

Estimation of Probability of Defaults (PD) for Low-Default Portfolios: An Actuarial Approach

Cristian Pacheco Granados

17 de mayo de 2025

Resumen

Este estudio replica el marco de estimación de la probabilidad de incumplimiento (PD) de Iqbal & Ali (2012), empleando convoluciones de distribuciones binomial y de Poisson y regresiones exponenciales para asegurar PDs estrictamente crecientes. Las PDs iniciales se re-estiman utilizando el teorema de Bayes con las probabilidades empíricas del portafolio. Se realizan convoluciones, restringiendo la distribución binomial a un único incumplimiento para facilitar la re-estimación bayesiana y la regresión exponencial subsiguiente. Una segunda convolución evalúa la alineación de las PDs ajustadas con la distribución real de incumplimientos del portafolio en cada calificación. Las PDs finales se derivan de una combinación ponderada de PDs regresionadas exponencialmente y ratios de incumplimiento empíricos, refinadas a través de una regresión exponencial final, con el objetivo de lograr resultados comparables al artículo original a pesar de su metodología no divulgada para el cálculo de las PDs.

1. Introducción

En el panorama actual de la gestión de riesgos, las instituciones financieras de todo el mundo reconocen la necesidad crítica de una gestión mejorada del riesgo de crédito. Los marcos de Basilea II y III del Comité de Supervisión Bancaria de Basilea han impulsado a muchas instituciones a adoptar el enfoque de calificaciones internas (IRB), siendo la determinación precisa de la probabilidad de incumplimiento (PD) un componente central. Este artículo aborda el desafío de la estimación de PD, particularmente para portafolios con baja tasa de incumplimiento (LDP), donde los métodos tradicionales que se basan en las tasas de incumplimiento realizadas a largo plazo a menudo resultan inadecuados.

Iqbal & Ali (2012) proponen un mecanismo dinámico que emplea técnicas actuariales, específicamente la convolución, para derivar las PD implícitas mediante la combinación de distribuciones de probabilidad. Este estudio replica su marco utilizando Python para todos los análisis. Sin embargo, dada la falta de detalle explícito del artículo original con respecto a la metodología precisa para obtener las PD, nuestra implementación incorpora varias adaptaciones clave. Específicamente, utilizamos regresiones exponenciales para asegurar PDs estrictamente crecientes, y re-estimamos las PDs iniciales utilizando el teorema de Bayes con las probabilidades empíricas del portafolio. Para facilitar la re-estimación bayesiana, se realizan convoluciones donde la distribución binomial se restringe a un único incumplimiento. Se utilizan convoluciones adicionales para validar las PDs ajustadas con la distribución real de incumplimientos de la cartera. Finalmente, las PDs se calculan como un promedio ponderado de las PDs obtenidas mediante regresión exponencial y los ratios de incumplimiento empíricos, refinados con una regresión exponencial final, para lograr resultados comparables al artículo original.

2. Modelo

Nuestro modelo utilizará información sencilla del portafolio. Como se explicó en la sección anterior, el modelo solo emplea el número total de clientes y el número total de incumplimientos en cada calificación. Una de nuestras principales preocupaciones es utilizar el peso del incumplimiento de cada calificación dentro del portafolio de incumplimientos, lo cual se obtendrá simplemente aplicando el teorema de Bayes. Esto producirá la probabilidad de incumplimiento en cada calificación para el próximo cliente que forme parte del portafolio. A partir del siguiente portafolio, podemos descubrir la estimación bayesiana.

Grade	Number of Obligors	Number of Defaults
AAA	34	1
AA	56	1
A	119	3
BBB	257	2
BB	191	2
B	102	6
CCC	50	3
CC	34	1
C	12	2
Sum of	855	21

Cuadro 1: Table 1.1

Teorema de Bayes:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$P(A|B)$ = Probabilidad de estar en la calificación i dado que ha ocurrido un incumplimiento

$P(B|A)$ = Probabilidad de incumplimiento dado que el cliente está en calificación i

$P(A)$ = Probabilidad de un cliente de estar en una calificación específica

$P(B)$ = Probabilidad de un Default

Despejando Queda:

$$P(A|B) = \frac{\frac{ND_i}{NO_i} \cdot \frac{NO_i}{855}}{\frac{21}{855}} = \frac{ND_i}{21}$$

Grade	Number of Obligors	Number of Defaults	Bayesian Estimates
AAA	34	1	4.76 %
AA	56	1	4.76 %
A	119	3	14.29 %
BBB	257	2	9.52 %
BB	191	2	9.52 %
B	102	6	28.57 %
CCC	50	3	14.29 %
CC	34	1	4.76 %
C	12	2	9.52 %
Sum of	855	21	

Cuadro 2: Table 1.2

Con la función de densidad de probabilidad de la distribución binomial definida:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

n = número total de defaults en el portafolio.
k = número de defaults en una calificación específica.
p = probabilidad estimada mediante el teorema de Bayes.

nos proporciona las herramientas para modelar la probabilidad de cómo se distribuyen los 21 defaults totales de nuestro portafolio entre las diferentes calificaciones crediticias. Al considerar cada calificación como un grupo de ensayos (donde un 'éxito' es un default), y conociendo la probabilidad de default individual para pertenecer a cada calificación (p), podemos utilizar esta fórmula para calcular la probabilidad de observar una cantidad específica de defaults (k) dentro de cada grupo, contribuyendo así a la distribución total de los 21 defaults en el portafolio como se ve en la tabla.

k	AAA	AA	A	BBB	BB	B	CCC	CC	C
0	0.35894	0.35894	0.03927	0.12224	0.12224	0.00085	0.03927	0.35894	0.12224
1	0.37688	0.37688	0.13746	0.27021	0.27021	0.00717	0.13746	0.37688	0.27021
2	0.18844	0.18844	0.22910	0.28444	0.28444	0.02868	0.22910	0.18844	0.28444
3	0.05967	0.05967	0.24183	0.18962	0.18962	0.07266	0.24183	0.05967	0.18962
4	0.01342	0.01342	0.18137	0.08982	0.08982	0.13080	0.18137	0.01342	0.08982
5	0.00228	0.00228	0.10277	0.03214	0.03214	0.17789	0.10277	0.00228	0.03214
6	0.00030	0.00030	0.04568	0.00902	0.00902	0.18975	0.04568	0.00030	0.00902
7	0.00003	0.00003	0.01631	0.00203	0.00203	0.16264	0.01631	0.00003	0.00203
8	0.00000	0.00000	0.00475	0.00037	0.00037	0.11385	0.00475	0.00000	0.00037
9	0.00000	0.00000	0.00114	0.00005	0.00005	0.06578	0.00114	0.00000	0.00005
10	0.00000	0.00000	0.00022	0.00000	0.00000	0.03157	0.00022	0.00000	0.00000

Cuadro 3: Probabilidades Binomiales por Calificación Crediticia

En los cálculos previos de probabilidades de default, nos enfocamos en la proporción de defaults observados en el portafolio total, sin distinguir entre las diferentes calificaciones crediticias ni la cantidad de clientes en cada una. Ahora, vamos a refinar nuestro análisis incorporando la composición del portafolio por calificación y las tasas de default inherentes a cada grupo. Para modelar la distribución de la frecuencia de estos eventos de default, la distribución de Poisson se presenta como la herramienta estadística más apropiada con función

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

λ = frecuencia de default en cada calificación.
x = número de defaults incrementales en la calificación específica.

n	AAA	AA	A	BBB	BB	B	CCC	CC	C
0	0.97102	0.98230	0.97511	0.99225	0.98958	0.94287	0.94176	0.97102	0.84648
1	0.02856	0.01754	0.02458	0.00772	0.01036	0.05546	0.05651	0.02856	0.14108
2	0.00042	0.00016	0.00031	0.00003	0.00005	0.00163	0.00170	0.00042	0.01176
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00003	0.00003	0.00000	0.00065
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00003
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Cuadro 4: Probabilidades de Poisson por Calificación Crediticia

3. Metodología para el Cálculo de Probabilidades de Incumplimiento (PDs)

1. Ajuste de Probabilidades Iniciales:

- Debido a que los pesos de cada calificación se incorporan en las probabilidades de la distribución de Poisson (λ), se considera un cliente por cada λ de incumplimiento.
- En la aplicación del *teorema de Bayes*:
 - Se modifica la probabilidad previa $P(A)$ (probabilidad de pertenecer a una calificación) a una distribución equitativa, ya que la influencia de los pesos de cada calificación se considera mediante las probabilidades de Poisson (λ).
 - Se ajustan dos probabilidades de $P(B)$ (probabilidad general de incumplimiento) al promedio de los λ de cada calificación, para representar mejor el riesgo global del portafolio, de igual forma usamos la probabilidad general de incumplimiento del portafolio.

$$P(B)_1 = \frac{\sum \lambda_i}{n} = 0,04507$$

$$P(B)_2 = \frac{\text{total incumplimientos}}{\text{total clientes}} = ,02456$$

2. Convolución Binomial-Poisson con Incumplimiento Unitario:

- Se realiza la convolución de las distribuciones binomial y Poisson, fijando el número de incumplimientos en la distribución binomial a 1.
- El resultado de esta convolución representa la distribución de probabilidad de que ese único incumplimiento ocurra en cada una de las calificaciones, lo que equivale a $P(A|B)$ en Bayes.

Calificación	Probabilidad
AAA	0,14194273565254403
AA	0,10767392690454687
A	0,20729745215059067
BBB	0,1196197572455968
BB	0,1276727609684301
B	0,35864093782353706
CCC	0,2741959972860557
CC	0,14194273565254403
C	0,3585770355148263

Cuadro 5: Nuevas $P(A|B)$

- La distribución de probabilidad generada por la convolución refleja la influencia de los λ de Poisson, que modelan el comportamiento de incumplimiento específico de cada calificación.

3. Actualización de PDs mediante el Teorema de Bayes:

- Se actualizan las probabilidades de incumplimiento ($PDs = P(B|A)$) utilizando el *teorema de Bayes*:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

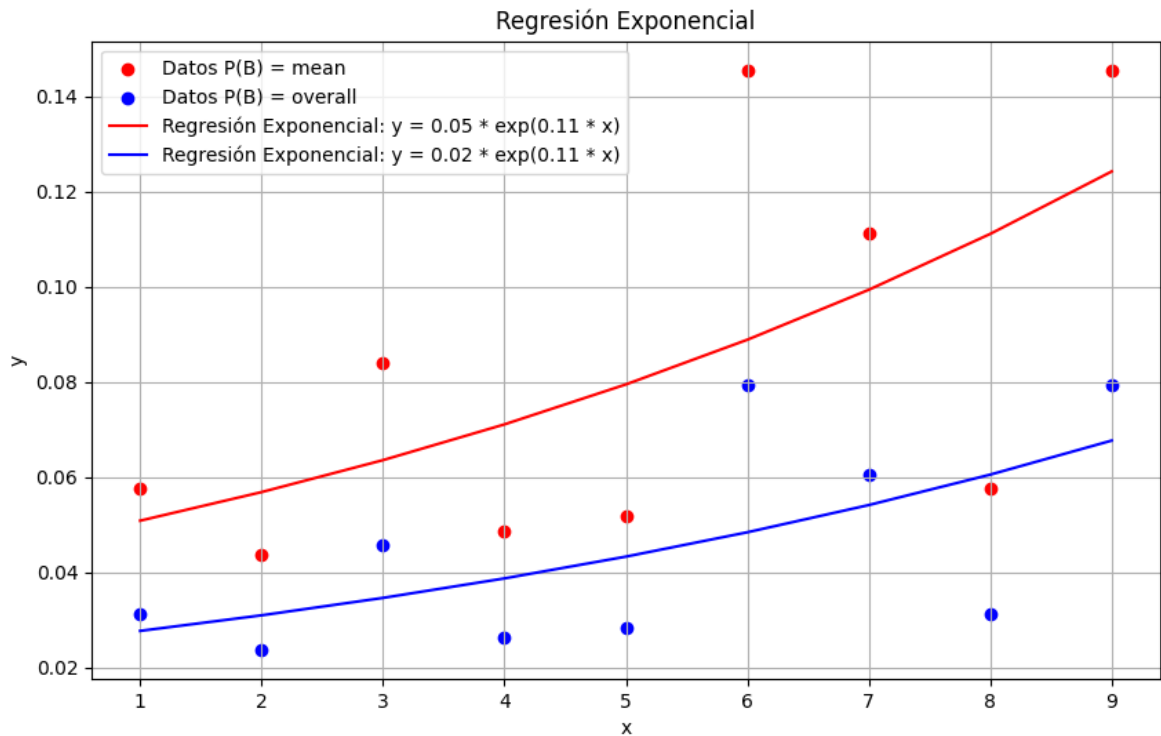
- Esta actualización incorpora la información sobre la distribución de incumplimientos capturada por la convolución.

new_pds_2	new_pds
0,031376815249509735	0,05757683621754331
0,02380160489468931	0,04367623341752687
0,045823647317498994	0,08408694813382414
0,026442262127974032	0,04852187144096779
0,028222399793021397	0,05178844562863056
0,07927852309783452	0,1454770506080621
0,060611746768496524	0,11122328983351691
0,031376815249509735	0,05757683621754331
0,07926439732433002	0,14545112964250054

Cuadro 6: Comparación $P(B|A)$ con $P(B)_1$ y $P(B)_2$

4. Suavizado de PDs Actualizadas:

- Las PDs actualizadas se suavizan mediante una regresión exponencial.
- Estas PDs suavizadas se guardan como "PDs exponenciales".



5. Convolución Binomial-Poisson con Incumplimientos Totales:

- Se realiza una nueva convolución de las distribuciones binomial y Poisson, esta vez utilizando el número total de incumplimientos reales del portafolio (21).
- Se extrae la probabilidad correspondiente al número real de incumplimientos observado en cada calificación a partir del resultado de esta convolución.
- Estas probabilidades se interpretan como "pesos" que indican la relevancia de las PDs exponenciales en comparación con las PDs empíricas, que tan bien se ajustan las probabilidades exponenciales a la realidad de nuestro portafolio.

Calificación	Proba ($P(B)_1$)	Proba ($P(B)_2$)	ND
aaa_conv	0,37555	0.37626	1
aa_conv	0,37533	0.37617	1
a_conv	0,24086	0.24134	3
bbb_conv	0,28310	0.28379	2
bb_conv	0,28290	0.28370	2
b_conv	0,18856	0.18913	6
ccc_conv	0,24019	0.24103	3
cc_conv	0,37294	0.37520	1
c_conv	0,28169	0.28318	2

Cuadro 7: Probabilidades en convolución por calificación

6. Combinación Lineal de PDs:

- Se calcula una combinación lineal de las PDs exponenciales y las PDs empíricas de cada calificación, utilizando los "pesos" obtenidos en el paso anterior.

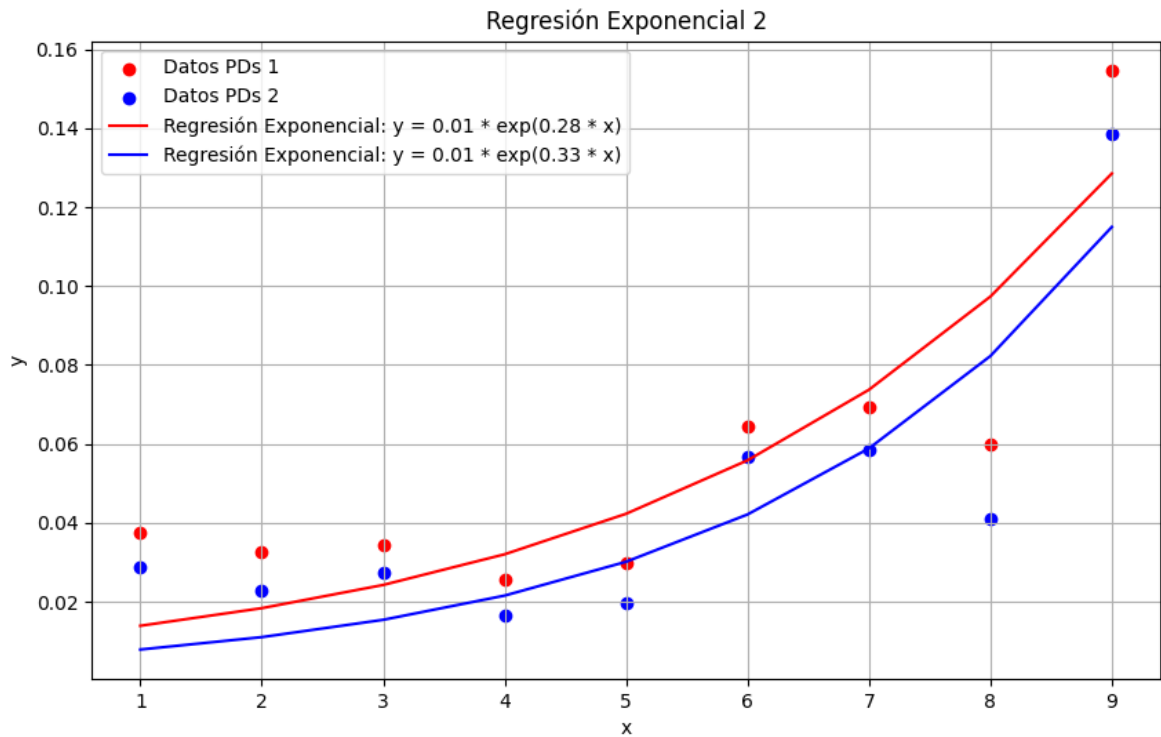
$$PDs = \alpha \cdot \text{new_pds} + (1 - \alpha) \cdot \text{original_lambda}$$

Calificación	PDs con $P(B)_1$	PDs con $P(B)_2$
AAA	0,037466	0,028773
AA	0,032497	0,022796
A	0,034451	0,027487
BBB	0,025702	0,016567
BB	0,029992	0,019788
B	0,064487	0,056857
CCC	0,069451	0,058588
CC	0,059869	0,041089
C	0,154702	0,138636

Cuadro 8: Weigth PDs

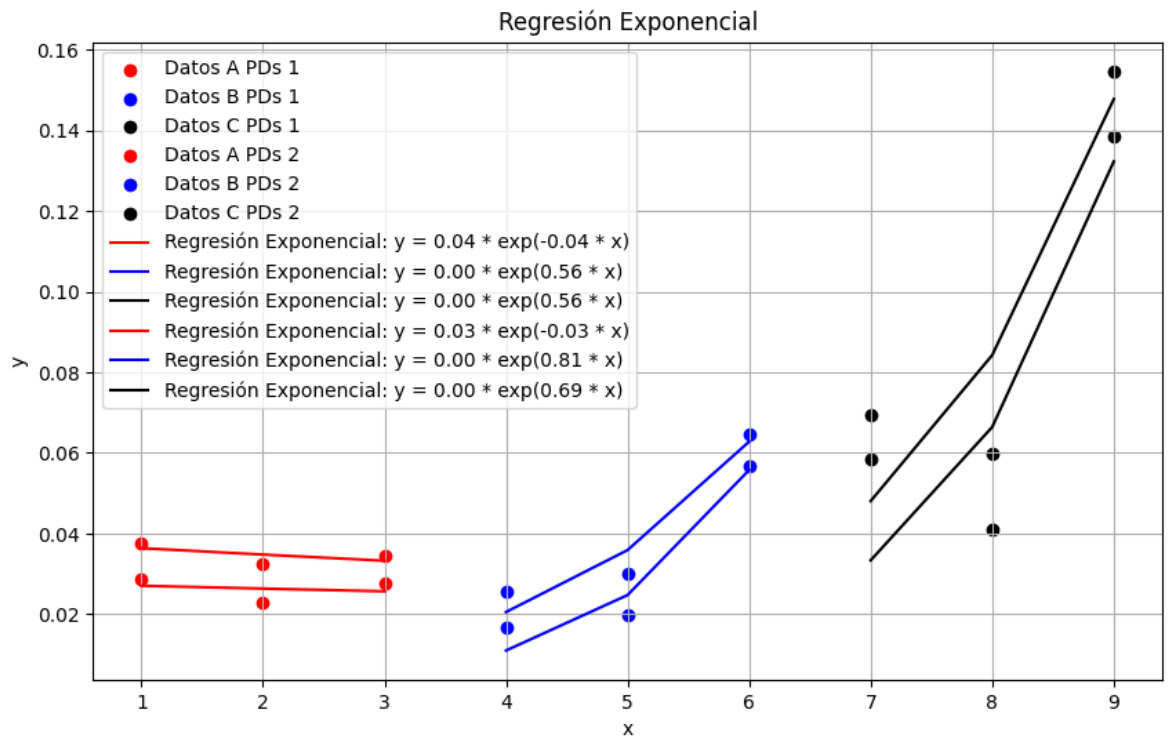
7. Regresión Exponencial Final:

- Se aplica una regresión exponencial final a las PDs resultantes de la combinación lineal para asegurar que sean *estrictamente crecientes*.



8. Ajuste por Correlación entre Calificaciones:

- Se realizan tres regresiones exponenciales separadas para las calificaciones agrupadas como A, B y C, con el fin de modelar la correlación entre calificaciones cercanas.



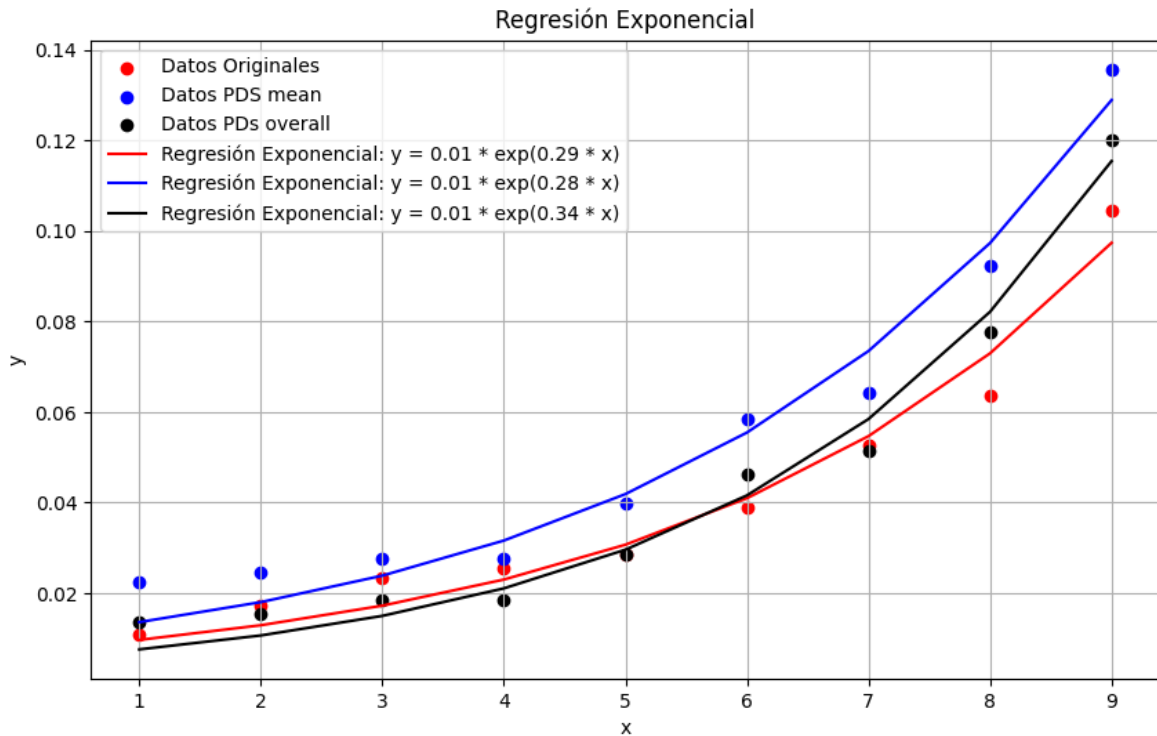
- Se asignan nuevos pesos para combinar las PDs de las calificaciones dentro de cada grupo y con las PDs exponenciales anteriores. con resultados en la tabla

Calificación	PDs 1 Finales	PDs 2 Finales	PDs Paper Original
AAA	2.231 %	1.355 %	1.08 %
AA	2.450 %	1.555 %	1.74 %
A	2.762 %	1.842 %	2.33 %
BBB	2.773 %	1.846 %	2.55 %
BB	3.990 %	2.855 %	2.85 %
B	5.843 %	4.611 %	3.90 %
CCC	6.413 %	5.139 %	5.28 %
CC	9.244 %	7.762 %	6.36 %
C	13.569 %	12.003 %	10.46 %

Cuadro 9: Comparación de PDs Finales

9. Resultados y Validación:

- La metodología produce PDs que son similares a las del artículo original, aunque no idénticas debido a la falta de especificidad en el artículo sobre su método de cálculo exacto.
- Se considera que las PDs resultantes se ajustan adecuadamente a las características del portafolio analizado.



4. Justificación

1. Elección de Probabilidades Iniciales para la Reestimación Bayesiana

- $P(A)$ (Probabilidad de estar en la calificación i): Se modifica a una distribución equitativa (uniforme).
- $P(B)$ (Probabilidad general de incumplimiento): Se ajusta al promedio de los λ (lambdas) de cada calificación.

Justificación:

$P(A)$ (Equitativa):

- **Problema con $P(A)$ empírica inicial:** Si se usan las proporciones empíricas de clientes en cada calificación como en $P(A)$ inicial, las calificaciones con mayor número de clientes dominarían la actualización bayesiana, incluso si su comportamiento de incumplimiento no es tan informativo. Esto oscurece la información valiosa de las calificaciones con menos clientes pero con patrones de incumplimiento más fuertes.
- **Ventaja de la distribución equitativa:** Al establecer $P(A)$ como uniforme, le das a cada calificación la misma voz en la actualización bayesiana. El foco se desplaza a la información contenida en los incumplimientos, es decir, en cómo se distribuyen los incumplimientos dado que ocurren ($P(A|B)$). Esto es especialmente importante en LDPs, donde cada incumplimiento es una pieza de información valiosa.
- **Relación con las λ :** Dado que las intensidades de incumplimiento (λ) de Poisson ya capturan las diferencias en el riesgo base de cada calificación, no es necesario que $P(A)$ también lo haga inicialmente.

$P(B)$ (Promedio de λ):

- **Limitación de la tasa de incumplimiento global:** Usar la tasa de incumplimiento global del portafolio como $P(B)$ puede ser demasiado simplista. No refleja las variaciones en la intensidad de riesgo entre las calificaciones.
- **Beneficio del promedio de λ :** El promedio de las λ proporciona una estimación más ponderada del riesgo general del portafolio, ya que cada λ representa la intensidad de incumplimiento esperada para una calificación específica. Esto le da más peso a las calificaciones con λ más altas, que contribuyen más al riesgo general.

2. Convolución con un Único Incumplimiento para Obtener $P(A|B)$

- Se Realiza una convolución entre la distribución binomial y la de Poisson, pero con una modificación clave: fijar el número de incumplimientos en la distribución binomial a 1.
- Utilizar el resultado de esta convolución como una estimación de $P(A|B)$.

Justificación:

El Propósito de $P(A|B)$:

- $P(A|B)$ representa la probabilidad de que un incumplimiento pertenezca a una calificación específica (A), dado que ha ocurrido un incumplimiento (B) en el portafolio.
- Esencialmente, nos dice cómo se distribuye la probabilidad de un evento de incumplimiento entre las diferentes calificaciones.

Limitaciones de los Enfoques Tradicionales:

- Si simplemente observamos las tasas de incumplimiento históricas de cada calificación, no capturamos completamente la interdependencia entre ellas ni cómo contribuye cada calificación al riesgo general de incumplimiento del portafolio.

La Innovación de la Convolución con un Único Incumplimiento:

- Al forzar la distribución binomial a tener un solo incumplimiento, estamos diseñando un experimento mental: "Si ocurriera *un solo* incumplimiento en el portafolio, ¿cuál es la probabilidad de que provenga de cada calificación?".
- Esto permite aislar la contribución relativa de cada calificación al riesgo de incumplimiento, controlando por el hecho de que ha ocurrido un incumplimiento.

Cómo la Convolución Captura la Sofisticación:

- La convolución combina la información de dos distribuciones de probabilidad:
 - La distribución de Poisson (con sus λ) captura la frecuencia esperada de incumplimientos para cada calificación, reflejando su riesgo inherente.
 - La distribución binomial (con $n = 1$) actúa como un interruptor que te permite observar la probabilidad de que ese único incumplimiento se active en cada calificación.
- El resultado de la convolución es una nueva distribución de probabilidad que representa $P(A|B)$, donde cada valor refleja la probabilidad de que el único incumplimiento provenga de una calificación específica, teniendo en cuenta tanto su frecuencia de incumplimiento esperada (λ) como su contribución relativa al riesgo general.
- Esto es más sofisticado que simplemente usar las tasas de incumplimiento históricas porque:
 - Incorpora la estructura de riesgo del portafolio a través de las λ de Poisson.
 - Proporciona una estimación de $P(A|B)$ que está condicionada a la ocurrencia de un incumplimiento, lo cual es fundamental para la actualización bayesiana.

3. Regresión Exponencial para Suavizar las PDs y Asegurar Crecimiento Estricto

- Aplicar varias y diferentes regresiones exponenciales a las PDs obtenidas después de la actualización bayesiana.

Justificación

Necesidad de Suavizado:

- Las PDs obtenidas directamente de la actualización bayesiana pueden ser irregulares o ruidosas, especialmente en portafolios con baja tasa de incumplimiento. Esto se debe a la volatilidad que puede introducir la pequeña cantidad de datos de incumplimiento.
- El suavizado es necesario para obtener una curva de PDs más estable y generalizable, que no esté demasiado influenciada por fluctuaciones aleatorias en los datos.

Imposición de Crecimiento Estricto:

- Un principio fundamental en la modelización del riesgo de crédito es que las PDs deben ser estrictamente crecientes con respecto al deterioro de la calidad crediticia.
- Esto refleja la lógica de que los prestatarios con peor calificación son inherentemente más propensos a incumplir.
- La regresión exponencial garantiza este crecimiento estricto, ya que la función exponencial es inherentemente creciente.

Ajuste a la Forma Exponencial Común de las PDs:

- Empíricamente, se observa que las curvas de PDs tienden a tener una forma aproximadamente exponencial.
- La regresión exponencial captura esta relación no lineal y proporciona un buen ajuste a la forma típica de las curvas de PDs.

4. Justificación de las Combinaciones Lineales de PDs

- Combina linealmente las PDs exponenciales (suavizadas y modeladas) con las PDs empíricas (observadas en los datos) y luego ajusta por correlación entre calificaciones.

Justificación

El Desafío de la Estimación de PDs:

- La estimación de PDs es un equilibrio entre el modelado para capturar relaciones subyacentes y suavizar el ruido y la realidad empírica para asegurar que el modelo se ajuste a los datos específicos.
- Ningún enfoque único es perfecto. Los modelos pueden ser demasiado simplificados o no capturar todas las complejidades, mientras que los datos empíricos pueden ser ruidosos o insuficientes, especialmente en portafolios de baja tasa de incumplimiento.

Combinación de PDs Exponenciales y PDs Empíricas:

- **PDs Exponenciales (Modeladas):**
 - Capturan la forma general esperada de la curva de PDs (crecimiento exponencial).
 - Pueden no ajustarse perfectamente a los detalles específicos del portafolio.
- **PDs Empíricas (Observadas):**
 - Reflejan la experiencia real de incumplimiento en el portafolio.
 - Pueden ser ruidosas o volátiles, especialmente con pocos datos.
- **Combinación Lineal:**
 - La combinación lineal (promedio ponderado) permite asignar pesos a las PDs exponenciales y empíricas, dando más importancia a la fuente de información que se considera más confiable.
 - Los pesos se determinan objetivamente a partir de la convolución con el número total de incumplimientos, lo que proporciona una base sólida para la combinación.
 - Esto logra un equilibrio entre la estabilidad del modelo y la relevancia empírica.

Ajuste por Correlación entre Calificaciones:

- **Limitación de las PDs Independientes:**
 - Las PDs calculadas hasta ahora asumen implícitamente que los incumplimientos de diferentes calificaciones son independientes.
 - En realidad, las calificaciones de crédito tienden a estar correlacionadas, especialmente las calificaciones adyacentes o dentro del mismo grupo (por ejemplo, A, B, C).
- **Modelado de Correlación:**
 - Agrupar las calificaciones y ajustar las PDs dentro de cada grupo permite capturar la dependencia entre ellas.
 - La regresión exponencial dentro de los grupos fuerza a las PDs a moverse de manera más coherente, reflejando la correlación, proporcionando una estimación de riesgo más precisa y realista.

Información Total del Portafolio:

- En resumen, el uso de combinaciones lineales y ajustes por correlación tiene como objetivo capturar la mayor cantidad posible de información relevante del portafolio:
 - Tendencias generales de riesgo (exponencialidad).
 - Experiencia empírica específica.
 - Dependencias entre calificaciones.
- Esto conduce a una estimación de PDs más completa y precisa, que es esencial para una gestión eficaz del riesgo.

5. Conclusión

El modelo propuesto ofrece una metodología robusta para estimar las probabilidades de incumplimiento (PDs), particularmente adecuada para portafolios de baja tasa de incumplimiento donde los métodos tradicionales a menudo tienen dificultades. Al integrar técnicas actuariales (convoluciones), principios bayesianos y modelado estadístico (regresiones), el modelo aborda eficazmente los desafíos de los datos limitados y el posible ruido. Las características clave incluyen el uso de convoluciones de incumplimiento único para refinar las estimaciones de PD, la regresión exponencial para garantizar PDs estrictamente crecientes y una combinación de datos empíricos y basados en modelos, mejorada por ajustes de correlación, para capturar una visión integral del riesgo del portafolio. Este enfoque multifacético minimiza la pérdida de información y evita la subestimación del riesgo, proporcionando una evaluación más confiable y precisa del riesgo crediticio en entornos de baja tasa de incumplimiento.

Referenciass

Iqbal, Nabil, and Syed Afraz Ali. .^Estimation of Probability of Defaults (PD) for Low-Default Portfolios: An Actuarial Approach.”2012 Enterprise Risk Management Symposium. April 18-20, 2012. <https://www.academia.edu/download/90701271/mono-2012-as12-1-iqbal.pdf>