

CAPITULO 2

El número de siniestros

2.1 Introducción

(a) **Número de siniestros.** Tanto el número de siniestros como la intensidad de los mismos son, en general, estocásticos en las aplicaciones prácticas y nuestro modelo final se formulará conforme con ello. Sin embargo, para introducir los conceptos pertinentes en forma gradual en el presente capítulo solamente se estudia el número k de siniestros que surge de un colectivo de riesgos durante un período de tiempo. El caso general, en el cual también se estudia la naturaleza estocástica de la intensidad de los siniestros individuales se considerará en el capítulo 3.

El comportamiento de la variable k , que representa el número de siniestros, puede describirse en términos de su distribución de probabilidades, que está determinada por las probabilidades

$$p_k = \text{Prob}\{k = k\} \quad (k = 0, 1, 2, \dots) \quad (2.1.1)$$

de que ocurran exactamente k siniestros en el período previsto.

El colectivo de riesgos consiste en unidades de riesgo individuales, tales como casas, edificios y fábricas con seguro contra incendio, o las personas que tienen seguro de vida. Los eventos primarios son los accidentes que inciden en forma aleatoria en las unidades dando lugar a los siniestros. En la teoría del riesgo individual el modelo del proceso de riesgo se basa en la consideración de estas unidades como entidades separadas. La probabilidad de que uno o más siniestros tengan lugar en un período está determinada para cada unidad, así como la distribución de las intensidades de los

siniestros. La variable k del número de siniestros de un grupo completo se obtiene como la suma de las variables del número de siniestros de las unidades de riesgo individual y, por consiguiente, el importe acumulado de siniestros es la suma de los importes acumulados de siniestros de las unidades de riesgo.

Históricamente, la teoría individual apareció como una primera fase en el desarrollo de la teoría del riesgo. Sin embargo, en la práctica no es conveniente para manejar grandes colectivos de riesgos, por lo tanto, con frecuencia se la reemplaza por un enfoque colectivo. Un modelo se desarrolla directamente para la variable k del número de siniestros y para el importe acumulado de siniestros para todo el colectivo, sin considerar las unidades de riesgo individuales. Hoy en día, este método se reconoce, con frecuencia, como más apropiado en la práctica y lo utilizamos a menos que se especifique explícitamente lo contrario.

Los enfoques individual y colectivo no son necesariamente mutuamente excluyentes. En las secciones 3.3.2 y en el Apéndice A, sección A1 se demostrará sucintamente que, en supuestos naturales, el modelo individual conduce al modelo colectivo. Es una cuestión de preferencia si, al introducir los fundamentos de la teoría del riesgo, se comienza por las unidades de riesgo individuales y los eventos que inciden en forma aleatoria en ellos o se encara directamente la descripción colectiva del proceso.

(b) Poisson y otras leyes. En la sección 2.2 se muestra cómo, bajo ciertas condiciones ideales y restrictivas, la variable k del número de siniestros se distribuye conforme con Poisson. Las propiedades de las variables con distribución de Poisson se tratan en la sección 2.3.

Una generalización de la distribución de Poisson, la denominada distribución de Poisson mixta, se introducirá en la sección 2.4. En la práctica, la distribución de Poisson mixta proporciona un modelo más realista del

número de siniestros que la distribución de Poisson pura. En la sección 2.5 se introduce un tipo especial de distribuciones de Poisson mixtas con propiedades técnicas provechosas, basadas en la ley de distribución binomial negativa.

2.2 La distribución de Poisson

(a) La ley de Poisson. Los siniestros se desarrollan como una secuencia de sucesos de una forma aleatoria tal que no es posible pronosticar el momento exacto de ocurrencia de uno de ellos ni el número total exacto durante un período de observación.

Si podemos suponer que los siniestros son independientes entre sí, entonces el número de siniestros en un período determinado se distribuye conforme con Poisson.

Estas propiedades descriptas vagamente deben expresarse en forma rigurosa para que sean apropiadas para el desarrollo matemático. Para esto consideremos el número acumulado $k(t)$ de siniestros que tienen lugar durante un período de 0 a t como una función del tiempo t . Entonces $k(t)$ es un proceso estocástico. Se postula que este proceso de número de siniestros debe cumplir con las tres condiciones siguientes.

1. El número de siniestros que tengan lugar en cualquiera de dos intervalos de tiempo disjuntos son independientes (**independencia de los incrementos**).
2. No puede ocurrir más de un siniestro por el mismo suceso (**exclusión en los siniestros múltiples**).

3. La probabilidad de que un siniestro ocurra en un punto determinado en el tiempo es igual a cero (**exclusión de puntos de tiempo especiales**).

Por lo tanto el número de siniestros que tienen lugar en cualquier intervalo de tiempo fijo está **distribuido conforme con Poisson**. La demostración se da en el Apéndice A. La condición (3) se la puede reemplazar por la condición equivalente:

3(a) El número de siniestros esperados $n(t) = E(k(t))$ es una función continua de t .

Si k es una variable con distribución de Poisson, entonces las probabilidades (2.1.1) que corresponden a diferentes valores de k se obtienen de la conocida fórmula

$$p_k = p_k(n) = e^{-n} \cdot \frac{n^k}{k!} \quad (2.2.1)$$

donde n es un número real positivo, conocido como el parámetro de Poisson, siendo igual al valor esperado de k , es decir, $n=E(k)$.

La ley de distribución de Poisson (2.2.1) define una clase de distribuciones caracterizadas por el parámetro n , simbolizada por $\text{Poisson}(n)$.

(b) **Análisis de la condición (1).** Con frecuencia, en la práctica, la condición de independencia es satisfecha sólo aproximadamente. Habitualmente existen factores de contexto que pueden hacer que siniestros en diferentes subperiodos estén correlacionados. Los factores típicos de esta índole pueden ser las condiciones meteorológicas y las económicas.

EJEMPLO 1. Si tiene lugar un gran número no usual de incendios forestales en la primera semana de julio (N.T. ello referido al hemisferio norte), esto puede ser consecuencia de un verano particularmente seco, lo que incrementa, en gran medida, la propensión al riesgo. En este caso es bastante probable que el número de incendios forestales durante la segunda semana de julio sea bastante alto. Por otro lado, si no hay incendios forestales en la primera semana, probablemente no habrá muchos durante la segunda semana.

EJEMPLO 2. En los seguros de enfermedad y de vida las enfermedades contagiosas y las epidemias pueden provocar correlación entre los siniestros en diferentes períodos de tiempo.

La presencia de factores de contexto tales como los descriptos más arriba no debe impedir la aplicación del modelo de Poisson, en tanto que éste se generalice en la forma apropiada. La influencia de los factores de contexto puede, con frecuencia, cuantificarse por medio de la introducción de una variable (o proceso) que controle los cambios en la propensión al riesgo. De esta manera llegamos a la denominada distribución de Poisson mixta, de la que nos ocuparemos en la sección 2.4.

(c) **Análisis de la condición (2).** En ciertos ramos del seguro es posible que surjan más de un siniestro a partir del mismo hecho, por ejemplo, en un choque entre dos automóviles. En tal caso ambos siniestros están muy relacionados y no se cumple la condición de independencia (2). Una manera de evitar el problema de siniestros múltiples es considerar todos los siniestros que surjan del mismo hecho como partes de un único siniestro. También debería considerarse esta convención técnica cuando se construye la función de distribución de la intensidad de los siniestros, es decir, cuando los datos pertinentes se obtienen de los registros de los siniestros reales, donde los componentes de los siniestros múltiples primero deben ser acumulados.

(d) **Análisis de la condición (3).** Un aspecto característico de los siniestros es que tienen lugar en forma aleatoria de tal manera que los momentos

exactos de ocurrencia son impredecibles. Por lo tanto la condición (3), como regla, se satisface automáticamente.

Sin embargo, debe notarse que en algunas formas especiales de seguros pueden existir hechos que se contradicen con la condición (3), por ejemplo, en un seguro dotal el pago se realiza con la edad de vencimiento. Los casos especiales de este tipo deben considerarse por separado.

2.3 Propiedades de las variables de Poisson

Las variables aleatorias con distribución de Poisson tienen muchas propiedades útiles. Algunas de estos aspectos serán analizados en esta sección. Las demostraciones se encuentran en el Apéndice A y en los ejercicios.

(a) Aditividad. La suma de variables de Poisson independientes tiene distribución de Poisson. Más precisamente si k_1, k_2, \dots, k_m son variables aleatorias independientes distribuidas conforme con Poisson, entonces su suma

$$k = k_1 + k_2 + \dots + k_m \quad (2.3.1)$$

también tiene distribución de Poisson, donde el parámetro de Poisson n es igual a la suma de los parámetros de Poisson n_j de los sumandos k_j

$$n = n_1 + n_2 + \dots + n_m \quad (2.3.2)$$

Este resultado se puede deducir directamente de las condiciones de la sección 2.2(a). Si se combinan dos o más colectivos de riesgos independientes, por ejemplo líneas de negocios o secciones de una cartera, y si los siniestros ocurren en forma independiente en cada uno de estos

grupos, entonces, naturalmente lo mismo es cierto para el colectivo de riesgos combinado.

Una demostración técnica simple de aditividad puede obtenerse utilizando la f.g.m. $M(s) = \exp(n \cdot (e^s - 1))$ (véase Ejercicio 2.3.1) de una variable con distribución de Poisson (n), o su f.g.c. (véase sección 1.4(d))

$$\psi(s) = \ln M(s) = n \cdot (e^s - 1). \quad (2.3.3)$$

Teniendo en cuenta la aditividad de las f.g.c. de las variables independientes (véase 1.4.20) obtenemos para la suma (2.3.1) la f.g.c.

$$\psi(s) = n_1(e^s - 1) + n_2(e^s - 1) + \dots + n_m(e^s - 1) = n \cdot (e^s - 1) \quad (2.3.4)$$

que es la f.g.c. de una variable con distribución de Poisson (n). La demostración se deduce de la correspondencia uno a uno de la f.d. y la f.g.c. (cf. sección 1.4(d)).

(b) Propiedades. Los cumulantes de la variable con distribución de Poisson (n), se obtienen como derivadas de la f.g.c. (véase 1.4.19))

$$\kappa_j = \psi^{(j)}(0) = n \quad (2.3.5)$$

para todas las j . Por lo tanto, la media, el desvío estándar, la asimetría y la curtosis de una variable con distribución de Poisson (n), son (véase (1.4.25) y el Ejercicio 2.3.2):

$$\alpha_j = \kappa_j = n \quad (\text{media})$$

$$\sigma = k_2^{1/2} = \sqrt{n} \quad (\text{desvío estándar}) \quad (2.3.6)$$

$$y = k_3 / \sigma^3 = 1 / \sqrt{n} \text{ (asimetría)}$$

$$\gamma_2 = k_4 / \sigma^4 = 1 / n \text{ (curtosis)}$$

Nótese que la media y la varianza de una variable con distribución de Poisson son iguales; este hecho puede utilizarse en las aplicaciones para probar la hipótesis de Poisson según los datos referidos al número de siniestros (véase sección 2.6).

(c) Fórmula de recurrencia. Con frecuencia la forma más fácil de calcular las probabilidades p_k de una distribución de Poisson es utilizando la fórmula de recurrencia

$$p_k = \frac{n}{k} \cdot p_{k-1} \quad (2.3.7)$$

con el valor inicial $p_0 = e^{-n}$. (Debido a problemas de desborde en las aplicaciones por computadora puede ser necesario efectuar cambios de escala o utilizar logaritmos cuando n es grande.) Esta regla de recurrencia es una propiedad clave cuando se desarrollan los algoritmos por recurrencia para las distribuciones de siniestros (véase sección 4.1).

(d) Cálculo numérico y aproximaciones. Simbolicemos con F la función de distribución (acumulativa) de una variable de Poisson(n) k . Entonces

$$F(k) = F_n(k) = \Pr\{X \leq k\} = \sum_{i=0}^k e^{-n} \frac{n^i}{i!}. \quad (2.3.8)$$

Para una k grande el cálculo de $F(k)$ es trabajoso. Por lo tanto, para muchas aplicaciones, se requieren fórmulas aproximadas para proporcionar una mayor velocidad con precisión aceptable.

(e) Aproximación normal. De acuerdo con el conocido **teorema del límite central** del cálculo de probabilidades, la función de distribución F_m de una variable aleatoria \mathbf{X}_m suma de m variables independientes idénticamente distribuidas, con media μ y el desvío estándar σ ($<\infty$), tiende en forma asintótica a la f.d. normal N a medida que m tiende a infinito. Aplicado a la variable del número de siniestros tenemos

$$F_m(x) = \Pr\{\mathbf{x} \leq x\} \rightarrow N(x) = \int_{-\infty}^x \exp(-u^2/2) du \text{ cuando } m \rightarrow \infty, \quad (2.3.9)$$

dónde \mathbf{x} y x son cantidades estandarizadas

$$\mathbf{x} = (\mathbf{X}_m - \mu_m) / \sigma_m \quad x = (X - \mu_m) / \sigma_m. \quad (2.3.10)$$

Acá $\mu_m = m \cdot \mu$ es la media de \mathbf{X}_m y $\sigma_m = \sigma \cdot \sqrt{m}$ su desvío estándar (véase Kendall y Stuart, 1977). Este aspecto también se ve en la Figura 2.3.1.

Las fórmulas para el cálculo de N y su inversa se presentan en el Apéndice D.

Si \mathbf{X}_m está asintóticamente normalmente distribuida, puede esperarse que $N(x)$ pueda servir como una aproximación para $F_m(x)$ si m es grande. Como (para n entero) una distribución de Poisson(n) es, por aditividad, la distribución de una suma de n variables de Poisson(1) independientes, se mantienen las premisas del teorema del límite central. Esto indica la aproximación

$$F(k) = N\left(\frac{k - n}{\sqrt{n}}\right) \quad (2.3.11)$$

que se obtiene de (2.3.9) y de (2.3.10) sustituyendo las propiedades correspondientes (2.3.6).

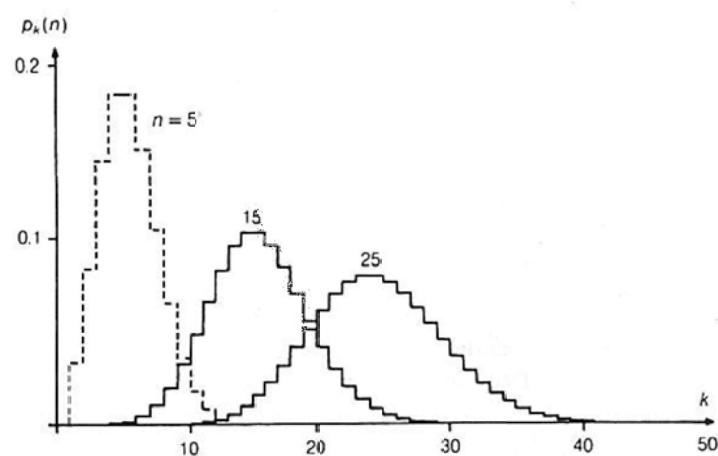


Figura 2.3.1. Probabilidades de Poisson (los valores discretos de probabilidad $p_k(n)$ se vinculan como curvas escalonadas, de tal manera que se puedan visualizar las distribuciones).

Tabla 2.3.1 Comparación de las aproximaciones para la distribución de Poisson. La función que aparece es $F(k)$ para $k \leq n$ y $1 - F(k)$ para $k > n$

<i>n</i>	<i>k</i>	Exacta	Normal	Ansccombe	Feitzer y Pratt
10	0	0.000045	0.000783	0.000034	0.000044
	2	0.002769	0.005706	0.002672	0.002763
	8	0.332820	0.263545	0.332775	0.332833
	10	0.583040	0.500000	0.582704	0.583059
	12	0.208444	0.263545	0.208786	0.208432
	18	0.007187	0.000252	0.007137	0.007187
	23	0.000120	0.000020	0.000115	0.000120
100	80	0.022649	0.022750	0.022643	0.022649
	90	0.171385	0.158655	0.171405	0.171386
	100	0.526562	0.500000	0.526551	0.526563
	110	0.147137	0.158655	0.147161	0.147137
	120	0.022669	0.022750	0.022665	0.022669
	130	0.001707	0.001350	0.001703	0.001707
	140	0.000064	0.000032	0.000064	0.000064
	145	0.000010	0.000003	0.000010	0.000010
	905	0.001215	0.001332	0.001214	0.001215
	937	0.023172	0.023173	0.023172	0.023172
1000	968	0.159596	0.155786	0.159599	0.159596
	1032	0.152095	0.155786	0.152097	0.152095
	1063	0.023155	0.023173	0.023155	0.023155
	1095	0.001446	0.001332	0.001446	0.001446

Desafortunadamente, la precisión de esta aproximación es pobre, con excepción de los valores elevados de n , como podrá observarse en la Tabla 2.3.1.

(f) La aproximación de Anscombe (Anscombe, 1960). Ésta también utiliza la f.d. normal N pero primero el argumento se transforma convenientemente, así como aparece en la siguiente expresión

$$F(k) \approx N\left(\frac{2}{3} \left(k + \frac{5}{8} \right)^{\frac{2}{3}} \cdot n^{-\frac{1}{6}} - \frac{3}{2} \frac{1}{\sqrt{n}} + \frac{1}{24} \frac{1}{\sqrt{n}} \right) \quad (2.3.12)$$

Como se puede observar en la Tabla 2.3.1, esta fórmula proporciona una aproximación muy precisa.

(g) La aproximación de Peizer y Pratt (Molenaar, 1974). Ésta es otra aproximación que se basa en una f.d. normal transformada y es aún más precisa que la aproximación de Anscombe, a pesar de ser un poco más engorrosa:

$$F(k) \approx N\left(\left[\frac{k-n}{\sqrt{n}} + \frac{1}{\sqrt{n}} \left(\frac{2}{3} + \frac{0,022}{k+1} \right) \right] \cdot \sqrt{1+T(z)} \right). \quad (2.3.13)$$

donde

$$z = \frac{k+0,5}{n}; \quad T(z) = \frac{1-z^2 + 2z \ln(z)}{1-z^2}, \quad y \quad T(1) = 0$$

Una ventaja especial de esta fórmula es que también es válida para valores de n bastante pequeños.

Nótese que (2.3.13) tiene la misma variable que la normal simple (2.3.11) como el término principal dentro de los corchetes, pero ésta es ajustada conjuntamente por términos de corrección lineales y multiplicativos.

(h) Comparaciones. La Tabla 2.3.1 muestra algunos resultados ilustrativos obtenidos por medio de los tres métodos de aproximación descriptos más arriba.

La Tabla 2.3.1 indica la excelente precisión de las fórmulas de Anscombe y de Peizer y Pratt, mientras que la aproximación normal da un ajuste pobre a menos que el parámetro de Poisson sea bastante grande, digamos de un orden de intensidad de por lo menos 1.000. Esto es previsible, ya que la distribución normal es simétrica y no puede, por lo tanto,逼近arse en forma satisfactoria a ninguna distribución asimétrica.

Las pruebas numéricas muestran que el error absoluto máximo de la aproximación de Anscombe es menor a 10^{-4} para $n \geq 35$, siendo el error absoluto mayor en el entorno de la media. El límite correspondiente para la fórmula de Peizer y Pratt es $n \geq 6$.

La ventaja de la aproximación de Anscombe por sobre la de Peizer y Pratt es que es más apropiada para la simulación de números aleatorios, ya que la transformación inversa del argumento de N puede resolverse fácilmente en forma explícita (sección 5.3(b) para detalles).

OBSERVACION Si una fórmula de aproximación no posee la precisión requerida donde n es pequeña, debe utilizarse el desarrollo exacto (2.3.8), que no es tan inconveniente para los valores pequeños de n .

Ejercicio 2.3.1 Obtener la f.g.m. y el f.g.c. de la variable aleatoria con distribución de Poisson (n).

Ejercicio 2.3.2 Probar (2.3.5) y (2.3.6).

Ejercicio 2.3.3 Demostrar que las probabilidades de Poisson p_k cumplen la fórmula de recurrencia (2.3.7).

Ejercicio 2.3.4 Hallar el modo de Poisson (n), es decir el valor de k en el que p_k alcanza el máximo.

Ejercicio 2.3.5 Sea k una variable con distribución de Poisson(n). Demostrar que

$$E[k \cdot (k - 1) \cdots \cdot (k - i)] = n^{i+1} \quad (i = 1, 2, 3, \dots).$$

Calcular la asimetría de k utilizando directamente estas expresiones.

Ejercicio 2.3.6 El número de siniestros observados que surgieron de una cartera en dos años consecutivos fue de 9.025 y 10.131. ¿Estas cifras fundamentan el supuesto de que el parámetro de Poisson se mantuvo igual?

2.4 Variable Poisson mixta del número de siniestros

(a) La volatilidad de la propensión al riesgo. En la sección 2.2(b) se especificó que las premisas de la ley de Poisson estándar, con frecuencia no son válidas en la práctica, debido a que los factores de contexto tales como las condiciones meteorológicas y económicas, etc., influyen en la intensidad de los siniestros. Este aspecto importante se ilustra en los ejemplos que se exhiben en las Figuras 2.4.1 y 2.4.2.

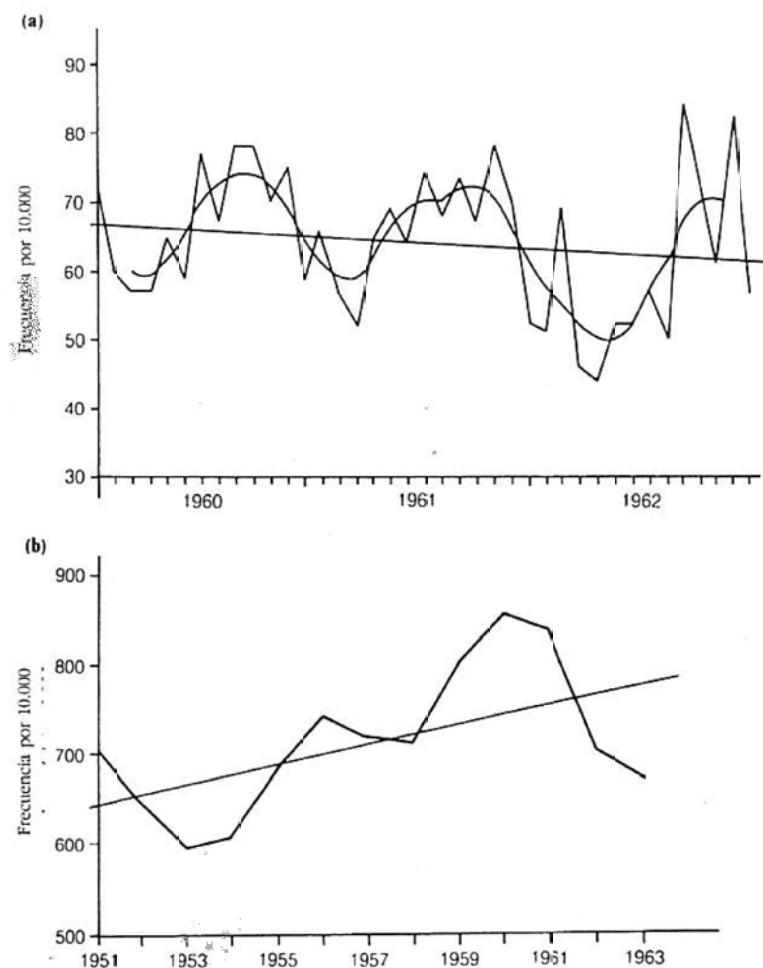


Figura 2.4.1 Frecuencias de siniestros en un período de cuatro semanas (a) y anual (b) para motocicletas. Los datos anteriores están uniformados utilizando un promedio móvil con ponderaciones (1,2,3,2,1). Estadísticas del Reino Unido proporcionadas por Beard para "Risk Theory" (Beard y otros, 1984).

El gráfico (a) de la Figura 2.4.1 muestra las variaciones estacionales en las propensiones al riesgo, superpuestas con una tendencia descendente. El gráfico (b) muestra un ciclo a largo plazo combinado con una tendencia ascendente.

La serie cronológica de seguros de accidentes del trabajo de la Figura 2.4.2 muestra un comportamiento similar.

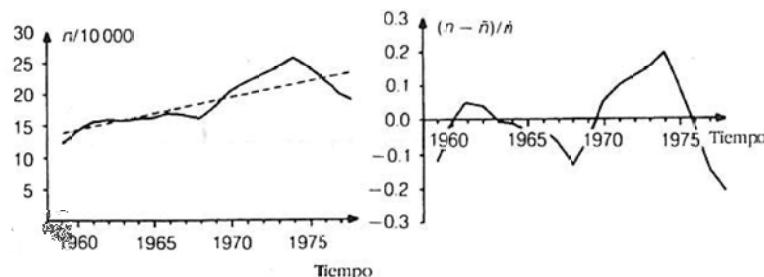


Figura 2.4.2. El número de accidentes en los seguros de accidentes del trabajo: datos conjuntos de los aseguradores finlandeses. El gráfico de la derecha muestra los desvíos relativos desde la línea de tendencia \bar{n} del gráfico de la izquierda.

Los ciclos que aparecen en la Figura 2.4.2 están muy correlacionados con los períodos de auge y las recesiones económicas generales. Durante un período de auge, la industria funciona a su capacidad máxima y las horas extras son frecuentes. Por consiguiente es previsible que se incremente el número de accidentes. Por otro lado, durante los períodos de recesión se reducen las horas de trabajo, lo que se refleja en la frecuencia de siniestros.

Como lo indican las cifras anteriores, el proceso de los siniestros puede componerse de (1) **tendencias**, (2) variaciones periódicas (no regulares) de propensión al riesgo, normalmente denominadas **ciclos**, (3) **oscilaciones de corto plazo** de la propensión al riesgo y (4) **fluctuaciones aleatorias puras**.

- (1) Las tendencias implican un cambio lento de las probabilidades de siniestros. Son causadas, por ejemplo, por cambios en los métodos de construcción de edificios afectando la frecuencia de los incendios, por los cambios de las condiciones del tráfico afectando la propensión de accidentes o mejorando la tasa de mortalidad de los seguros de vida.
- (2) Los ciclos de las propensiones a los siniestros surgen, por ejemplo, debido a los ciclos económicos nacionales o internacionales. Nótese la diferencia entre los ciclos de siniestros y los ciclos de suscripción, que se tratan en la sección 12.3. El último es un concepto más amplio, que se origina también de los cambios y los sesgos en las primas y, especialmente, por el aspecto usual de que los cambios en la propensión de los siniestros sólo pueden considerarse en las tasas de primas con un desfase de un período. Los ciclos que se considerarán son específicamente respecto de los ciclos de siniestros en los capítulos 2 al 6. En el capítulo 12 se introduce un enfoque más general, cuando el horizonte de tiempo se extiende para cubrir varios años.
- (3) Las oscilaciones a corto plazo son variaciones estacionales u otros efectos similares de corta duración.
- (4) La fluctuación aleatoria pura es del tipo que se describió mediante la ley de Poisson en secciones anteriores.

(b) Variable de ponderación. Los ejemplos anteriores muestran que la propensión al riesgo, es decir, la potencia de los hechos que provocan siniestros, puede estar sujeta a grandes variaciones.

En tanto que estas variaciones sean determinísticas, tales como las relacionadas con los horarios diurnos y nocturnos, las condiciones (1)-(3) de la sección 2.2(a) usualmente se satisfacen y la ley de Poisson es aplicable.

Cuando la variación del número de siniestros es aleatoria, entonces la condición de independencia (1) no se cumple como, por ejemplo, se observó en los ejemplos de la sección 2.2(b). En tal situación la variación estocástica del número de siniestros puede, con frecuencia, interpretarse como cambios aleatorios de los parámetros de Poisson respecto de su nivel esperado n . El cambio se describirá mediante un factor multiplicativo q tal que $E(q)=1$. Se lo define de tal manera que si el número de siniestros durante el período hubiera sido del nivel esperado, el valor q de q habría sido igual a 1. Si $q > 1$, el número es mayor que el esperado y viceversa si $0 < q < 1$. Nótese que si el valor de q es fijo e igual a q , la condición (1) se cumple y la f.d. del número condicional de siniestros $F(k|q=q)$ es de Poisson ($n \cdot q$).

La variable q ($q > 0$) se denomina **variable de ponderación** y la correspondiente variable k del número de siniestros, **variable con distribución de Poisson mixta** (N.T. o "ponderada").

La f.d. de la variable de ponderación, es decir, la f.d. de ponderación, se indica

$$H(q) = \text{Prob}\{q \leq q\}. \quad (2.4.1)$$

Sean k una variable con distribución de Poisson mixta del número de siniestros y q la variable de ponderación correspondiente. Entonces, por definición, la distribución condicional $F(k|q=q)$ de k , es de Poisson ($n \cdot q$) para cualquier valor q de la variable de ponderación q . Las probabilidades del número de siniestros p_k de k pueden obtenerse de las probabilidades de Poisson condicionales $p_k(n \cdot q)$ dadas en (2.2.1) por medio de la fórmula de probabilidades condicionales (1.4.39)

$$p_k = E(p_k(n \cdot q)) = \int_0^{\infty} e^{-nq} \frac{(nq)^k}{k!} dH(q). \quad (2.4.2)$$

Esta fórmula tiene una explicación directa: se evalúa el resultado de Poisson (el integrando) para todos los estados alternativos posibles de los factores de contexto, es decir, para todos los valores de la variable de ponderación q , y luego se obtiene un promedio ponderado, utilizando la f.d. de ponderación para promediar. La f.d. F de la variable con distribución de Poisson mixta se transcribe como un promedio ponderado similar (1.4.39) a las f.d. de Poisson simples.

$$F(k) = E(F(k|q)) = \int_0^{\infty} F_{nq}(k) dH(q). \quad (2.4.3)$$

Debe recordarse que debido a la convención de escala el valor medio de q es igual a

$$E(q) = 1 \quad (2.4.4)$$

Como la media condicional de la variable k de Poisson mixta, para $q = q$, es $n \cdot q$, siendo la distribución condicional Poisson($n \cdot q$), tenemos, de acuerdo con (1.4.32),

$$E(k) = E(E(k|q)) = E(n \cdot q) = n \cdot E(q) \quad (2.4.5)$$

Hemos verificado que la media de una variable de Poisson mixta k es igual a la media del parámetro de Poisson de variación aleatoria.

Nótese que la definición anterior, y las consideraciones resultantes siguientes son lo suficientemente generales como para que no haya

necesidad de especificar las causas del fenómeno de ponderación. Pueden ser ciclos o variaciones de corto plazo o la combinación de éstos. Sin embargo, los parámetros pertinentes que especifican la distribución de ponderación pueden depender, en gran medida, de los fenómenos del contexto.

Existen dos formas de considerar un ciclo potencial. Por un lado, se puede hacer un pronóstico determinístico a partir de la amplitud y de la fase del ciclo y el parámetro de Poisson n se evalúa conforme con ello. Entonces, la variable de ponderación q describe sólo una oscilación de corto plazo. Alternativamente, q puede representar un efecto conjunto de los ciclos y la oscilación de corto plazo en un punto previsto en el tiempo t . En el último caso, por supuesto, el rango de variación de q puede ser mucho mayor.

OBSERVACIÓN. La deducción de (2.4.2) y de (2.4.3) posee una analogía directa con el conocido modelo de urna del cálculo de probabilidades. Una urna contiene billetes de lotería, cada uno de los cuales posee un valor para la variable de ponderación q (es más fácil pensar que q posee valores discretos para un número de resultados finitos). Los billetes para cada valor de q se incluyen en proporción a los incrementos $dH(q)$ de la f.d. supuesta. Se extrae un billete de la urna y se le asigna el valor q que muestra. Las probabilidades de que $\{k=k\}$ y $\{k \leq k\}$ pueden encontrarse utilizando las reglas para la multiplicación y la suma de probabilidades. Este método se ilustra en el ejemplo que aparece más adelante en la sección 2.4(c).

(c) Un ejemplo. Consideremos el número de incendios forestales en julio (N.T.: en el hemisferio norte). Se supone que tiene distribución de Poisson mixta con un valor medio a largo plazo de $n = 100$.

Para caracterizar las variaciones meteorológicas, el número de siniestros promedio a largo plazo n puede multiplicarse por una variable q de ponderación. Para simplificar, supongamos que las condiciones meteorológicas posibles en julio se clasifican en 5 clases, $i = 1, \dots, 5$, y como figura en la Tabla 2.4.1, sean b_i las probabilidades respectivas de cada tipo

de condición meteorológica. (Verifíquese que $\sum q_i b_i = 1$ como lo requiere (2.4.4)). Aplicando (2.4.3) obtenemos la f.d. del número de incendios.

$$F(k) = \sum_i F_{npi}(k) b_i. \quad (2.4.6)$$

La f.d. Poisson mixta F de la Tabla 2.4.2 puede obtenerse por medio de cálculos simples. Para comparar, también se dan los correspondientes valores de Poisson.

Tabla 2.4.1 Fluctuación del valor de la variable de ponderación de acuerdo con el clima

Tipo de clima i	q_i	b_i
1. Muy seco	3,00	0,05
2. Seco	1,75	0,20
3. Normal	0,80	0,40
4. Húmedo	0,60	0,25
5. Muy húmedo	0,30	0,10

Tabla 2.4.2 Distribución de siniestros.

k	$F(k)$	$F(k q=1)$
50	0,13	0,00
70	0,47	0,01
100	0,74	0,53
150	0,76	1,00
200	0,94	1,00
300	0,98	1,00

El desvío estándar se incrementa de $\sigma_k = 1$ a $\sigma_k = 65,4$ debido a la ponderación. Esto puede calcularse utilizando la fórmula para los momentos centrados.

$$\alpha_j = \sum_i \alpha_j(nq_i) \cdot h_i \quad (2.4.7)$$

que se obtiene de la misma forma directa que (2.4.6) (véase 1.4.10), (1.4.22) y (1.4.25)).

(d) La función generatriz de momentos de una variable con distribución de Poisson mixta k puede obtenerse (véase (1.4.39)) tomando un promedio ponderado de las f.g.m. de Poisson:

$$M(s) = E(M(s|q)) = \int_0^\infty e^{nq(e^s-1)} dH(q). \quad (2.4.8)$$

Utilizando las esperanzas condicionales (cf. sección 1.4(g)) se obtienen las siguientes expresiones

$$\begin{aligned} M(s) &= E[M(s|q)] = E[\exp(nq(e^s-1))] \\ &= E[\exp(q(n(e^s-1)))] = M_q(n(e^s-1)) \\ &= M_q(\varphi(s)) \end{aligned} \quad (2.4.9)$$

donde $\varphi(s) = n \cdot (e^s - 1)$ es la f.g.c. (2.3.3) de la variable Poisson (n).

Tomando logaritmos, la f.g.c. ψ de la variable con distribución de Poisson mixta puede escribirse de la siguiente manera

$$\psi(s) = \ln M_q(\varphi(s)) = \psi_q(\varphi(s)), \quad (2.4.10)$$

siendo ψ_q la f.g.c. de la variable de ponderación.

(e) Características. Los cumulantes $\kappa_k(k)$ de una variable de Poisson mixta k , expresados en términos de los cumulantes de la variable de ponderación q , se obtienen de las derivadas de la fórmula de la f.g.c. $\psi(s) = \psi_q(\varphi(s))$ (véase (2.4.10)). Nótese que $\psi'_q(0) = E(q) = 1$, $\varphi'(0) = n$ (2.3.5), y $\varphi(0) = 0$. Entonces, por ejemplo, $\kappa_1(k) = E(k) = \psi'_q(\varphi(0)) \cdot \varphi'(0) = 1 \cdot n = n$, de acuerdo con (2.4.5). De igual manera, las derivadas superiores de la f.g.c. dan (Ejercicio 2.4.3)

$$\kappa_2(k) = n + n^2 \cdot \kappa_2(q)$$

$$\kappa_3(k) = n + 3 \cdot n^2 \cdot \kappa_3(q) + n^3 \cdot \kappa_3(q). \quad (2.4.11)$$

Por (1.4.25) las propiedades claves de la variable de Poisson mixta k son

$$\text{Media } \mu_k = n$$

$$\text{Varianza } \sigma_k^2 = n + n^2 \cdot \sigma_q^2$$

$$\text{Asimetría } \gamma_k = \frac{n+3 \cdot n^2 \cdot \sigma_q^2 + n^3 \cdot \gamma_q \cdot \sigma_q^3}{\sigma_k^3} \quad (2.4.12)$$

Para la curtosis, véase Ejercicio 2.4.6.

(f) Cómo obtener la f.d. de ponderación. La elección de la f.d. H de ponderación debe, por supuesto, basarse sobre la experiencia (o expectativas) relativa al contexto en consideración. Existen tres técnicas que se utilizan comúnmente para introducir el efecto de ponderación.

- (1) La f.d. H se expresa en **forma analítica**. Esto se ejemplifica en la sección 2.5.
- (2) La f.d. H se da en **forma tabular**, dividiendo el rango relevante de q en intervalos. El ejemplo anterior de la sección 2.4(c) presenta este tratamiento.
- (3) La fórmula exacta de la f.d. H no se especifica, pero sólo se dan las propiedades claves, en particular el desvío estándar σ_q y la asimetría γ_q . Este **método de los momentos** es un enfoque posible, y en realidad natural, cuando se utilizan métodos aproximados para el cálculo del número y del importe de los siniestros, como se hará regularmente en los capítulos que siguen. Las fórmulas pertinentes se basan sobre las propiedades de las distribuciones, tales como la media, el desvío estándar, la asimetría, etc. (comparar también (2.4.11) y (2.4.12)), pero no requieren conocimiento de la f.d. H .

En la práctica, la información concerniente al fenómeno de ponderación es escasa, a veces permite estimar solamente el rango de variación. Por lo tanto el método analítico (1) podría ser apropiado, en particular si se tiene experiencia de situaciones similares para sustentar la elección de la función. Si existe gran cantidad de datos estadísticos, entonces es preferible el método por tabulación (2). Si los datos son escasos, pero existe algún indicio de asimetría, entonces se sugiere el método de los momentos (3), especialmente si se supone que la asimetría es grande, como ocurre con frecuencia, por ejemplo, en las líneas sensibles a las condiciones meteorológicas (tormentas, nieve, inundaciones) o los períodos de auge y las recesiones de la economía nacional (seguro de crédito).

Nótese que el uso de una función de ponderación analítica restringe la forma de la distribución. Por ejemplo, en el caso Pólya, al que nos referimos en la sección 2.5, la f.d. de ponderación tiene un solo parámetro disponible.

Por lo tanto las otras alternativas, con frecuencia, son preferibles. Una ventaja del método de los momentos es que las propiedades necesarias se pueden obtener directamente a partir de datos empíricos y por experiencia previa. Por lo tanto no es necesario crear una función de ponderación analítica.

Una presentación gráfica tal como la de la Figura 2.4.2 puede contribuir a la construcción de la f.d. de ponderación. Si se requiere una distribución transversal y sus características en un punto de tiempo t , los desvíos de la línea de tendencia son apropiados, como se muestra en el gráfico de la derecha de la Figura 2.4.2. Si se necesita una estructura a largo plazo, también deberá considerarse la tendencia.

Pentikäinen y Rantala (1982, sección 2.3) presentaron los datos empíricos en el efecto de ponderación. Para la mayoría de las clases de seguros se encontró que el desvío estándar σ_q de la fluctuación a corto plazo era del orden de la intensidad de 0,02 a 0,08. Sin embargo para seguros de crédito y forestales se obtuvieron valores mucho más altos.

Nótese nuevamente, como se especificó en la sección 2.4(a), que si la fluctuación en un punto de tiempo dado t es de interés, tanto la variación de corto plazo como los ciclos de largo plazo pueden afectar simultáneamente el rango de la variación total y el desvío estándar "de ponderación" puede ser mucho mayor, como bien puede verse, por ejemplo, a partir de la Figura 2.4.2. Este problema se desarrollará con más detalle en el Capítulo 9.

(g) Aditividad de las variables de Poisson mixtas. Existen dos situaciones diferentes donde surge la cuestión de la aditividad de las variables de Poisson mixtas. En primer lugar, podemos tener dos o más clases de seguros, que suponemos que tienen una variable Poisson mixta del número de siniestros, o en segundo lugar, podemos tener una situación en la que se consideran períodos consecutivos, y donde el número de siniestros

que tienen lugar en cada período tiene distribución Poisson mixta. Nótese que nos estamos refiriendo al mismo problema matemático en cada caso.

La cuestión es si la suma \mathbf{k} de las variables de Poisson mixtas \mathbf{k}_i tiene una distribución de Poisson mixta, y si éste fuera el caso, ¿cuál es la variable de ponderación? Nótese que las variables de ponderación, con frecuencia, pueden correlacionarse unas con otras. Por ejemplo, si pensamos acerca de los períodos consecutivos, entonces las circunstancias de los períodos cercanos se correlacionan, especialmente cuando la ponderación es a causa de factores de contexto que se modifican relativamente despacio.

La respuesta es "sí", en tanto que los sumandos \mathbf{k}_i con distribución de Poisson(n_i) mixta sean mutuamente independientes o, más generalmente, dependan entre sí sólo a través de sus variables de ponderación \mathbf{q}_i . Nótese que las variables de ponderación pueden depender entre ellas. Para probar esta aditividad de las variables de Poisson mixtas, ahora consideraremos el caso de dos variables \mathbf{k}_i ($i = 1, 2$), pero es fácil observar que la demostración puede generalizarse para el caso de más de dos \mathbf{k}_i .

Sean \mathbf{k}_1 y \mathbf{k}_2 dependientes entre ellas sólo a través de las variables de ponderación. De acuerdo con la independencia condicional y con (1.4.20) y (2.3.3) la f.g.c. condicional $\psi_{\mathbf{k}}$ de \mathbf{k} es

$$\psi_{\mathbf{k}}(\mathbf{q}_1 = q_1, \mathbf{q}_2 = q_2) = (n_1 q_1 + n_2 q_2)(e^s - 1), \quad (2.4.13)$$

siendo la f.d. condicional del número de siniestros de \mathbf{k}_i Poisson($n_i q_i$). Denominando $n = n_1 + n_2 = E(n_1 q_1 + n_2 q_2)$ y

$$\mathbf{q} = \frac{n_1 \cdot \mathbf{q}_1 + n_2 \cdot \mathbf{q}_2}{n} \quad (2.4.14)$$

obtenemos

$$\Psi_{\mathbf{k}}(\mathbf{q}) = n \cdot \mathbf{q} \cdot (e^s - 1). \quad (2.4.15)$$

Por lo tanto, la suma $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$ es una variable con distribución de Poisson mixta con el promedio ponderado (2.4.14) de las distintas variables de ponderación como variable de ponderación.

OBSERVACIÓN. Aunque la suma de las variables de (número de siniestros) Poisson mixtas (condicionalmente) independientes es también de Poisson mixta, el resultado correspondiente no siempre se cumple para las variables de Poisson mixtas compuestas del importe de siniestros, como se podrá observar en la sección 3.2(d).

Ejercicio 2.4.1 Calcular el desvío estándar de la distribución de Poisson mixta dada en el ejemplo de la sección 2.4(c).

Ejercicio 2.4.2 Supongamos que el número de siniestros es una variable con distribución de Poisson mixta con media $n = 10.000$ y el desvío estándar $\sigma = 1.000$. Hallar el desvío estándar de la variable de ponderación.

Ejercicio 2.4.3 Calcular los cumulantes κ_j , $j = 1, 2, 3$ de una variable de Poisson mixta \mathbf{k} conociendo los cumulantes de la variable de ponderación.

Ejercicio 2.4.4 Deducir las características (2.4.12).

Ejercicio 2.4.5 Demostrar que el desvío estándar de una variable con distribución de Poisson mixta, con propiedades $n, \sigma_{\mathbf{q}} > 0$, es siempre mayor que el desvío estándar de la distribución de Poisson(n). ¿Se cumple lo mismo para la asimetría?

Ejercicio 2.4.6 Demostrar que la curtosis de una variable con distribución de Poisson mixta es

$$\gamma_2 = \frac{n + 7n^2 \cdot \sigma_q^2 + 6n^3 \cdot \gamma_q \cdot \sigma_q^3 + n^4 \cdot \gamma_{2,q} \cdot \sigma_q^4}{\sigma_q^4}$$

2.5 El caso Pólya: distribución binomial negativa

(a) Variable de ponderación con distribución Gamma La distribución Gamma(r, α), con frecuencia, se utiliza como la distribución de la variable de ponderación q (2.4.1). Su densidad es

$$f(x) = \frac{\alpha^r}{\Gamma(r)} e^{-\alpha x} x^{r-1} \quad (x \geq 0) \quad (2.5.1)$$

donde r y α son constantes positivas y

$$\Gamma(r) = \int_0^\infty e^{-u} u^{r-1} du \quad (2.5.2)$$

es la función Γ de Euler.

El valor medio de la variable Gamma (r, α) es r/α . De acuerdo con la convención de escala (2.4.4) debería ser igual a 1 cuando se la utiliza como la variable de ponderación, es decir, los parámetros r y α deberían ser iguales. Si simbolizamos por h al valor corriente, la f.d. H de la variable de ponderación q es Gamma(h, h), es decir,

$$H(q) = \frac{1}{\Gamma(h)} \int_0^{hq} e^{-z} z^{h-1} dz, \quad (2.5.3)$$

y sus propiedades principales (Apéndice B, sección B.1(a)) son

$$E(q) = 1; \quad \sigma_q = \frac{1}{\sqrt{h}}; \quad \gamma_q = \frac{2}{\sqrt{h}}; \quad \gamma_{2,q} = \frac{6}{h}. \quad (2.5.4)$$

La Figura 2.5.1 da algunos ejemplos sobre la forma de las densidades de la distribución Gamma, H' . A medida que se incrementa el parámetro h , la masa de probabilidades se concentra. Cuando $h \rightarrow \infty$ esta distribución de Poisson mixta se acerca a la correspondiente distribución de Poisson simple. De hecho éste es un resultado obvio, ya que la distribución de ponderación se comprime hasta una probabilidad puntual en $q = 1$, el efecto de ponderación se elimina. Una demostración se encuentra en el Apéndice B.

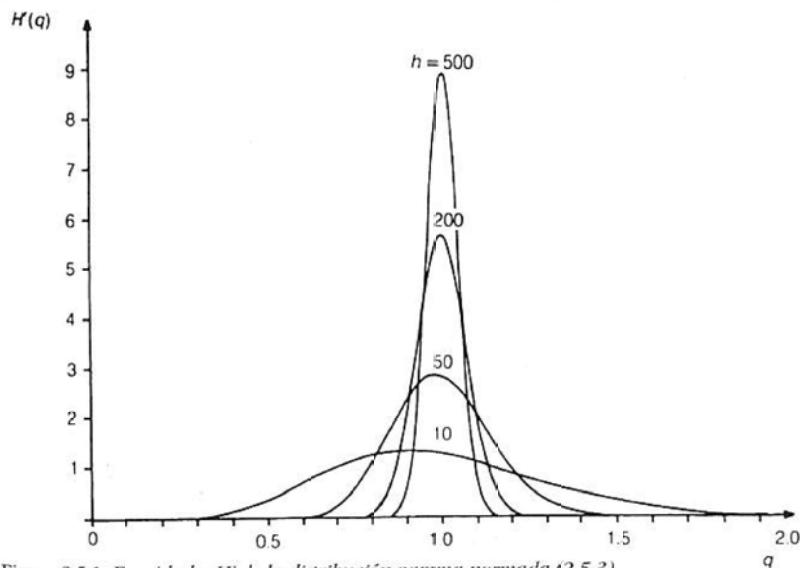


Figura 2.5.1 Densidades H' de la distribución gamma normada (2.5.3).

Nótese que si se conoce σ_q , el parámetro b se puede resolver a partir de (2.5.4) como $1/\sigma_q^2$. Esta relación es muy útil cuando se requiere una estimación para b , porque con frecuencia σ_q puede evaluarse, por lo menos en forma aproximada, a partir de los datos disponibles.

En la sección 3.3.5 se considerará la distribución gamma de tres parámetros.

(b) La distribución binomial negativa. Cuando utilizamos la f.d. gamma normada (2.5.3) como una distribución de ponderación, las probabilidades puntuales (2.4.2) de la distribución de Poisson compuesta mixta pueden obtenerse en la forma (demostración en el Apéndice B, sección B.1(b))

$$p(k) = \binom{h+k-1}{k} p^h \cdot (1-p)^k, \quad (2.5.5)$$

donde $p = \frac{h}{n+h}$

y

$$\binom{r}{s} = \frac{(r+s)!}{r! s!} = \frac{\Gamma(r+s+1)}{\Gamma(r+1) \cdot \Gamma(s+1)} \quad (2.5.6)$$

es el coeficiente binomial (generalizado a valores de las variables no-enteras a través de la función Γ de Euler (2.5.2)).

La distribución de probabilidades definida en (2.5.5), de hecho, se conoce en el cálculo de probabilidades como **binomial negativa**, también conocido como la **distribución de Pólya**.

La f.g.m. de una variable k del número de siniestros con distribución de Pólya es

$$M(s) = \left(\frac{h}{h+n-n \cdot e^s} \right)^h \quad (2.5.7)$$

y las propiedades claves (Apéndice B, sección B.1(d)) son

$$\mu_k = E(k) = n$$

$$\begin{aligned} \sigma_k^2 &= n + \frac{n^2}{h} \\ \gamma_k &= \frac{n + \frac{3n^2}{h} + \frac{2n^3}{h^2}}{\sigma_k^3} \end{aligned} \quad (2.5.8)$$

Las probabilidades binomiales negativas pueden calcularse más fácilmente por medio de la fórmula de recurrencia siguiente

$$p_k = \left(a + \frac{b}{k} \right) \cdot p_{k-1}, \quad k = 1, 2, \dots, \quad (2.5.9)$$

que se deduce inmediatamente de (2.5.5) (véase Apéndice B), donde

$$p_0 = p^h; \quad p = \frac{h}{n+h}; \quad a = 1-p; \quad b = (h-1)a. \quad (2.5.10)$$

Recuérdese que la distribución de Poisson (n) satisface una fórmula de recurrencia similar (2.3.7) donde $a=0$ y $b=n$. Una ventaja de la fórmula de recurrencia (2.5.9) es que conduce a una fórmula de recurrencia para el cálculo de la distribución del importe acumulado de siniestros, como podrá verse en la sección 4.1.

(c) **Análisis.** La distribución binomial negativa es razonablemente conveniente para las aplicaciones, lo que la ha hecho muy popular. Por otro lado, como sólo se dispone de un parámetro libre, b , para la distribución de ponderación, con frecuencia el ajuste no es siempre bueno. No es posible adaptar la distribución de ponderación con distribución gamma de tal manera que, por ejemplo, tanto el desvío estándar como la asimetría asuman algunos valores deseados. Esto es una desventaja, especialmente si la distribución de ponderación "real" es muy asimétrica.

El enfoque que utiliza la distribución binomial negativa del número de siniestros, especialmente en el contexto de la teoría del riesgo, se asocia con el nombre de Pólya. El, junto con Eggenberger (1923), la dedujeron para el modelo de la contaminación. Con frecuencia nos referimos abreviadamente a los modelos basados sobre la ley de Poisson y la binomial negativa, como el **Caso de Poisson** y el **Caso de Pólya**.

Ejercicio 2.5.1 Calcular y graficar en el mismo diagrama la función de probabilidades de Poisson p_k y las correspondientes probabilidades de Pólya p_k^* para $n = 5$ y $b = 10$.

Ejercicio 2.5.2 ¿Para qué valor de k , p_k , según está dado en (2.5.5), alcanza el máximo?

Ejercicio 2.5.3 Los ~~números~~ acumulados de siniestros de dos carteras estocásticamente independientes tienen distribución de Pólya. Probar que si se fusionan las carteras, la distribución conjunta es nuevamente del tipo

Pólya, siempre que los parámetros n y h sean iguales para ambas carteras originarias.

Ejercicio 2.5.4 Demostrar que si k es de Poisson mixta con distribución de ponderación exponencial, entonces k tiene distribución geométrica

(2.6) Variaciones de la propensión a riesgo dentro de una cartera

(a) **Propensión del riesgo individual.** Las distribuciones de Poisson mixtas también pueden aplicarse en otros contextos además de los recientemente descriptos. Un ejemplo es la situación en la que se considera la propensión al riesgo de las unidades de riesgo individuales de la cartera asegurada, por ejemplo, el problema puede ser estimar la f.d. del número de siniestros que surgen a partir de una única pieza de un automóvil. El proceso físico puede justificar la presunción de una ley de Poisson para los accidentes, pero el parámetro de riesgo n , el número de siniestros esperados por auto, puede variar para diferentes autos de acuerdo con el tipo, uso, tiempo de exposición, kilometraje, etc. del auto y la habilidad del conductor. Puede suponerse que cada unidad de riesgo i , o el auto en este ejemplo, cuenta con un parámetro de frecuencia siniestral $n_i = nq_i$, que es el número esperado de siniestros para esta unidad. Aquí n es un valor promedio y q_i un coeficiente que indica el desvío por unidad a partir de n . Sea H la f.d. que describe la variación de los valores q_i (se puede suponer que existe tal f.d. aun cuando puede ser desconocida o sólo estimada vagamente en la práctica). Esta función caracteriza la distribución del riesgo dentro de la cartera (o dentro de una parte en particular de la cartera que se considera, por ejemplo una clase de automóviles).

La distribución de la variable del número de siniestros k de una unidad individual que se selecciona al azar de una cartera puede obtenerse tomando, en primer lugar, la probabilidad de que el parámetro de riesgo q esté en el intervalo $(q, q+dq)$, suponiendo la ley de Poisson para el valor del

parámetro nq , e integrando luego con respecto a q . La obtención de las probabilidades p_k es análoga a la aplicada en la obtención de (2.4.2). De una unidad de tiempo a la otra, sólo el medio físico es diferente en la variación del parámetro de Poisson n ; en este caso es la variación de una unidad de riesgo a la siguiente. Por lo tanto, (2.4.2) es fácilmente aplicable

$$p_k = \text{Prob} \{ \mathbf{k} = k \} = \int_0^{\infty} p_k(nq) dH(q), \quad (2.6.1)$$

donde $p_k(nq)$ es la probabilidad de Poisson estándar (2.2.1). En este sentido $H(q)$ generalmente se denomina **función de estructura** (véase Ammeter, 1948; Bühlmann, 1970). Este concepto es muy importante en muchas aplicaciones, incluyendo la teoría de credibilidad (sección 6.5) en particular, y en la tarificación en general.

(b) Ejemplo. Consideremos el ejemplo dado en la Tabla 2.6.1. Las estadísticas se toman de Johnson y Hey (1971) y relacionan los siniestros conforme con las pólizas de todo riesgo de vehículos en el Reino Unido en 1968. Las 421.240 pólizas se clasificaron de acuerdo con el número de siniestros de 1968, el número promedio de siniestros por póliza es 0,13174 y la varianza 0,13852. La columna con el encabezamiento "Poisson" establece la distribución que habría resultado si la ocurrencia de siniestros hubiera seguido la ley de Poisson con $n = 0,13174$, es decir, el número esperado de siniestros por póliza en un año.

Tabla 2.6.1. Pólizas con cobertura de todo riesgo de vehículos de acuerdo con el número de siniestros

<i>k</i>	Observadas	Poisson	Binomial negativa	Dos "Poisson"
0	370412	369246	370460	370460
1	46545	48644	46411	46418
2	3935	3204	4045	4036
3	317	141	301	306
4	28	5	21	20
5	3	-	1	1

Obviamente, la distribución de Poisson posee una cola más corta que los datos, observación confirmada por la prueba χ^2 . En otras palabras, se rechaza la hipótesis de que la propensión al riesgo es la misma para todas las pólizas.

La insuficiencia de la ley de Poisson también podría anticiparse a partir del hecho de que la varianza es mayor que la media, mientras que si la ley de Poisson fuera válida tendrían que ser iguales. (2.3.6).

La columna encabezada por "Binomial negativa" establece la distribución de acuerdo con esta ley, con parámetros $n = 0,13174$ y $h = 2,555$, encontrándose el último por el método de máxima verosimilitud. El valor de χ^2 es 6,9 que da una probabilidad de 0,14 para 4 grados de libertad, de tal manera que la representación es aceptable. Existen indicios de que la distribución binomial negativa represente en forma insuficiente el extremo de la cola y en algunas aplicaciones sería deseable mejorar el modelo. El modelo se puede usar sin problemas para las aplicaciones que no poseen un grado significativo de asimetría.

OBSERVACIÓN. Sería útil comentar que las diferencias en la estructura de las colas de las distribuciones anteriores, aunque son significativas, conciernen sólo a un número muy pequeño de accidentes y no necesariamente tienen un efecto importante en las finanzas de los negocios, las que están determinadas por un volumen muy grande de casos en la parte superior de la Tabla 2.6.1.

Otro enfoque es aproximar la función de estructura $H(q)$ por una f.d. discreta, suponiendo valores q_1, q_2, \dots, q_r con probabilidades b_1, b_2, \dots, b_r . En realidad, esto significa que la f.d. está compuesta por r términos Poisson. Cuanto mayor número de parámetros libres haya, mayor será la posibilidad de lograr ~~un~~ ajustamiento razonable, aun en carteras heterogéneas.

En este caso una distribución de dos términos da un resultado satisfactorio. Los valores de los parámetros $q_1 = 0,65341$ y $q_2 = 2,1293$, con probabilidades $b = 0,76519$ y $1-b = 0,23481$, pueden encontrarse a partir de la ecuación

$$h \cdot p_k(n \cdot q_1) + (1-h) \cdot p_k(n \cdot q_2) = p_k \quad (2.6.2)$$

igualando la media y el desvío estándar. La última columna de la Tabla 2.6.1 muestra esta distribución.

(c) Análisis. Para la mayoría de las aplicaciones que trataremos en ~~este~~ libro, la variación interna en el grupo no es importante. El comportamiento del riesgo se considerará en el acumulado para un período dado o dividido en períodos consecutivos, en general en años calendario. La heterogeneidad se toma en cuenta por medio de una asignación adecuada de los valores para las distribuciones pertinentes y para los parámetros del modelo. Cuando éstos se deducen de datos reales, esto sucede en forma automática.

Téngase en cuenta que nuestra notación y la mayoría de las fórmulas presentadas en la sección 2.4 de las variables de Poisson mixtas son formalmente, exactamente las mismas que aquellas que muchos autores utilizan para consideraciones sobre variables de estructura. Esto es natural, debido a que las estructuras reales de los procesos son similares, a pesar de que las interpretaciones físicas son diferentes. Por lo tanto, debe tenerse cuidado cuando se lee bibliografía sobre la teoría del riesgo, mas aún ya que con frecuencia no se hace una distinción clara de la terminología. En realidad, debido a las estructuras similares, el término **variable de estructura** se ha utilizado tanto en el contexto de un efecto de ponderación relacionado con el tiempo como para la heterogeneidad genuina de las estructuras. Desafortunadamente, éste también fue el caso en *Risk Theory* de Beard y otros (1984), el precursor de este libro. La terminología ha cambiado.

(d) Referencias. Existe una vasta bibliografía acerca de las funciones de estructura. El inconveniente para encontrar los parámetros q_i y b_i para la aproximación discreta fue tratado por D'Hooge y Goovaerts (1976). Gossiaux y Lemaire (1981) estudiaron el ajuste de los métodos anteriores y aplicaron las estadísticas de los accidentes automovilísticos. Loimaratana y otros. (1980) presentó un enfoque de análisis grupal como una solución al mismo problema.