Sarriko-On

Estadística actuarial: modelos estocásticos

ISBN: 978-84-693-1205-6

Eva Ferreira
María Araceli Garín

03-10

Facultad de Ciencias Económicas y Empresariales
Ekonomia eta Enpresal-Zientzien Fakultatea

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Estadística Actuarial: Modelos Estocásticos

Eva Ferreira y María Araceli Garín

Departamento de Economía Aplicada III
Facultad de Ciencias Económicas y Empresariales
Bilbao

Fichero .pdf disponible en http://ekasi.ehu.es

Prólogo

Este documento constituye una redacción y adaptación de las notas de clase, apuntes y ejercicios, elaborada por las autoras y profesoras a partir de distintos textos y que han servido de guía para la docencia de la asignatura, desde sus comienzos en el curso 97/98.

El objetivo del mismo ha sido poner a disposición de los alumnos una herramienta adicional, además de la bibliografía recomendada en el programa. Muchos de los conceptos y ejercicios que se presentan han sido tomados de la literatura; donde ello ha sido posible, hemos indicado la fuente. Sin embargo, es posible que la fuente citada no sea original.

Son destinatarios de estas notas de clase, los alumnos de la asignatura Estadística Actuarial: modelos estocásticos, actualmente impartida dentro de la Licenciatura en Ciencias Actuariales y Financieras, futuro grado en Finanzas y Seguros.

El contenido se estructura en cinco capítulos. El primero contiene una colección de ejercicios propuestos, con la idea de revisar mediante su resolución, una serie de conceptos básicos del cálculo de probabilidades, que serán necesarios en el resto de los temas. Los restantes capítulos contienen una serie de conceptos teóricos en relación a la aplicación de la Teoría de la Probabilidad al estudio de las situaciones de azar en las Ciencias Actuariales. Finaliza cada capítulo con una colección de ejercicios propuestos.

Índice general

Capítulo 1

Introducción. Probabilidad en los modelos estocásticos actuariales

Se describe a continuación la Tarea 1, en la que se enumeran un conjunto de ejercicios de ensayo de las competencias correspondientes a este tema; en particular a la primera de las formuladas en el programa de la asignatura:

> Identificar los elementos matemáticos que caracterizan a las distribuciones de probabilidad tanto discretas como continuas, habituales en el ámbito de la empresa aseguradora: función de probabilidad, función de distribución, momentos, funciones características, etc, operar correctamente con ellos y conocer sus propiedades.

1.1. Tarea 1

1. Consideremos una moneda trucada, de manera que la probabilidad de cara $P(\text{cara}) = 0{,}3$, y por lo tanto $P(\text{cruz}) = 0{,}7$. Calcula la probabilidad de obtener 3 caras en 5 lanzamientos al aire de dicha moneda. ¿Porqué no puedes utilizar la fórmula de Laplace?

2. Una compañía de seguros de automóviles clasifica a sus asegurados en cuatro grupos de edad. La siguiente tabla recoge la proporción de asegurados dentro de cada grupo de edad, junto con la probabilidad de tener un accidente.

Grupo de edad	Proporción de asegurados	Prob. de accidente
18-25	0.10	0.07
25-45	0.40	0.04
45-60	0.30	0.02
+60	0.20	0.05

Se elige un asegurado al azar de la compañía:

a) Probabilidad de que tenga un accidente.

b) Si sabemos que el asegurado ha tenido un accidente, obtener la probabilidad de que pertenezca a cada uno de los grupos.

3. Una compañía de seguros de automóviles tiene clasificados a sus asegurados en dos grupos de edad. En el grupo de los más jóvenes, J, están el 30 % de los clientes, mientras que el 70 % restante se encuentra en el grupo S. Los contratos con la compañía tienen una vigencia anual. La probabilidad de que un asegurado del grupo J tenga un siniestro es del 75 %, mientras que esa probabilidad para un asegurado en el grupo S se reduce a un 32 %. La probabilidad de que un cliente de la compañía tenga un primer siniestro es independiente de que tenga un segundo o un tercer siniestro a lo largo del año, tanto para el grupo J como para el grupo S. Si elegimos a un cliente al azar de la compañía, calcular:

   a) Probabilidad de que tenga exáctamente un siniestro.

   b) Si el asegurado ha tenido un siniestro, calcular la probabilidad de que pertenezca al grupo J.

   c) Suponiendo que el asegurado elegido al azar ha dado cuenta de un siniestro, calcular la probabilidad de que tenga un segundo siniestro antes de finalizar el año de contrato.¡Cuidado!, el resultado de esta cuestión no es evidente, el ser independiente la ocurrencia de posteriores siniestros dado el primero, en cada grupo J ó S no hace que lo sea para un individuo elegido al azar del total.

   d) ¿ Cómo es $P(A_2/A_1)$=probabilidad de un segundo accidente dado un primero, en relación a $P(A_1)$= probabilidad de un primer accidente? Interpreta la relación de desigualdad entre ambas probabilidades.

4. De acuerdo con un reciente estudio de mercado, el 40 % de los varones mayores de 30 años poseen su propio coche y su propia casa, el 60 % su propia casa y el 70 % su propio coche.

   a) Calcula la probabilidad de que un varón mayor de 30 años posea al menos casa propia o coche propio.

   b) Calcula la probabilidad de que no posea ninguna de las dos cosas.

   c) Si escogemos dos varones al azar e independientemente, calcula la probabilidad de que ninguno de los dos posea casa ni coche propios.

   d) Si sabemos que un determinado varón dispone de coche propio, calcula la probabilidad de que también posea casa propia.

5. En una Comunidad Autónoma el 60 % de los hogares está asegurado contra incendios. Con objeto de estudiar con detalle la situación de una determinada provincia, se seleccionan al azar e independientemente 10 hogares. Calcular:

   a) Probabilidad de que alguno de los hogares esté asegurado contra incendios.

   b) Probabilidad de que al menos dos hogares estén asegurados contra incendios.

   c) Se define la v.a. X como el número de hogares entre los 10, que están asegurados contra incendios. Observa que la v.a. X puede tomar los valores $\{0, 1, ..., 10\}$. Obten la función de cuantía de X.

   d) Obtén el número esperado de hogares que tienen seguro contra incendios.

6. De un total de 9872 pólizas de seguros de automóvil contratadas en una compañía de seguros, 1083 presentaron reclamaciones durante el año 2004. Disponemos de la tabla de

datos que nos da el número de reclamaciones efectuadas X, así como la cuantía total de las reclamaciones Y (en euros).

X Y	< 1200	≥ 1200
1	937	28
2 ó más	103	15

Si escogemos una póliza al azar entre las 9872, calcula las siguientes probabilidades:

a) Que haya efectuado una única reclamación.

b) Si se sabe que ha reclamado, la probabilidad de que la cuantía de dicha reclamación haya sido igual o superior a 1200 euros.

c) Que haya efectuado más de una reclamación con una cuantía total igual o superior a 1200 euros.

d) ¿Son independientes el número de reclamaciones y la cuantía total?

7. Se sabe que el coste total (en cientos de euros) de un determinado tipo de reclamación es una variable aleatoria con función de distribución $F(y) = 1 - e^{-0{,}16y}, y \geq 0$.

   a) Obtén la función de densidad de la v.a. Y.

   b) Calcula la probabilidad de que el valor de una reclamación esté comprendido entre 500 y 1000 euros.

   c) Calcula el valor esperado de una reclamación de este tipo.

   d) Si se realizan 5 reclamaciones de forma independiente. Calcula la probabilidad de que ninguna de las cinco supere los 1000 euros. ¿Podrías calcular la probabilidad de que la suma de ellas no supere los 5000 euros? ¿Cuál es la dificultad?

8. Se sabe que el coste total (en cientos de euros) de un determinado tipo de reclamación es una variable aleatoria, Y, con función de distribución $N(m = 6{,}25, \sigma^2 = 39{,}06)$.

   a) Obtén la función de densidad de la v.a. Y.

   b) Calcula la probabilidad de que el valor de una reclamación esté comprendido entre 500 y 1000 euros.

   c) Calcula el valor esperado de una reclamación de este tipo.

   d) Si se realizan 5 reclamaciones de forma independiente. Calcula la probabilidad de que ninguna de las cinco supere los 1000 euros. ¿Podrías calcular la probabilidad de que la suma de ellas no supere los 5000 euros?

9. Se sabe que el coste total (en euros) de un determinado tipo de reclamación es una variable aleatoria, Y, con distribución $U(a, b)$ de media 6.25 y varianza 39.06.

   a) Obtén la función de densidad de la v.a. Y.

   b) Calcula la probabilidad de que el valor de una reclamación esté comprendido entre 500 y 1000 euros.

   c) Calcula el valor esperado de una reclamación de este tipo.

d) Si se realizan 5 reclamaciones de forma independiente. Calcula la probabilidad de que ninguna de las cinco supere los 1000 euros. ¿Podrías calcular la probabilidad de que la suma de ellas no supere los 5000 euros?

10. Se sabe que el número anual de siniestros en automóvil correspondientes a las pólizas de una determinada compañía, X, sigue una distribución binomial de parámetros $p = 0{,}005$ y $n = 2000$.

    a) Obtén la función de cuantía de X. Calcula las siguientes probabilidades: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$,$P(X = 3)$, $P(X = 4)$ y $P(X = 5)$.

    b) Calcula la probabilidad de que $X \in [5, 15]$, ¿cuál es la dificultad?

11. Se sabe que el número anual de siniestros en automóvil correspondientes a las pólizas de una determinada compañía, X, sigue una distribución de Poisson $P(\lambda = 10)$.

    a) Obtén la función de cuantía de X. Calcula las siguientes probabilidades: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$,$P(X = 3)$, $P(X = 4)$ y $P(X = 5)$.

    b) Calcula la probabilidad de que $X \in [5, 15]$.

12. En Irak se encuentran algunas de las terminales petrolíferas más grandes del mundo. Una compañía se dedica a asegurar buques. Su cartera incluye medio millón de toneladas aseguradas de un número elevado de petroleros que viajan fundamentalmente a esa zona del Golfo Pérsico. La valoración de la situación es la siguiente: En época de conflicto, la probabilidad de que se produzca un bombardeo es del 50 %. La probabilidad de que un buque sea atacado si se produce un bombardeo es de $\frac{1}{3}$, mientras que si no se produce tal bombardeo, la probabilidad de que un buque sea atacado es 0.

    La probabilidad de hundimiento de un buque si es atacado es de $\frac{8}{10}$. Mientras que la probabilidad de que un buque no padezca daño alguno si es atacado es de $\frac{2}{10}$. Se prescinde por simplicidad, de considerar otras posibilidades (pérdida parcial del buque, etc.). En caso de hundimiento, la indemnización a pagar por parte de la compañía aseguradora asciende a \$1300 dólares por tonelada asegurada.

    a) En estas condiciones y en época de conflicto, calcula:

       1) La probabilidad de que un buque sea atacado.
       2) La probabilidad de que un buque sea hundido.
       3) La probabilidad, si se produce un bombardeo, de que un buque sea atacado.

    b) Si la prima de seguro es de \$150 por tonelada, calcula el rendimiento (primas cobradas − indemnizaciones pagadas) esperado de la compañía sobre la cartera mencionada de medio millón de toneladas aseguradas.

    c) Prescindiendo de todos los gastos inherentes al negocio asegurador, ¿qué prima debería fijar la compañía (en \$ por tonelada) para esperar beneficios nulos, antes de saber si se producirá el bombardeo?

    d) Calcula el mismo rendimiento esperado, después de saber que habrá bombardeo.

13. Una compañía de seguros conoce la distribución que sigue el coste de un determinado siniestro (en cientos de euros). Dicho coste se distribuye según la función de densidad:

$$f(y) = \frac{k}{(1+y)^3}, \quad y \geq 0$$

y $f(y) = 0$ si $y < 0$.

   a) Hallar el valor de k.

   b) Obtener la función de distribución. Calcular la probabilidad de que el coste de un siniestro sea superior a 100 euros.

   c) Determinar los cuartiles de la distribución del coste del siniestro. Interpretar su obtención.

   d) Calcular el coste medio por siniestro.

14. A una reunión internacional de Estadística Actuarial acuden n personas, de las cuales n_1 hablan exclusivamente inglés, n_2 hablan exclusivamente francés y el resto hablan los dos idiomas. Si se eligen dos congresistas al azar, calcula la probabilidad de que se entiendan.

Ahora que ya lo sabes hacer...Observa y recuerda:

1. *A lo largo de la asignatura, verás que cualquier compañía de seguros ha de tener completa información del comportamiento de la siniestralidad entre sus clientes. Esta información aparecerá descrita en tablas de datos para variables similares a las introducidas en el ejercicio 5, X : número de siniestros de sus clientes, e $Y|X$ cuantía por reclamación condicionada a la ocurrencia de X siniestros. Lógicamente X será una v.a. discreta, similar a las descritas en el Capítulo 2 de este texto, mientras que $Y|X$ será una v.a. continua, cuyo comportamiento posible será estudiado en el Capitulo 3. A partir de ambas variables, nuestra labor será procesar la información y obtener la que denominaremos "cuantía total por reclamaciones", que será una variable aleatoria nueva, cuya distribución será la distribución "incondicionada" de la variable Y.*

2. *Va a ser muy habitual tener que calcular probabilidades para sumas de v.a. independientes, por ejemplo a la hora de manejar cuantías totales por reclamación de varios siniestros. En ocasiones, cualquiera de las funciones características nos identifica la nueva distribución como una conocida y manejable. En otras, esto no es así, como en el ejercicio de la cuantía uniforme, y habrá que definir la función de distribución de la suma de v.a. independientes.*

Capítulo 2

Modelos de distribuciones de probabilidad discretas. Modelos relacionados con el número de siniestros

En el contexto de la *Estadística Actuarial* entenderemos por *siniestro* a cada realización de un *experimento aleatorio*, ésto es un fenómeno que provoca un riesgo. El individuo afectado, pretenderá cubrirse de las consecuencias económicas de la ocurrencia del siniestro. El *método* que da esta cobertura, se conoce habitualmente como *seguro*. El vendedor del seguro (y deseablemente también el tomador), ha de estudiar bien este fenómeno aleatorio, en términos que poder estimar el precio del contrato, en particular, la *prima* que supone este riesgo.

Así, en el estudio de estos fenómenos aleatorios, serán importantes dos variables aleatorias, a saber: el *número de siniestros* y la *cuantía por reclamación*. El comportamiento de la primera nos dará idea de la frecuencia con la que se producen dichos siniestros. La segunda, nos ayuda a valorar en términos económicos la ocurrencia de los mismos.

Este primer tema, lo vamos a dedicar a presentar tres de las distribuciones más habituales a la hora de modelizar el comportamiento de la v.a. *número de siniestros*. Como veremos se trata de distribuciones discretas, que dependiendo de la situación se ajustan mejor o peor al fenómeno estudiado. También hay que decir, que ninguna de ellas es de nueva creación, todas son distribuciones de probabilidad muy conocidas y estudiadas, pero que han tenido su aplicación y desarrollo particular por parte de los actuarios, de forma relativamente reciente.

2.1. Distribución de Poisson

En esta sección, partiendo de la definición general de la distribución de Poisson, estudiaremos distintos casos particulares, como son las distribuciones de Poisson *truncadas*, *aleatorizadas* en algunos de sus valores o *trasladadas*.

Definición 2.1 *Decimos que la v.a.* X *sigue una distribución de Poisson de parámetro* λ, $X \in$

$\mathcal{P}(\lambda)$, cuando su soporte o conjunto de valores es el conjunto de los enteros positivos, y su función de cuantía viene dada por

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k \in \mathbf{Z}^+ \tag{2.1}$$

La distribución de Poisson, puede ser introducida también como caso particular de la distribución binomial $b(p, n)$, donde el número de repeticiones **independientes** del experimento aleatorio n es muy alto, y la probabilidad de ocurrencia del suceso estudiado p es constante y muy pequeña. En este caso, *el número de ocurrencias del suceso a lo largo de las n repeticiones* sigue una distribución de Poisson de parámetro $\lambda = np$.

En distintos textos de probabilidad pueden encontrarse muy variadas introducciones a esta distribución, también conocida como *ley de los sucesos raros* y sin embargo hay algo en lo que todas coinciden, y que debemos recalcar; es una distribución que modeliza repeticiones **independientes** del experimento, con **probabilidad constante** de ocurrencia de los dos posibles resultados; hipótesis que debemos tener en cuenta cuando trabajemos con esta distribución.

Como propiedades que debemos recordar, tenemos las siguientes:

1. Su función de distribución es, $F(x) = 0$, si $x < 0$, y $F(x) = \dfrac{e^{-\lambda}\sum_{k=0}^{[x]}\lambda^k}{k!}$, si $x \geq 0$.

   En particular $F(0) = P(X \leq 0) = e^{-\lambda}$.
2. Su función característica, es, $\psi_X(u) = e^{\lambda(e^{iu}-1)}$.
3. Su función generatriz es, $\alpha_X(u) = e^{\lambda(e^u-1)}$.
4. Su función cumulativa es, $\mu_X(u) = \lambda(e^u - 1)$.
5. La media y la varianza son iguales, e iguales a λ. $E(X) = V(X) = \lambda$.

2.1.1. Distribuciones de Poisson truncadas

A veces, extensiones de la distribución de Poisson, nos sirven para modelizar la *v.a. número de siniestros*, donde sabemos que, por ejemplo, los primeros valores de la variable entre 0 y s, o bien los últimos a partir de un valor l, están excluidos. Surgen así las distribuciones de Poisson truncadas, cuyas funciones de cuantía obtendremos a continuación:

En todos los casos, partimos de una v.a. $X \in \mathcal{P}(\lambda)$.

1. Exclusión del valor 0.

   Sea Y la v.a. definida como $X|X > 0$. En este caso, el soporte de Y es el conjunto $\{1, 2, 3, ..., \}$, y su función de cuantía viene dada por:

$$P(Y = k) = P(X = k|X > 0) = \frac{P(X = k \cap X > 0)}{P(X > 0)} = \frac{e^{-\lambda}\lambda^k}{k!(1 - e^{-\lambda})},$$

donde $k \geq 1$. No es difícil comprobar la forma de la media y varianza de Y.

$$\begin{aligned} E(Y) &= \frac{\lambda}{(1-e^{-\lambda})} \\ V(Y) &= \frac{\lambda}{(1-e^{-\lambda})} - \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})^2} \\ \alpha_Y(u) &= \frac{e^{-\lambda}}{1-e^{-\lambda}}[e^{\lambda e^u} - 1] \end{aligned}$$

2. Truncada por la derecha

   Sea Y la v.a. definida como $X|X \leq l$. En este caso, el soporte de Y es el conjunto $\{0, 2, 3, ..., l-1, l\}$, y su función de cuantía viene dada por:

$$P(Y=k) = P(X=k|X \leq l) = \frac{P(X=k \cap X \leq l)}{P(X \leq l)} = \frac{\lambda^k}{k! \sum_{j=0}^{l} \frac{\lambda^j}{j!}},$$

   donde $0 \leq k \leq l$.

3. Truncada por la izquierda

   Sea Y la v.a. definida como $X|X \geq s$. En este caso, el soporte de Y es el conjunto $\{s, s+1, s+2, s+3, ...\}$, y su función de cuantía viene dada por:

$$P(Y=k) = P(X=k|X \geq s) = \frac{P(X=k \cap X \geq s)}{P(X \geq s)} = \frac{\lambda^k}{k! \sum_{j=s}^{\infty} \frac{\lambda^j}{j!}},$$

   donde $s \leq k$.

4. Doblemente truncada

   Sea Y la v.a. definida como $X|s \leq X \leq l$. En este caso, el soporte de Y es el conjunto $\{s, s+1, s+2, ..., l-1, l\}$, y su función de cuantía viene dada por:

$$\begin{aligned} P(Y=k) &= P(X=k|s \leq X \leq l) = \frac{P(X=k \cap s \leq X \leq l)}{P(s \leq X \leq l)} = \\ &= \frac{\lambda^k}{k! \sum_{j=s}^{l} \frac{\lambda^j}{j!}}, \end{aligned}$$

   donde $s \leq k \leq l$.

2.1.2. Distribución de Poisson generalizada

Se conoce con este nombre a la distribución de una variable, sea Y, que tiene el mismo soporte que una variable de Poisson, X, pero que añade un proceso de aleatorización en las probabilidades de los valores 0 y 1, es decir:

$$\begin{aligned} P(Y=0) &= P(X=0)+\theta P(X=1) \\ P(Y=1) &= P(X=1)-\theta P(X=1) \\ P(Y=k) &= P(X=k), \quad k\geq 2 \end{aligned}$$

donde θ es un parámetro real. Es fácil demostrar que en este caso,

$$\begin{aligned} E(Y) &= E(X)-\theta P(X=1)=\lambda(1-\theta e^{-\lambda}) \\ E(Y^2) &= E(X^2)-\theta P(X=1)=\lambda+\lambda^2-\theta\lambda e^{-\lambda} \\ V(Y) &= E(Y^2)-(E(Y))^2=\lambda-\theta\lambda e^{-\lambda}(1+\theta\lambda e^{-\lambda}-2\lambda) \end{aligned}$$

2.1.3. Distribución de Poisson desplazada

Se conoce con este nombre a la distribución de una variable, sea Z, que se obtiene de trasladar al origen una variable de Poisson truncada por la izquierda. Sea Y definida como $Y = X|X \geq r$ la v.a. de Poisson truncada por la izquierda. Así, la función de cuantía de Y es

$$P(Y=k)=\frac{\lambda^k}{k!\sum_{j=r}^{\infty}\frac{\lambda^j}{j!}},$$

para $k \geq r$. Si trasladamos al origen la variable Y, obtenemos la v.a. de Poisson desplazada $Z = Y - r$, que tiene como función de cuantía:

$$P(Z=k)=P(Y=r+k)=\frac{\lambda^{r+k}}{(r+k)!\sum_{j=r}^{\infty}\frac{\lambda^j}{j!}},$$

con $k \geq 0$.

2.2. Distribución binomial negativa

Al igual que en el caso de la distribución de Poisson, hay varias formas de introducir una distribución binomial negativa. Presentaremos aqui, una definición que parte de nuevo de repeticiones sucesivas e **independientes** de un experimento aleatorio que da lugar a dos posible resultados, en general, A: éxito y A^c: fracaso.

Así, en cada repetición, tendríamos definida una variable binaria

$X_n = 1$, si se produce un éxito, con $p = P(X_n = 1)$, y

$X_n = 0$, si se produce un fracaso, con $q = 1 - p = P(X_n = 0)$.

De manera que en las sucesivas repeticiones independientes, tendríamos definida una sucesión de v.a. independientes $\{X_n\}_{n\in\mathcal{N}}$.

En el estudio de este experimento nos pueden interesar distintas variables o indicadores, a saber:

1. Número de éxitos conseguidos después de n repeticiones del experimento. Esta variable, Y, corresponde a la suma de las n primeras variables de la sucesión anterior, $Y = X_1 + ... + X_n$. Su distribución nos es conocida; se trata de una *binomial* de parámetros n, p. Recordar que en casos particulares, n grande y p muy pequeño, podemos pensar que la distribución de Y es de Poisson de parámetro $\lambda = np$.

2. Si lo que queremos es estudiar la v.a. número de fracasos hasta obtener el primer éxito, veríamos que se trata de una v.a. con distribución *geométrica* de parámetro p.

3. Si somos más ambiciosos, y lo que queremos es estudiar el número de fracasos antes de obtener m éxitos, tendremos, en principio que esta situación generaliza la anterior, y por lo tanto la distribución *geométrica* será el caso particular de ésta cuando $m = 1$. Así, la variable objeto de estudio en este caso, Y_m, definida como *número de fracasos hasta obtener m éxitos* sigue una distribución conocida como binomial negativa de parámetros m y p.

   Es claro que el soporte de Y_m será el conjunto de los enteros positivos, $\mathbf{Z}^+$. Si queremos obtener su función de cuantía, tendremos que analizar el suceso $\{Y_m = k\}$, o lo que es lo mismo, ha habido k fracasos hasta obtener m éxitos. Es claro que

   $$\{Y_m = k\} = \{S_{m+k-1} = m - 1, X_{m+k} = 1\}$$

   Es decir, hemos realizado $m + k$ repeticiones. Entre las $m + k - 1$ primeras repeticiones se han producido los $m - 1$ éxitos y los k fracasos. La última repetición, ha dado lugar a un éxito, y por lo tanto hemos parado.

   Así, si $S_{m+k-1} = X_1 + X_2 + ... + X_{m+k-1}$ es la suma de las primeras $m + k - 1$ v.a. independientes de la sucesión $\{X_n\}$, sabemos que su distribución es binomial de parámetros $m + k - 1$, p. Por otro lado, esta v.a. es independiente de X_{m+k}, por lo que la probabilidad del suceso anterior, es

   $$\begin{aligned} P(Y_m = k) &= P(S_{m+k-1} = m - 1)P(X_{m+k} = 1) = \\ &= \begin{pmatrix} m+k-1 \\ m-1 \end{pmatrix} p^{m-1} q^k p = \\ &= \begin{pmatrix} m+k-1 \\ k \end{pmatrix} p^m q^k = \\ &= \begin{pmatrix} -m \\ k \end{pmatrix} p^m (-q)^k \end{aligned}$$

Proposición 2.1 *Se verifica la siguiente igualdad:*

$$\begin{pmatrix} m+k-1 \\ k \end{pmatrix} = \begin{pmatrix} m+k-1 \\ m-1 \end{pmatrix} = \begin{pmatrix} -m \\ k \end{pmatrix} (-1)^k$$

donde $k \in \mathbf{Z}^+$.

Demostración: Ver Apéndice A.

Definición 2.2 *Decimos que la v.a. X sigue una distribución binomial negativa de parámetros m, p, y lo denotamos por $X \in bn(m,p)$, cuando su soporte o conjunto de valores es el conjunto de los enteros positivos, y su función de cuantía viene dada por*

$$P(X=k) = \left(\begin{array}{c} m+k-1 \\ k \end{array} \right) p^m q^k = \left(\begin{array}{c} -m \\ k \end{array} \right) p^m (-q)^k, k \in \mathbf{Z}^+ \tag{2.2}$$

Para cualquier valor m, si expresamos la condición que debe cumplir la función de cuantía anterior por el hecho de serlo, tenemos que:

$$\begin{aligned} 1 &= \sum_{k=0}^{\infty} P(X=k) = p^m \sum_{k=0}^{\infty} \left(\begin{array}{c} -m \\ k \end{array} \right) (-q)^k = \\ &= p^m \sum_{k=0}^{\infty} \left(\begin{array}{c} m+k-1 \\ k \end{array} \right) q^k = p^m \sum_{k=0}^{\infty} \left(\begin{array}{c} m+k-1 \\ m-1 \end{array} \right) q^k \end{aligned}$$

de donde

$$\sum_{k=0}^{\infty} \left(\begin{array}{c} -m \\ k \end{array} \right) (-q)^k = \sum_{k=0}^{\infty} \left(\begin{array}{c} m+k-1 \\ k \end{array} \right) q^k = p^{-m} = (1-q)^{-m} \tag{2.3}$$

expresión ésta que nos será de utilidad en adelante.

A partir de aqui, podemos obtener las siguientes características estadísticas de la distribución binomial negativa.

1. Su función característica es:

$$\begin{aligned} \psi_X(u) &= E(e^{iuX}) = \sum_{k=0}^{\infty} e^{iuk} P(X=k) = \\ &= \sum_{k=0}^{\infty} e^{iuk} \left(\begin{array}{c} m+k-1 \\ k \end{array} \right) p^m q^k = \\ &= p^m \sum_{k=0}^{\infty} \left(\begin{array}{c} m+k-1 \\ k \end{array} \right) (e^{iu}q)^k = \\ &= p^m (1-e^{iu}q)^{-m} \end{aligned}$$

2. Su función generatriz es, $\alpha_X(u) = p^m(1-e^u q)^{-m}$.

3. Su función cumulativa es, $\mu_X(u) = mlnp - mln(1-e^u q)$.

4. La media y la varianza son, respectivamente:

$$E(X) = \frac{mq}{p} \quad V(X) = \frac{mq}{p^2}$$

En la introducción que hemos presentado, hemos tratado de modelizar el número de repeticiones necesarias del experimento aleatorio, para que se dé un suceso determinado. Por ejemplo, podríamos pensar en el número de lanzamientos sucesivos de un dado hasta obtener cinco seises; el número de lanzamientos de una moneda hasta obtener diez caras, etc. En nuestro contexto, querremos modelizar el número de exposiciones al riesgo hasta que se produzcan m siniestros, donde p es la probabilidad de ocurrencia de dicho siniestro.

Debemos recordar dos aspectos del experimento estudiado hasta ahora: la probabilidad de ocurrencia del siniestro es **constante** y las repeticiones se producen con **independencia**.

Por ota parte, hay otra observación que señalar de la distribución binomial negativa. De la introducción realizada, se puede inferir el caracter de entero positivo del parámetro m, restricción que formalmente no es necesaria en la definición de la variable. Es más, serán habituales, ejemplos y casos prácticos en los que a partir de la tabla de datos y dados los valores de los estadísticos muestrales, media y varianza, pensemos en ajustar el modelos a una distribución binomial negativa, donde la estimación de dichos parámetros nos proporcione en particular, valores de m fraccionales.

En este caso, tendremos que tener cuidado con el cálculo del número combinatorio

$$\binom{m+k-1}{k} = \frac{(m+k-1)(m+k-2)\cdots(m+k-k)}{k!}$$

donde además sabemos que

$$\binom{m+k-1}{k} = \binom{m+k-1}{m-1} = 1$$

siempre que $m+k-1 < k$ (esto sucede cuando $0 < m < 1$ y $k = 0$). Recordar además que $0! = 1$ y si $0 < x < 1$, $x! = 1$.

Proposición 2.2 *Sea X una v.a. con distribución binomial negativa de parámetros m, p; $X \in bn(m,p)$. Entonces, se verifica la siguiente fórmula recurrente:*

$$p_k = P(X = k) = \frac{(m+k-1)q}{k} P(X = k-1) = \frac{(m+k-1)q}{k} p_{k-1}$$

Demostración: Queda como ejercicio.

2.3. Distribución de Polya-Eggenberger

En las dos secciones anteriores, hemos recalcado la **independencia** como característica de las sucesivas repeticiones del experimento aleatorio considerado. Esto conlleva, como sabemos, el hecho de que la **probabilidad de ocurrencia** de cada posible resultado permanece **constante** a lo largo de la serie de repeticiones.

Sin embargo, hay fenómenos aleatorios en particular en los que es inadmisible esta hipótesis, dado que existe lo que habitualmente es conocido como *contagio*. Para introducir este efecto, y modelizar el comportamiento de la v.a. X: *número de siniestros*, en esta situación, vamos utilizar el siguiente experimento con la denominada *urna de Polya*.

Consideremos una urna, que inicialmente tiene N bolas, de las cuales a son blancas y b negras, de manera que $a + b = N$. Cada vez que se extrae una bola se mira su color y se introduce de nuevo a la urna, junto con c bolas del mismo color.

Observa, que el caso particular $c = 0$, replica una extracción con reemplazamiento, en la que se mantiene la hipótesis de independencia entre resultados de distintas extracciones; mientras que si $c = -1$, estaríamos ante el caso de extracción sin reemplazamiento, y consiguiente pérdida de independencia entre resultados de distintas extracciones.

Denotemos por X a la v.a. definida como *número de bolas blancas obtenidas en n extracciones* realizadas según el procedimiento de reemplazamiento que hemos descrito.

Claramente, si $c = 0$, la distribución de la v.a. X es binomial de parámetros n, $p = \frac{a}{N}$; $X \in b(n, \frac{a}{n})$.

Analicemos en general el caso $c \neq 0$. Trataremos de obtener la función de cuantía de X, en el caso $c \neq 0$. Para ello, pensemos en el suceso bl_i: se obtiene bola blanca en la i-ésima extracción y en su complementario, ng_i: se obtiene bola negra en la i-ésima extracción.

Mediante el teorema de la intersección, podemos calcular la probabilidad de obtener bola blanca en las r primeras extracciones. Para ello, necesitamos calcular previamente las siguientes probabilidades condicionadas:

$$
\begin{aligned}
P(bl_1) &= \frac{a}{N} \\
P(bl_2|bl_1) &= \frac{a+c}{N+c} \\
P(bl_3|bl_1 \cap bl_2) &= \frac{a+2c}{N+2c} \\
... &= ... \\
P(bl_r|bl_1 \cap bl_2 \cap ... \cap bl_{r-1}) &= \frac{a+(r-1)c}{N+(r-1)c}
\end{aligned}
$$

Así,

$$
\begin{aligned}
& P(bl_1 \cap bl_2 \cap ... \cap bl_{r-1} \cap bl_r) = \\
= \ & P(bl_1)P(bl_2|bl_1)P(bl_3|bl_1 \cap bl_2) \cdots P(bl_r|bl_1 \cap bl_2 \cap ... \cap bl_{r-1}) = \\
= \ & \frac{a}{N}\frac{(a+c)}{(c+N)} \cdots \frac{(a+(r-1)c)}{(N+(r-1)c)}
\end{aligned}
$$

Del mismo modo, podemos calcular la probabilidad de obtener s bolas negras, en las s siguientes extracciones. Para ello, de nuevo nos hacen falta las siguientes probabilidades condicionadas:

$$\begin{aligned}
P(ng_{r+1}|bl_1 \cap bl_2 \cap ... \cap bl_r) &= \frac{b}{N+rc} \\
P(ng_{r+2}|bl_1 \cap bl_2 \cap ... \cap bl_r \cap ng_{r+1}) &= \frac{(b+c)}{N+(r+1)c} \\
... &= ... \\
P(ng_n|bl_1 \cap bl_2 \cap ... \cap bl_r \cap ng_{r+1} ... \cap ng_{r+(s-1)}) &= \frac{b+(s-1)c}{N+(n-1)c}
\end{aligned}$$

Por tanto, en n extracciones, la probabilidad de obtener las r primeras bolas blancas y las s siguientes negras, donde $r+s=n$, es:

$$\begin{aligned}
& P(bl_1 \cap bl_2 \cap ... \cap bl_{r-1} \cap bl_r \cap ng_{r+1} ... \cap ng_{r+s}) = \\
= \ & P(bl_1)P(bl_2|bl_1) \cdots P(bl_r|bl_1 \cap bl_2 \cap ... \cap bl_{r-1}) P(ng_{r+1}|bl_1 \cap bl_2 \cap ... \cap bl_r) ... \\
... \ & P(ng_{r+s}|bl_1 \cap bl_2 \cap ... \cap bl_r \cap ng_{r+1} ... \cap ng_{r+(s-1)}) = \\
= \ & \frac{a}{N}\frac{(a+c)}{(N+c)} \cdots \frac{(a+(r-1)c)}{(N+(r-1)c)}\frac{b}{(N+rc)}\frac{(b+c)}{(N+(r+1)c)} \cdots \frac{(b+(s-1)c)}{(N+(n-1)c)}
\end{aligned}$$

Sin embargo, en las n extracciones, el número de formas en que se pueden colocar r bolas blancas y s negras es $\binom{n}{r} = \binom{n}{s}$; y la probabilidad de cada ordenación de esta n-tupla es igual a la de la ordenación fijada, de r primeras blancas y $s = n-r$ siguientes negras, que ya hemos calculado.

De esta forma, la función de cuantía de la v.a X, que calcula, con este modelo de extracción, la probabilidad de obtener r bolas blancas en $n = r+s$ extracciones, es:

$$P(X=r) = \binom{n}{r} \frac{a}{N}\frac{(a+c)}{(N+c)} \cdots \frac{(a+(r-1)c)}{(N+(r-1)c)}\frac{b}{(N+rc)}\frac{(b+c)}{(N+(r+1)c)} \cdots \frac{(b+(s-1)c)}{(N+(n-1)c)}$$

Dividiendo numerador y denominador en la expresión anterior por $c \neq 0$, se obtiene:

$$P(X=r) = \binom{n}{r} \frac{\frac{a}{c}}{\frac{N}{c}}\frac{(\frac{a}{c}+1)}{(\frac{N}{c}+1)} \cdots \frac{(\frac{a}{c}+(r-1))}{(\frac{N}{c}+(r-1))}\frac{\frac{b}{c}}{(\frac{N}{c}+r)}\frac{(\frac{b}{c}+1)}{(\frac{N}{c}+(r+1))} \cdots \frac{(\frac{b}{c}+(s-1))}{(\frac{N}{c}+(n-1))}$$

Reordenando los términos, se obtiene

$$P(X=r) = \frac{\binom{\frac{a}{c}+r-1}{r}\binom{\frac{b}{c}+s-1}{s}}{\binom{\frac{N}{c}+n-1}{n}}$$

Definición 2.3 *Decimos que la v.a X sigue una distribución de Polya- Eggenberger de parámetros p, q y δ, cuando el soporte de valores de la variables es el conjunto de enteros positivos $\{0, 1, 2..., n\}$, y su función de cuantía viene dada por*

$$P(X = k) \quad = \quad \frac{\begin{pmatrix} \frac{p}{\delta} + k - 1 \\ k \end{pmatrix} \begin{pmatrix} \frac{q}{\delta} + s - 1 \\ s \end{pmatrix}}{\begin{pmatrix} \frac{1}{\delta} + n - 1 \\ n \end{pmatrix}}$$

donde p representa la probabilidad inicial de siniestro, $q = 1 - p$, δ es la probabilidad de contagio, y $s = n - k$.

Para llegar a la expresión de la función de cuantía anterior, hemos renombrado los parámetros que aparecían en el experimento de la *urna de Polya*. Así la probabilidad inicial de siniestro es la probabilidad inicial de bola blanca $p = \frac{a}{N}$, $q = \frac{b}{N}$ y $\delta = \frac{c}{N}$ es la probabilidad de contagio.

Proposición 2.3 *Sea X una v.a. con distribución de Polya de parámetros, p, δ. Entonces, se verifica la siguiente fórmula recurrente:*

$$p_k = P(X = k) = \frac{(\frac{p}{\delta} + k - 1)}{k} \frac{s + 1}{(\frac{q}{\delta} + s)} P(X = k - 1)$$

donde, $s = n - k$ y

$$p_0 = P(X = 0) = \frac{\begin{pmatrix} \frac{q}{\delta} + n - 1 \\ n \end{pmatrix}}{\begin{pmatrix} \frac{1}{\delta} + n - 1 \\ n \end{pmatrix}}$$

Demostración: Queda como ejercicio.

En cuanto a las características estadísticas de esta distribución, tenemos:

1. La media es:

$$E(X) = np$$

2. La varianza es:

$$V(X) = npq \frac{1 + n\delta}{1 + \delta}$$

Es decir el efecto contagio, modelizado en esta distribución produce una media independiente de dicho efecto, e igual a la media en una distribución binomial, en la que el contagio es nulo. Sin embargo la varianza aumenta a medida que aumenta la probabilidad de contagio.

Este tipo de distribuciones se emplea en los estudios estadísticos médicos, sobre posibles contagios de enfermedades. Nosotros la utilizaremos para modelizar propagación de siniestros en contratos de seguros de incendios, por ejemplo.

Ejemplo 2.1 *A partir de datos relativos a 1000 edificios, cada uno con 10 viviendas, se ha estimado para la v.a. X: número de reclamaciones por incendio/edificio, los siguientes estadísticos muestrales $\overline{x} = 0{,}3$, y $s^2 = 1{,}164$. ($n = 10$).*

1. *Estima los parámetros, bajo la hipótesis de que la distribución de X es Polya. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.*
2. *Estima los parámetros, bajo la hipótesis de que la distribución de X es binomial. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.*
3. *Estima los parámetros, bajo la hipótesis de que la distribución de X es Poisson. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.*
4. *Estima los parámetros, bajo la hipótesis de que la distribución de X es binomial negativa. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.*

2.4. Tarea 2

Se enumeran a continuación un conjunto de ejercicios de ensayo de las competencias correspondientes a este tema; en particular a la segunda de las formuladas en el programa de la asignatura:

> Describir y analizar situaciones sencillas de riesgo en problemas relacionados con el mundo actuarial o de las finanzas y expresarlas formalmente en términos de variables aleatorias y sus distribuciones.

1. En una compañía de seguros, la media mensual de reclamaciones es de 1.5 por cada 100 pólizas contratadas. Consideremos la v.a. X: número de reclamaciones mensuales.
   a) Bajo la hipótesis de que X sigue una distribución de Poisson, calcula la probabilidad de que se dé al menos una reclamación a lo largo de un més cualquiera.
   b) Calcula la misma probabilidad que en el apartado anterior, bajo la hipótesis de que X sigue una distribución binomial con $n = 10$. Repite los cálculos para $n = 100$ y $n = 1000$.
2. Demuestra que la media, varianza y función generatriz de la v.a. $Y = X|X > 0$, con $X \in P(\lambda)$, son las dadas en la Sección 2.1.1.1.
3. En una compañía de seguros se ha estimado que la probabilidad de reclamación en pólizas anuales contra el riesgo A es de 0.08, mientras que en las contratadas contra el riesgo B es de 0.05. La compañía decide ofertar una póliza combinada para cubrir ambos riesgos. Considera que la compañía ha contratado 1000 pólizas de este tipo en un año, que las reclamaciones por ambos riesgos son independientes y que el titular de cada póliza efectúa una reclamación como mucho.
   a) Calcula la probabilidad de reclamación sobre contratos de esta póliza combinada.

b) Si se modeliza la v.a. X: número de reclamaciones como una v.a. de Poisson, estima el valor del parámetro λ.

c) Calcula la probabilidad aproximada de que es ese año haya habido más de 140 reclamaciones.

4. Un cierto siniestro tiene probabilidad de ocurrir $p = 0{,}1$. Sea Y, la variable aleatoria que recoge el número de exposiciones al riesgo hasta que se producen tres siniestros, teniendo en cuenta que terminan las exposiciones después del último siniestro.

   a) Si X es una variable con distribución binomial negativa de parámetros $p = 0{,}1$ y $m = 3$. ¿Cuál es la relación que observas entre X e Y?

   b) Calcula la probabilidad $P(Y = 3)$

   c) Calcula el número medio de exposiciones al riesgo, $E(Y)$.

   d) Calcula la varianza, $V(Y)$.

5. Demuestra que si X sigue una distribución binomial negativa de parámetros p y m,

$$P(X = k) = p_k = p_{k-1}(1 - p)(m + k - 1)/k$$

6. Sea X el número de reclamaciones por incendio en un edificio de cuatro viviendas, una v.a. con binomial negativa de media $EX = 0{,}04$, y varianza $V(X) = 0{,}0504$. Calcula los primeros valores de la función de cuantía de X: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

7. Sea X el número de reclamaciones por incendio en un edificio, una v.a. con distribución de Poisson de parámetro $\lambda = 0{,}04$.

   a) Calcula los primeros valores de la función de cuantía de X: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

   b) Calcula la media y la varianza de X.

   c) Si X_1 y X_2 son dos v.a. independientes entres sí y con distribuciones de Poisson de parámetros λ_1 y λ_2, obtén la distribución de $X_1 + X_2$ a partir de su función característica.

   d) Obtén en general la distribución de la suma $Z = X_1 + ... + X_n$, donde $X_1, ..., X_n$ son v.a. independientes con distribución de Poisson de parámetros $\lambda_1, ..., \lambda_n$, respectivamente.

8. Sea X el número de reclamaciones por incendio en un edificio, una v.a. con distribución de Polya de parámetros $p = 0{,}01$, $\delta = 0{,}1$ (recuerda que es el parámetro que modeliza la probabilidad de contagio en el siniestro), y $n = 4$, lo cual modeliza que el edificio en cuestión tiene 4 viviendas. Calcula los primeros valores de la función de cuantía de X: $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

9. Considera la siguiente tabla de datos sobre el número de siniestros (X: número de reclamaciones por incendio), relativos a 1000 edificios.

x_i	0	1	2	3	4
n_i	975	15	6	3	1

Con la tabla de datos anterior, puedes observar que la media muestral es $\overline{x} = 0{,}04$ y la varianza muestral es $s^2 = 0{,}0792$. Sabiendo además que cada edificio consta de 4 viviendas:

a) Estima los parámetros correspondientes si planteamos como hipótesis que la distribución de X es de Polya. Calcula los valores de la función de cuantía de X, en este caso, $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

b) Estima los parámetros correspondientes si planteamos como hipótesis que la distribución de X es binomial negativa. Calcula los valores de la función de cuantía de X, en este caso, $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

c) Estima los parámetros correspondientes si planteamos como hipótesis que la distribución de X es de Poisson. Calcula los valores de la función de cuantía de X, en este caso, $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, $P(X \geq 4)$.

d) A la vista de los resultados de las cuestiones anteriores, ¿qué distribución te parece más adecuada para modelizar el comportamiento aleatorio de X?

10. A partir de datos relativos a 1000 edificios, cada uno con 10 viviendas, se ha estimado para la v.a. X: número de reclamaciones por incendio/edificio, los siguientes estadísticos muestrales $\overline{x} = 0{,}3$, y $s^2 = 1{,}164$.

a) Estima los parámetros, bajo la hipótesis de que la distribución de X es Polya. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.

b) Estima los parámetros, bajo la hipótesis de que la distribución de X es binomial. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.

c) Estima los parámetros, bajo la hipótesis de que la distribución de X es Poisson. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.

d) Estima los parámetros, bajo la hipótesis de que la distribución de X es binomial negativa. Calcula en este caso, las probabilidades $P(X = 0)$, $P(X = 1)$, $P(X = 2)$ y $P(X = 3)$.

Capítulo 3

Modelos de distribuciones de probabilidad continuas. Modelos relacionados con el coste de cada siniestro

Lo que cuesta cada siniestro, o en adelante, lo que denominaremos *cuantía por reclamación*, también es una variable aleatoria, en este caso de carácter continuo, pues modeliza una variable económica medida en dinero, que como la variable tiempo, se considera continua.

Estudiaremos en este tema distintas distribuciones continuas, con sus características estadísticas más relevantes.

3.1. Distribución log-normal

También conocida como distribución logarítmico-normal. Es lo que se conoce en estadística como una distribución de colas anchas, es decir, su densidad refleja una característica que tambien corrobora la varianza de esta distribución, hay probabilidad alta de que la variable tome valores alejados de la media. Nos servirá para modelizar siniestros con grandes costes o cuantías por reclamación.

La habitual definición de una v.a. X con distribución log-normal, es la de aquella variable, cuyo logaritmo neperiano $Y = lnX$ sigue una distribución normal $\mathcal{N}(m, \sigma^2)$. Así para obtener la función de densidad de X, necesitaremos un argumento basado en la transformación de variables.

Pensemos primero en el caso de una v.a. $T \in \mathcal{N}(0, 1)$. Sabemos que la función de densidad es:

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

donde $t \in \mathbf{R}$.

La v.a. $Y \in \mathcal{N}(m, \sigma^2)$, está definida como una transformación lineal creciente de T, $Y = \sigma T + m$. Utilizando, la fórmula de transformación ya conociada y dada por

$$f_Y(y) = f_t(t(y))|t'(y)| \tag{3.1}$$

donde f_t es la función de densidad de T, y sabiendo que $t(y) = \frac{y-m}{\sigma}$, y $t'(y) = \frac{1}{\sigma} > 0$, obtenemos la función de densidad de Y.

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

donde $y \in \mathbf{R}$.

De nuevo, si $Y = lnX$, nuestra variable de interés es $X = e^Y$. Por por tanto utilizando la expresión (3.1), obtenemos la función de densidad de X.

$$\begin{aligned} f(x) &= f_Y(y(x)) \cdot |y'(x)| = \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(y(x)-m)^2}{2\sigma^2}} |y'(x)| \end{aligned}$$

donde $y(x) = lnx$, e $|y'(x)| = |\frac{1}{x}|$.

Observa, que la nueva variable $X = e^Y$, no puede tomar valores negativos, y por lo tanto su soporte es el conjunto de los reales positivos, $x \in \mathbf{R}^+$.

Definición 3.1 *Decimos que la v.a. X tiene distribución log-normal, si su soporte es el conjunto $\mathbf{R}^+$, y su función de densidad viene dada por*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{\frac{-(lnx-m)^2}{2\sigma^2}}, x \geq 0$$

Sus características estadísticas más relevantes, se deducen al igual que su función de densidad, de su relación con la distribución normal.

1. Para calcular los valores de su función de distribución, $F(x)$, emplearemos la tablas de la $N(0,1)$, mediante la siguiente relación:

$$F(x) = P(X \leq x) = P(e^Y \leq x) = P(Y \leq lnx) = \Phi(\frac{lnx - m}{\sigma})$$

2. Su media resulta de realizar la integral

$$E(X) = \int_0^\infty x f(x) dx = \int_0^\infty x \frac{1}{\sqrt{2\pi}\sigma x} e^{\frac{-(lnx-m)^2}{2\sigma^2}} dx$$

   Sin más que hacer el cambio de variable $t = \frac{lnx-m}{\sigma}$, y por tanto $x = e^{\sigma t + m}$, $dx = \sigma e^{\sigma t + m} dt$, la integral anterior se puede expresar como

$$E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^\infty e^{\frac{-t^2}{2}} \sigma e^{\sigma t + m} dt = e^m \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{\frac{-(t^2 - 2\sigma t)}{2}} dt$$

Completando el exponente a un cuadrado completo $\frac{(t^2-2\sigma t+\sigma^2-\sigma^2)}{2} = \frac{(t-\sigma)^2}{2} - \frac{\sigma^2}{2}$, se obtiene

$$E(X) = e^{m+\frac{\sigma^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-h^2}{2}} dh = e^{m+\frac{\sigma^2}{2}}$$

donde $h = t - \sigma$, y $dh = dt$.

3. Operando de forma similar, en la integral de $E(X^2)$, se puede demostrar que la varianza de la variable X es,

$$V(X) = e^{2m+\sigma^2}(e^{\sigma^2} - 1)$$

3.2. Distribución de Pareto

Al igual que la distribución anterior, es una ley de colas anchas por lo que la utilizaremos para modelizar cuantías elevadas.

La distribución de Pareto surge al considerar que la probabilidad de que una v.a. X tome una valor superior a un x determinado, tiene la siguiente forma funcional:

$$P(X > x) = (\frac{k}{x})^\alpha$$

si $x \geq k$, y $P(X > x) = 1$, en otro caso. En la expresión anterior, k y α son dos parámetros reales.

A partir de la consideración anterior, es sencillo obtener la función de distribución de la variable X,

$$F(x) = P(X \leq x) = 1 - P(X > x) = 1 - (\frac{k}{x})^\alpha$$

si $x \geq k$, y $F(x) = 0$, en otro caso.

Definición 3.2 *Decimos que la v.a. X tiene distribución de Pareto de parámetros k, α, si su soporte es el subconjunto de números reales $[k, \infty)$ y su función de densidad viene dada por*

$$f(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}$$

Normalmente k es un parámetro prefijado de antemano, a partir de los datos iniciales de la aplicación. El parámetro α es un parámetro de ajuste, cuyo valor se estima o determina a posteriori. Como veremos más adelante, para algunos valores de este parámetro de ajuste, no existen momentos de esta variable.

En cuanto a las características estadísticas,

1. Su media,

$$E(X) = \frac{\alpha k}{(\alpha - 1)}$$

se puede calcular como la integral anterior siempre que $\alpha > 1$.

2. Su varianza,

$$V(X) = \frac{\alpha k^2}{(\alpha-2)(\alpha-1)^2}$$

siempre que $\alpha > 2$.

Hemos visto que la existencia de los dos primeros momentos depende del valor del parámetro α. Es por ello, que a menudo conviene estimar su valor a partir de los datos, para tener una mayor información del comportamiento de la variable.

Como método de estimación, se puede emplear tanto el de momentos como el de máxima verosimilitud, aunque los estimadores que nos proporcionan no coinciden.

1. Estimación por momentos.

   Partiendo de una m.a.s. de la distribución, $\{x_1, ..., x_n\}$, igualamos el momento de orden uno poblacional, $E(X)$, con el correspondiente momento muestral $\overline{x}$.

   $$\frac{\alpha k}{(\alpha-1)} = \overline{x}$$

   de donde

   $$\hat{\alpha}_M = \frac{\overline{x}}{(\overline{x}-k)}$$

2. Estimación por máxima verosimilitud.

   Partiendo de una m.a.s. de la distribución, $\{x_1, ..., x_n\}$, construimos la función de verosimilitud:

   $$L(\vec{x}, \alpha) = (\frac{\alpha}{k})^n \prod_{i=1}^{n} (\frac{k}{x_i})^{\alpha+1}$$

   Tomando logaritmos,

   $$lnL(\vec{x}, \alpha) = n(ln\alpha - lnk) + \sum_{i=1}^{n} (\alpha+1)(lnk - lnx_i)$$

   derivando respecto a α e igualando a 0, despejamos el estimador máximo verosímil

   $$\hat{\alpha}_{MV} = \frac{1}{(\overline{lnx} - lnk)}$$

   donde $\overline{lnx} = \frac{1}{n}\sum_{i=1}^{n} lnx_i$.

Considerando las expresiones de los estimadores anteriores obtenidas para una muestra de tamaño uno, se puede intuir su relación.

3.2.1. Distribuciones de Pareto de segundo y tercer tipo

1. La distribución de Pareto de segundo tipo, no es más que la de primer tipo, trasladada al origen. Es decir, si Y es una v.a. con distribución de Pareto (de primer tipo) y parámetros k, α, se define como distribución de Pareto de segundo tipo a la distribución de la v.a. $X = Y - k$.

   La relación entre las funciones de distribución, es

   $$F(x) = P(X \leq x) = P(Y - k \leq x) = P(Y \leq x + k) = 1 - (\frac{k}{x+k})^{\alpha}$$

   si $x \geq 0$, y $F(x) = 0$, en otro caso.

2. La distribución de Pareto de tercer tipo, es una generalización de las anteriores, en la que además de la traslación se produce un cambio en la estructura de varianza. Se define como distribución de Pareto de tercer tipo a la distribución de la v.a. X, cuya función de distribución es:

   $$F(x) = P(X \leq x) = 1 - (\frac{k}{x+k})^{\alpha} e^{-bx}$$

   si $x \geq 0$, y $F(x) = 0$, en otro caso. Esta distribución depende de tres parámetros, α, k y b.

3.2.2. Distribución de Burr

Es otra distribución transformada de la distribución de Pareto, que se emplea habitualmente para estudiar las pérdidas en carteras de seguros.

Sea X una v.a. de Pareto de segundo tipo y Z transformación definida como $Z^{\beta} = X$, con $\beta > 0$.

La relación entre ambas funciones de distribución, teniendo en cuenta que $Z = X^{\frac{1}{\beta}}$, es:

$$F(z) = P(Z \leq z) = P(X^{\frac{1}{\beta}} \leq z) = P(X \leq z^{\beta}) = 1 - (\frac{k}{z^{\beta}+k})^{\alpha}$$

si $z \geq 0$ y $F(z) = 0$, en otro caso.

Proposición 3.1 *Sea Z una v.a. con distribución de Burr de parámetros k, α, β. El momento ordinario de orden r de la v.a. Z es:*

$$E(Z^r) = k^{\frac{r}{\beta}} \frac{\Gamma(\alpha - r/\beta)\Gamma(r/\beta + 1)}{\Gamma(\alpha)}$$

siempre que $r < \alpha\beta$.

Demostración: Queda como ejercicio.

3.3. Distribución Gamma

Aunque la hemos presentado en el resumen del primer tema, la incluímos de nuevo aqui como modelo para ciertas distribuciones de pérdida o cuantía de siniestros.

Recordemos la siguiente definición:

Definición 3.3 *Se dice que la v.a X sigue una distribución Gamma de parámetros $a > 0$, $r > 0$, y se denota por $\gamma(a, r)$, si su función de densidad viene dada por*

$$f(x) = \frac{a^r}{\Gamma(r)} x^{r-1} e^{-ax}, \quad x \geq 0 \tag{3.2}$$

y $f(x) = 0$, en otro caso; donde $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$

Nota: Si $r \in \mathbf{Z}^+$, se verifica $\Gamma(r) = (r-1)!$.

Además como principales características estadísticas tenemos:

1. Su media es:

$$E(X) = \frac{r}{a} \tag{3.3}$$

2. Su varianza es:

$$V(X) = \frac{r}{a^2} \tag{3.4}$$

3. Su función generatriz de momentos es:

$$\alpha_X(u) = (1 - \frac{u}{a})^{-r} \tag{3.5}$$

4. Su función característica es:

$$\psi_X(u) = (1 - \frac{iu}{a})^{-r} \tag{3.6}$$

5. Su función cumulativa es:

$$\mu_X(u) = -r \cdot ln(1 - \frac{u}{a}) \tag{3.7}$$

3.3.1. Distribución exponencial

La distribución exponencial se puede introducir como el caso particular de la distribución γ cuando el parámetro $r = 1$, en este caso:

Definición 3.4 *Se dice que la v.a X sigue una distribución exponencial de parámetro $a > 0$, y se denota por $\epsilon(a)$, si su función de densidad viene dada por*

$$f(x) = ae^{-ax}, \quad x \geq 0 \tag{3.8}$$

y $f(x) = 0$, en otro caso.

Sus principales características, se obtienen también como caso particular de la distribución γ.

1. Su media es:

$$E(X) = \frac{1}{a} \tag{3.9}$$

2. Su varianza es:

$$V(X) = \frac{1}{a^2} \tag{3.10}$$

3. Su función generatriz de momentos es:

$$\alpha_X(u) = (1 - \frac{u}{a})^{-1} \tag{3.11}$$

4. Su función característica es:

$$\psi_X(u) = (1 - \frac{iu}{a})^{-1} \tag{3.12}$$

5. Su función cumulativa es:

$$\mu_X(u) = -ln(1 - \frac{u}{a}) \tag{3.13}$$

Observar además que de la forma de la función característica, tanto de la distribución γ, como en particular de la exponencial, se deduce el siguiente resultado:

Proposición 3.2 *La suma de v.a. independientes con distribución γ del mismo parámetro a, es una v.a. con distribución γ del mismo parámetro a y donde su parámetro r se obtiene como suma de dichos parámetros en las distribuciones γ que hemos sumado.*

Demostración: Evidente si calculamos la función característica de la suma de v.a. independientes con distribución γ.

3.3.2. Distribución exponencial desplazada

La distribución exponencial se puede generalizar, trasladándola del origen. Surge así, la distribución *exponencial desplazada.*

Sea X una v.a. con distribución exponencial de parámetro a. $X \in \mathcal{E}(a)$. Consideremos la v.a. transformación lineal de X, $Y = X + k$, $k > 0$. La función de densidad de Y, de acuerdo con la expresión (3.1), vendrá dada por:

$$f_Y(y) = f_X(y-k) = ae^{-a(y-k)} \tag{3.14}$$

si $y \geq k$; siendo $f_Y(y) = 0$, en otro caso.

Así, tenemos la siguiente definición:

Definición 3.5 *Se dice que la v.a X sigue una distribución exponencial desplazada de parámetros $a > 0$, $k > 0$ y se denota por $\mathcal{E}(a,k)$, si su función de densidad viene dada por*

$$f(x) = ae^{-a(x-k)}, \quad x \geq k > 0 \tag{3.15}$$

y $f(x) = 0$, en otro caso.

Sus principales características, son evidentes a partir de su relación con una v.a. exponencial. De la propia definición, se sigue que:

1. Su media es:

$$E(X) = \frac{1}{a} + k \tag{3.16}$$

2. Su varianza es la misma que la de la exponencial de parámetro a:

$$V(X) = \frac{1}{a^2} \tag{3.17}$$

3. Su función característica es:

$$\psi_X(u) = e^{iuk}(1 - \frac{iu}{a})^{-1} \tag{3.18}$$

3.4. Distribución de Weibull

Es otra generalización de la distribución exponencial, en la que intervienen dos parámetros, el parámetro $\theta > 0$, que corresponde en la definición de la distribución exponencial a $\frac{1}{a}$, y se denomina parámetro de escala; y el parámetro $\alpha > 0$, denominado comúnmente parámetro de forma.

Esta distribución se usa habitualmente para modelar la compensación de las reclamaciones de los trabajadores. Proporciona una adecuada respresentación de la distribución de la pérdida.

Se define como sigue:

Definición 3.6 *Se dice que la v.a X sigue una distribución de Weibull de parámetros $\alpha > 0$, $\theta > 0$ y se denota por $W(\alpha, \theta)$, si su función de densidad viene dada por*

$$f(x) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} e^{-(\frac{x}{\theta})^\alpha}, \quad x \geq 0 \tag{3.19}$$

y $f(x) = 0$, en otro caso.

Recordando la definición de la función $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx$, para $r > 0$, se deducen sus principales características estadísticas:

1. Su media es:

$$E(X) = \frac{\alpha}{\theta^\alpha} \int_0^\infty x x^{\alpha-1} e^{-(\frac{x}{\theta})^\alpha} dx = \theta \Gamma(\frac{1}{\alpha} + 1) \tag{3.20}$$

2. Su varianza es:

$$V(X) = \theta^2 [\Gamma(\frac{2}{\alpha} + 1) - \Gamma^2(\frac{1}{\alpha} + 1)] \tag{3.21}$$

3.5. Distribución Beta

Pensemos en una póliza que cubre dos tipos de siniestros independientes, y que la cuantía de cada uno de ellos sigue una distribución $\gamma(a, r_i)$, $i = 1, 2$. En este caso la v.a. $Y_1 = \frac{X_1}{X_1 + X_2}$ representa la fracción de pérdida asociada al primer siniestro, mientras que $Y_2 = \frac{X_2}{X_1 + X_2}$, respresenta la fracción asociada al segundo siniestro.

Veremos que la distribución de Y_1 es una distribución Beta de parámetros r_1, r_2, mientras que la de Y_2 es una distribución Beta de parámetros r_2, r_1.

Dada la independecia entre X_1 y X_2 y la forma de su distribución γ del mismo parámetro a, de acuerdo con la Proposición 3.2, sabemos que $X = X_1 + X_2 \in \gamma(a, r_1 + r_2)$. Es decir, veremos que una v.a. con distribución Beta surge como cociente de dos v.a. con distribuciones γ del mismo parámetro a **no independientes**.

Para obtener la función de densidad de la v.a. transformación $Z_1 = X_1/X$. Para ello, definimos una aplicación $\vec{Z}$ de $\mathbf{R}^2$ en $\mathbf{R}^2$ que haga corresponder a

$$(X_1, X) - - \rightarrow (Z_1, Z_2)$$

donde $Z_1 = X_1/X$ y $Z_2 = X$. Transformación que podemos expresar en forma matricial como:

$$\vec{Z} = \begin{pmatrix} \frac{1}{X} & 0 \\ 0 & 1 \end{pmatrix} \vec{X} = A\vec{X},$$

con $||A|| = \frac{1}{|X|} \neq 0$ y

$$\vec{X} = A^{-1}\vec{Z} = \begin{pmatrix} X & 0 \\ 0 & 1 \end{pmatrix} \vec{Z},$$

de donde $X_1 = Z_1 X_1 = Z_1 Z_2$ y $X_2 = Z_2$.

El jacobiano de esta transformación, es decir el determinante de la matriz de derivadas de $\vec{X}$ con respecto a $\vec{Z}$, que en valor absoluto,

$$||\frac{D(\vec{X})}{D(\vec{Z})}|| = || \begin{matrix} Z_2 & Z_1 \\ 0 & 1 \end{matrix} || = |Z_2|$$

es la generalización de la expresión $|x'(z)|$, en la fórmula (3.1), que ahora nos proporciona la función de densidad conjunta de $\vec{Z} = (Z_1, Z_2)$ a partir de la de (X_1, X),

$$\begin{aligned} f_{\vec{Z}}(z_1, z_2) &= f_{\vec{X}}(x_1(z_1, z_2), x(z_1, z_2))||\frac{D(\vec{X})}{D(\vec{Z})}|| = \\ &= f_{\vec{X}}(z_1 z_2, z_2)|z_2| \end{aligned}$$

La función de densidad que nos interesa es la marginal de la primera componente $Z_1 = X_1/X$. Luego,

$$\begin{aligned} f_1(z_1) &= \int_{-\infty}^{\infty} f(z_1, z_2) dz_2 = \\ &= \int_{-\infty}^{\infty} f_{\vec{X}}(z_1 z_2, z_2)|z_2| dz_2 = \\ &= \int_{0}^{\infty} f_{\vec{X}}(z_1 z_2, z_2) z_2 dz_2 \end{aligned}$$

si $z_2 = x_2 \geq 0$.

Pero para poder calcular la integral, tenemos que determinar previamente cuál es la función de densidad conjunta de la v.a. bidimensional (X_1, X), cuyas componentes son **no independientes**. Para calcular su función de densidad, es más sencillo calcular primero su función de distribución:

$$\begin{aligned} F(x_1, x) &= P(X_1 \leq x_1, X \leq x) = P(X_1 \leq x_1, X_1 + X_2 \leq x) = \\ &= P(X_1 \leq x_1, X_2 \leq x - x_1) = F_1(x_1) F_2(x - x_1) \end{aligned}$$

puesto que X_1 y X_2 sí son independientes. Por lo tanto,

$$f_{\vec{X}}(x_1, x) = f_1(x_1) f_2(x - x_1)$$

donde f_1 es la función de densidad de $X_1 \in \gamma(a, r_1)$ y f_2 es la función de densidad de $X_2 \in \gamma(a, r_2)$.

Sustituyendo las funciones de densidad por sus expresiones correspondientes, se obtiene:

$$\begin{aligned} f_{\vec{X}}(x_1, x) &= \frac{a^{r_1}}{\Gamma(r_1)} x_1^{r_1 - 1} e^{-a x_1} \frac{a^{r_2}}{\Gamma(r_2)} (x - x_1)^{r_2 - 1} e^{-a(x - x_1)} = \\ &= \frac{a^{r_1} a^{r_2}}{\Gamma(r_1)\Gamma(r_2)} x_1^{r_1 - 1} (x - x_1)^{r_2 - 1} e^{-a x} \end{aligned}$$

Y por tanto

$$f_{\vec{X}}(z_1 z_2, z_2) = \frac{a^{r_1} a^{r_2}}{\Gamma(r_1)\Gamma(r_2)} (z_1 z_2)^{r_1 - 1} (z_2 - z_1 z_2)^{r_2 - 1} e^{-a z_2}$$

De donde la función de densidad de Z_1, resulta ser:

$$\begin{aligned}
f_1(z_1) &= \int_{-\infty}^{\infty} f(z_1, z_2)dz_2 = \\
&= \int_0^{\infty} f_{\vec{X}}(z_1z_2, z_2)z_2dz_2 = \\
&= \int_0^{\infty} \frac{a^{r_1}a^{r_2}}{\Gamma(r_1)\Gamma(r_2)}(z_1z_2)^{r_1-1}(z_2 - z_1z_2)^{r_2-1}e^{-az_2}z_2dz_2 = \\
&= \frac{a^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)}z_1^{r_1-1}(1-z_1)^{r_2-1}\int_0^{\infty} e^{-az_2}z_2^{r_1+r_2-1}dz_2 = \\
&= \frac{a^{r_1+r_2}}{\Gamma(r_1)\Gamma(r_2)}z_1^{r_1-1}(1-z_1)^{r_2-1}\frac{\Gamma(r_1+r_2)}{a^{r_1+r_2}} = \\
&= \frac{\Gamma(r_1+r_2)}{\Gamma(r_1)\Gamma(r_2)}z_1^{r_1-1}(1-z_1)^{r_2-1}
\end{aligned}$$

si $0 \leq z_1 \leq 1$, y $f_1(z_1) = 0$, en otro caso.

Así, formalmente, tenemos la siguiente definición:

Definición 3.7 *Se dice que la v.a X sigue una distribución Beta de parámetros $r_1 > 0$, $r_2 > 0$ y se denota por $\mathcal{B}(r_1, r_2)$, si su función de densidad viene dada por*

$$\begin{aligned}
f(x) &= \frac{\Gamma(r_1+r_2)}{\Gamma(r_1)\Gamma(r_2)}x^{r_1-1}(1-x)^{r_2-1} = \\
&= \frac{1}{B(r_1, r_2)}x^{r_1-1}(1-x)^{r_2-1}
\end{aligned}$$

si $0 \leq x \leq 1$, y $f(x) = 0$, en otro caso.

De imponer la condición de que la función de densidad integra la unidad, se deduce que

$$B(r_1, r_2) = \int_0^1 x^{r_1-1}(1-x)^{r_2-1}dx = \frac{\Gamma(r_1)\Gamma(r_2)}{\Gamma(r_1+r_2)}$$

y, $\Gamma(r) = \int_0^{\infty} x^{r-1}e^{-x}dx$, para $r > 0$.

En cuanto a sus carácterísticas estadísticas:

1. Su media es:

$$E(X) = \frac{r_1}{r_1 + r_2} \tag{3.22}$$

2. Su varianza es:

$$V(X) = \frac{r_1r_2}{(r_1+r_2)^2(r_1+r_2+1)} \tag{3.23}$$

3.6. Tarea 3

Se enumeran a continuación un conjunto de ejercicios de ensayo de las competencias correspondientes a este tema; en particular seguimos con la segunda de las formuladas en el programa de la asignatura:

> Describir y analizar situaciones sencillas de riesgo en problemas relacionados con el mundo actuarial o de las finanzas y expresarlas formalmente en términos de variables aleatorias y sus distribuciones.

1. En una compañía de seguros se ha estimado que la cuantía por reclamación por responsabilidad del profesional médico, es una v.a. con distribución de Pareto con parámetros k, α. Se sabe además que la empresa aseguradora cubre las reclamaciones de cuantía menor que un cierto valor $c > k$, y que subcontrata a su vez a una empresa reaseguradora para cubrir las reclamaciones con cuantías superiores a c.

   a) Obtén la función de densidad truncada de X, es decir, la de la v.a. $Y = X|X > c$.

   b) Demuestra que la v.a. Y, sigue una distribución de Pareto de parámetros $k = c$, α.

2. La cuantía por reclamación, X, (en miles de euros) en la compañía aseguradora *Gentilicia* es una v.a. con distribución de Pareto de parámetros $k = 3$ (miles de euros) y $\alpha = 2{,}6$. La compañía decide reasegurarse con en banco *Providencial*, de manera que dicho banco cubra las reclamaciones por encima de los 6000 euros.

   a) Obtener la función de densidad de la v.a. X.

   b) Calcular su media y su varianza.

   c) Calcular la probabilidad de que en una reclamación, la cuantía supere los 6000 euros.

   d) Si sabemos que la cuantía de una reclamación ha superado los 6000 euros, calcular la probabilidad de que supere los 9000.

   e) Obtener la función de densidad de las cuantías reclamadas al banco *Providencial*.

3. Para una v.a. no negativa X con función de distribución $F(x, \theta)$, se define la "*prima de riesgo bajo el principio de prima de riesgo ajustada*", como:

   $$P(\theta) = \int_0^{\infty} [P(X > x)]^{\frac{1}{\rho}} dx = \int_0^{\infty} [1 - F(x, \theta)]^{\frac{1}{\rho}} dx$$

   Demostrar que:

   a) Si $X \in \exp(\theta)$, entonces $P(\theta) = \frac{\rho}{\theta}$.

   b) Si $X \in Pareto(\theta, \alpha)$, con α conocido, entonces $P(\theta) = \frac{\rho\theta}{(\alpha-\rho)}$, $\alpha > \rho$.

4. Un reaseguro "*excess loss*" es una modalidad de reaseguro por la que el reasegurador cubre la parte del seguro que excede del pleno fijado por la cedente para cada siniestro. En esta modalidad, el reasegurador asume la parte de cada siniestro que supera una determinada cantidad. Esta cantidad recibe el nombre de *cesión* en los seguros vida y *deducible* en los seguros no vida. Si denotamos por M al pleno o deducible fijado para cada siniestro, y X es la v.a. que representa la cuantía por reclamación, pueden suceder dos cosas:

a) Que $X \leq M$, entonces el responsable es el asegurador y no hay reaseguro.

b) Que $X > M$, existe reaseguro que cubre $X - M$. Por tanto, el asegurador paga $Y = \text{mín}\{X, M\}$, mientras que el reasegurador paga $Z = \text{máx}\{0, X - M\}$.

Si $F_Y(y)$ es la función de distribución de la v.a. $Y = \text{mín}\{X, M\}$, se verifica:

$$F_Y(y) = \left\{ \begin{array}{ll} F(x), & x \leq M \\ 1, & x \geq M \end{array} \right.$$

De donde $E(Y^n) = \int_0^\infty [\text{mín}\{X, M\}]^n f(x)dx$, y teniendo en cuenta que

$$Y = \text{mín}\{X, M\} = \left\{ \begin{array}{lr} x, & 0 \leq x \leq M \\ M, & x \geq M \end{array} \right.$$

y dado que $E([\text{mín}\{X, M\}]^n) = \int_0^M x^n f(x)dx + M^n[1 - F(M)]$, Se obtiene que los pagos medios del asegurador son: $E(Y) = \int_0^M xf(x)dx + M[1 - F(M)]$,

Si la v.a X sigue una distribución log-normal (μ, σ), calcula $E(Y^n)$ y $E(Y)$, para un reaseguro *excess loss*.

5. Considera una v.a. continua de la que sabemos que $X > k$, con $k = 2$, su media es $E(X) = 3$.

   a) Bajo la hipótesis de que X sigue una distribución de Pareto de primer tipo, calcula el valor del parámetro α.

   b) Bajo la hipótesis de que X sigue una distribución exponencial desplazada, calcula su varianza.

   c) Analiza las diferencias en varianza de las posibilidades anteriores.

   d) Analiza las diferencias entre las probabilidades de que dicha variables tome valores superiores a 10, bajo ambas condiciones.

   e) Representa gráficamente, $f(2), f(3), f(4)$ y $f(10)$, en ambas situaciones, para darte una idea de cómo decrecen ambas distribuciones.

6. Sea X una v.a con distribución $\gamma(a, r)$ y c una constante positiva. Demuestra que la v.a cX sigue una distribución $\gamma(\frac{a}{c}, r)$.

   Aplicación: Si la v.a. cuantía por reclamación sigue una distribución γ, y estaba expresada en años anteriores en pesetas, ¿cuál es la distribución de la misma variable si la expresamos en euros?

7. La cuantía por reclamación de cada uno de n siniestros independientes, sigue una distribución exponencial de parámetro a.

   a) Calcula la distribución de la v.a. *cuantía total*, definida como la suma de las cuantías de los n siniestros.

   b) Calcula la distribución de la v.a. *cuantía media*, definida como la media aritmética de las n cuantías. Calcula su media y su varianza.

8. Una opción de compra (*call*) sobre un activo es un contrato por el cual el poseedor de la opción adquiere el derecho a comprar el activo (*subyacente*) en una fecha de vencimiento (T) y a un precio (p) fijado de antemano (*precio de ejercicio*).

   Considera un inversor de una *call* que paga hoy 1,7 euros por el derecho a comprar en T un activo $BVBV$ a 28 euros. Obviamente si el activo vale en T, menos de 28 euros, el inversor no ejercerá su derecho a compra y perderá el euro con siete; mientras que si el activo en T vale más, el inversor ejercerá la opción de compra y comprará el activo.

   Sobre la distribución del precio X del activo $BVBV$ en T, sabemos que

$$\begin{aligned} P(X < 28) &= 0{,}3 \\ E(X|X > 28) &= \int_{28}^{\infty} xf(x)dx = 21{,}6 \end{aligned}$$

   a) Define la v.a. B:*beneficio proporcionado por la opción*, y determina sus valores en función del valor del activo en T.

   b) Calcula en beneficio esperado proporcionado por la opción, $E(B)$.

   c) Si el precio del activo BVBV hoy es de 27.64 euros, ¿cuál ha de ser su valor esperado en T, para que el beneficio esperado por cada euro sea el mismo, invirtiendo en opciones que directamente en el mismo activo subyacente?

Capítulo 4

Distribuciones compuestas

En los dos temas anteriores, hemos visto por separado el comportamiento de dos variables aleatorias, a saber, número de siniestros y cuantía por reclamación. Sin embargo, en la práctica, éstas son dos variables que aparecen simultáneamente, y que tenemos que estudiar de forma conjunta. La *cuantía por reclamación, condicionada a la ocurrencia de un siniestro*, puede tener características distintas de la *cuantía por reclamación, condicionada a la ocurrencia de k siniestros* o en general de la *cuantía por reclamación* incondicionada al número de siniestros, por ejemplo.

Además haciendo extensivo este estudio, veremos que a la hora de modelizar el comportamiento de una v.a., como una distribución concreta, ésta puede depender de parámetros, que a su vez sean variables aleatorias. Este caso también se enmarca dentro del estudio de las *distribuciones de probabilidad compuestas*.

4.1. Convolución de variables

Hasta ahora, cuando queríamos estudiar la distribución de la suma de k v.a., $X_1 + ... + X_k$, independientes, realizábamos un razonamiento basado en la función característica, generatriz de momentos o cumulativa.

En esta sección vamos a ver un razonamiento alternativo basado en sus funciones de cuantía o densidad, dependiendo de si se trata de v.a. discretas o continuas.

Consideremos, sin pérdida de generalidad dos v.a. X e Y, independientes entre sí, discretas, que en particular tienen su soporte en el conjunto de los enteros no negativos, $\mathbf{Z}^+$. Así, sus respectivas funciones de cuantía vendrán dadas por:

$$\begin{aligned} P(X = j) &= a_j \\ P(Y = k) &= b_k \end{aligned}$$

para $j, k = 0, 1, 2, ...$ Entonces la función de cuantía conjunta, que nos da la probabilidad del suceso $(X = j, Y = k)$, será

$$P(X = j, Y = k) = P(X = j)P(Y = k) = a_j b_k$$

dada la independencia entre ambas variables. Consideramos ahora la v.a. suma, $Z = X + Y$, donde para todo $r \in \mathbf{Z}^+$, el suceso

$$\{Z = r\} = \{X + Y = r\} = \cup_{j=0}^{r} \{X = j, Y = r - j\}$$

De nuevo por la independencia entre X e Y,

$$\begin{aligned} P(Z = r) = P(X + Y = r) &= \sum_{j=0}^{r} P(X = j, Y = r - j) = \\ &= \sum_{j=0}^{r} a_j b_{r-j} = c_r \end{aligned} \tag{4.1}$$

para $r \in \mathbf{Z}^+$.

A la sucesión $\{c_r\} = \sum_{j=0}^{r} a_j b_{r-j}$, se le denomina convolución de las sucesiones $\{a_r\}$, $\{b_r\}$, y se denota como $\{c_r\} = \{a_r\} * \{b_r\}$. Esta sucesión nos da la distribución de la v.a. Z, suma de las variables independientes X e Y a través de la expresión (4.1).

Ejemplo 4.1 *Sean X e Y v.a. independientes con distribución de Poisson de parámetro λ. Es decir cada una de ellas tiene función de cuantía dada por:*

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k \in \mathbf{Z}^+$$

La función de cuantía de $Z = X + Y$ es la convolución de ambas y viene dada por:

$$\begin{aligned} P(Z = r) &= P(X + Y = r) = \sum_{j=0}^{r} P(X = j)P(Y = r - j) = \\ &= \sum_{j=0}^{r} \frac{e^{-\lambda}\lambda^j}{j!} \frac{e^{-\lambda}\lambda^{r-j}}{(r-j)!} = \frac{e^{-2\lambda}\lambda^r}{r!} \sum_{j=0}^{r} \frac{r!}{j!(r-j)!} = \\ &= \frac{e^{-2\lambda}\lambda^r}{r!} \sum_{j=0}^{r} \binom{r}{j} = \frac{e^{-2\lambda}\lambda^r}{r!}(1+1)^r = \frac{e^{-2\lambda}(2\lambda)^r}{r!} \end{aligned}$$

para $r = 0, 1, 2, ...$, es decir Z tiene distribución de Poisson de parámetro 2λ.

En el caso de v.a. independientes y continuas, denotemos a sus respectivas funciones de distribución por $F(x)$ y $G(y)$. De nuevo consideremos la v.a. suma, $Z = X + Y$. La función de distribución de Z, se puede calcular como:

$$H(z) = P(Z \leq z) = P(X + Y \leq z)$$

Esta es la probabilidad del recinto $\{X+Y \leq z\}$, en el plano (X, Y), y se puede obtener integrando sobre él la función de densidad conjunta de la v.a (X, Y). Por la independencia, dicha función de densidad es producto de marginales, es decir:

$$P(X + Y \leq z) = \int\int_{x+y\leq z} f(x)g(y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f(x)g(y)dydx$$

Así,

$$\begin{aligned} H(z) &= P(X+Y \leq z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f(x)g(y)dydx = \\ &= \int_{-\infty}^{\infty} f(x)(\int_{-\infty}^{z-x} g(y)dy)dx = \int_{-\infty}^{\infty} f(x)G(z-x)dx \end{aligned}$$

o análogamente, intercambiando el papel de las variables X e Y,

$$\begin{aligned} H(z) &= P(X+Y \leq z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f(x)g(y)dxdy = \\ &= \int_{-\infty}^{\infty} g(y)(\int_{-\infty}^{z-y} f(x)dx)dy = \int_{-\infty}^{\infty} g(y)F(z-y)dy \end{aligned}$$

Es decir,

$$\begin{aligned} H(z) &= \int_{-\infty}^{\infty} f(x)G(z-x)dx = \\ &= \int_{-\infty}^{\infty} g(y)F(z-y)dy \end{aligned}$$

Admitiendo la derivabilidad de las funciones de distribución y por lo tanto la existencia de las funciones de densidad, se obtiene la función de densidad de Z,

$$\begin{aligned} h(z) &= H'(z) = \int_{-\infty}^{\infty} f(x)\frac{\partial G(z-x)}{z}dx \\ &= \int_{-\infty}^{\infty} g(y)\frac{\partial F(z-y)}{z}dy \end{aligned}$$

de donde

$$h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)dx = \int_{-\infty}^{\infty} g(y)f(z-y)dy \tag{4.2}$$

La expresión (4.2) generaliza la presentada para el caso de distribuciones discretas en (4.1).

Ejemplo 4.2 *Sean X e Y v.a. independientes con distribución exponencial de parámetro a (también se suele decir de media $\frac{1}{a}$). Es decir cada una de ellas tiene función de densidad dada por:*

$$f(x) = ae^{-ax}, x \in \mathbf{R}^+$$

La v.a. $Z = X + Y$ tendrá como función de densidad la convolución de ambas. Por lo tanto:

$$\begin{aligned} h(z) &= \int_{-\infty}^{\infty} f(x)g(z-x)dx = a^2\int_0^z e^{-ax}e^{-a(z-x)}dx = \\ &= a^2e^{-az}\int_0^z dx = a^2ze^{-az}, \quad z \in \mathbf{R}^+ \end{aligned}$$

función de densidad que corresponde a una $\gamma(a,2)$, resultado que ya esperábamos.

Cualquiera de los dos ejemplos de aplicación de las fórmulas de convolución presentadas ya nos eran familiares, y podíamos llegar a ellos a través del cálculo de las funciones características. Sin embargo hay veces que la propiedad de convolución nos permite obtener funciones de cuantía o de densidad de determinados tipos de sumas, que no son fáciles de abordar mediante el uso de la función característica.

4.2. Algunas distribuciones compuestas

Consideremos una v.a. X cuya función de probabilidad, ya sea función de cuantía o de densidad, dependa de un parámetro θ_0; es decir $P(x;\theta_0)$. Ahora bien, si el parámetro θ es además una v.a. con su función de probabilidad denotada por $P(\theta)$. Realmente cuando escribamos $P(x;\theta_0)$ nos estaremos refiriendo la la probabilidad de que X el valor x condicionada o sabiendo que el valor de θ es fijo y vale θ_0, es decir $P(x|\theta=\theta_0)$.

Definición 4.1 *Se denomina distribución compuesta de X en relación a θ, a la distribución de X incondicionada al valor de θ, que se obtiene directamente del teorema de la partición, como*

$$P(X=x)=\sum_{\theta}P(x|\theta)P(\theta) \tag{4.3}$$

si X y θ son v.a. discretas y como

$$f(x)=\int_{\theta}f(x|\theta)f(\theta)d\theta \tag{4.4}$$

si tanto X como θ son v.a. continuas.

Como veremos en el siguiente ejemplo también son frecuentes las combinaciones de v.a. discretas y continuas.

Ejemplo 4.3 *Sea X una v.a. con distribución de Poisson de parámetro λ, donde dicho parámetro se comporta a su vez como una v.a. $\gamma(a,r)$.*

Es decir:

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

y

$$f(\lambda)=\frac{a^r}{\Gamma(r)}\lambda^{r-1}e^{-a\lambda}, \quad \lambda\geq 0$$

y $f(\lambda)=0$, en otro caso; donde $\Gamma(r)=\int_0^\infty x^{r-1}e^{-x}dx$.

Por lo tanto la distribución de X incondicionada a λ es:

$$P(x)=\int_0^\infty P(x|\lambda)f(\lambda)d\lambda=\frac{a^r}{\Gamma(r)k!}\int_0^\infty e^{-\lambda(a+1)}\lambda^{k+r-1}d\lambda \tag{4.5}$$

Sabiendo que

$$\frac{(a+1)^{k+r}}{\Gamma(k+r)} \int_0^\infty \lambda^{k+r-1} e^{-(a+1)\lambda} d\lambda = 1$$

puesto que es la función de densidad de una v.a. $\gamma(a+1, k+r)$ *y sabiendo además que* $\Gamma(k+r) = (k+r-1)\Gamma(k+r-1)$*; en la expresión (4.5), si* $k+r \in \mathbf{Z}^+$*, tenemos:*

$$P(x) = \frac{a^r}{\Gamma(r)k!} \frac{\Gamma(k+r)}{(a+1)^{k+r}} = \left(\begin{array}{c} r+k-1 \\ k \end{array} \right) \left(\frac{a}{a+1} \right)^r \left(\frac{1}{a+1} \right)^k$$

que corresponde a una función de cuantía de una v.a. binomial negativa de parámetros $m = r$, $p = \frac{a}{(a+1)}$.

Es decir otra forma alternativa de introducir la distribución *binomial negativa* es como la distribución compuesta de una v.a. Poisson, donde el parámetro λ sigue una distribución γ.

4.2.1. La distribución binomial compuesta

Sea X una v.a. con distribución binomial de parámetros (N, p), es decir $X \in B(N, p)$. Su función de cuantía viene dada por:

$$P(X = k; N) = \left(\begin{array}{c} N \\ k \end{array} \right) p^k (1-p)^{N-k}, \quad k \in \{0, ..., N\}$$

Si el número de repeticiones o exposiciones al riesgo, N, es también una v.a., la expresión anterior es en realidad la probabilidad condicionada al valor del parámetro N, es decir:

$$P(X = k; N) = P(X = k | N = n) = \frac{P(X = k, N = n)}{P(N = n)}$$

Si queremos obtener la distribución incondicionada de X, tendremos que conocer cuál es la distribución de la v.a. N. Consideremos el caso en el que N es una v.a. con distribución de Poisson de parámetro λ, así

$$P(N = n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad n = 0, 1, 2, ...$$

Entonces la distribución de X *incondicionada* a N, será:

$$\begin{aligned}
P(X=k) &= \sum_n P(k|N=n)P(N=n) = \quad (k \leq n) \\
&= \sum_{n=k}^{\infty} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} \frac{e^{-\lambda}\lambda^n}{n!} = \\
&= e^{-\lambda}\lambda^k \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} \frac{p^k(1-p)^{n-k}\lambda^{n-k}}{n!} = \\
&= \frac{e^{-\lambda}(p\lambda)^k}{k!} \sum_{n=k}^{\infty} \frac{(1-p)^{n-k}\lambda^{n-k}}{(n-k)!} = \\
&= \frac{e^{-\lambda}(p\lambda)^k}{k!} \sum_{n=0}^{\infty} \frac{(q\lambda)^n}{(n)!} = \frac{e^{-\lambda}(p\lambda)^k}{k!} e^{q\lambda} = \frac{e^{-p\lambda}(p\lambda)^k}{k!}, \quad k = 0, 1, ...
\end{aligned}$$

función de cuantía que corresponde a una distribución de Poisson de parámetro $p\lambda$. Es decir, la distribución binomial compuesta con una Poisson de parámetro λ es una Poisson $\lambda \cdot p$.

4.2.2. La distribución de Poisson compuesta

Sea X una v.a de Poisson de parámetro λ. Es decir

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

donde a su vez el parámetro λ es una v.a., por ejemplo continua y con función de densidad $f(\lambda)$.

La distribución compuesta de X, vendrá dada por:

$$P(X = k) = \int_0^{\infty} P(X = k|\lambda) f(\lambda) d\lambda = \int_0^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} f(\lambda) d\lambda$$

En el caso particular de que λ siga una distribución γ, ya vimos en el ejemplo 4.3 que la distribución compuesta de X es una binomial negativa.

4.3. La teoría de los valores extremos

Consideremos de forma general que la cuantía por reclamación por siniestro es una v.a. S que tiene como función de distribución $F(x)$. Así, para un valor cualquiera s, $F(s) = P(S \leq s)$, nos da la probabilidad de que la cuantía o el coste del siniestro sea inferior o igual a s.

El número de siniestros por su parte es otra v.a. X.

En muchos casos prácticos nos interesa conocer la probabilidad de que, a lo largo de un tiempo determinado, el coste mayor de los correspondientes a los siniestros ocurridos no supere una cierta cantidad s.

Es decir nos interesa buscar la distribución de probabilidad del *máximo de las cuantías* (o coste de la reclamación de mayor cuantía), *incondicionada al número de siniestros que han podido ocurrir.*

Partimos de la hipótesis de que los siniestros ocurren de manera independiente y la cuantía de todos ellos, S, tiene la misma distribución, con función de distribución $F(\cdot)$.

Denotaremos por $\phi(s)$ a la función de distribución buscada, esto es, la función de distribución de *la cuantía con mayor coste* ($= \max S$).

Claramente, de acuerdo con el teorema de la partición,

$$\phi(s) = P(\max S \leq s) = \sum_{j=0}^{\infty} P(\max S \leq s | X = j) \cdot P(X = j)$$

Además,

$P(\max S \leq s | X = j) = P(\max\{S_1, ..., S_j\} \leq s) = P(S_1 \leq s \cap ... \cap S_j \leq s) = [P(S \leq s)]^j$

de donde

$$\begin{aligned} \phi(s) &= P(X = 0) + P(X = 1)P(S \leq s) + P(X = 2)P(S \leq s)^2 + ... = \\ &= \sum_{j=0}^{\infty} P(X = j)P(S \leq s)^j = \sum_{j=0}^{\infty} P(X = j)F(s)^j \end{aligned}$$

Ahora bién, esta suma ha aparecido ya en probabilidad. Pensemos en una v.a. discreta X, para la que queremos calcular su función generatriz de probabilidad (también denominada función generatriz de momentos factoriales), $G_X(u)$, definida como, $G_X(u) = E(u^X)$. Si X es discreta:

$$G(u) = \sum_{j=0}^{\infty} P(X = j)u^j$$

Es decir,

$$\phi(s) = G(F(s))$$

para cualesquiera v.a. X y S y siempre la v.a. X tenga función generatriz de probabilidad.

Ejemplo 4.4 *Considera que el número de siniestros X es una v.a de Poisson de parámetro λ. Obtener la distribución del coste de la reclamación de mayor cuantía.*

En este caso:

$$\begin{aligned} \phi(s) &= \sum_{j=0}^{\infty} P(X = j)(F(s))^j = \\ &= \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!} F(s)^j = \sum_{j=0}^{\infty} \frac{e^{-\lambda}(\lambda F(s))^j}{j!} = \\ &= e^{-\lambda} e^{\lambda F(s)} = e^{-\lambda(1-F(s))} \end{aligned}$$

Siendo $F(s)$ la función de distribución de la cuantía en s.

Nota: Observa que la función generatriz de probabilidad de una v.a. $X \in \mathcal{P}(\lambda)$, es $G_X(u) = e^{-\lambda(1-u)}$, $\lambda > 0$.

En similares condiciones a las descritas anteriormente, en ocasiones nos interesará conocer la distribución de la *cuantía total incondicionada al número de siniestros*. Si denotamos por Z a la v.a. cuantía total, y por $F_Z(\cdot)$ a su función de distribución, tenemos:

$$\begin{aligned} F_Z(z) &= \sum_{j=0}^{\infty} P(Z \leq z|X = j)P(X = j) = \\ &= \sum_{j=0}^{\infty} P(S_1 + ... + S_j \leq z)P(X = j) = \\ &= \sum_{j=0}^{\infty} P(Z_j \leq z)P(X = j) = \sum_{j=0}^{\infty} F^{*j}(z)P(X = j) \end{aligned}$$

donde $Z_j = S_1 + ... + S_j$ es la cuantía total condicionada a la ocurrencia de j siniestros, es decir la suma de las cuantías de los j-ésimos primeros siniestros y F^{*j} es su función de distribución.

Entonces, la distribución de la *cuantía total* es la distribución compuesta de la cuantía con respecto al número de siniestros.

A veces, es sencillo obtener las funciones de distribución anteriores, y a partir de ellas, las de densidad. En otras ocasiones, simplemente no es necesario obtener la forma de su función de densidad y basta con obtener una relación de sus primeros momentos, media y varianza.

4.4. Relación entre momentos condicionados y no condicionados

Consideremos dos v.a. discretas X e Y. La función de cuantía condicionada de $X|Y = y$, viene dada por:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

Si ambas v.a. son continuas, la correspondiente función de densidad condicionada es:

$$f(x|Y = y) = \frac{f(x, y)}{f(y)}$$

Así, siempre que los denominadores sean distintos de cero, las ditribuciones condicionadas están bién definidas y podemos obtener sus momentos. Por simplicidad haremos las cuentas en el caso particular de v.a. discretas, auque las relaciones que obtengamos serán también válidas en el caso de v.a. continuas o mixtas.

En el caso de v.a. discretas, se obtiene la esperanza de la distribución condicionada de $X|Y = y$, como

$$E(X|Y = y) = \sum_x xP(X = x|Y = y)$$

De modo que su relación con la esperanza de la v.a. X incondicionada es:

$$\begin{aligned}
E(X) &= \sum_x xP(X=x) = \sum_x x \sum_y P(X=x, Y=y) = \\
&= \sum_x \sum_y xP(X=x|Y=y)P_Y(y) = \\
&= \sum_y \sum_x xP(X=x|Y=y)P_Y(y) = \\
&= \sum_y P_Y(y) \sum_x xP(X=x|Y=y) = \\
&= \sum_y P_Y(y)E(X|Y=y) = E_Y[E(X|Y)]
\end{aligned}$$

Así mismo si la varianza de la distribución de X condicionada a $Y=y$, es:

$$V(X|Y=y) = \sum_x x^2 P(X=x|Y=y) - [E(X|Y)]^2$$

entonces la varianza incondicionada de X se puede expresar en función de los dos momentos de la distribución condicionada como sigue:

Proposición 4.1 *Sean X e Y dos v.a. cualesquiera, tales que existen y son finitos los valores los momentos condicionados, $E(X|Y=y)$ y $V(X|Y=y)$. Entonces, se verifica:*

$$V(X) = E_Y[V(X|Y)] + V_Y[E(X|Y)]$$

Demostración:

$$\begin{aligned}
V(X) &= \sum_x x^2 P(X=x) - [E(X)]^2 = \\
&= \sum_y P_Y(y) \sum_x x^2 P(X=x|Y=y) - [E_Y(E(X|Y))]^2 = \\
&= E_Y(E(X^2|Y)) - [E_Y(E(X|Y))]^2
\end{aligned}$$

sumamos y restamos $E_Y([E(X|Y)]^2)$,

$$\begin{aligned}
V(X) &= E_Y(E(X^2|Y)) - E_Y([E(X|Y)]^2) + E_Y([E(X|Y)]^2) - [E_Y(E(X|Y))]^2 = \\
&= E_Y[E(X^2|Y) - [E(X|Y)]^2] + V_Y(E(X|Y)) = \\
&= E_Y(V(X|Y)) + V_Y(E(X|Y))
\end{aligned}$$

Las relaciones anteriores son válidas en cualquier contexto de probabilidad. En particular nos resultarán de gran utilidad para determinar valores esperados y varianzas de distribuciones incondicionadas, sin necesidad de obtener de forma explícita la forma completa de la distribución compuesta.

Ejemplo 4.5 *Considera que el número de siniestros que suceden en determinado tipo de riesgo sigue una distribución de Poisson de parámetro λ, y que la cuantía por reclamación de cada uno de ellos sigue una distribución exponencial de parámetro a. Bajo la hipótesis de independencia tanto entre cuantías como entre siniestros, vamos a obtener la media y la varianza de la v.a.* cuantía total.

En primer lugar si llamamos $Z_k = Y_1 + ... + Y_k$ a la cuantía total para un número de siniestros k, donde Y_i representa la cuantía por reclamación en cada siniestro, sabemos que la distribución de Z_k es $\gamma(a, k)$. Por lo tanto $E(Z_k) = \frac{k}{a}$ y $V(Z_k) = \frac{k}{a^2}$.

Por otro lado si el número de siniestros X, sigue una distribución de Poisson, tenemos que la media incondicionada de Z, es

$$E(Z) = E_X(E(Z|X = k)) = E_X(\frac{k}{a}) = \frac{1}{a}\lambda$$

y la varianza incondicionada de Z, será

$$\begin{aligned} V(Z) &= E_X(V(Z|X = k)) + V_X(E(Z|X = k)) = E_X(\frac{k}{a^2}) + V_X(\frac{k}{a}) = \\ &= \frac{1}{a^2}\lambda + \frac{1}{a^2}\lambda = \frac{2\lambda}{a^2} \end{aligned}$$

4.5. Mixtura de distribuciones normales

Vamos a ver en esta sección distintos casos de distribuciones compuestas que se obtienen a partir de distribuciones normales donde los parámetros varían.

Consideremos una v.a X que sigue una distribución normal de parámetros (m, σ^2), tal que el parámetro σ a su vez toma el valor σ_1 con probabilidad p y el valor σ_2 con probabilidad $q = 1-p$.

Tenemos por lo tanto que la distribución de X condicionada al valor de su varianza es conocida; su función de densidad es:

$$f(x|\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-m)^2}{2\sigma^2}}, \quad x \in \mathbf{R}$$

Lo que queremos es determinar *la distribución compuesta de X*; es decir su distribución incondicionada.

Su función de densidad, $f(x)$, se obtiene como:

$$\begin{aligned} f(x) &= f(x|\sigma = \sigma_1)P(\sigma = \sigma_1) + f(x|\sigma = \sigma_2)P(\sigma = \sigma_2) = \\ &= p\frac{1}{\sqrt{2\pi}\sigma_1} e^{\frac{-(x-m)^2}{2\sigma_1^2}} + (1-p)\frac{1}{\sqrt{2\pi}\sigma_2} e^{\frac{-(x-m)^2}{2\sigma_2^2}} \end{aligned}$$

En ocasiones esta función de densidad no es manejable y lo que nos interesa es conocer simplemente sus valores típicos: media, varianza, asimetría y curtosis.

En cuanto a la media y la varianza, utilizando las relaciones vistas en secciones anteriores, se obtiene:

$$\begin{aligned}
E(X) &= E_\sigma(E(X|\sigma)) = E_\sigma(m) = m \\
V(X) &= E_\sigma(V(X|\sigma)) + V_\sigma(E(X|\sigma)) = \\
&= E_\sigma(\sigma^2) + V_\sigma(m) = {\sigma_1}^2 p + \sigma_2^2(1-p) + 0 = p\sigma_1^2 + (1-p)\sigma_2^2
\end{aligned}$$

Si lo que nos interesa es información sobre la asimetría y la curtosis, podríamos calcular los cumulantes k_3 y k_4. Para ello debemos obtener primero la función cumulativa, $\mu_X(u) = ln\alpha_X(u)$.

La función generatriz de momentos de X, $\alpha_X(u)$, es fácil de obtener a partir de la relación de las distribuciones condicionadas con la distribución normal.

En particular, la función generatriz de momentos $\alpha_X(u)$, es:

$$\begin{aligned}
\alpha_X(u) &= E(e^{uX}) = \int e^{ux} f(x)dx = p\int e^{ux} f(x|\sigma=\sigma_1)dx + \\
&+ (1-p)\int e^{ux} f(x|\sigma=\sigma_2)dx = \\
&= pe^{\frac{\sigma_1^2 u^2}{2}+mu} + (1-p)e^{\frac{\sigma_2^2 u^2}{2}+mu} = \\
&= e^{mu}[pe^{\frac{\sigma_1^2 u^2}{2}} + (1-p)e^{\frac{\sigma_2^2 u^2}{2}}]
\end{aligned}$$

Utilizando del desarrollo en serie de la función, e^x.

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ...$$

e igualando los coeficientes de términos de igual grado en el desarrollo en serie de la función generatriz,

$$\alpha_X(u) = 1 + \alpha_1 u + \alpha_2 \frac{u^2}{2!} + \alpha_3 \frac{u^3}{3!} + ... \tag{4.6}$$

podemos obtener los momentos ordinarios α_k para la variable X. Concretamente:

$$\begin{aligned}
\alpha_X(u) &= pe^{\frac{\sigma_1^2 u^2}{2}+mu} + (1-p)e^{\frac{\sigma_2^2 u^2}{2}+mu} = \\
&= e^{mu}[pe^{\frac{\sigma_1^2 u^2}{2}} + (1-p)e^{\frac{\sigma_2^2 u^2}{2}}] = \\
&= e^{mu}[p(1 + \frac{\sigma_1^2 u^2}{2} + \frac{\sigma_1^4 u^4}{2^2 2!} + \frac{\sigma_1^6 u^6}{2^3 3!} + ...) + \\
&+ (1-p)(1 + \frac{\sigma_2^2 u^2}{2} + \frac{\sigma_2^4 u^4}{2^2 2!} + \frac{\sigma_2^6 u^6}{2^3 3!} + ...)] = \\
&= e^{mu}[1 + \frac{p\sigma_1^2 + (1-p)\sigma_2^2}{2}u^2 + \frac{p\sigma_1^4 + (1-p)\sigma_2^4}{2^2 2!}u^4 + \frac{p\sigma_1^6 + (1-p)\sigma_2^6}{2^3 3!}u^6 + ...]
\end{aligned}$$

Tomando logaritmo neperiano en la anterior expresión, se obtiene:

$$\begin{aligned}\mu_X(u) &= ln\alpha_X(u) = mu + ln(1 + \frac{p\sigma_1^2 + (1-p)\sigma_2^2}{2}u^2 + \frac{p\sigma_1^4 + (1-p)\sigma_2^4}{2^2 2!}u^4 + \\ &+ \frac{p\sigma_1^6 + (1-p)\sigma_2^6}{2^3 3!}u^6 + ...) = mu + ln(1+Y)\end{aligned}$$

donde llamamos

$$Y = \frac{p\sigma_1^2 + (1-p)\sigma_2^2}{2}u^2 + \frac{p\sigma_1^4 + (1-p)\sigma_2^4}{2^2 2!}u^4 + \frac{p\sigma_1^6 + (1-p)\sigma_2^6}{2^3 3!}u^6 + ...$$

Utilizando ahora el desarrollo en serie de la función $ln(1+x)$, que es:

$$ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}x^k}{k} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ...$$

trataremos de buscar los coeficientes de los términos en $\frac{u^3}{3!}$ y $\frac{u^4}{4!}$. Observando que en el desarrollo de Y, aparecen únicamente potencias pares, se deduce que los términos en u^3, tienen coeficiente nulo, y por la tanto, $k_3 = 0$.

Calculando los coeficientes de los primeros términos de $Y - \frac{1}{2}Y^2$ e identificando todos aquellos que acompañan a $\frac{u^4}{4!}$, se obtiene:

$$\begin{aligned}k_4 &= 4!\left\{\frac{1}{2^2 2!}[p\sigma_1^4 + (1-p)\sigma_2^4] - \frac{1}{2}(\frac{p\sigma_1^2 + (1-p)\sigma_2^2}{2})^2\right\} = \\ &= 4!\frac{1}{8}(1-p)p(\sigma_1^2 - \sigma_2^2)^2 = \\ &= 3p(1-p)(\sigma_1^2 - \sigma_2^2)^2\end{aligned}$$

Así, los cumulantes nos proporcionan los valores típicos:

$$\begin{aligned}k_1 &= E(X) = E_\sigma(E(X|\sigma)) = E_\sigma(m) = m \\ k_2 &= V(X) = E_\sigma(V(X|\sigma)) + V_\sigma(E(X|\sigma)) = \\ &= E_\sigma(\sigma^2) + V_\sigma(m) = {\sigma_1}^2 p + \sigma_2^2(1-p) + 0 = p\sigma_1^2 + (1-p)\sigma_2^2 \\ k_3 &= \mu_3 = E(X-m)^3 = 0 \\ k_4 &= \mu_4 - 3\sigma^4 = E(X-m)^4 - 3\sigma^4 = 3p(1-p)(\sigma_1^2 - \sigma_2^2)^2\end{aligned}$$

4.6. Tarea 4

Se enumeran a continuación un conjunto de ejercicios de ensayo de las competencias correspondientes a este tema; en particular la segunda y tercera de las formuladas en el programa de la asignatura:

Describir y analizar situaciones sencillas de riesgo en problemas relacionados con el mundo actuarial o de las finanzas y expresarlas formalmente en términos de variables aleatorias y sus distribuciones.

Modelizar la incertidumbre de las situaciones sencillas de azar en el ámbito del negocio financiero y asegurador, mediante formulaciones matemáticas basadas en la teoría de la probabilidad. Deducir soluciones en cuanto a la prima de riesgo, función de ingreso, riesgo de ruina, etc.

1. Sean X e Y dos v.a. independientes con distribución exponencial de parámetros a y b, $a \neq b$, respectivamente.

   a) Demuestra que la función de densidad de la suma, $Z = X + Y$ es

   $$f(z) = \frac{ab}{(b-a)}(e^{-az} - e^{-bz}), \quad z \geq 0$$

   b) Para $a = 3$ y $b = 4$, calcula la probabilidad de que $Z > 8$.

   c) Calcula $P(Z < 10|Z > 8)$.

2. Considera una compañía de seguros en la que el número de siniestros es una v.a discreta, X, con distribución binomial negativa de parámetros m, p; mientras que las cuantías por reclamación de los distintos siniestros, son v.a. Y, independientes y con distribución $\gamma(a, r)$.

   a) Obtener la distribución de la cuantía total, incondicionada al número de siniestros, o lo que es lo mismo la distribución compuesta de la cuantía total.

   b) Obtener la media y la varianza de la distribución compuesta de la cuantía total.

3. Considera que el número de siniestros es una v.a. con distribución binomial negativa, de parámetros m, p, y que la cuantía por reclamación de cada uno de ellos, es una v.a. exponencial de parámetro a. Bajo la hipótesis de independencia tanto entre cuantías como entre siniestros, calcula la media y la varianza de la cuantía total por ese tipo de riesgo.

4. Calcula la media y la varianza de una v.a. X con distribución binomial negativa de parámetros $m = r$, $p = \frac{a}{a+1}$, teniendo en cuenta que X es la distribución compuesta de una v.a. de Posison de parámetro λ donde el parámetro sigue una distribución $\gamma(a, r)$.

5. Considera una compañía de seguros que contrata 200 pólizas donde el número de siniestros X y la cuantía por reclamación en cada uno de ellos, se pueden modelizar como en el Ejemplo 3.5 de las notas de clase, donde la media de la exponencial es 20 euros y el parámetro de la distribución de Poisson es 0.05. Si la prima que se paga por cada póliza es de 6 euros, calcula:

   a) La probabilidad de que en una determinada póliza se produzca una reclamación y que además la cuantía por reclamación sea superior a 6 euros.

   b) La probabilidad de que la cuantía más elevada, para una póliza en concreto, sea menor que 60 euros.

c) La probabilidad de que la cuantía más elevada, sobre el total de las 200 pólizas, sea menor que 60 euros.

d) Media y varianza de la cuantía total por póliza.

e) Media y varianza de la cuantía total por reclamaciones de las 200 pólizas.

f) La probabilidad aproximada de que el beneficio neto de la compañía sea superior a 600 euros.

g) La probabilidad aproximada de que la compañía pierda dinero en esta operación con las 200 pólizas.

6. Cierta compañía tiene 100 pólizas contratadas en 100 edificios. 25 de ellos poseen 5 viviendas, otros 25 poseen 15 viviendas y el resto posee 10 viviendas. Se ha estimado que el número de reclamaciones anuales por incendio y edificio sigue una ley de Polya con parámetro de contagio $\delta = 0{,}2$, y que la probabilidad inicial de incendio en una vivienda es $p = 0{,}01$.

   a) Sabiendo que cierto edificio tiene 10 viviendas, calcula la probabilidad de que se efectúen en él 2 reclamaciones por incendio.

   b) Idem con un edificio de 5 viviendas.

   c) Calcula medias y varianzas del número de reclamaciones, condicionadas al número de viviendas por edificio.

   d) Calcula la media y la varianza del número de reclamaciones por edificio, incondicionada al número de viviendas que posea.

   e) Calcula la probabilidad aproximada de que el total de reclamaciones no sea superior a 10.

7. Cierto siniestro tiene una probabilidad de ocurrir de 0,005, en cada exposición al riesgo del tomador de la póliza. Se sabe que el número de exposiciones al riesgo sigue una distribución de Poisson de parámetro $\lambda = 5$.

   a) Calcula la distribución compuesta del número de siniestros.

   b) Calcula la probabilidad de que no haya ningún siniestro si ha habido 20 exposiciones al riesgo.

   c) Calcula la probabilidad incondicionada de que no haya ningún siniestro.

8. Una compañía aseguradora dispone de dos tipos de asegurados A y B. El 30 % son de tipo A y el resto de tipo B. Para los del tipo A, el número de siniestros sigue una distribución de Poisson de parámetro $\lambda_A = 0{,}5$, y para los de tipo B de parámetro $\lambda_B = 0{,}3$.

   a) Obtén la función de cuantía del número de siniestros incondicionada al tipo de asegurado.

   b) Calcula las probabilidades correspondientes de 0,1 y 2 siniestros.

   c) Obtén la media y la varianza de la distribución incondicionada.

d) Si la empresa tiene 1000 pólizas, por lo tanto 300 de tipo A y 700 de tipo B, y hay independencia entre siniestros, calcula de forma aproximada la probabilidad de que el número total de siniestros no exceda de 40. ¿En qué te basas para realizar esta aproximación?

9. Considera una póliza de una comunidad que cubre dos tipos de siniestros, A y B. El número de reclamaciones por el siniestros A sigue una distribución de Poisson de parámetro $\lambda = 0{,}01$, mientras que el número de reclamaciones de tipo B sigue una distribución binaria de parámetro $p = 0{,}001$. Bajo la hipótesis de independencia entre ambos tipos de siniestros.

   a) La cuantía que se paga por cada reclamación del tipo A es de 10000 euros, y de 100000 euros por cada una de tipo B. Calcula el valor esperado de la cuantía total por póliza.

   b) Obtén la varianza de la cuantía total por póliza.

   c) Si se tienen 100 pólizas iguales, estima el valor de la prima para que se alcance un beneficio final total al menos de 30000 euros con probabilidad 0.8. (Utiliza el T.C.L.)

   d) Resuelve de nuevo los tres apartados anteriores, bajo la hipótesis de que la cuantía por reclamación es una v.a. exponencial de media 10000 para las pólizas de tipo A y de media 100000 para las de tipo B.

10. Una empresa aseguradora sabe que la cuantía por reclamación de determinado tipo de póliza tiene una distribución de Pareto de parámetros $k = 10$ euros y $\alpha = 2{,}1$. La empresa cubre las reclamaciones de hasta 1000 euros de cuantía, y cuantías superiores a dicha cantidad son cubiertas por otra compañía reaseguradora.

    a) Para cada reclamación, calcula la probabilidad de que la compañía reaseguradora no tenga que cubrir su cuantía.

    b) Siendo el número de reclamaciones una v.a. de Poisson de parámetro $\lambda = 10$. Calcula la probabilidad de que la compañía reaseguradora no tenga que cubrir ninguna reclamación, incondicionada al número de las mismas.

11. Considera dos tipos de siniestros A y B independientes. La probabilidad de reclamar por A es 0.1 y la de reclamar por B es 0.01, y como mucho se admite una reclamación por cada uno de ellos. La cuantía por reclamación, para ambos tipos de siniestro sigue una distribución exponencial de media 100 euros. Una compañía oferta una póliza compuesta para cubrir ambos siniestros y cobra una prima por póliza de 30 euros. Se pide:

    a) Distribución del número de reclamaciones por póliza. Media y varianza.

    b) Probabilidad de que una póliza efectúe una reclamación y que su cuantía sea superior a los 5 euros.

    c) Media y varianza de la cuantía incondicionada al número de reclamaciones.

    d) Si la empresa posse 100 pólizas independientes, utiliza el T.C.L. para obtener la distribución de la cuantía total por reclamaciones.

    e) Probabilidad de que los beneficios de la compañía superen los 1000 euros.

12. Consideremos un activo tal que su precio mañana es una v.a con distribución normal $N(m = 3, \sigma = 2)$ con probabilidad 0.7 y sigue una distribución normal $N(m = 5, \sigma = 1)$ con probabilidad 0.3. Obtén la distribución incondicionada del precio del activo mañana, su media, varianza, asimetría y curtosis.

13. Considera un activo cuyo rendimiento es normal $N(m = 3, \sigma_1^2 = 30)$ con probabilidad p y sigue una distribución normal $N(m = 3, \sigma_2^2 = 100)$ con probabilidad $1 - p$. Obtén media y la varianza incondicionada del rendimiento del activo.

Capítulo 5

Decisión y riesgo

Considera un juego en el que el concursante puede ganar 100 euros con probabilidad $\frac{1}{5}$, 0 euros con probabilidad $\frac{2}{5}$ o perder 40 euros con probabilidad $\frac{2}{5}$. La ganancia esperada en este juego será:

$$\frac{1}{5}(100) + \frac{2}{5}(0) + \frac{2}{5}(-40) = 4$$

En general, cualquier juego de este tipo se puede representar por una v.a. X con una distribución de probabilidad específica. Hay que entender que un valor positivio de la variable representa una ganancia monetaria real para la persona y que un valor negativo de X representa una pérdida (o ganancia negativa). La ganancia esperada del juego es simplemente el valor de $E(X)$.

Aunque dos juegos muy distintos pueden tener la misma ganancia esperada, una persona que deba elegir entre ambos, tendrá preferencia por uno de ellos.

Por ejemplo un juego alternativo al anterior, podía ser aquel en el que en el que el concursante puede ganar 400 euros con probabilidad $\frac{1}{2}$, o perder 392 euros con probabilidad $\frac{1}{2}$. En este caso, la ganacia esperada en este segundo juego es la misma que en el anterior, 4 euros, aunque existe la misma probabilidad de que el concursante gane 400 o de que pierda 392. Veremos que esta segunda situación comporta un riesgo mayor.

5.1. Funciones de utilidad

Las *funciones de utilidad* surgen dentro de la *teoría de la utilidad* durante las décadas de 1930 y 1940 para describir preferencias de una persona entre juegos como los que acabamos de mostrar.

De acuerdo con esta teoría una persona preferirá un juego X para el que la esperanza de una cierta función $U(X)$ sea un máximo, antes que un juego para el que simplemente $E(X)$ sea un máximo. La función U se denomina *función de utilidad* y representa el valor que asigna cada persona a cada cantidad posible x, o lo que es lo mismo, el valor real para la persona de ganar la cantidad x.

Si la función de utilidad para el concursante anterior es U,

$$E[U(X)] = \frac{1}{5}U(100) + \frac{2}{5}U(0) + \frac{2}{5}U(-40)$$

y

$$E[U(Y)] = \frac{1}{2}U(400) + \frac{1}{2}U(-392)$$

La persona preferirá el juego para el que la *utilidad esperada de ganancia*, sea mayor. Formalmente, la función de utilidad de una persona se define como una función con la siguiente propiedad: Cuando la persona debe elegir entre dos juegos X e Y, preferirá X a Y, si $E[U(X)] > E[U(Y)]$, y será indiferente a X e Y si $E[U(X)] = E[U(Y)]$. Cuando la persona esté eligiendo entre más de dos juegos, elegirá aquél, X, para el que el valor $E[U(X)]$ sea máximo.

En un contexto general, las funciones de utilidad se suelen clasificar como lineales, cóncavas y convexas.

En el caso de funciones de utilidad lineales, por ejemplo, $U(x) = ax + b$, donde $a, b \in \mathbf{R}^+$, se cumple que para cualquier v.a. X: $E[U(X)] = aE(X) + b$. Por tanto para cualesquiera juegos X e Y, la desigualdad $E[U(X)] > E[U(Y)]$, se cumple si y sólo si se cumple $E(X) > E(Y)$.

En otras palabras, una persona con una función de utilidad lineal, elegirá un juego donde la ganancia esperada sea mayor.

Analicemos el caso de una función de utilidad logarítmica, por ejemplo, $U(x) = log(x + 400)$ para $x > -400$.Una persona que tenga esta función de utilidad no puede elegir un juego en el que exista alguna posibilidad de que su ganancia sea -400 o menos. Para los juegos X e Y definidos antes:

$$\begin{aligned} E[U(X)] &= \frac{1}{5}log(500) + \frac{2}{5}log(400) + \frac{2}{5}log(360) = \\ &= \frac{1}{5}2{,}698 + \frac{2}{5}2{,}6020 + \frac{2}{5}2{,}556 = 2{,}6028 \end{aligned}$$

y

$$E[U(Y)] = \frac{1}{2}log(800) + \frac{1}{2}log(8) = \frac{1}{2}2{,}9030 + \frac{1}{2}0{,}9030 = 1{,}903$$

Además la utilidad de no aceptar ningún juego es $U(0) = log(400) = 2{,}6020$. Puesto que $E[U(X)] > E[U(0)] > E[U(Y)]$, la persona elegiría el juego X. Si dicho juego no estuviera disponible, preferiría no jugar, antes de aceptar el juego Y.

Si la función de utilidad es cuadrática, por ejemplo, $U(x) = x^2$, para $x \geq 0$. Tendríamos que para los juegos definidos antes:

$$\begin{aligned} E[U(X)] &= \frac{1}{5}(100)^2 + \frac{2}{5}(0)^2 + \frac{2}{5}(-40)^2 = \\ &= \frac{1}{5}10000 + \frac{2}{5}0 + \frac{2}{5}1600 = 2640 \end{aligned}$$

y

$$E[U(Y)] = \frac{1}{2}(400)^2 + \frac{1}{2}(-392)^2 = \frac{1}{2}160000 + \frac{1}{2}153664 = 236832$$

Por lo tanto la persona elegiría el segundo juego, Y, antes que el primero X, y antes que no jugar, pues $E[U(0)] = 0$.

En el contexto actuarial, y en general en el económico, consideraremos que las funciones de utilidad son cóncavas, (como $-x^2$, $logx$, etc) o lo que es los mismo, que existe *aversión al riesgo*. Cuando la función de utilidad de una persona es lineal se dice que es *indiferente al riesgo*, mientras que si es convexa, se dice que es *aficionada al riesgo*.

5.2. Función de riesgo

El riesgo es considerado una media de desutilidad, puesto que la compañía preferirá situaciones menos arriesgadas a situaciones más arriesgadas.

Así conocer la función de riesgo, nos permitirá conocer el orden de preferencias de la compañía.

Una forma de definir el riesgo, es la basada en considerarlo como la dispersión de la variable que nos define el juego. La definición que consideraremos aquí, será la desviación típica, raíz cuadrada de la varianza, en nuestro caso de la cuantía por reclamación de los siniestros.

Así,

$$r = SD(X) = \sqrt{E(X - E(X))^2}$$

será el valor que trataremos de minimizar, cuando el valor esperado del juego, $E(X)$ sea dado.

La decisión en un contexto de riesgo medido en términos de desviaciones, será la de elegir aquel juego que con el mismo valor esperado tiene menor desviación típica como medida de la dispersión del riesgo.

En el contexto que carteras de seguros, fondos de inversión, etc, donde se trata con eventos que tienen probabilidades pequeñas de ocurrir, pero donde la cuantía toma en general valores altos (pérdidas elevadas), veremos que la medida del riesgo definida anteriormente, conduce a valores de éste muy elevados.

En el caso de los dos juegos planteados al principio del tema, para describir el primero, podemos utilizar la v.a. X, tal que

$$P(X = 100) = \frac{1}{5}, \quad P(X = 0) = \frac{2}{5}, \quad P(X = -40) = \frac{2}{5}$$

donde, la ganacia esperada es $E(X) = 4$ euros con un riesgo de $r_1 = SD(X) = 51{,}22$ euros.

En el caso del segundo juego, si empleamos una v.a. Y para describirlo, tenemos:

$$P(Y = 400) = \frac{1}{2}, \quad P(Y = -392) = \frac{1}{2}$$

donde, la ganacia esperada es la misma, $E(Y) = 4$ euros con un riesgo mucho mayor, $r_2 = SD(Y) = 396$ euros.

Estas dos situaciones ilustran el hecho de que, a igual beneficio esperado, en una situación de mayor riesgo hay una probabilidad más alta de obtener una mayor pérdida.

Dado el carácter introductorio de esta sección, no nos plantearemos de momento el diseño de carteras de seguros o fondos de inversión óptimos, sino que veremos, dada una deteminada cartera, de qué forma puede cubrirse la compañía contra estos riesgos elevados.

Dicho de otro modo, en el caso de carteras de seguros, calcularemos las primas que debe cobrar la compañía para cubrirse en esta situación de riesgo.

5.3. Utilidad con nivel de riesgo ϵ

Consideremos una función de utilidad definida sobre el espacio de posibles resultados Ω, es decir $\Omega = \{S\}$, tal que la v.a. X describe la utilidad de S, $X = U(S)$. Así podemos hablar de $P(X = U(S) \leq a)$ para un cierto valor $a \in \mathbf{R}$.

Definición 5.1 *Para un valor ϵ prefijado, se denomina utilidad con nivel de riesgo ϵ, al valor a tal que*

$$P(X = U(S) \leq a) \geq \epsilon$$

y tal que

$$P(X = U(S) \geq a) \leq 1 - \epsilon$$

El valor a es tal que la función de distribución de la utilidad, $H = F_X$, en él, es al menos ϵ. Es decir $a = H(S, \epsilon)$.

El siguiente resultado, conocido como *indiferencia con respecto al azar*, nos indica que da lo mismo trabajar con S que con $X = U(S)$.

Proposición 5.1 *Se verifica la siguiente igualdad:*

$$H(U(S)) = U(H(S))$$

Demostración:

Llamemos $a_1 = H(U(S), \epsilon)$ i.e. si $X = U(S)$,

$$P(U(X) \leq a_1) \geq \epsilon,$$

Si la función de utilidad, U es estrictamente monótona,

$$P(U(X) \leq a_1) = P(X \leq U^{-1}(a_1)) \geq \epsilon,$$

donde $a = U^{-1}(a_1)$ es la utilidad con nivel de riesgo ϵ, i.e. $a = H(S, \epsilon)$ luego,

$$a_1 = H(U(S), \epsilon) = U(a) = U(H(S, \epsilon))$$

A continuación estudiaremos un caso particular con gran interés en los estudios actuariales, la situación de ruina.

5.4. Probabilidad de ruina

Consideremos una compañía que se pone en marcha con una reserva económica inicial que denomiraremos W. Si X es la v.a. que representa la capacidad económica de la empresa en un determinado momento, la probabilidad de ruina, es la probabilidad de que la compañía deba cesar su actividad por no disponer de reserva alguna. Matemáticamente,

$$P(X < -W) = P(X + W < 0)$$

Cada momento o periodo de estudio se denomina *unidad de riesgo*, y si $X < -W$, decimos que se *ha producido la ruina* en dicha unidad de riesgo.

Consideremos conocida la distribución de X, en particular su función de distribución, F_X, así la probabilidad de ruina, π, es

$$\pi = F_X(-W) = P(X < -W) = P(X + W < 0)$$

Dada la definición de X, es habitual su caracter continuo. Mirando la expresión anterior, se puede interpretar que la reserva económica inicial con signo menos, es la utilidad con nivel de riesgo π. También se puede decir que dicha reserva es la desutilidad a ese nivel de riesgo.

Como nuestro interés se centrará en conocer la probabilidad de ruina, para ello, tendremos las siguientes opciones:

1. Conocer la función de distribución de la capacidad económica, F_X y calcular $\pi = F_X(-W)$.

2. Si no conocemos F_X de forma exacta o su expresión resulta muy complicada de manipular, y estamos en situación de emplear el T.C.L.; aproximar la distribución de X a $N(m, \sigma)$, $\pi \approx \phi(\frac{-W-m}{\sigma})$.

3. Si no conocemos F_X ni estamos en situación de emplear el T.C.L.; utilizaremos la cota que nos proporciona el siguiente resultado.

Proposición 5.2 *Sea $R(X)$ una función monótona decreciente, tal que $E(e^{R(X)}) \leq 1$. La probabilidad de ruina es siempre menor o igual que $e^{-R(-W)}$. Es decir,*

$$\pi \leq e^{-R(-W)}$$

Demostración:

$$\begin{aligned} 1 &\geq E(e^{R(X)}) = \int_{-\infty}^{\infty} e^{R(x)} dF_X(x) \geq \int_{-\infty}^{-W} e^{R(x)} dF_X(x) \geq \\ &\geq e^{R(-W)} \int_{-\infty}^{-W} dF_X(x) = e^{R(-W)} P(X < -W) = e^{R(-W)} \pi \end{aligned}$$

Despejando, de la desigualdad, se obtiene el resultado

$$\pi \leq e^{-R(-W)}$$

Veremos ahora la utilidad de la cota proporcionada por este resultado. Si identificamos la función $R(x)$, podremos acotar la probabilidad de ruina de forma sencilla.

Recordemos la expresión de la función generatriz de momentos de una v.a. continua X, esto es,

$$\alpha_X(u) = E(e^{uX}) = \int_{-\infty}^{\infty} e^{ux} f(x) dx$$

Para analizar el crecimiento o decrecimiento de esta función, analizamos la forma de la integral

$$\alpha_X(u) = \int_{-\infty}^{0} e^{ux} f(x) dx + \int_{0}^{\infty} e^{ux} f(x) dx$$

Sabemos que $\alpha_X(0) = 1$ y $\alpha'_X(0) = m > 0$, luego es creciente en 0. Así, si las probabilidades $P(X < 0)$ y $P(X > 0)$ son no nulas a la vez, $\alpha_X(u) \to \infty$, cuando $u \to \infty$ o cuando $u \to -\infty$, ya que el valor de la exponencial crece en ambos casos con u.

Por otra parte,

$$\alpha''_X(u) = E(X^2 e^{uX}) = \int_{-\infty}^{\infty} e^{ux} x^2 f(x) dx > 0, \quad \forall u$$

Esto quiere decir que $\alpha_X(u)$ tiene una representación parabólica con un único vértice (mínimo), en algún valor de $u < 0$, si el valor de la media es positivo. En este caso, $\alpha'_X(0) = m > 0$, y la función está creciendo en 0, con lo que el vértice es anterior, y por lo tanto negativo.

Como además sabemos que $\alpha(0) = 1$, ha de existir un valor $u^* < 0$ tal que $\alpha(u^*) = 1$.

Tomemos ahora como función $R(x) = u^* x = -|u^*| x$. Se trata de una función que es claramente decreciente de x, y además $E(e^{R(x)}) = E(e^{ux^*}) = \alpha(u^*) = 1$.

Utilizando esta función para la v.a. X, el resultado anterior, nos lleva a que

$$\pi \leq e^{-R(-W)} = e^{-|u^*| W}$$

expresión que es cierta para una v.a. X tal que $P(X < 0)$ y $P(X > 0)$ son no nulas, o lo que es lo mismo en cada unidad de riesgo hay probabilidad de beneficios positivos o negativos.

Definición 5.2 *Se denomina nivel de la probabilidad de ruina al valor α, calculado como*

$$\pi \leq \alpha = e^{-|u^*|W}$$

Lo interesante en las aplicaciones de la definición anterior, no es calcular el nivel de probabilidad de ruina dado el valor de u^*, sino al revés. Lo habitual es, fijar el nivel de probabilidad de ruina de una compañía, α, su reserva inicial W, y calcular el valor de u^*, donde

$$|u^*| = -\frac{ln\alpha}{W}, \quad u^* < 0$$

Este resultado nos proporciona el valor de u^*. Valor que no tiene sentido por sí mismo. Lo que sí es interesante es conocer cual es el valor de ingreso que tiene que generar una compañía para poder mantener pequeño su nivel de probabilidad de ruina. Es decir todavía tenemos que calcular este ingreso a partir del valor u^*.

Sin embargo, antes de obtener la función de ingreso, veamos las siguientes observaciones:

- El nivel de probabilidad de ruina, α, es bastante conservador, en el sentido de que $\pi << \alpha$. Como veremos en ejemplos en los que se puede calcular explícitamente π, este valor siempre es mucho menor que α.
- Por eso, la función de ingreso siempre nos proporcionará valores de los ingresos que mantienen acotada la probabilidad de ruina en el nivel deseado.

5.5. Función de ingreso

Sean X y Z dos v.a. que representan respectivamente el *montante de siniestralidad*, es decir, la cuantía total por reclamaciones, y *el beneficio* en la unidad de riesgo.

Si denotamos por I al ingreso, obtenemos, $Z = I - X$, donde I representa una cantidad fija, a provenir por la reserva inicial más las primas por pólizas.

Definición 5.3 *Llamamos recargo de seguridad al valor esperado del beneficio, $m_z = E(Z)$, calculado como*

$$m_z = E(Z) = I - m_X$$

Despejando el valor del ingreso, I, y sustituyendo en Z, el beneficio se puede expresar como una transformación lineal del montante de siniestralidad:

$$Z = m_Z + m_X - X$$

Si denotamos por $\alpha_X(u)$ y $\alpha_Z(u)$ a sus respectivas funciones generatrices de momentos, la relación entre ambas es:

$$\alpha_Z(u) = e^{(m_Z + m_X)u}\alpha_X(-u)$$

Si denotamos por $\mu_X(u) = ln\alpha_X(u)$ a la función cumulativa de X, se tiene que $\alpha_X(u) = e^{\mu_X(u)}$, de donde

$$\alpha_Z(u) = e^{(m_Z+m_X)u}e^{\mu_X(-u)}$$

Si además W representa la reserva inicial de la compañía, tenemos que existe $u^* < 0$, que cumple, $\alpha_Z(u^*) = 1$, y por tanto

$$(m_Z + m_X)u^* + \mu_X(-u^*) = 0$$

de donde

$$\frac{\mu_X(|u^*|)}{|u^*|} = m_Z + m_X$$

y, por lo tanto

$$m_Z = \frac{\mu_X(|u^*|)}{|u^*|} - m_X$$

Definición 5.4 *El cociente dado por*

$$\frac{\mu_X(u)}{u}$$

recibe el nombre de función de ingreso.

Por lo tanto, el recargo de seguridad, m_Z, resulta ser el valor de la función de ingreso en el punto u^* menos la cuantía media por siniestros, m_X.

Si sustituimos u^* por su valor, en términos del nivel de la probabilidad de ruina α y de la reserva inicial de la compañía, obtenemos el valor del recargo de seguridad, como

$$m_Z = \frac{\mu_X(\frac{|ln\alpha|}{W})}{\frac{|ln\alpha|}{W}} - m_X = I(W) - m_X$$

donde el ingreso se puede expresar también en función del nivel de probabilidad de ruina y de la reserva inicial

$$I(W) = \frac{\mu_X(\frac{|ln\alpha|}{W})}{\frac{|ln\alpha|}{W}}$$

Esta es la expresión del nivel de ingreso necesario para garantizar la supervivencia de la empresa a un nivel de probabilidad de ruina α, y con una reserva inicial W.

Observando la expresión anterior, podemos deducir que a mayor reserva inicial, el ingreso se puede disminuir y así obtener primas más competitivas.

Por tanto, la fusión de compañías es beneficiosa, y permite afrontar mejor situaciones más arriesgadas.

5.5.1. Desarrollo en serie de la función de ingreso

El numerador de la función de ingreso es una función cumulativa. Del desarrollo el serie de ésta, visto en, (**??**), donde

$$\begin{aligned} k_1 &= E(X) = m_X \\ k_2 &= V(X) = \sigma_X^2 \\ k_3 &= \mu_3 = E(X-m)^3 \\ k_4 &= \mu_4 - 3\sigma^4 = E(X-m)^4 - 3\sigma^4 \end{aligned}$$

se obtiene:

$$\frac{\mu_X(u)}{u} = m_X + \sigma_X^2 \frac{u}{2!} + \mu_3 \frac{u^2}{3!} + ...$$

Así el valor del recargo de seguridad, para un nivel de probabilidad de ruina α, y con reserva inicial W, $m_Z = I(W) - m_X$, admite el siguiente desarrollo en serie:

$$m_Z = \sigma_X^2 \frac{|ln\alpha|}{2!W} + \mu_3 \frac{|ln\alpha|^2}{3!W^2} + ...$$

5.5.2. Función de ingreso de variables compuestas

Consideremos ahora que la v.a. montante por siniestralidad corresponde a la cuantía por reclamación de k siniestros independientes e idénticamente distribuidos; es decir, $Y_1 + ... + Y_k$, donde el número de siniestros es de nuevo una v.a. X con función generatriz de momentos $\alpha_X(u)$.

La función generatriz de momentos de la distribución condicionada de $Y|X = k$, es:

$$\alpha_{Y|X=k}(u) = \alpha_{Y_1+...+Y_k}(u) = (\alpha_{Y_i}(u))^k = (\alpha_Y(u))^k$$

dado que las cuantías por las k reclamaciones son independientes y tienen la misma distribución.

Por lo tanto la función generatriz de momentos de la distribución compuesta de Y en relación a X, es:

$$\begin{aligned} \alpha_Y(u) &= E(e^{uY}) = E_X(E(e^{uY}|X=k)) = E_X(\alpha_{Y|X=k}(u)) = \\ &= E_X[\alpha_Y(u)^k] = E_X(e^{\mu_Y(u)k}) = \alpha_X(\mu_Y(u)) \end{aligned} \tag{5.1}$$

con $\mu_Y(u)$, función cumulativa común a las cuantías por reclamación Y_i.

La expresión (5.1) nos indica que la función generatriz de momentos de la cuantía total incondicionada al número de siniestros, coincide con la función generatriz de momentos de la v.a. número de siniestros, X, evaluada en $\mu_Y(u)$, función cumulativa de la v.a. Y cuantía por reclamación.

De ahí, se obtiene en particular la función de ingreso de la distribución compuesta de Y, como

$$\frac{\mu_Y(u)}{u} = \frac{\mu_X(\mu_Y(u))}{u}$$

Ejemplo 5.1 *Consideremos una compañía de seguros, donde el número de siniestros sigue una distribución de Poisson de parámetro λ, y la cuantía por reclamación es una v.a. con distribución exponencial de parámetro a. Vamos a obtener la función de ingreso de la compañía para un nivel prefijado de riesgo α y una reserva inicial de W.*

La v.a. cuantía total para X reclamaciones es $\mathcal{Y} = Y_1 + ... + Y_X$. Si las cuantía por reclamación son independientes, la función cumulativa de $\mathcal{Y}|X$ es la suma de las funciones cumulativas de las Y_i.

La función cumulativa de la distribución compuesta de la cuantía total, $\mathcal{Y}$, es:

$$\mu_{\mathcal{Y}}(u) = \mu_X(\mu_Y(u))$$

En primer lugar recordamos la función generatriz de momentos de una v.a. de Poisson

$$\alpha_X(u) = e^{\lambda(e^u - 1)}$$

Luego, su función cumulativa es:

$$\mu_X(u) = \lambda(e^u - 1)$$

Para la distribución exponencial de parámetro a:

$$\alpha_Y(u) = \frac{a}{(a-u)}, \forall u < a$$

y

$$\mu_Y(u) = ln\left(\frac{a}{(a-u)}\right), \forall u < a$$

Por tanto

$$\mu_{\mathcal{Y}}(u) = \mu_X(\mu_Y(u)) = \lambda(e^{\mu_Y(u)} - 1) = \lambda(\frac{a}{(a-u)} - 1) = \lambda\frac{u}{(a-u)}$$

Finalmente la función de ingreso de $\mathcal{Y}$, es:

$$I(W) = \frac{\lambda}{a - \frac{|ln\alpha|}{W}}$$

y tiene sentido siempre que $W > \frac{|ln\alpha|}{a}$.
Por ejemplo si la cuantía media es de 500 euros, $\alpha = 0{,}05$, se obtiene que $W > 1497,86$ euros.

5.5.3. Función de integración

En este apartado nos ocuparemos de T unidades de riesgo. En cada una de ellas, t, consideraremos que se dispone de una reserva inicial W_t, que el nivel de probabilidad de ruina α y que las v.a. X_t correspondientes al montante total de siniestralidad son independientes. Denotaremos además por μ_t a la función cumulativa de la v.a. X_t para la unidad de riesgo t.

Definición 5.5 *Llamaremos función de integración y la denotaremos por $\nu_T(u)$ a la función de ingreso correspondiente al conjunto de las T unidades de riesgo integradas en una sola, es decir:*

$$\nu_T(u) = \frac{\mu_T(u)}{u} = \frac{\sum_{t=1}^{T} \mu_t(u)}{u}$$

por la independencia de las v.a X_t.

Así si denotamos por W a la reserva inicial total, y denotamos por W_t a las reservas iniciales en cada unidad de riesgo, t. Se obtiene la función de ingreso en términos de la reserva inicial total, W, como

$$\begin{aligned} I_T(W) &= \frac{\mu_T(\frac{|ln\alpha|}{W})}{\frac{|ln\alpha|}{W}} = \sum_{t=1}^{T} \frac{\mu_t(\frac{|ln\alpha|}{W})}{\frac{|ln\alpha|}{W}} \\ &\leq \sum_{t=1}^{T} I_t(W_t) \end{aligned}$$

Es decir, el ingreso total necesario para garantizar la supervivencia de la compañía con nivel de probabilidad de ruina α considerando las unidades de riesgo integradas, es menor que la suma de los ingresos necesarios para obtener la misma garantía, considerando la unidades de riesgo por separado.

Como resumen, podemos decir que la función de ingreso puede ser interpretada como una buena representación de la desutilidad de la compañía. A menor ingreso necesario, es decir primas más bajas, mayor utilidad. Es decir, la función de ingreso con signo menos, proporciona una función de utilidad que depende de la estructura de siniestralidad y donde la característica de cada compañía viene determinada por su reserva inicial.

Con esta interpretación obtenemos compañías todas ellas aversas al riesgo, utilidad cóncava, y tanto más aversas cuanto menos reserva inicial tengan, lo que es coherente con lo estudiado aqui.

5.6. VaR: Value at Risk, Valor en Riesgo

El tratamiento del riesgo basado en acotar la probabilidad de ruina que hemos visto en la sección anterior, resulta en ocasiones muy conservador, pero adecuado por otra parte en el cálculo de primas de pólizas de seguros.

En esta sección vamos a presentar un tratamiento alternativo, basado en calcular la pérdida máxima que se está dispuesto a asumir. Se trata de un tema muy relevante tanto para compañías de seguros como para entidades financieras.

Vamos a comenzar por mostrar cómo evaluar el VaR, en el contexto de carteras de activos financieros.

En este contexto, consideraremos un activo de precio P_t y con rendimiento definido por la expresión:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Estos rendimientos son v.a. aleatorias continuas, y es habitual la hipótesis de que siguen una distribución normal, de media m y varianza σ^2.

En este caso se puede calcular la probabilidad de que este rendimiento esté por debajo de una cantidad determinada.

Definición 5.6 *Se define el* valor en riesgo, VaR, *como la pérdida máxima que se puede dar con un nivel de confianza* $0 < \alpha < 1$.

Es decir, $P(R_t < -VaR) = \alpha$.

Bajo la hipótesis de normalidad de los rendimientos, se puede obtener la habitual expresión de VaR, como sigue:

$$P(R_t < -VaR) = P(\frac{R_t - m}{\sigma} < \frac{-VaR - m}{\sigma}) = \alpha$$

de donde

$$\frac{-VaR - m}{\sigma} = -t_\alpha$$

y por lo tanto

$$VaR = \sigma t_\alpha - m$$

Si la distribución de los rendimientos no es normal, la obtención de VaR se realiza del mismo modo, pero tomado el cuantil de la distribución correspondiente.

En el caso de querer calcular el VaR de una cartera compuesta por n activos, los cálculos se complican y una forma de simplificarlos es mantener la hipótesis de normalidad en el comportamiento de todos los rendimientos.

Así, el rendimiento de una cartera en t, se define como

$$R_t = \sum_{i=1}^{n} \omega_{it} R_{it}$$

donde ω_{it} representa la proporción de capital destinada a la compra del activo i en t, y por lo tanto, $\sum_{i=1}^{n} \omega_{it} = 1$, para todo t. Por su parte R_{it} representa el rendimiento del activo i en t.

La primera dificultad es encontrar la distribución del rendimiento de la cartera en t. Para ello, necesitamos la distribución conjunta de los rendimientos de los activos.

En caso de que las proporciones no varíen con t es relativamente sencillo, puesto que la podemos tomar a su vez como un activo y es una distribución univariante. Sin embargo, el caso más realista pasa por considerar que los inversores revisan con relativa frecuencia la composición de su cartera y por tanto, dichas proporciones varían con t.

Como ya hemos dicho consideraremos que la distribución conjunta es normal. Donde en particular, para cada activo la media es m_i y la varianza σ_i^2. Además $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, respresenta el término de covarianza, y ρ_{ij} correlación entre el rendimiento del activo i y el de j.

Al ser la distribución conjunta normal, la distribución marginal de los rendimientos individuales de los activos es $N(m_{it}, \sigma^2_{it})$ y la distribución del rendimiento de la cartera en t es también normal $N(m_t, \sigma^2_t)$, donde $m_t = \sum_{i=1}^n \omega_{it} m_{it}$ y $\sigma^2_t = \sum_{i,j} \omega_{it}\omega_{jt}\sigma_{ij}$

De este modeo el VaR de la cartera, se obtiene utilizando el cuantil correspondiente de la distribución normal anterior:

$$\begin{aligned} VaR &= \sigma_t t_\alpha - m_t = \\ &= \sqrt{\sum_{i,j} \omega_{it}\omega_{jt}\sigma_{ij} t_\alpha} - \sum_{i=1}^n \omega_{it} m_{it} \end{aligned}$$

Ejemplo 5.2 *Consideremos los rendimientos de los activos BBVA, REPSOL, e IBERIA durante el año 2000. Del estudio descriptivo de dichos rendimiento se han obtenido los siguientes valores típicos:*

$m_B = 0{,}00046$, $m_R = -0{,}00121$, $m_I = -0{,}00012$,

$\sigma_B = 0{,}019779$, $\sigma_R = 0{,}021331$, $\sigma_I = 0{,}016025$,

$\rho_{BI} = 0{,}02814$, $\rho_{BR} = 0{,}02234$, $\rho_{RI} = 0{,}01637$,

Partiendo de la hipótesis de normalidad en las distribuciones, obtenemos los valores individuales del VaR, para cada uno de los tres activos. Así, con $\alpha = 1\,\%$, $t_{0{,}01} = 2{,}33$, obtenemos

$$\begin{aligned} VaR_B &= 0{,}019779 \cdot 2{,}33 - 0{,}00046 = 0{,}04562 \\ VaR_R &= 0{,}021331 \cdot 2{,}33 + 0{,}00121 = 0{,}05091 \\ VaR_I &= 0{,}016025 \cdot 2{,}33 - 0{,}00012 = 0{,}03748 \end{aligned}$$

Si confeccionamos una cartera con igual proporción de los tres activos, es decir, donde su rendimiento es

$$R = \frac{1}{3}R_B + \frac{1}{3}R_R + \frac{1}{3}R_I$$

En este caso el rendimiento medio de la cartera es $m = -0{,}00029$ y la desviación $\sigma = 0{,}012559$. A partir de aqui, dada la hipótesis de normalidad, se sigue que el VaR de la cartera es:

$$VaR = 0{,}012559 \cdot 2{,}33 - 0{,}00029 = 0{,}02955$$

Vemos que el VaR de la cartera, o pérdida máxima a un nivel 0,01, es menor cuando diversificamos el riesgo invirtiendo en activos que tienen poca correlación.

5.7. Tarea 5

Se enumeran a continuación un conjunto de ejercicios de ensayo de las competencias correspondientes a este tema; en particular la tercera de las formuladas en el programa de la asignatura:

> Modelizar la incertidumbre de las situaciones sencillas de azar en el ámbito del negocio financiero y asegurador, mediante formulaciones matemáticas basadas en la teoría de la probabilidad. Deducir soluciones para el cálculo de la prima de riesgo, y la función de ingreso acotando la probabilidad de ruina. Utilizar un tratamiento alternativo basado en el cálculo de la pérdida máxima, para medir el riesgo de una cartera.

1. Considera una empresa aseguradora que se plantea asegurar dos tipos de siniestros. En el primero, el número de siniestros, X_1 es una v.a. de Poisson $\mathcal{P}(\lambda_1)$, con cuantía por reclamación fija e igual a C. En el segundo el número de siniestros, X_2 es también una v.a. de Poisson $\mathcal{P}(\lambda_2)$, con cuantía por reclamación con distribución exponencial de parámetro $a = \frac{1}{C}$. La función de utilidad de la empresa viene dada por:

$$U(m, \sigma^2) = -(Am + B\sigma^2)$$

   donde m es la media de cuantía total por reclamaciones, (montante medio de siniestralidad), σ^2 es la varianza y A y B son parámetros positivos fijos de la empresa.

   a) ¿Qué tipo de seguro le conviene más a la compañía?
   b) ¿En qué situación estamos si $\lambda_1 = \lambda_2$?
   c) Considera ahora que la empresa quiere diseñar una póliza combinada para cubrir ambos tipos de siniestro, y sabe además que si se produce reclamación en uno de los dos tipos, se produce en todas las pólizas de ese tipo. Si la utilidad sigue siendo la anterior y hay una porporción de pólizas p del primer tipo.
      1) Calcula la media y la varianza de del beneficio de la cartera.
      2) Calcula la proporción p que le proporciona utilidad máxima a la empresa.

2. Consideremos una compañía en la que el número de siniestros N, es una v.a. con distribución binomial $b(n, p)$, donde n es el número de pólizas contratadas y p la probabilidad de reclamación en cada una de ellas. La cuantía por reclamación en cada una de ellas es una v.a. Y con distribución exponencial de parámetro a. Obtener la función de ingreso de la cuantía total por reclamaciones, para un nivel de riesgo α y una reserva inicial W.

3. Obtén la función generatriz de momentos para una v.a. X binomial negativa de parámetros m y p.

4. Obtén la función generatriz de momentos para una v.a. X $\gamma(a, r)$.

5. Obtén la función de ingreso de la cuantía total por reclamaciones, si el número de siniestros X es una v.a. binomial negativa y la cuantía por reclamación Y es constante e igual a a.

6. Obtén la función de ingreso de la cuantía total por reclamaciones, si el número de siniestros X es una v.a. binomial negativa y la cuantía por reclamación Y_i, en cada siniestro es $\gamma(a, r)$. Además Y_i, son v.a. independientes e i.d. para cada siniestro. Compara el resultado con el obtenido en el ejercicio anterior.

7. Considera dos compañías que cubren respectivamente dos tipos de siniestros. En la primera el número de reclamaciones es Poisson ($\lambda_1 = 1$) y la cuantía por reclamación es exponencial de parámetro $a_1 = 0{,}1$. En la segunda, el número de reclamaciones es Poisson ($\lambda_2 = 10$) y la cuantía por reclamación es exponencial de parámetro $a_2 = 1$.

   a) ¿Cuál de las dos compañías aborda una situación más arriesgada?
   b) Obtén la función de ingreso, en ambas compañías para un nivel de riesgo $\alpha = 0{,}05$.
   c) Si la reserva inicial para la primera compañía es $W_1 = 50$ y para la segunda es $W_2 = 70$, obtén el valor de las funciones de ingreso respectivas.
   d) ¿Qué pasaría si se fusionasen ambas compañías? ¿Cuál sería su función de ingreso integrada? ¿Y su valor considerando la reserva inicial como suma de las dos reservas $W_1 + W_2$?

8. Una compañía tiene 100 pólizas contratadas. El número de reclamaciones por cada póliza es una v.a Poisson de media 0,01, y la cuantía por reclamación es exponencial de media 2 (en miles de euros).

   a) Calcula la probabilidad de que una determinada póliza presente una reclamación. Calcula la probabilidad de que entre las 100 pólizas se presenten 20 reclamaciones.
   b) Sabiendo que una póliza ha presentado una reclamación, calcula la probabilidad de que la cuantía de la misma sea superior a 3 mil euros.
   c) Calcula la media y la varianza de la cuantía total por póliza, incondicionada al numero de reclamaciones.
   d) Si sabemos que en la compañía se han presentado un total de 20 reclamaciones, ¿cuál es la distribución de la cuantía total?
   e) Obtén la función cumulativa de la cuantía total incondicionada al número de reclamaciones.
   f) Obtén la función de ingreso de la compañía para un nivel de 0.05 de probabilidad de ruina.

9. En una empresa aseguradora el número de reclamaciones por póliza sigue una distribución de Poisson de parámetro λ, donde λ puede tomar los valores, 0.09 (asegurados tipo A), con probabilidad 0.5 y 0.005 (asegurados de tipo B), con la misma probabilidad 0.5. Las primas se calculan como $k \cdot E(N)$, donde k es una constante igual para todas las pólizas en todos los años y $E(N)$ es la media de la v.a. número de reclamaciones.

   Un asegurado es de tipo A y le cobran una prima de 75 u.m. Después del primer año, durante el que no ha presentado ninguna reclamación se le revisa la prima. Obtener el nuevo valor de la prima, a partir de las siguientes cuestiones:

   a) Obtener el valor de k.

b) Calcula las probabilidades de que el asegurado sea de tipo A o B condiciconadas a que no ha presentado ninguna reclamación.

c) Calcula la media del número de reclamaciones de este asegurado, para el segundo año, condicionada a que no ha presentado ninguna reclamación.

d) Calcula la prima para el segundo año.

e) Repite los apartados *b*), *c*) y *d*) suponiendo que el asegurado ha presentado una reclamación.

10. En una entidad bancaria se quiere confeccionar una cartera formada por dos activos, con rendimientos aleatorios R_1 y R_2, y el activo de rendimiento seguro $S = 5\,\%$. Se sabe que las distribuciones de los rendimientos R_1 y R_2 son independientes, siendo $R_1 \equiv N(15\,\%, \sigma_1 = 10\,\%)$ y $R_2 \equiv N(7\,\%, \sigma_2 = 5\,\%)$. La cartera se confecciona de manera que $R_A = 0{,}25R_1 + 0{,}25R_2 + 0{,}5S$.

   a) Analiza el rendimiento medio, varianza y probabilidad de que el rendimiento supere el 10 %.

   b) Rendimiento mínimo al nivel del 5 % ($VaR_{0,95}$).

   c) ¿Qué rendimiento puede llegar a ofrecer el banco si se quiere asegurar una probabilidad igual a 0.8 de obtener un rendimiento neto positivo?

11. El número de reclamaciones recibidas sigue una distribución binomial negativa de media 5 y varianza 20, y la cuantía por reclamación es gamma de media 2 u.m. y varianza 8 u.m.^{2}. Calcular:

   a) La estructura de probabilidad para el siniestro más elevado.

   b) Media y varianza de la cuantía total por reclamaciones.

   c) Estructura de probabilidad de la cuantía total.

   d) Función de ingreso de la compañía para un nivel de 0.05 de probabilidad de ruina.

   e) Reserva para que el ingreso necesario por póliza no deba exceder las 20 u.m.

12. Una compañía asegura dos tipos de siniestros A y B. El número de reclamaciones por póliza, X_A, correspondientes al siniestro de tipo A, sigue una ley de Poisson de parámetro $\lambda_A = 0{,}01$ y la cuantía por reclamación, Y_A sigue una ley de Pareto de forma

$$f(y) = \frac{3 \cdot 10^3}{y^4}, \quad y \geq 10$$

   El número de reclamaciones por póliza, X_B, correspondientes al siniestro de tipo B, sigue una ley de Poisson de parámetro $\lambda_B = 0{,}01$ y la cuantía por reclamación, Y_B sigue una ley exponencial de media 15.

   a) ¿Cuál de los dos siniestros es más arriesgado? ¿Qué prima debe ser más cara? Atendiendo a la función de ingreso, ¿cuál es el valor mínimo que se debe cobrar para cada una de las pólizas?

   b) La compañía contrata 100 pólizas, 60 de tipo A y 40 de tipo B. Utilizando el T.C.L., obtén la probabilidad de que la cuantía total por reclamaciones no supere las 20 unidades.

Apéndice A

Binomios negativos

Proposición A.1 *Se verifica la siguiente igualdad:*

$$\binom{m+k-1}{k} = \binom{m+k-1}{m-1} = \binom{-m}{k}(-1)^k$$

donde $k \in \mathbf{Z}^+$.

Demostración:

Basta con desarrollar el término de la izquierda:

$$\begin{aligned}
\binom{m+k-1}{k} &= \frac{(m+k-1)(m+k-2)\cdots(m+k-k)(m-1)!}{k!(m-1)!} \\
&= \frac{(-1)(-m-k+1)(-1)(-m-k+2)\cdots(-1)(m-k+k)(m-1)!}{k!(m-1)!} \\
&= (-1)^k \frac{(-m-(k-1))(-m-(k-2))\cdots(-m-1)(-m)}{k!} = \\
&= (-1)^k \frac{(-m)(-m-1)\cdots(-m-(k-2))(-m-(k-1))(-m-k)!}{k!(-m-k)!} = \\
&= (-1)^k \frac{(-m)!}{k!(-m-k)!} = (-1)^k \binom{-m}{k} \bullet
\end{aligned}$$

De la Proposición A.1, si la leemos de derecha a izquierda, se deduce la definición de un número combinatorio en el que el número superior es negativo. Así, para cualquier $m > 0$ y cualquier k, si queremos calcular

$$\binom{-m}{k}$$

lo haremos,

$$\binom{-m}{k} = (-1)^k \frac{(m+k-1)(m+k-2)\cdots(m+k-k)}{k!}$$

De la demostración, se induce la **identidad** del factorial de un entero negativo. Así, situándonos en las dos últimas lineas de la demostración:

$$(-m)! = (-m)(-m-1)\cdots(-m-(k-2))(-m-(k-1))(-m-k)!$$

Pero observa que este producto es infinito, y **no sirve** como procedimiento de cálculo.

Si queremos calcular números combinatorios, con el número superior fraccional, debemos recordar la forma de cálculo general:

$$\binom{m+k-1}{k} = \frac{(m+k-1)(m+k-2)\cdots(m+k-k)}{k!}$$

donde además sabemos que

$$\binom{m+k-1}{k} = \binom{m+k-1}{m-1} = 1$$

siempre que $m+k-1 < k$ (esto sucede por ejemplo cuando $0 < m < 1$ y $k = 0$). Recordar además que $0! = 1$.

Bibliografía

[Bárcena et al., 2003] Bárcena, M., K., F., Ferreira, E., and Garín, M. (2003). *Elementos de Probabilidad y Estadística Descriptiva*. Servicio Editorial de la UPV|EHU, Bilbao.

[Boland, 2007] Boland, P. (2007). *Statistical and probabilistic methods in actuarial science*. Chapman & Hall/CRC, London.

[Hossack et al., 1983] Hossack, I., Pollard, J., and Zennwirth, B. (1983). *Introductory statistics and applications in general insurance*. Cambridge University Press, USA.

[López Cachero, 1996] López Cachero, M. (1996). *Estadística para actuarios*. Fundación Mapfre Estudios, Madrid.

[Sarabia et al., 2006] Sarabia, J., Gómez Déniz, E., and Vázquez Polo, F. (2006). *Estadística Actuarial: teoría y aplicaciones*. Prentice Hill, Madrid.

[Vegas Pérez, 1981] Vegas Pérez, A. (1981). *Estadística. Aplicaciones econométricas y actuariales*. Pirámide, Madrid.