibración:

$$u_j = \Phi^{-1}(p_j) \ \text{ y } \ \tilde{\rho}_j = \{r : \Phi(u_j, u_j; r) = p_j^2 + \rho_j \cdot p_j(1 - p_j)\} \ \ \forall j \in \{1, \ldots, m\}$$

La calibración de $\tilde{\rho}$ es mucho más compleja por todas las combinaciones posibles que puede haber de grupos, por lo que para este caso se considera una $\tilde{\rho}$ fija e igual a 0.005.

Así, para poder simular una realización de la variable $L$ se necesita primeramente calibrar los parámetros $\{u_j\}_{j=1}^m$ y $\{\tilde{\rho}\}_{j=1}^m$, luego simular para cada grupo las $n_j$ *variables latentes* para obtener sus correspondientes *variables indicadoras*, para finalmente multiplicarlas por sus respectivos montos y realizar la sumatoria.

**Modelo colectivo**

En este caso, la variable aleatoria de pérdida agregada se define como:

$$L = \sum_{j=1}^{m} L_j \text{ con: } L_j = \sum_{k=1}^{K_j} f_{D_k}^{(j)} \text{ con: } K_j = \sum_{k=1}^{n_j} I_k^{(j)}$$

donde $I_k^{(j)} \sim \text{Bernoulli}(p_j)$ , $\{D_1, \ldots, D_{K_j}\} \subset \{1, \ldots, n_j\}$ ; $j = 1, \ldots, m$ y $\{D_1, \ldots, D_{K_j}\}$ es una muestra sin remplazo de $K_j$ elementos del conjunto $\{1, \ldots, n_j\}$, donde todos los elementos del conjunto comparten la misma probabilidad de ser seleccionados.

Una forma eficiente de simular el número de incumplimientos de cada grupo ($K_j$) es a través de la suma de variables indicadoras de *variables aleatorias latentes*, donde estas últimas variables se modelan con un *modelo de factor* utilizando dos factores (ver ecuación 1), de tal forma que si se condicionan los dos factores ($Z_0$ y $Z_j$) de las $n$ *variables latentes* ($X_i$'s) entonces se tiene que dichas *variables latentes* son mutuamente independientes, propiedad que se le da el nombre de *independencia condicional*, por lo que la suma de las variables indicadoras de estas *variables latentes* condicionadas en sus dos *factores* se tiene que se distribuye *Binomial* (por ser una sumatoria de v.a. *Bernoullis* independientes), con parámetros dados por el número de acreditados del grupo que se trate ($n_j$) y por la probabilidad condicional al valor de los *factores* que correspondan. Esto es:

$$K_j | Z_0 = z_0, Z_j = z_j \sim Binomial(n_j, p(z_0, z_j)),$$

con

$$\begin{aligned} p(z_0, z_j) &= \mathbb{P}\{I_i^{(j)} c = 1 | Z_0 = z_0, Z_j = z_j\} \\ &= \mathbb{P}\{X_i^{(j)} \leq u_j | Z_0 = z_0, Z_j = z_j\} \\ &= \Phi\left(\frac{u_j - \sqrt{\tilde{\rho}}\, z_0 - \sqrt{\tilde{\rho}_j - \tilde{\rho}}\, z_j}{\sqrt{1 - \tilde{\rho}_j}}\right) \end{aligned}$$

Por lo tanto, para poder simular una realización de la variable $L$ se necesita primeramente calibrar los parámetros $\{u_j\}_{j=1}^m$, $\{\tilde{\rho}\}_{j=1}^m$ y $\rho$ (exactamente de la misma forma en que se hace bajo el *modelo individual*), luego simular el *factor sistémico* y los *factores sectoriales* para poder simular la v.a.'s $\{K_j\}_{j=1}^m$ condicionadas en los valores de los *factores* simulados, para con ello simular la *muestra sin remplazo* de $K_j$ elementos tomados del conjunto $\{f_1, \ldots, f_{n_j}\}$ y sumar dichos valores para obtener la pérdida agregada del grupo ($L_j$), realizándose lo anterior para cada grupo (i.e. $j = 1, \ldots, m$) para finalmente sumar las pérdidas agregadas de cada uno de los grupos.

**Técnicas de reducción de varianza**

En la teoría de simulación, la reducción de varianza es un procedimiento que se utiliza para aumentar la precisión de la estimación de un parámetro de interés utilizando el mismo número de simulaciones. En este proyecto se utiliza la técnica de *variables antitéticas*. Para entender el procedimiento utilizado, en los siguientes párrafos, se presentan algunas definiciones y resultados.

**Definición 1 (Variables antitéticas)** *Se dice que la variable aleatoria $X'$ es una* ***variable antitética*** *de la variable aleatoria $X$ si cumple que:*

- *$X$ tiene la misma distribución que $X'$*
- $\operatorname{corr}(X, X') < 0$

Muchas veces se utiliza la simulación para estimar algún parámetro $\theta$ de la variable aleatoria $X$, siendo la mayoría de estas veces que $\theta = \mathbb{E}[X]$, es decir, el parámetro de interés es la varianza. En el caso de este trabajo, el parámetro de interés es la esperanza de las pérdidas agregadas de la cartera, así que es importante enunciar el siguiente resultado.

**Teorema 1** *Sean $X$ una variable aleatoria, $X'$ variable antitética de $X$, $\theta^* = \frac{1}{n}\sum_{i=1}^{n} x_i$ donde $\{x_i\}_{i=1}^{n}$ es un conjunto de realizaciones independientes de $X$, $\hat{\theta^*} = \frac{1}{n}\sum_{i=1}^{\frac{n}{2}}(x_i + x_i')$ donde $\{x_i'\}_{i=1}^{n}$ es un conjunto de realizaciones independientes de $X'$. Entonces* $\operatorname{Var}(\hat{\theta}^*) < \operatorname{Var}(\theta^*)$.

**Teorema 2** *Sean $X$ una variable aleatoria tal que $X \sim \text{Normal}(\mu, \sigma)$. Se sabe que $X$ se puede expresar como $X = \mu + \sigma Z$ donde $Z \sim \text{Normal}(0, 1)$. Entonces una variable antitética de $X$ es $X' = \mu - \sigma Z$.*

Entonces, al usar variables antitéticas, se tiene que con el mismo número de simulaciones, se puede tener una mayor precisión en la estimación. Un algoritmo para estimar la esperanza de una variable aleatoria mediante simulación es:

- Definir tamaño de la simulación $n \in \mathbb{N}$
- Simular $\{x_i\}_{i=1}^{\frac{n}{2}} \sim F_X$ iid.
- Obtener $\hat{\theta^*} = \frac{1}{n}\sum_{i=1}^{\frac{n}{2}}(x_i + x_i')$

# Metodología

Como ya se mencionó, se realizaron simulaciones de las pérdidas agregadas de una cartera crediticia cuyos acreditados pertenecen a los siguientes grupos homogéneos de riesgo: *industrial*, *construcción*, *comercio* y *servicios*, con probabilidades de incumplimiento y correlaciones intragrupo definidos en el **cuadro 3**, con distintos montos. Se utiliza un modelo de dos factores donde la correlación de cualquier pareja de variables latentes que pertenezcan a grupos diferentes es de **0.5 %**, es decir, $r\hat{h}o = 0.005$, de acuerdo con la notación utilizada anteriormente.

Se prueban los dos modelos: el individual y el colectivo, y además se utiliza una técnica de reducción de varianza mediante variables antitéticas para el modelo individual.

| Grupo | **industrial** | **construc.** | **comercio** | **servicios** |
|---|---|---|---|---|
| $p_j$ | 0.70 % | 0.90 % | 0.65 % | 0.60 % |
| $\rho_j$ | 0.09 % | 0.04 % | 0.05 % | 0.07 % |

Cuadro 3: Componentes de Riesgo de cada grupo.

### Tamaño de la muestra

Antes de proceder con el desarrollo de los modelos, fue necesario calcular el tamaño de la simulación requerido para cada uno de los grupos. Se utilizó la fórmula

$$n_j = \left(\frac{\sigma z_{1-\frac{\alpha}{2}}}{c}\right)^2,$$

para $j = 1, 2, 3, 4$, donde $c = 10$ es el nivel de precisión, $1 - \alpha = 0.95$ es el nivel de confianza y $z_{1-\frac{\alpha}{2}}$ es el cuantil $1 - \frac{\alpha}{2}$ de la v.a. $Z \sim N(0, 1)$.

### Modelo individual

Los pasos llevados a cabo para cada grupo ($j = 1, 2, 3, 4$) son los siguientes (ver códigos 0, 1 y 2 en el anexo):

1. Calibrar los parámetros $u_j$ y $\tilde{\rho}_j$ como sigue:

$$u_j = \Phi^{-1}(p), \tag{2}$$

$$\tilde{\rho}_j = \left\{r : \Phi(u_j, u_j; r) = p_j^2 + \rho_j p_j (1 - p_j)\right\}. \tag{3}$$

2. Simular las variables $\{x_i^{(j)}\}_{i=1}^{n_j}$, con

$$x_i^{(j)} = \sqrt{\tilde{\rho}} \cdot z_0 + \sqrt{\tilde{\rho}_j - \tilde{\rho}} \cdot z_j + \sqrt{1 - \tilde{\rho}_j} \cdot \varepsilon_i, \quad \text{donde } \{z_j\}_{j=1}^{m} \text{ y } \{\varepsilon_i\}_{i=1}^{n_j} \overset{iid}{\sim} N(0, 1).$$

3. Obtener $I_i^{(j)} = 1_{x_i^{(j)} \leq u_j}$.
4. Calcular

$$L_j = \sum_{k=1}^{n_j} I_k^{(j)} f_k^{(j)}.$$

5. Obtener

$$L = \sum_{j=1}^{m} L_j.$$

### Modelo colectivo

Los pasos llevados a cabo para cada grupo ($j = 1, 2, 3, 4$) son los siguientes (ver códigos 0 y 3 en el anexo):

1. Calibrar los parámetros $u_j$ y $\tilde{\rho}_j$ de la misma forma que en (2) y (3).
2. Simular $\{z_i^{(j)}\}_{i=0}^{m}$, con $z_j \overset{iid}{\sim} N(0, 1)$.

3. Calcular

$$p(z_0, z_j) = \Phi\left(\frac{u_j - \sqrt{\tilde{\rho}} \cdot z_0 - \sqrt{\tilde{\rho}_j - \tilde{\rho}} \cdot z_j}{\sqrt{1-\tilde{\rho}}}\right).$$

4. Simular $k_j \overset{iid}{\sim} Binomial\big(n_j, p(z_0, z_j)\big)$
5. Obtener $\{D_1, \ldots, D_{k_j}\}$ una muestra sin reemplazo de $K_j$ elementos del conjunto $\{1, \ldots, n_j\}$, cuyos elementos comparten la misma probabilidad de ser seleccionados.
6. Calcular

$$L_j = \sum_{k=1}^{k_j} f_{D_k}^{(j)}.$$

7. Obtener

$$L = \sum_{j=1}^{m} L_j.$$

**Modelo individual con reducción de varianza**

Para poder implementar el modelo individual usando una técnica de reducción de varianza con variables antitéticas, es necesario primero encontrar la variable antitética de nuestra variable de interés. Debido a que se usa un modelo de variable latente, la variable aleatoria para la cual se debe de encontrar la variable antitética es precisamente la variable latente, $X_i^{(j)}$ para $i = 1, \ldots, n_j$ y $j = 1, 2, 3, 4$. Como $X_i^{(j)}$ está definida de la siguiente forma:

$$X_i^{(j)} = \sqrt{\tilde{\rho}} \cdot z_0 + \sqrt{\tilde{\rho}_j - \tilde{\rho}} \cdot z_j + \sqrt{1-\tilde{\rho}_j} \cdot \varepsilon_i, \quad \text{donde } \{z_j\}_{j=1}^{m} \text{ y } \{\varepsilon_i\}_{i=1}^{n_j} \overset{iid}{\sim} N(0,1), \tag{4}$$

Sea entonces $\hat{X}_i^{(j)} = -X_i^{(j)}$. Es fácil ver que $\hat{X}_i^{(j)}$ es una variable antitética de $X_i^{(j)}$, pues tienen la misma distribución y se cumple que $\operatorname{corr}(\hat{X}_i^{(j)}, X_i^{(j)}) < 0$.

Entonces, los pasos que se deben llevar a cabo para cada grupo ($j = 1, 2, 3, 4$) para simular la esperanza usando variables antitéticas son los siguientes (ver códigos 0 y 4 en el anexo):

1. Calibrar los parámetros $u_j$ y $\tilde{\rho}_j$ de la misma forma que en (2) y (3).
2. Simular las variables $\{x_i^{(j)}\}_{i=1}^{\frac{n_j}{2}}$, con

$$x_i^{(j)} = \sqrt{\tilde{\rho}} \cdot z_0 + \sqrt{\tilde{\rho}_j - \tilde{\rho}} \cdot z_j + \sqrt{1-\tilde{\rho}_j} \cdot \varepsilon_i, \quad \text{donde } \{z_j\}_{j=1}^{m} \text{ y } \{\varepsilon_i\}_{i=1}^{\frac{n_j}{2}} \overset{iid}{\sim} N(0,1).$$

3. Obtener $I_i^{(j)} = 1_{x_i^{(j)} \leq u_j}$.
4. Calcular

$$L_j = \sum_{k=1}^{n_j} I_k^{(j)} f_k^{(j)}.$$

5. Obtener

$$L = \sum_{j=1}^{m} L_j.$$

# Resultados

## Modelo individual

El tiempo transcurrido fue de 0.6638 segundos. Los resultados obtenidos se muestran a continuación:

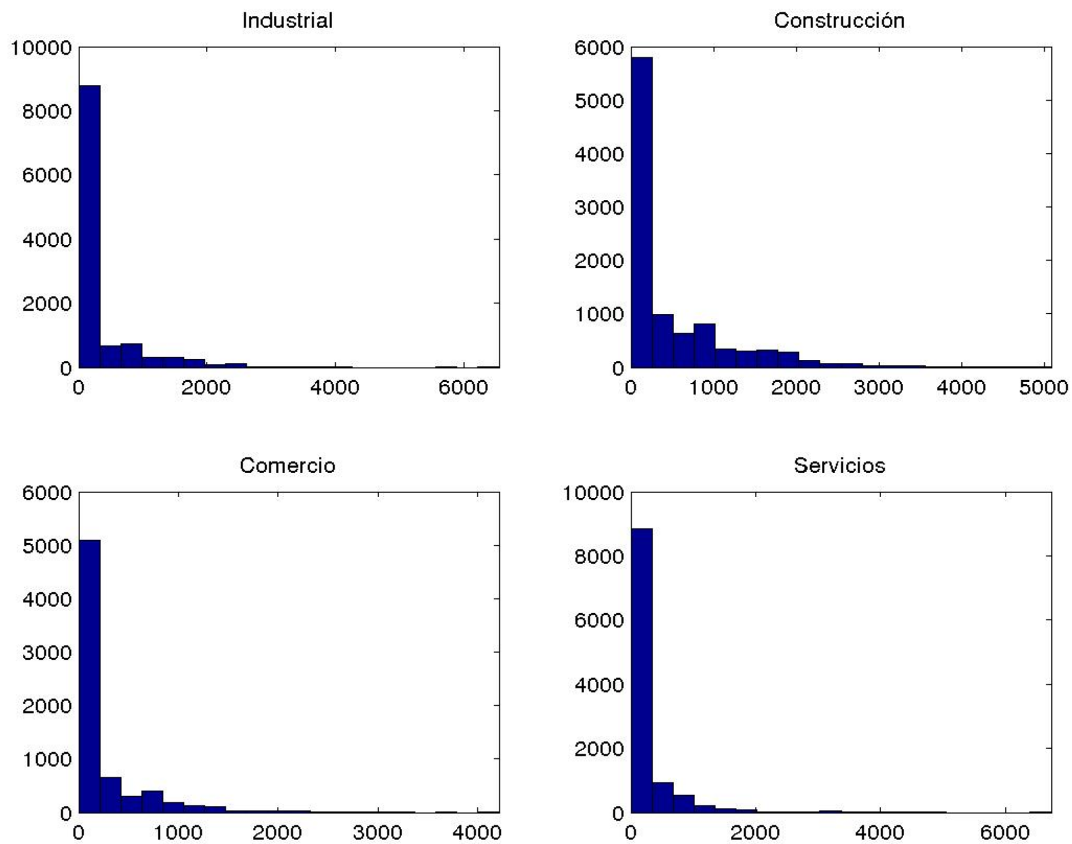


Figura 1: Histogramas del modelo individual.

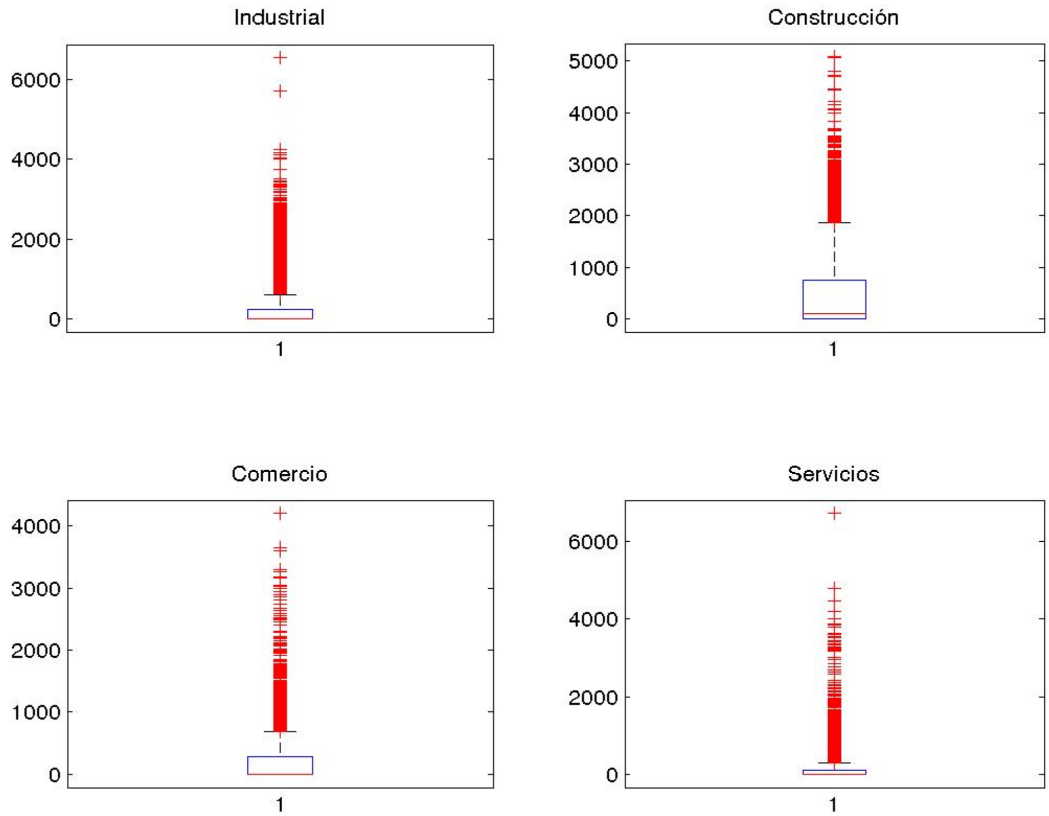


Figura 2: Diagramas de caja del modelo individual.

| | **Industrial** | **Construcción** | **Comercio** | **Servicios** |
|---|---|---|---|---|
| Media | 267.0894 | 456.3001 | 222.9059 | 183.0590 |
| Varianza | 301,018.03 | 465,114.70 | 187,045.33 | 201,086.22 |
| Kurtosis | 12.7177 | 6.9613 | 13.8913 | 27.4552 |
| Sesgo | 2.7845 | 1.8954 | 2.8896 | 4.1931 |
| Cuartil 1 | 0 | 0 | 0 | 0 |
| Cuartil 2 | 0 | 95.35 | 0 | 0 |
| Cuartil 3 | 243.00 | 748.10 | 274.05 | 116.00 |
| Cuartil 4 | 6,550 | 5,088 | 4,210 | 6,735 |
| Cuantil 95 | 1523 | 1879 | 1107 | 999 |

Cuadro 4: Resultados del modelo individual.

**Modelo colectivo**

El tiempo transcurrido fue de 2.3642 segundos. Los resultados obtenidos se muestran a continuación:

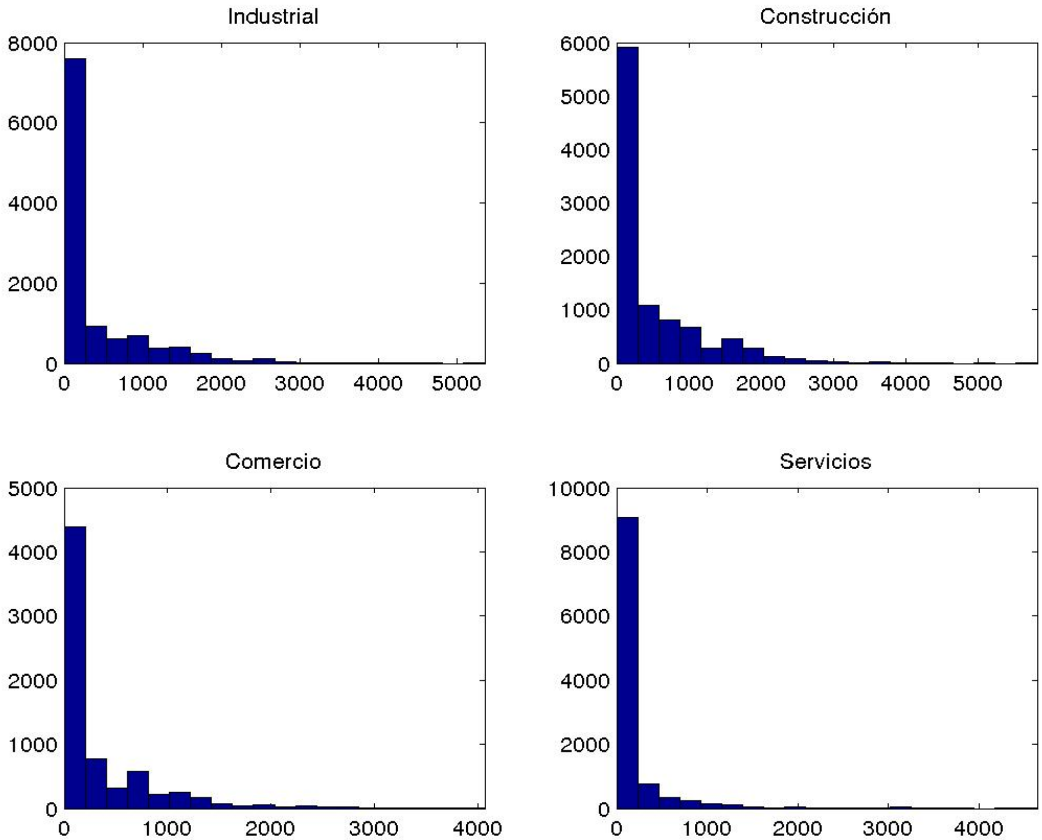


Figura 3: Histogramas del modelo colectivo.

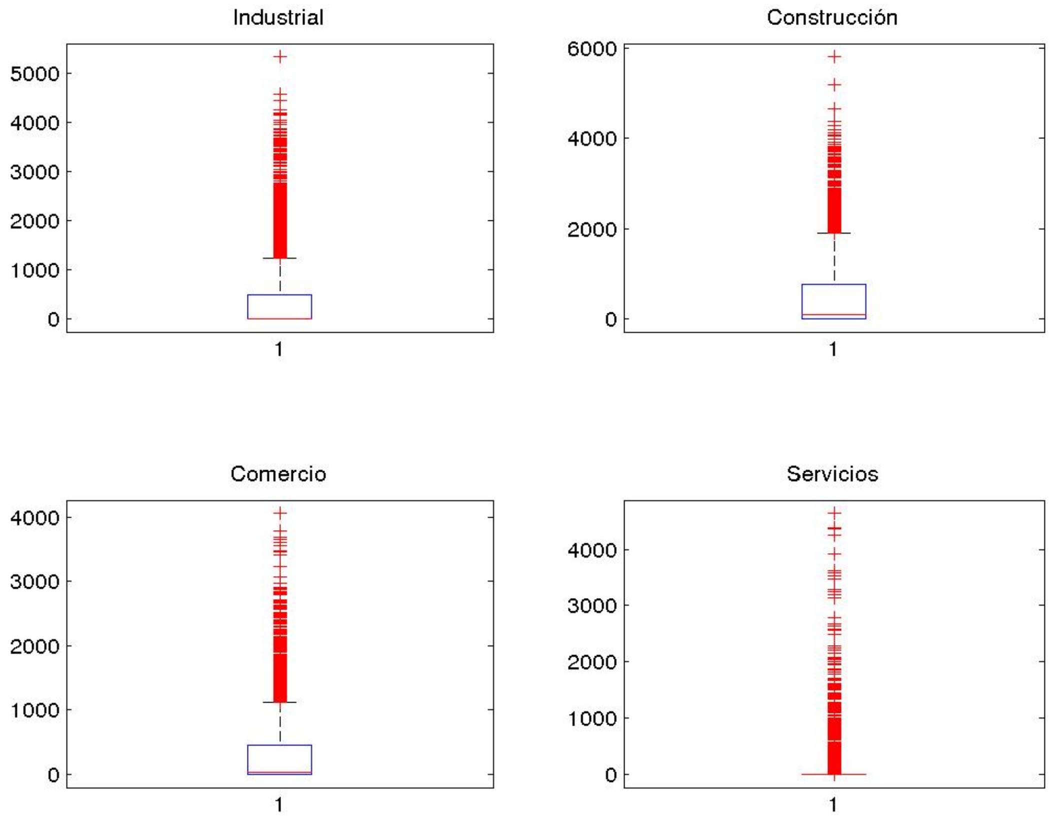


Figura 4: Diagramas de caja del modelo colectivo.

| | **Industrial** | **Construcción** | **Comercio** | **Servicios** |
|---|---|---|---|---|
| Media | 267.0894 | 456.3001 | 222.9059 | 183.0590 |
| Varianza | 301,018.03 | 465,114.70 | 187,045.33 | 201,086.22 |
| Kurtosis | 12.7177 | 6.9613 | 13.8913 | 27.4552 |
| Sesgo | 2.7845 | 1.8954 | 2.8896 | 4.1931 |
| Cuartil 1 | 0 | 0 | 0 | 0 |
| Cuartil 2 | 0 | 96.00 | 29.00 | 0 |
| Cuartil 3 | 490.00 | 761.00 | 448.00 | 0 |
| Cuartil 4 | 5348.00 | 5817.00 | 4065.00 | 4642.00 |
| Cuantil 95 | 1745.15 | 1846.00 | 1370.00 | 828.00 |

Cuadro 5: Resultados del modelo colectivo.

**Modelo individual con reducción de varianza**

El tiempo transcurrido fue de 0.6570 segundos. Los resultados obtenidos se muestran a continuación:

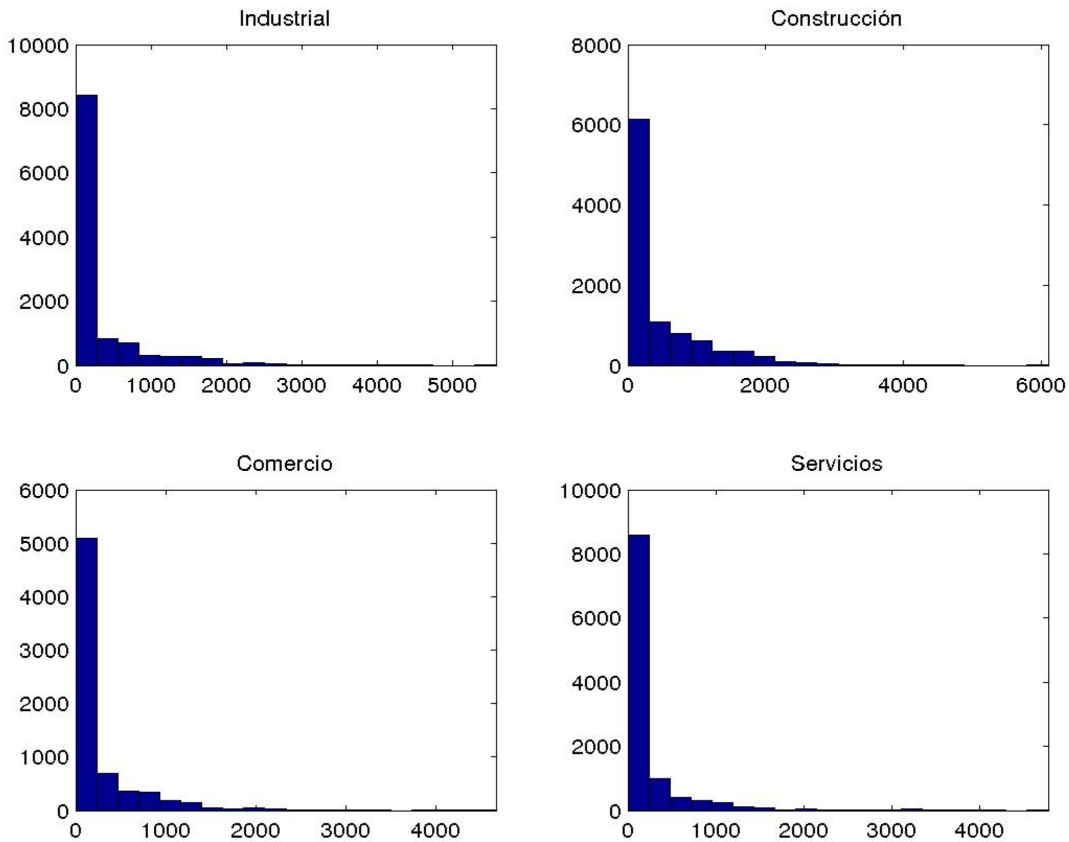


Figura 5: Histogramas del modelo individual con reducción de varianza.

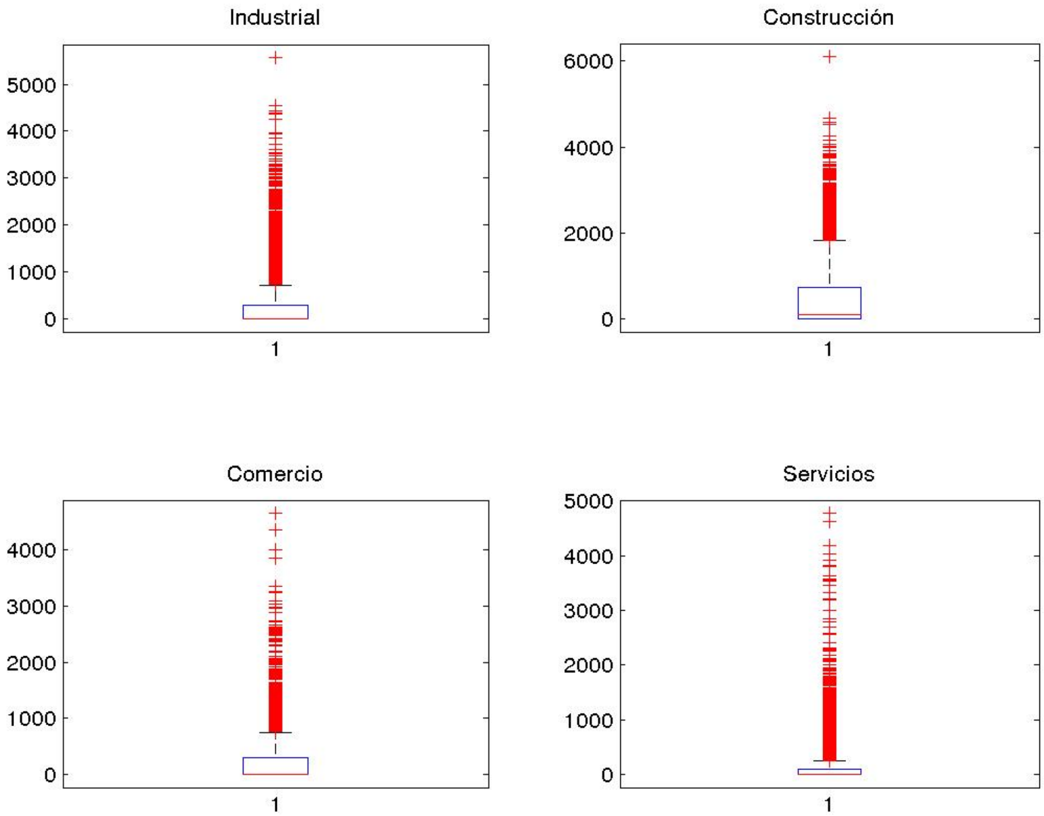


Figura 6: Diagramas de caja del modelo individual con reducción de varianza.

| | **Industrial** | **Construcción** | **Comercio** | **Servicios** |
|---|---|---|---|---|
| Media | 266.03 | 442.05 | 226.89 | 183.59 |
| Varianza | 282,316 | 448,527 | 182,652 | 196,497 |
| Kurtosis | 10.8 | 7.64 | 11.72 | 25.62 |
| Sesgo | 2.61 | 1.99 | 2.68 | 4.11 |
| Cuartil 1 | 0 | 0 | 0 | 0 |
| Cuartil 2 | 0 | 95 | 0 | 0 |
| Cuartil 3 | 273 | 736 | 280.00 | 140 |
| Cuartil 4 | 4452.00 | 5571.00 | 3455.00 | 4772.00 |
| Cuantil 95 | 1526 | 1846.00 | 1158.00 | 999.00 |

Cuadro 6: Resultados del modelo individual con reducción de varianza.

**Comparación de los modelos**

| | **Teórico** | **Individual simple** | **Individual con red. de varianza** | **Colectivo** |
|---|---|---|---|---|
| Media | 1,119.10 | 1,109.30 | 1,118 | 797.35 |
| **Tiempo de ejecución** | - | 0.6638 | 0.6570 | 2.3642 |

Cuadro 7: Comparación de resultados.

En la tabla anterior podemos ver que el modelo colectivo es, en general, el peor en términos tanto de eficiencia como de precisión de las estimaciones. En cuanto a los modelos individuales, en ambos se obtuvieron resultados empíricos bastante cercanos a los teóricos y el tiempo de ejecución fue muy parecido, esto porque se simuló el mismo número en ambos casos.

La media teórica fue obtenida a partir de la definición de la variable aleatoria $L$, y se llega a que

$$\mathbb{E}\left[L\right] = \sum_{j=1}^{4} p_j \sum_{k=1}^{n_j} f_k^{(j)} = 1119.10.$$

Si se quisiera calcular la probabilidad de que las pérdidas agregadas de la cartera excedan 2.5 veces la media teórica, entonces habría que simular muchas veces $L$ y hacer el cociente del número de veces que la $L$ simulada, $L_{sim}$ fue mayor que $2.5 \cdot 1,119.1 = 2,797,750$. El código implementado se puede ver en el anexo (código 5).

# Anexos

A continuación se presenta el código de Matlab utilizado para las simulaciones anteriores.

## Código 0

```
%% Declaración de variables
clear
load('grupos')

rng(48151623);

% Tamaño de la simulación
tam = zeros(4,1);
tam(1) = length(industrial);
tam(2) = length(construccion);
tam(3) = length(comercio);
tam(4) = length(servicios);

z = norminv(0.975);
n_sim = zeros(4,1);
n_sim(1) = ceil((z*std(industrial)/10)^2);
n_sim(2) = ceil((z*std(construccion)/10)^2);
n_sim(3) = ceil((z*std(comercio)/10)^2);
n_sim(4) = ceil((z*std(servicios)/10)^2);

m = 4;
p = [0.007,0.009,0.0065,0.0060]';
rho_dif = 0.005;
r0 = .5;
rho = [.0009,.0004,.0005,.0007]';

% calibrar u
for i=1:m
    u(i) = norminv(p(i));
end

% calibrar ro
for j = 1:m
    fun = @(r) (mvncdf([u(j); u(j)],[0; 0],[1 r; r 1]) - (p(j)^2 + rho(j)*p(j)*(1-p(j))));
    rho_tilde(j) = fzero(fun,r0);
end
```

## Código 1

```
%% Individual
L_ind = zeros(n_sim(1),1);
L_cons = zeros(n_sim(2),1);
L_com = zeros(n_sim(3),1);
L_serv = zeros(n_sim(4),1);

tic;
for i = 1:max(n_sim)
    [ I_ind, I_cons, I_com, I_serv ] = individual(m, u,rho_tilde, tam, rho_dif);
    L_ind(i) = I_ind'*industrial;
    L_cons(i) = I_cons'*construccion;
    L_com(i) = I_com'*comercio;
    L_serv(i) = I_serv'*servicios;
end
K = [sum(I_ind);sum(I_cons);sum(I_com);sum(I_serv)]
L_ind = L_ind(1:n_sim(1));
L_cons = L_cons(1:n_sim(2));
L_com = L_com(1:n_sim(3));
L_serv = L_serv(1:n_sim(4));
L = sum(L_ind) + sum(L_cons) + sum(L_com) + sum(L_serv)

% ------ PLOTS ------
figure(1)
subplot(2,2,1)
hist(L_ind)
title('Industrial')

subplot(2,2,2)
hist(L_cons)
title('Construcción')

subplot(2,2,3)
hist(L_com)
title('Comercio')

subplot(2,2,4)
hist(L_serv)
title('Servicios')

figure(2)
subplot(2,2,1)
boxplot(L_ind)
title('Industrial')
```

```
subplot(2,2,2)
boxplot(L_cons)
title('Construcción')

subplot(2,2,3)
boxplot(L_com)
title('Comercio')

subplot(2,2,4)
boxplot(L_serv)
title('Servicios')

% ------ ESTADÍSTICOS ------
mu(1) = mean(L_ind); mu(2) = mean(L_cons);
mu(3) = mean(L_com); mu(4) = mean(L_serv);
vari(1) = var(L_ind); vari(2) = var(L_cons);
vari(3) = var(L_com); vari(4) = var(L_serv);
k(1) = kurtosis(L_ind); k(2) = kurtosis(L_cons);
k(3) = kurtosis(L_com); k(4) = kurtosis(L_serv);
sesgo(1) = skewness(L_ind); sesgo(2) = skewness(L_cons);
sesgo(3) = skewness(L_com); sesgo(4) = skewness(L_serv);
quart_ind = quantile(L_ind,[0.25, 0.5, 0.75, 1]);
quart_cons = quantile(L_cons,[0.25, 0.5, 0.75, 1]);
quart_com = quantile(L_com,[0.25, 0.5, 0.75, 1]);
quart_serv = quantile(L_serv,[0.25, 0.5, 0.75, 1]);
quart = [quart_ind; quart_cons; quart_com; quart_serv];

t = toc
```

### Código 2

```
function [ I_ind, I_cons, I_com, I_serv ] = individual(m, u, rho_tilde, tam, rho_dif )
% Este programa simula las variables binarias de impago en un modelos de
% pérdida agregada en riesgo de crédito con varios grupos homogéneos de
% riesgo.
% ------------------------------------------------------------------------
% In: m, número de grupos homogéneos
%     u, vector que contiene las u's calibradas
%     rho_tilde, vector que contiene las rho_tildes's calibradas
%     tam, vector que contiene el número total de créditos de cada grupo
%     rho_dif, correlación extragrupo
% Out: I_ind, vector de simulaciones para el grupo industrial
%      I_cons, vector de simulaciones para el grupo construcción
%      I_com, vector de simulaciones para el grupo comercio
```

```
%       I_serv, vector de simulaciones para el grupo servicios
%
% ---------------------------------------------------------------------------
% calcular las xi
Z0 = randn;
Z = randn(1,m);
for i = 1:m
    x = zeros(tam(i),1);
    for j = 1:tam(i)
        eps = randn;
        x(j) = sqrt(rho_dif)*Z0 + sqrt(rho_tilde(i) - rho_dif)*Z(i)...
               + sqrt(1- rho_tilde(i))*eps;
    end
    switch i
        case 1
            I_ind = x < u(i);
        case 2
            I_cons = x < u(i);
        case 3
            I_com = x < u(i);
        case 4
            I_serv = x < u(i);
    end
end
```

### Código 3

```
%% Colectivo
Z0 = randn;
Z = randn(m,1);
p2 = zeros(m,1);
for i = 1:m
    z = (u(i) - sqrt(rho_dif)*Z0 - sqrt(rho_tilde(i)-rho_dif)*Z(i))/(sqrt(1-rho_tilde(i)));
    p2(i) = normcdf(z);
end
p2

L_ind = zeros(n_sim(1),1);
L_cons = zeros(n_sim(2),1);
L_com = zeros(n_sim(3),1);
L_serv = zeros(n_sim(4),1);
tic;
for i=1:n_sim(1)
    K = binornd(tam(1),p2(1));
    f_ind = randsample(industrial,K);
```

```
    L_ind(i) = sum(f_ind);
end

for i=1:n_sim(2)
    K = binornd(tam(2),p2(2));
    f_cons = randsample(construccion,K);
    L_cons(i) = sum(f_cons);
end

for i=1:n_sim(3)
    K = binornd(tam(3),p2(3));
    f_com = randsample(comercio,K);
    L_com(i) = sum(f_com);
end

for i=1:n_sim(4)
    K = binornd(tam(4),p2(4));
    f_serv = randsample(servicios,K);
    L_serv(i) = sum(f_serv);
end
L = sum(L_ind) + sum(L_cons) + sum(L_com) + sum(L_serv)


% ------ PLOTS ------
figure(1)
subplot(2,2,1)
hist(L_ind)
title('Industrial')

subplot(2,2,2)
hist(L_cons)
title('Construcción')

subplot(2,2,3)
hist(L_com)
title('Comercio')

subplot(2,2,4)
hist(L_serv)
title('Servicios')

figure(2)
subplot(2,2,1)
boxplot(L_ind)
title('Industrial')
```

```
subplot(2,2,2)
boxplot(L_cons)
title('Construcción')

subplot(2,2,3)
boxplot(L_com)
title('Comercio')

subplot(2,2,4)
boxplot(L_serv)
title('Servicios')

% ------ ESTADÍSTICOS ------
mu(1) = mean(L_ind); mu(2) = mean(L_cons);
mu(3) = mean(L_com); mu(4) = mean(L_serv);
vari(1) = var(L_ind); vari(2) = var(L_cons);
vari(3) = var(L_com); vari(4) = var(L_serv);
k(1) = kurtosis(L_ind); k(2) = kurtosis(L_cons);
k(3) = kurtosis(L_com); k(4) = kurtosis(L_serv);
sesgo(1) = skewness(L_ind); sesgo(2) = skewness(L_cons);
sesgo(3) = skewness(L_com); sesgo(4) = skewness(L_serv);
quart_ind = quantile(L_ind,[0.25, 0.5, 0.75, 1]);
quart_cons = quantile(L_cons,[0.25, 0.5, 0.75, 1]);
quart_com = quantile(L_com,[0.25, 0.5, 0.75, 1]);
quart_serv = quantile(L_serv,[0.25, 0.5, 0.75, 1]);
quart = [quart_ind; quart_cons; quart_com; quart_serv];
t = toc
```

## Código 4

```
% Individual con reducción de varianza

function [ I_ind, I_cons, I_com, I_serv ] = individual_antitetico(u, rho_tilde, tam, rho_dif )
% Este programa simula las variables binarias de impago en un modelos de
% pérdida agregada en riesgo de crédito con varios grupos homogéneos de
% riesgo usando reducción de varianza mediante variables antitéticas
%
% In:
%       u, vector que contiene las u's calibradas
%       rho_tilde, vector que contiene las rho_tildes's calibradas
%       tam, vector que contiene el número total de créditos de cada grupo
%       rho_dif, correlación extragrupo
% Out:
%       I_ind, vector de simulaciones para el grupo industrial
```

```
%       I_cons, vector de simulaciones para el grupo construcción
%       I_com, vector de simulaciones para el grupo comercio
%       I_serv, vector de simulaciones para el grupo servicios
%
% -------------------------------------------------------------------------

%calcular las xi
m = 4;
tam2 = ceil(tam/2);
Z0=randn;
Z=randn(1,m);
for k=1:m
    x=zeros(tam2(k),1);
    for l=1:tam2(k)
        eps=randn;
        x(l)=sqrt(rho_dif)*Z0+sqrt(rho_tilde(k)-rho_dif)*Z(k)+sqrt(1- rho_tilde(k))*eps;
    end
    x = [x; -x];
    x = x(1:tam(k));
    switch k
        case 1
            I_ind = x < u(k);
        case 2
            I_cons = x < u(k);
        case 3
            I_com = x < u(k);
        case 4
            I_serv = x < u(k);
    end
end
```

## Código 5

```
% Simula probabilidad de que L exceda 2.5 veces la media

L_vec = [];

for j = 1:100
for i = 1:max(n_sim)
    [ I_ind, I_cons, I_com, I_serv ] = individual(u,rho_tilde, tam, rho_dif);
    L_ind(i) = I_ind'*industrial;
    L_cons(i) = I_cons'*construccion;
    L_com(i) = I_com'*comercio;
    L_serv(i) = I_serv'*servicios;
```

```
end
K = [sum(I_ind);sum(I_cons);sum(I_com);sum(I_serv)];
L_ind = L_ind(1:n_sim(1));
L_cons = L_cons(1:n_sim(2));
L_com = L_com(1:n_sim(3));
L_serv = L_serv(1:n_sim(4));
L = sum(L_ind) + sum(L_cons) + sum(L_com) + sum(L_serv);
L_vec = [L_vec; L];
end
```

# Referencias


[1] Sheldon M. Ross. *Simulation*. Elsevier Academic Press, Amsterdam, Boston, 2006.