

Capítulo 4

Procesos de Poisson

4.1. Distribución Exponencial

Definición 4.1 Una variable aleatoria T tiene distribución exponencial con parámetro λ , $T \sim \text{Exp}(\lambda)$, si su función de distribución está dada por

$$F_T(t) = P(T \leq t) = 1 - e^{-\lambda t}, \quad \text{para } t \geq 0.$$

Equivalentemente, T tiene densidad $f_T(t)$ dada por

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{para } t \geq 0, \\ 0 & \text{para } t < 0. \end{cases}$$

Esta distribución tiene el siguiente valor esperado:

$$\begin{aligned} E[T] &= \int_{-\infty}^{\infty} t f_T(t) dt = \int_0^{\infty} t \lambda e^{-\lambda t} dt \\ &= -te^{-\lambda t} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}. \end{aligned}$$

De manera similar podemos calcular $E[T^2]$ integrando por partes,

$$\begin{aligned} E[T^2] &= \int_{-\infty}^{\infty} t^2 f_T(t) dt = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt \\ &= -t^2 e^{-\lambda t} \Big|_0^{\infty} + \int_0^{\infty} 2te^{-\lambda t} dt = \frac{2}{\lambda^2} \end{aligned}$$

y por lo tanto, la varianza de T es

$$\text{Var}(T) = E[T^2] - (E[T])^2 = \frac{1}{\lambda^2}.$$

4.1.1. Falta de Memoria

Una de las propiedades fundamentales de la distribución exponencial es la siguiente:

$$P(T > t + s | T > t) = P(T > s).$$

Para demostrar esta propiedad usamos la definición de probabilidad condicional

$$P(T > t + s | T > t) = \frac{P(T > t + s)}{P(T > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(T > s).$$

4.1.2. Mínimo de Variables Exponenciales

Sean $S \sim \mathcal{Exp}(\lambda)$ y $T \sim \mathcal{Exp}(\mu)$ variables independientes. Tenemos en primer lugar

$$\begin{aligned} P(\min(S, T) > t) &= P(S > t, T > t) \\ &= P(S > t)P(T > t) = e^{-(\lambda+\mu)t}, \end{aligned}$$

es decir, $\min(S, T)$ tiene distribución exponencial de parámetro $\lambda + \mu$. El mismo cálculo muestra que para una colección de variables independientes T_1, \dots, T_n con $T_i \sim \mathcal{Exp}(\lambda_i)$, $1 \leq i \leq n$,

$$\begin{aligned} P(\min(T_1, \dots, T_n) > t) &= P(T_1 > t, \dots, T_n > t) \\ &= \prod_{i=1}^n P(T_i > t) = \prod_{i=1}^n e^{-\lambda_i t} = e^{-(\lambda_1 + \dots + \lambda_n)t} \end{aligned} \quad (4.1)$$

En consecuencia, el mínimo de varias variables independientes con distribuciones exponenciales tiene distribución exponencial con parámetro igual a la suma de los parámetros.

Veamos ahora con qué probabilidad una variable exponencial es menor que otra. Sean $S \sim \mathcal{Exp}(\lambda)$ y $T \sim \mathcal{Exp}(\mu)$, tenemos

$$\begin{aligned} P(T > S) &= \int_0^\infty P(T > s | S = s) f_S(s) ds \\ &= \int_0^\infty \lambda e^{-\lambda s} e^{-\mu s} ds = \frac{\lambda}{\lambda + \mu} \int_0^\infty (\lambda + \mu) e^{-(\lambda + \mu)s} ds \\ &= \frac{\lambda}{\lambda + \mu}. \end{aligned}$$

Para varias variables, el resultado es el siguiente

$$\begin{aligned} P(T_i = \min(T_1, \dots, T_n)) &= P(T_i < T_1, \dots, T_i < T_{i-1}, T_i < T_{i+1}, \dots, T_i < T_n) \\ &= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}. \end{aligned}$$

Para demostrar esta propiedad llamemos $S = T_i$ y sea U el mínimo de T_j , $j \neq i$. Por (4.1) sabemos que U es exponencial con parámetro $\mu = (\lambda_1 + \dots + \lambda_n) - \lambda_i$. Usando el resultado para dos variables

$$P(T_i = \min(T_1, \dots, T_n)) = P(S < U) = \frac{\lambda_i}{\lambda_i + \mu} = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}.$$

Sea I el índice (aleatorio) de la menor de las variables exponenciales, hemos demostrado que

$$P(I = i) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}.$$

Lema 4.1 I y $V = \min(T_1, \dots, T_n)$ son independientes.

Demostración. Calculamos la siguiente probabilidad conjunta

$$\begin{aligned} P(I = i, V > t) &= P(T_i > t, T_j > T_i, \forall j \neq i) = \int_t^\infty P(T_j > s, \forall j \neq i) f_{T_i}(s) ds \\ &= \int_t^\infty \lambda_i e^{-\lambda_i s} \prod_{j \neq i} e^{-\lambda_j s} ds = \lambda_i \int_t^\infty e^{-s(\sum_j \lambda_j)} ds \\ &= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n} e^{-t(\sum_j \lambda_j)} = P(I = i)P(V > t). \end{aligned}$$

Veamos a continuación cómo se distribuye una suma de exponenciales. ■

Teorema 4.1 Sean T_1, T_2, \dots v.a.i.i.d. con distribución exponencial de parámetro λ . La suma $\tau_n = T_1 + \dots + T_n$ tiene distribución $\Gamma(n, \lambda)$, es decir, la densidad está dada por

$$f_{\tau_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \quad \text{para } t \geq 0$$

y 0 en otro caso.

Demostración. Haremos la prueba por inducción. Para $n = 1$, $\tau_1 = T_1$ tiene distribución exponencial de parámetro λ , que concuerda con la densidad de la fórmula anterior.

Supongamos ahora que la fórmula es cierta para n . Tenemos $\tau_{n+1} = \tau_n + T_{n+1}$ y por independencia

$$\begin{aligned} P(\tau_{n+1} \leq t) &= \int_0^t P(\tau_n + T_{n+1} \leq t | \tau_n = s) f_{\tau_n}(s) ds \\ &= \int_0^t P(T_{n+1} \leq t - s) f_{\tau_n}(s) ds \end{aligned}$$

Usamos ahora la distribución exponencial para el primer factor y la fórmula inductiva para el segundo obtenemos

$$\begin{aligned} \int_0^t (1 - e^{-\lambda(t-s)}) \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds &= \frac{\lambda^n}{(n-1)!} \int_0^t e^{-\lambda s} s^{n-1} ds - \frac{\lambda^n}{(n-1)!} \int_0^t e^{-\lambda t} s^{n-1} ds \\ &= \frac{\lambda^n}{(n-1)!} \left[\frac{1}{n} s^n e^{-\lambda s} \Big|_0^t + \int_0^t \lambda \frac{s^n}{n} e^{-\lambda s} ds - \frac{t^n}{n} e^{-\lambda t} \right] \\ &= \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^n}{n!} ds. \end{aligned}$$

■

Como consecuencia del teorema anterior, teniendo en cuenta que la distribución $\Gamma(n, \lambda)$ se obtiene como suma de v.a.i.i.d. con distribución exponencial de parámetro λ , vemos que

$$E[\tau_n] = \frac{n}{\lambda}, \quad \text{Var}(\tau_n) = \frac{n}{\lambda^2}.$$

También es posible demostrar que la función de distribución de τ_n se puede escribir de la siguiente manera:

$$P(\tau_n \leq x) = 1 - \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} e^{-\lambda x} = \sum_{i=n}^{\infty} \frac{(\lambda x)^i}{i!} e^{-\lambda x}$$

Observación 4.1 Tenemos los siguientes casos especiales de la distribución Gamma: $\Gamma(1, \lambda)$ es la distribución exponencial de parámetro λ mientras que $\Gamma(k, 2)$ es la distribución Ji-cuadrado con $2k$ grados de libertad, χ_{2k}^2 . Además, si $X \sim \Gamma(n, \lambda)$ entonces $cX \sim \Gamma(n, \lambda/c)$.

4.2. La Distribución de Poisson

Definición 4.2 Una variable aleatoria X tiene distribución de Poisson de parámetro $\lambda > 0$ si toma valores en el conjunto $\{0, 1, 2, \dots\}$, con probabilidad dada por

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

Calculemos la función generadora de probabilidad de una variable de este tipo:

$$\phi_X(s) = E[s^X] = \sum_{k=0}^{\infty} s^k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(s\lambda)^k}{k!} = e^{\lambda(s-1)}.$$

A partir de esta expresión podemos obtener los momentos de la distribución:

$$\begin{aligned} E[X] &= \left. \frac{d\phi}{ds} \right|_{s=1} = \left. \lambda e^{\lambda(s-1)} \right|_{s=1} = \lambda, \\ E[X(X-1)] &= \left. \frac{d^2\phi}{ds^2} \right|_{s=1} = \left. \lambda^2 e^{\lambda(s-1)} \right|_{s=1} = \lambda^2, \\ E[X^2] &= E[X(X-1)] + E[X] = \lambda^2 + \lambda, \\ \text{Var}(X) &= E[X^2] - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \end{aligned}$$

Si $X \sim \mathcal{Pois}(\lambda)$ e $Y \sim \mathcal{Pois}(\mu)$ son independientes entonces la suma tiene f.g.p.

$$\phi_{X+Y}(s) = \phi_X(s)\phi_Y(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)}$$

y vemos que $X + Y$ tiene distribución de Poisson con parámetro $\lambda + \mu$.

Lema 4.2 Sea $N \sim \mathcal{Pois}(\lambda)$ y condicional a N , M tiene distribución binomial con parámetros N y p . Entonces la distribución (incondicional) de M es Poisson con parámetro λp .

Demostración. Podemos considerar M como la suma de una cantidad aleatoria N de variables de Bernoulli con probabilidad de éxito p :

$$M = X_1 + \cdots + X_N$$

donde X_i , $i \geq 1$ tiene distribución de Bernoulli con probabilidad de éxito p . La f.g.p. de una variable de Bernoulli es

$$\phi_X(s) = E[s^X] = q + sp$$

y ya hemos visto que la f.g.p. de N es $\phi_N(t) = e^{\lambda(t-1)}$. Por lo tanto, la f.g.p. de M es la composición de estas dos:

$$\phi_M(s) = \phi_N(\phi_X(s)) = e^{\lambda(q+sp-1)} = e^{\lambda(sp-p)} = e^{\lambda p(s-1)}$$

que es la f.g.p. de una v.a. de Poisson con parámetro λp . ■

4.3. El Proceso de Poisson

Definición 4.3 Sean T_1, T_2, \dots v.a.i.i.d. con distribución exponencial de parámetro λ , $\tau_0 = 0$ y $\tau_n = T_1 + \cdots + T_n$ para $n \geq 1$. Definimos el *proceso de Poisson de parámetro o intensidad* λ por

$$N(s) = \max\{n : \tau_n \leq s\}, \quad s \geq 0.$$

Las variables T_n representan los intervalos de tiempo entre eventos sucesivos (llegadas de clientes a una cola, de llamadas a una central telefónica, de pacientes a la emergencia de un hospital, etc.) y en consecuencia $\tau_n = T_1 + \cdots + T_n$ es el instante en el que ocurre el n -ésimo evento y $N(s)$ es el número de eventos que han ocurrido hasta el instante s . Llamaremos *tiempos de llegada* del proceso a las variables τ_n , $n \geq 1$.

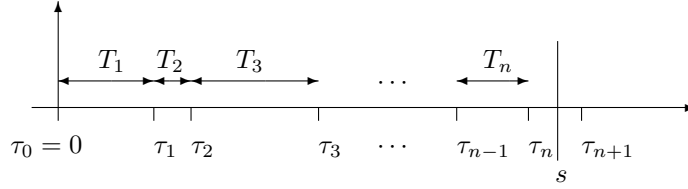


Figura 4.1

Para ver por qué $N(s)$, $s \geq 0$, recibe este nombre, calculemos su distribución: $N(s) = n$ si y sólo si $\tau_n \leq s < \tau_{n+1}$, es decir, el n -ésimo evento ocurre antes del instante s pero el $(n+1)$ -ésimo ocurre después de s . Usando la ley de la probabilidad total, condicionando respecto al instante en el cual ocurre τ_n , obtenemos

$$\begin{aligned} P(N(s) = n) &= P(\tau_{n+1} > s > \tau_n) = \int_0^s P(\tau_{n+1} > s | \tau_n = t) f_{\tau_n}(t) dt \\ &= \int_0^s P(T_{n+1} > s - t) f_{\tau_n}(t) dt. \end{aligned}$$

Usando ahora el resultado del teorema 4.1 obtenemos

$$\begin{aligned} &= \int_0^s \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda(s-t)} dt \\ &= \frac{\lambda^n}{(n-1)!} e^{-\lambda s} \int_0^s t^{n-1} dt = \frac{(\lambda s)^n}{n!} e^{-\lambda s}. \end{aligned}$$

Por lo tanto hemos demostrado el siguiente resultado

Lema 4.3 $N(s)$ tiene distribución de Poisson de parámetro λs .

Veamos algunas propiedades del proceso que acabamos de definir.

Lema 4.4 $N(t+s) - N(s)$, $t \geq 0$ es un proceso de Poisson de parámetro λ y es independiente de $N(r)$, $0 \leq r \leq s$.

Demostración. Supongamos que $N(s) = n$ y que el n -ésimo evento ocurrió en el instante τ_n . Sabemos que el intervalo de tiempo para el siguiente evento debe satisfacer $T_{n+1} > s - \tau_n$, pero por la propiedad de falta de memoria de la distribución exponencial

$$P(T_{n+1} > s - \tau_n + t | T_{n+1} > s - \tau_n) = P(T_{n+1} > t) = e^{-\lambda t}.$$

Esto muestra que la distribución del tiempo de espera hasta el primer evento después de s es exponencial de parámetro λ y es independiente de T_i , $1 \leq i \leq n$. Por otro lado T_{n+1}, T_{n+2}, \dots son independientes de T_i , $1 \leq i \leq n$ y por lo tanto también de τ_i , $1 \leq i \leq n$. Esto muestra que los intervalos entre eventos que ocurren después de s son v.a.i.i.d. con distribución exponencial de parámetro λ , y por lo tanto $N(t+s) - N(s)$ es un proceso de Poisson. ■

Como consecuencia de este resultado tenemos

Lema 4.5 $N(t)$ tiene incrementos independientes: Si $t_0 < t_1 < \dots < t_n$, entonces

$$N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$$

son independientes.

Demostración. El lema 4.5 implica que $N(t_n) - N(t_{n+1})$ es independiente de $N(r)$, $r \leq t_{n-1}$ y en consecuencia también de $N(t_{n-1}) - N(t_{n-2}), \dots, N(t_1) - N(t_0)$. El resultado sigue por inducción. ■

Combinando los dos lemas anteriores tenemos la mitad del siguiente resultado, que es una caracterización fundamental del proceso de Poisson.

Teorema 4.2 Si $\{N(s), s \geq 0\}$ es un proceso de Poisson de parámetro $\lambda > 0$, entonces

1. $N(0) = 0$.
2. $N(t+s) - N(s) \sim \text{Pois}(\lambda t)$.
3. $N(t)$ tiene incrementos independientes.

Recíprocamente, si 1, 2 y 3 valen, entonces $\{N(s), s \geq 0\}$ es un proceso de Poisson.

Demostración. Los lemas 4.3 y 4.4 demuestran la primera afirmación. Para ver el recíproco, sea τ_n el instante en el cual ocurre el n -ésimo evento. El primer evento ocurre después de t si y sólo si no ocurre ningún evento en $[0, t]$. Usando la fórmula para la distribución de Poisson

$$P(\tau_1 > t) = P(N(t) = 0) = e^{-\lambda t}$$

lo cual muestra que $\tau_1 = T_1 \sim \text{Exp}(\lambda)$. Para $T_2 = \tau_2 - \tau_1$ observamos que

$$\begin{aligned} P(T_2 > t | T_1 = s) &= P(\text{no ocurre ningún evento en } (s, s+t] | T_1 = s) \\ &= P(N(t+s) - N(s) = 0 | N(r) = 0 \text{ para } r < s, N(s) = 1) \\ &= P(N(t+s) - N(s) = 0) = e^{-\lambda t} \end{aligned}$$

por la propiedad de incrementos independientes, de modo que $T_2 \sim \text{Exp}(\lambda)$ y es independiente de T_1 . Repitiendo este argumento vemos que T_1, T_2, \dots son i.i.d. con distribución exponencial de parámetro λ . ■

Ejemplo 4.1

Un cable submarino tiene defectos de acuerdo a un proceso de Poisson de parámetro $\lambda = 0.1$ por km. (a) ¿Cuál es la probabilidad de que no haya defectos en los primeros dos kilómetros de cable? (b) Si no hay defectos en los primeros dos kilómetros, ¿cuál es la probabilidad de que tampoco los haya en el tercer kilómetro?

(a) $N(2)$ tiene distribución de Poisson de parámetro $(0.1)(2) = 0.2$. Por lo tanto

$$P(N(2) = 0) = e^{-0.2} = 0.8187.$$

(b) $N(3) - N(2)$ y $N(2) - N(0) = N(2)$ son independientes, de modo que

$$P(N(3) - N(2) = 0 | N(2) = 0) = P(N(3) - N(2) = 0) = e^{-0.1} = 0.9048$$

▲

Ejemplo 4.2

Los clientes llegan a una tienda de acuerdo con un proceso de Poisson de tasa $\lambda = 4$ por hora. Si la tienda abre a las 9 a.m. ¿Cuál es la probabilidad de que exactamente un cliente haya entrado antes de las 9:30 a.m. y que un total de cinco hayan entrado antes de las 11:30 a.m.?

Medimos el tiempo t en horas a partir de las 9 a.m. Queremos hallar $P(N(1/2) = 1, N(5/2) = 5)$, y para esto usaremos la independencia de los incrementos:

$$\begin{aligned} P(N(1/2) = 1, N(5/2) = 5) &= P(N(1/2) = 1, N(5/2) - N(1/2) = 4) \\ &= \left(\frac{e^{-4(1/2)} 4(1/2)}{1!} \right) \left(\frac{e^{-4(2)} [4(2)]^4}{4!} \right) \\ &= (2e^{-2}) \left(\frac{512}{3} e^{-8} \right) = 0.0155. \end{aligned}$$

▲

La importancia de la distribución de Poisson, y del proceso de Poisson en particular, se debe, al menos en parte, al siguiente resultado, que se conoce como la ley de los eventos raros.

Consideremos una cantidad grande n de ensayos de Bernoulli independientes con probabilidad de éxito p constante. Sea S_n el número de éxitos en los n ensayos. Sabemos que S_n tiene distribución Binomial de parámetros n y p :

$$P(S_n = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$

Supongamos ahora que el número de ensayos n tiende a infinito y p tiende a 0, de modo que $np = \lambda$. Veamos que ocurre con la distribución de S_n en este caso. Reemplacemos p por λ/n en la ecuación anterior

$$\begin{aligned} P(S_n = k) &= \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}. \end{aligned} \quad (4.2)$$

Veamos ahora el comportamiento de estos cuatro factores cuando $n \rightarrow \infty$. El primer factor no depende de n . En el segundo hay k factores en el numerador y k en el denominador y podemos escribirlo

$$\frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n}.$$

En virtud de que k está fijo es fácil ver que todos estos factores tienden a 1, y en consecuencia su producto también. El tercer factor converge a $e^{-\lambda}$. Finalmente, el último converge a 1 ya que $\lambda/n \rightarrow 0$ y la potencia k de este factor está fija. Reuniendo estos resultados vemos que la probabilidad (4.2) converge a

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

que es la distribución de Poisson de parámetro λ . El mismo resultado es cierto si en lugar de tener $np = \lambda$ tenemos que $p \rightarrow 0$ cuando $n \rightarrow \infty$ de modo que $np \rightarrow \lambda$.

En realidad la ley de eventos raros se cumple con mayor generalidad aún. Es posible suponer que los ensayos de Bernoulli no tienen una probabilidad de éxito común, como lo muestra el siguiente teorema. Primero enunciamos y demostramos un resultado auxiliar.

Lema 4.6 Sean S y T dos variables aleatorias y A un subconjunto medible de \mathbb{R} . Entonces

$$|P(S \in A) - P(T \in A)| \leq P(S \neq T).$$

Demostración.

$$\begin{aligned} P(S \in A) &= P(S \in A, S = T) + P(S \in A, S \neq T) = P(T \in A, S = T) + P(S \in A, S \neq T) \\ &= P(T \in A, S = T) + P(T \in A, S \neq T) - P(T \in A, S \neq T) + P(S \in A, S \neq T) \\ &= P(T \in A) - P(T \in A, S \neq T) + P(S \in A, S \neq T) \end{aligned}$$

Por lo tanto

$$P(S \in A) - P(T \in A) = P(S \in A, S \neq T) - P(T \in A, S \neq T) \leq P(S \in A, S \neq T) \leq P(S \neq T),$$

y de manera similar

$$P(T \in A) - P(S \in A) \leq P(S \neq T),$$

de modo que

$$|P(T \in A) - P(S \in A)| \leq P(S \neq T),$$

■

Teorema 4.3 (Le Cam) Sean X_m , $1 \leq m \leq n$, variables aleatorias independientes con

$$P(X_m = 1) = p_m, \quad P(X_m = 0) = 1 - p_m.$$

Sean

$$S_n = X_1 + \cdots + X_n, \quad \lambda_n = E[S_n] = p_1 + \cdots + p_n.$$

Entonces, para cualquier conjunto A ,

$$\left| P(S_n \in A) - \sum_{k \in A} e^{-\lambda_n} \frac{\lambda_n^k}{k!} \right| \leq \sum_{m=1}^n p_m^2$$

Demostación. Las variables X_m son independientes y tienen distribución de Bernoulli con parámetro p_m . Definimos variables independientes $Y_m \sim \mathcal{Pois}(p_m)$, y como la suma de variables Poisson independientes es Poisson, tenemos que $Z_n = Y_1 + \cdots + Y_n$ tiene distribución $\mathcal{Pois}(\lambda_n)$ y

$$P(Z_n \in A) = \sum_{k \in A} e^{-\lambda_n} \frac{\lambda_n^k}{k!}.$$

Por lo tanto queremos comparar $P(S_n \in A)$ y $P(Z_n \in A)$ para cualquier conjunto A de enteros positivos. Por el lema 4.6

$$|P(S_n \in A) - P(Z_n \in A)| \leq P(S_n \neq Z_n) = P\left(\sum_{m=1}^n X_m \neq \sum_{m=1}^n Y_m\right),$$

pero si S_n y Z_n difieren, al menos uno de los pares X_m y Y_m deben diferir también. En consecuencia

$$|P(S_n \in A) - P(Z_n \in A)| \leq \sum_{m=1}^n P(X_m \neq Y_m)$$

y para completar la demostración hay que ver que $P(X_m \neq Y_m) \leq p_m^2$. Para simplificar la notación sean $X \sim \text{Ber}(p)$ y $Y \sim \mathcal{Pois}(p)$ y veamos que $P(X \neq Y) \leq p^2$, o equivalentemente, que

$$1 - p^2 \leq P(X = Y) = P(X = Y = 0) + P(X = Y = 1).$$

Este resultado no depende de la distribución conjunta entre X y Y pues no hemos supuesto ninguna propiedad de independencia entre ellas. Lo importante es que las distribuciones marginales de X y Y sigan siendo las mismas. Escogemos la distribución conjunta de (X, Y) de la siguiente manera: Sea U una variable con distribución uniforme en $(0, 1]$ y sean

$$X = \begin{cases} 0 & \text{si } 0 < U \leq 1 - p \\ 1 & \text{si } 1 - p < U \leq 1. \end{cases}$$

$Y = 0$ si $0 < U < e^{-p}$ y para $k = 1, 2, \dots$

$$Y = k \quad \text{si} \quad \sum_{i=0}^{k-1} e^{-p} \frac{p^i}{i!} < U \leq \sum_{i=0}^k e^{-p} \frac{p^i}{i!}.$$

Es sencillo verificar que X y Y tienen las distribuciones marginales adecuadas (esto no es más que el método de la transformada inversa de generación de variables aleatorias, aplicado las distribuciones de Bernoulli y de Poisson). Como $1 - p \leq e^{-p}$ tenemos que $X = Y = 0$ si y sólo si $U \leq 1 - p$, de modo que

$$P(X = Y = 0) = 1 - p.$$

De manera similar $X = Y = 1$ si y sólo si $e^{-p} < U \leq (1+p)e^{-p}$ y

$$P(X = Y = 1) = pe^{-p}.$$

Sumando estas dos expresiones tenemos

$$P(X = Y) = 1 - p + pe^{-p} = 1 - p^2 + \frac{p^3}{2} + \dots \geq 1 - p^2.$$

■

Corolario 4.1 (Le Cam) Para cada n , sean $X_{n,m}$, $1 \leq m \leq n$, $n \geq 1$ variables aleatorias independientes con

$$P(X_{n,m} = 1) = p_{n,m}, \quad P(X_{n,m} = 0) = 1 - p_{n,m}.$$

Sean

$$S_n = X_{n,1} + \dots + X_{n,n}, \quad \lambda_n = E[S_n] = p_{n,1} + \dots + p_{n,n},$$

y $Z_n \sim \text{Pois}(\lambda_n)$. Entonces, para cualquier conjunto A ,

$$|P(S_n \in A) - P(Z_n \in A)| \leq \sum_{m=1}^n p_{n,m}^2$$

El teorema anterior nos da una cota para la diferencia entre la distribución de S_n y la distribución de Poisson de parámetro $\lambda_n = E[S_n]$, que podemos usar para obtener una versión general del teorema de convergencia de Poisson.

Corolario 4.2 Supongamos que en la situación del corolario anterior $\lambda_n \rightarrow \lambda < \infty$ y $\max_k p_{n,k} \rightarrow 0$, cuando $n \rightarrow \infty$, entonces

$$\max_A |P(S_n \in A) - P(Z_n \in A)| \rightarrow 0.$$

Demostración. Como $p_{n,m}^2 \leq p_{n,m}(\max_k p_{n,k})$, sumando sobre m obtenemos

$$\sum_{m=1}^n p_{n,m}^2 \leq \max_k p_{n,k} \sum_m p_{n,m}.$$

El primer factor de la derecha va a 0 por hipótesis. El segundo es λ_n que converge a $\lambda < \infty$ y en consecuencia el producto de los dos converge a 0. ■

4.4. Postulados para el Proceso de Poisson

Consideremos una sucesión de eventos que ocurren en $[0, \infty)$ como por ejemplo las emisiones de partículas por una sustancia radioactiva, la llegada de llamadas a una central telefónica, los accidentes que ocurren en cierto cruce carretero, la ubicación de fallas o defectos a lo largo de una fibra o las llegadas sucesivas de clientes a un establecimiento comercial. Sea $N((a, b])$ el número de eventos que ocurren en el intervalo $(a, b]$, es decir, si $\tau_1 < \tau_2 < \tau_3 \dots$ representan los instantes (o ubicaciones) de los sucesivos eventos, entonces $N((a, b])$ es el número de estos instantes τ_i que satisfacen $a < \tau_i \leq b$.

Proponemos los siguientes postulados:

1. El número de eventos que ocurren en intervalos disjuntos son variables aleatorias independientes: Para cualquier entero $m \geq 2$ y cualesquiera instantes $t_0 = 0 < t_1 < t_2 < \dots < t_m$, las variables aleatorias

$$N((t_0, t_1]), N((t_1, t_2]), \dots, N((t_{m-1}, t_m])$$

son independientes.

2. Para cualquier instante t y cualquier $h > 0$, la distribución de probabilidad de $N((t, t+h])$ depende sólo de la longitud del intervalo h y no del instante inicial t .
3. Hay una constante positiva λ para la cual la probabilidad de que ocurra al menos un evento en un intervalo de longitud h es

$$P(N((t, t+h]) \geq 1) = \lambda h + o(h), \quad \text{cuando } h \downarrow 0$$

(la notación $o(h)$ indica una función general indeterminada que representa el resto y satisface $o(h)/h \rightarrow 0$ cuando $h \downarrow 0$, es decir, que es de orden menor que h cuando $h \downarrow 0$). El parámetro λ se conoce como la intensidad del proceso.

4. La probabilidad de que haya dos o más eventos en un intervalo de longitud h es $o(h)$:

$$P(N((t, t+h]) \geq 2) = o(h), \quad \text{cuando } h \downarrow 0.$$

El número de sucesos que ocurren en intervalos disjuntos son independientes por 1, y 2 afirma que la distribución de $N((s, t])$ es la misma que la de $N((0, t-s])$. Por lo tanto, para describir la ley de probabilidad del sistema basta determinar la distribución de probabilidad de $N((0, t])$ para cualquier valor de t . Llamemos $N((0, t]) = N(t)$. Mostraremos que los postulados anteriores implican que $N(t)$ tiene una distribución de Poisson:

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad \text{para } k = 0, 1, \dots \quad (4.3)$$

Para demostrar (4.3) dividimos el intervalo $(0, t]$ en n subintervalos de igual longitud $h = t/n$ y definimos las siguientes variables de Bernoulli: $\xi_{n,i} = 1$ si hay al menos un evento en el intervalo $((i-1)t/n, it/n]$ y $\xi_{n,i} = 0$ si no, para $1 \leq i \leq n$. $S_n = \xi_{n,1} + \dots + \xi_{n,n}$ representa el número de subintervalos que contienen al menos un evento y

$$p_{n,i} = P(\xi_{n,i} = 1) = \frac{\lambda t}{n} + o\left(\frac{t}{n}\right)$$

según el postulado 3. Sea

$$\mu_n = \sum_{i=1}^n p_{i,n} = \lambda t + n o\left(\frac{t}{n}\right).$$

Usando el teorema 4.3 vemos que

$$\begin{aligned} \left| P(S_n = k) - \frac{\mu_n^k e^{-\mu_n}}{k!} \right| &\leq n \left[\frac{\lambda t}{n} + o^2\left(\frac{t}{n}\right) \right] \\ &= \frac{(\lambda t)^2}{n} + 2\lambda t o\left(\frac{t}{n}\right) + n o^2\left(\frac{t}{n}\right), \end{aligned}$$

Como $o(h) = o(t/n)$ es un término de orden menor que $h = t/n$ para n grande, se tiene que

$$n o(t/n) = t \frac{o(t/n)}{t/n} = t \frac{o(h)}{h}$$

tiende a 0 cuando n crece. Pasando al límite cuando $n \rightarrow \infty$ obtenemos que

$$\lim_{n \rightarrow \infty} P(S_n = k) = \frac{\mu^k e^{-\mu}}{k!}, \quad \text{con } \mu = \lambda t.$$

Para completar la demostración sólo falta ver que

$$\lim_{n \rightarrow \infty} P(S_n = k) = P(N((0, t]) = k)$$

pero S_n y $N((0, t])$ son diferentes si al menos uno de los subintervalos contiene dos o más eventos, y el postulado 4 impide esto porque

$$\begin{aligned} |P(N(t) = k) - P(S_n = k)| &\leq P(N(t) \neq S_n) \\ &\leq \sum_{i=1}^n P\left(N\left(\left(\frac{i-1}{n}t, \frac{it}{n}\right]\right) \geq 2\right) \\ &\leq n o\left(\frac{t}{n}\right) \\ &\rightarrow 0 \quad \text{cuando } n \rightarrow \infty. \end{aligned}$$

Tomando n suficientemente grande tenemos que

$$P(N((0, t]) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad \text{para } k \geq 0.$$

Esto completa la demostración de (4.3). ■

El proceso $N((a, b])$ se conoce como el Proceso Puntual de Poisson y sus valores, como hemos visto, se pueden calcular a partir de los del proceso $N(t)$:

$$N((s, t]) = N(t) - N(s)$$

Recíprocamente, $N(t) = N((0, t])$, de modo que ambos procesos son equivalentes, las diferencias son de enfoque, pero en algunos casos resulta útil considerar al proceso de una u otra manera.

A continuación presentamos sin demostración otra caracterización de los procesos de Poisson que resultará útil más adelante.

Teorema 4.4 $N(t), t \geq 0$ es un proceso de Poisson con intensidad λ si y sólo si

- a) Para casi todo ω , los saltos de $N(t, \omega)$ son unitarios.
- b) Para todo $s, t \geq 0$ se tiene que $E(N(t+s) - N(t) | N(u), u \leq t) = \lambda s$.

4.5. Distribuciones Asociadas a un Proceso de Poisson

Hemos visto que los intervalos de tiempo entre eventos sucesivos, $T_n, n \geq 0$ son v.a.i.i.d. con distribución exponencial de parámetro λ . Los instantes τ_n en los cuales ocurren los eventos, son sumas de las variables anteriores, y en consecuencia tienen distribución $\Gamma(n, \lambda)$.

Teorema 4.5 Sea $N(t), t \geq 0$ un proceso de Poisson de parámetro $\lambda > 0$. Para $0 < u < t$ y $0 \leq k \leq n$,

$$P(N(u) = k | N(t) = n) = \frac{n!}{k!(n-k)!} \left(\frac{u}{t}\right)^k \left(1 - \frac{u}{t}\right)^{n-k}, \quad (4.4)$$

Es decir, condicional a que para el instante t han ocurrido n eventos, la distribución del número de eventos que han ocurrido para el instante $u < t$ es binomial con parámetros n y (u/t) .

Demostración.

$$\begin{aligned} P(N(u) = k | N(t) = n) &= \frac{P(N(u) = k, N(t) = n)}{P(N(t) = n)} \\ &= \frac{P(N(u) = k, N(t) - N(u) = n - k)}{P(N(t) = n)} \\ &= \frac{[e^{-\lambda u} (\lambda u)^k / k!] [e^{-\lambda(t-u)} (\lambda(t-u))^{n-k} / (n-k)!]}{e^{-\lambda t} (\lambda t)^n / n!} \\ &= \frac{n!}{k!(n-k)!} \frac{u^k (t-u)^{n-k}}{t^n}. \end{aligned}$$

■

Ejemplo 4.3

Recordemos que la variable τ_n tiene distribución $\Gamma(n, \lambda)$ y por la observación 1 sabemos que $\lambda\tau_n/2 \sim \Gamma(n, 2) = \chi_{2n}^2$.

Si observamos un proceso de Poisson hasta que se registre un número prefijado m de eventos, el tiempo necesario τ_m puede usarse para construir intervalos de confianza para la intensidad λ del proceso, usando el hecho de que $\lambda\tau_m/2$ tiene distribución χ^2 con $2m$ grados de libertad. Sean $z_{\alpha/2}$ y $z_{1-\alpha/2}$ valores tales que si $Z \sim \chi_{2m}^2$, entonces $P(Z < z_{\alpha/2}) = P(Z > z_{1-\alpha/2}) = \alpha/2$. Tenemos

$$1 - \alpha = P(z_{\alpha/2} \leq \lambda\tau_m/2 \leq z_{1-\alpha/2}) = P\left(\frac{2z_{\alpha/2}}{\tau_m} \leq \lambda \leq \frac{2z_{1-\alpha/2}}{\tau_m}\right).$$

En consecuencia, $(2z_{\alpha/2}/\tau_m, 2z_{1-\alpha/2}/\tau_m)$ es un intervalo de confianza para λ a nivel $1 - \alpha$. ▲

Ejemplo 4.4

Sean N y M dos procesos de Poisson independientes con parámetros respectivos λ y μ . Sean n y m enteros, τ_n el tiempo de espera hasta el n -ésimo evento en el proceso N y γ_m el tiempo de espera hasta el m -ésimo evento en el proceso M . Las variables $\lambda\tau_n/2$ y $\mu\gamma_m/2$ son independientes y tienen distribuciones χ^2 con $2n$ y $2m$ grados de libertad, respectivamente. Por lo tanto, bajo la hipótesis de que $\lambda = \mu$, la variable $m\tau_n/n\gamma_m$ tiene distribución F con $2n$ y $2m$ grados de libertad, y podemos desarrollar una prueba de hipótesis para $\lambda = \mu$. ▲

4.6. Procesos de Poisson Compuestos

Asociamos ahora una variable aleatoria Y_i a cada evento de un proceso de Poisson. Suponemos que las variables Y_i , $i \geq 1$, son i.i.d y también son independientes del proceso. Por ejemplo, el proceso puede representar los carros que llegan a un centro comercial y las variables asociadas, el número de pasajeros que hay en cada uno de ellos; o el proceso puede representar los mensajes que llegan a un computador central para ser transmitidos via internet y las variables Y_i pueden representar el tamaño de los mensajes.

Es natural considerar la suma de las variables Y_i como una variable de interés:

$$S(t) = Y_1 + \cdots + Y_{N(t)}$$

donde ponemos $S(t) = 0$ si $N(t) = 0$. Ya hemos visto que para suma aleatorias, la media es el producto de las medias de N e Y , mientras que la varianza está dada por

$$\text{Var}(S(t)) = \text{E}[N(t)] \text{Var}(Y_i) + \text{Var}(N(t))(\text{E}[Y_i])^2.$$

En nuestro caso, $N(t) \sim \mathcal{Pois}(\lambda t)$ y por lo tanto, $\text{E}[N(t)] = \text{Var}(N(t)) = \lambda t$, de modo que la fórmula anterior es

$$\text{Var}(S(t)) = \lambda t(\text{Var}(Y_i) + (\text{E}[Y_i])^2) = \lambda t \text{E}[Y_i^2].$$

Ejemplo 4.5

El número de clientes de una tienda durante el día tiene distribución de Poisson de media 30 y cada cliente gasta un promedio de \$150 con desviación típica de \$50. Por los cálculos anteriores sabemos que el ingreso medio por día es $30 \cdot \$150 = \4.500 . La varianza del ingreso total es

$$30 \cdot [(\$50)^2 + (\$150)^2] = 750.000$$

Sacando la raíz cuadrada obtenemos una desviación típica de \$ 866,02. ▲

La función de distribución para el proceso de Poisson compuesto $S(t)$ puede representarse explícitamente si condicionamos por los valores de $N(t)$. Recordemos que la distribución de una suma de variables independientes es la convolución de las distribuciones: Si Y tiene f.d. G ,

$$G^{(n)}(y) = P(Y_1 + \cdots + Y_n \leq y) = \int_{-\infty}^{\infty} G^{(n-1)}(y-z) dG(z)$$

con

$$G^{(0)}(y) = \begin{cases} 1 & \text{para } y \geq 0, \\ 0 & \text{para } y < 0. \end{cases}$$

Ahora

$$\begin{aligned} P(S(t) \leq z) &= P\left(\sum_{k=1}^{N(t)} Y_k \leq z\right) \\ &= \sum_{n=0}^{\infty} P\left(\sum_{k=1}^{N(t)} Y_k \leq z \mid N(t) = n\right) \frac{(\lambda t)^n e^{-\lambda t}}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!} G^{(n)}(z). \end{aligned} \quad (4.5)$$

Ejemplo 4.6

Sea $N(t)$ el número de impactos que recibe un sistema mecánico hasta el instante t y sea Y_k el daño o desgaste que produce el k -ésimo impacto. Suponemos que los daños son positivos: $P(Y_k \geq 0) = 1$, y que se acumulan aditivamente, de modo que $S(t) = \sum_{k=1}^{N(t)} Y_k$ representa el daño total hasta el instante t . Supongamos que el sistema continúa funcionando mientras el daño total sea menor que un valor crítico a y en caso contrario falla. Sea T el tiempo transcurrido hasta que el sistema falla, entonces

$$\{T > t\} \quad \text{sí y sólo sí} \quad \{S(t) < a\}.$$

Teniendo en cuenta esta relación y (4.5) tenemos

$$P(T > t) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!} G^{(n)}(a).$$

Para obtener el tiempo promedio hasta que el sistema falle podemos integrar esta probabilidad:

$$\begin{aligned} E[T] &= \int_0^{\infty} P(T > t) dt = \sum_{n=0}^{\infty} \left(\int_0^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!} dt \right) G^{(n)}(a) \\ &= \lambda^{-1} \sum_{n=0}^{\infty} G^{(n)}(a), \end{aligned}$$

donde hemos intercambiado series e integrales porque todos los términos son positivos. Esta expresión se simplifica en el caso particular en el cual los daños tienen distribución exponencial de parámetro μ . En este caso la suma $\tau_n = Y_1 + \cdots + Y_n$ tiene distribución $\Gamma(n, \mu)$; sea $M(t)$ el proceso de Poisson asociado a estas variables i.i.d. exponenciales, entonces

$$\begin{aligned} G^{(n)}(z) &= P(\tau_n \leq z) = P(M(z) \geq n) \\ &= 1 - \sum_{k=0}^{n-1} \frac{(\mu z)^k e^{-\mu z}}{k!} = \sum_{k=n}^{\infty} \frac{(\mu z)^k e^{-\mu z}}{k!} \end{aligned}$$

y

$$\begin{aligned} \sum_{n=0}^{\infty} G^{(n)}(a) &= \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} \frac{(\mu a)^k e^{-\mu a}}{k!} = \sum_{k=0}^{\infty} \sum_{n=0}^k \frac{(\mu a)^k e^{-\mu a}}{k!} \\ &= \sum_{k=0}^{\infty} (k+1) \frac{(\mu a)^k e^{-\mu a}}{k!} = 1 + \mu a. \end{aligned}$$

Por lo tanto, cuando Y_i , $i \geq 1$, tienen distribución exponencial de parámetro μ ,

$$E[T] = \frac{1 + \mu a}{\lambda}.$$

▲

4.7. Descomposición de un Proceso de Poisson

En la sección anterior asociamos a cada evento de un proceso de Poisson una variable aleatoria Y_i , ahora vamos a usar estas variables para descomponer el proceso. Sea $N_j(t)$ el número de eventos del proceso que han ocurrido antes de t con $Y_i = j$. Si, por ejemplo, Y_i representa el número de personas en un carro que llega a un centro comercial, $N_j(t)$ representa el número de carros que han llegado antes del instante t con exactamente j personas dentro.

Veamos inicialmente el caso más sencillo, en el cual las variables Y_k son de Bernoulli:

$$P(Y_k = 1) = p, \quad P(Y_k = 0) = 1 - p,$$

para $0 < p < 1$ fijo y $k \geq 1$. Definimos ahora dos procesos, según el valor de las variables Y_k sea 0 ó 1:

$$N_1(t) = \sum_{k=1}^{N(t)} Y_k, \quad \text{y} \quad N_0(t) = N(t) - N_1(t).$$

Los valores de $N_1(t)$ sobre intervalos disjuntos son variables aleatorias independientes, $N_1(0) = 0$ y finalmente, el lema 4.2 nos dice que $N_1(t)$ tiene distribución de Poisson con media $\lambda p t$. Un argumento similar muestra que $N_0(t)$ es un proceso de Poisson con parámetro $\lambda(1-p)$. Lo que resulta más sorprendente es que N_0 y N_1 son procesos independientes. Para ver esto calculemos

$$\begin{aligned} P(N_0(t) = j, N_1(t) = k) &= P(N(t) = j+k, N_1(t) = k) \\ &= P(N_1(t) = k | N(t) = j+k) P(N(t) = j+k) \\ &= \frac{(j+k)!}{j!k!} p^k (1-p)^j \frac{(\lambda t)^{j+k} e^{-\lambda t}}{(j+k)!} \\ &= \left[\frac{e^{-\lambda p t} (\lambda p t)^k}{k!} \right] \left[\frac{e^{-\lambda(1-p)t} (\lambda(1-p)t)^j}{j!} \right] \\ &= P(N_1(t) = k) P(N_0(t) = j) \end{aligned}$$

para $j, k = 0, 1, 2, \dots$

Ejemplo 4.7

Los clientes entran a una tienda de acuerdo a un proceso de Poisson con intensidad de 10 por hora. De manera independiente, cada cliente compra algo con probabilidad $p = 0.3$ o sale de la tienda sin comprar nada con probabilidad $q = 1-p = 0.7$. ¿Cuál es la probabilidad de que durante la primera hora 9 personas entren a la tienda y que tres de estas personas compren algo y las otras 6 no?

Sea $N_1 = N_1(1)$ el número de clientes que hacen una compra durante la primera hora y $N_0 = N_0(1)$ el número de clientes que entran pero no compran nada. Entonces N_0 y N_1 son v.a.i. de Poisson con parámetros respectivos $(0.7)(10) = 7$ y $(0.3)(10) = 3$. Por lo tanto

$$P(N_0 = 6) = \frac{7^6 e^{-7}}{6!} = 0.149, \quad P(N_1 = 3) = \frac{3^3 e^{-3}}{3!} = 0.224.$$

y

$$P(N_0 = 6, N_1 = 3) = P(N_0 = 6)P(N_1 = 3) = (0.149)(0.224) = 0.0334. \quad \blacktriangle$$

En el caso general las variables Y_k toman valores sobre un conjunto numerable, por ejemplo sobre $\{0, 1, 2, \dots\}$, y el resultado correspondiente es el siguiente teorema, que no demostraremos.

Teorema 4.6 $N_j(t)$ son procesos de Poisson independientes con intensidad $\lambda P(Y_i = j)$.

4.8. Superposición de Procesos de Poisson

La situación inversa a la descomposición de un proceso de Poisson es la superposición de procesos. Ya que un proceso de Poisson puede descomponerse en procesos de Poisson independientes, es razonable esperar que el proceso inverso, la superposición de procesos de Poisson independientes, produzca un proceso de Poisson cuya intensidad sea la suma de las intensidades.

Teorema 4.7 Sean $N_1(t), \dots, N_k(t)$ procesos de Poisson independientes con parámetros $\lambda_1, \dots, \lambda_k$, entonces $N_1(t) + \dots + N_k(t)$ es un proceso de Poisson con parámetro $\lambda_1 + \dots + \lambda_k$.

Demostración. Haremos la demostración para el caso $k = 2$, el caso general se obtiene luego por inducción. Es inmediato que la suma tiene incrementos independientes y que $N_1(0) + N_2(0) = 0$. Para verificar que los incrementos tienen distribución de Poisson con parámetro igual a la suma de los parámetros observamos que si $Y = N_1(t+s) - N_1(s) \sim \text{Pois}(\lambda_1 t)$ y $Z = N_2(t+s) - N_2(s) \sim \text{Pois}(\lambda_2 t)$, entonces

$$\begin{aligned} N(t+s) - N(s) &= [N_1(t+s) - N_1(s)] + [N_2(t+s) - N_2(s)] \\ &= Y + Z \sim \text{Pois}((\lambda_1 + \lambda_2)t). \end{aligned} \quad \blacksquare$$

Ejemplo 4.8

Consideremos dos procesos de Poisson, uno con parámetro λ , que representa las llegadas a la meta del equipo rojo, y otro, independiente del anterior y con parámetro μ , que representa las llegadas del equipo verde. ¿Cuál es la probabilidad de que haya 6 llegadas rojas antes que 4 verdes?

Observamos que el evento en cuestión equivale a tener al menos 6 rojos en los primeros 9. Si esto ocurre, tenemos a lo sumo tres verdes antes de la llegada del sexto rojo. Por otro lado, si hay 5 o menos rojos en los primeros 9, entonces tendremos al menos 4 verdes.

Podemos ahora ver el problema en el marco de un proceso de Poisson general que incluye rojos y verdes, y tiene parámetro $\lambda + \mu$. Para cada llegada escogemos al azar el color lanzando una moneda con probabilidad $p = \lambda/(\lambda + \mu)$ para rojo. La probabilidad que nos interesa es

$$\sum_{k=6}^9 \binom{9}{k} p^k (1-p)^{9-k}.$$

En el caso particular en el cual ambos procesos iniciales tienen la misma intensidad $\lambda = \mu$, $p = 1/2$ y la expresión anterior es

$$\frac{1}{512} \sum_{k=6}^9 \binom{9}{k} = \frac{140}{512} = 0.273. \quad \blacktriangle$$

4.9. Procesos No Homogéneos

En el corolario 4.2 vimos qué ocurre si la probabilidad de cada evento individual no es homogénea. Si, en cambio, el parámetro del proceso, que representa la intensidad por unidad de tiempo con la cual ocurren los eventos, no es constante a lo largo del tiempo, tenemos un proceso no-homogéneo.

Definición 4.4 Decimos que $(N(t), t \geq 0)$ es un proceso de Poisson no homogéneo con tasa $\lambda(s)$, $s \geq 0$ si

1. $N(0) = 0$,
2. $N(t)$ tiene incrementos independientes,
3. $N(s+t) - N(s)$ tiene distribución de Poisson con media $\int_s^{s+t} \lambda(r) dr$.

En este caso los intervalos de tiempo entre eventos sucesivos, T_n , $n \geq 1$, ya no son independientes ni tienen distribución exponencial. Esta es la razón por la cual no usamos nuestra definición inicial para esta generalización. Veamos que esto es efectivamente cierto. Pongamos $\mu(t) = \int_0^t \lambda(s) ds$, entonces $N(t) \sim \mathcal{Pois}(\mu(t))$ y

$$P(T_1 > t) = P(N(t) = 0) = e^{-\mu(t)}.$$

Derivando obtenemos la densidad

$$f_{T_1}(t) = -\frac{d}{dt}P(T_1 > t) = \lambda(t)e^{-\int_0^t \lambda(s) ds} = \lambda(t)e^{-\mu(t)}$$

para $t \geq 0$. La relación anterior se puede generalizar de la siguiente manera

$$f_{T_1, \dots, T_n}(t_1, \dots, t_n) = \lambda(t_1)\lambda(t_1 + t_2) \cdots \lambda(t_1 + \dots + t_n)e^{-\mu(t_1 + \dots + t_n)},$$

lo cual muestra que, en general, las variables T_i no son independientes ni tienen distribución exponencial.

Ejemplo 4.9

Los clientes llegan a una tienda de acuerdo a un proceso de Poisson no-homogéneo con intensidad

$$\lambda(t) = \begin{cases} 2t & \text{para } 0 \leq t < 1, \\ 2 & \text{para } 1 \leq t < 2, \\ 4-t & \text{para } 2 \leq t \leq 4, \end{cases}$$

donde t se mide en horas a partir de la apertura. ¿Cuál es la probabilidad de que dos clientes lleguen durante las primeras dos horas y dos más durante las dos horas siguientes?

Como las llegadas durante intervalos disjuntos son independientes, podemos responder las dos preguntas por separado. La media para las primeras dos horas es $\mu = \int_0^1 2t dt + \int_1^2 2 dt = 3$ y por lo tanto

$$P(N(2) = 2) = \frac{e^{-3}(3)^2}{2!} = 0.2240.$$

Para las siguientes dos horas, $\mu = \int_2^4 (4-t) dt = 2$ y

$$P(N(4) - N(2) = 2) = \frac{e^{-2}(2)^2}{2!} = 0.2707.$$

La probabilidad que nos piden es

$$P(N(2) = 2, N(4) - N(2) = 2) = P(N(2) = 2)P(N(4) - N(2) = 2) = 0.0606$$

▲

4.9.1. Postulados para un proceso de Poisson no-homogéneo

Al igual que para el caso del proceso homogéneo, es posible demostrar que los siguientes postulados implican que el proceso de conteo $N(t)$ es un proceso de Poisson no-homogéneo con función de intensidad $\lambda(t)$, $t \geq 0$:

- (a) $N(0) = 0$.
- (b) $\{N(t), t \geq 0\}$ tiene incrementos independientes.
- (c) $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$.
- (d) $P(N(t+h) - N(t) \geq 2) = o(h)$.

Muestrear en el tiempo un proceso de Poisson ordinario a una tasa que depende del tiempo produce un proceso de Poisson no-homogéneo. Esto es similar a lo que vimos para la descomposición de un proceso de Poisson sólo que ahora la probabilidad de observar un evento del proceso original no es una constante p como ocurría antes, sino que depende del tiempo: $p(t)$.

Sea $\{N(t), t \geq 0\}$ un proceso de Poisson con intensidad constante λ y supongamos que un evento que ocurre en el instante t se observa con probabilidad $p(t)$, independientemente de lo que haya ocurrido antes. Llamemos $M(t)$ al proceso de los eventos que hemos logrado contar hasta el instante t , entonces $\{M(t), t \geq 0\}$ es un proceso de Poisson no-homogéneo con función de intensidad $\lambda(t) = \lambda p(t)$. Podemos verificar esta afirmación comprobando que se satisfacen los axiomas anteriores.

- (a) $M(0) = 0$.
- (b) El número de eventos que contamos en el intervalo $(t, t+h]$ depende únicamente de los eventos del proceso de Poisson N que ocurren en $(t, t+h]$, que es independiente de lo que haya ocurrido antes de t . En consecuencia el número de eventos observados en $(t, t+h]$ es independiente del proceso de eventos observados hasta el tiempo t , y por lo tanto M tiene incrementos independientes.
- (c) Condicionando sobre $N((t, t+h])$:

$$\begin{aligned} P(M((t, t+h]) = 1) &= P(M((t, t+h]) = 1 | N((t, t+h]) = 1)P(N((t, t+h]) = 1) \\ &\quad + P(M((t, t+h]) = 1 | N((t, t+h]) \geq 2)P(N((t, t+h]) \geq 2) \\ &= P(M((t, t+h]) = 1 | N((t, t+h]) = 1)\lambda h + o(h) \\ &= p(t)\lambda h + o(h) \end{aligned}$$

- (d) $P(M((t, t+h]) \geq 2) \leq P(N((t, t+h]) \geq 2) = o(h)$.

Hay un recíproco (parcial) para este resultado: todo proceso no-homogéneo de Poisson con intensidad acotada se puede obtener a partir de un proceso homogéneo muestreado en el tiempo. Para ver esto necesitamos la siguiente proposición que enunciamos sin demostración

Proposición 4.1 Sean $\{N(t), t \geq 0\}$ y $\{M(t), t \geq 0\}$ procesos de Poisson independientes no-homogéneos, con funciones de intensidad respectivas $\alpha(t)$ y $\beta(t)$ y sea $S(t) = N(t) + M(t)$. Entonces

- (a) $\{S(t), t \geq 0\}$ es un proceso de Poisson no-homogéneo con función de intensidad $\lambda(t) = \alpha(t) + \beta(t)$.
- (b) Dado que un evento del proceso S ocurre en el instante t entonces, independientemente de lo que haya ocurrido antes de t , el evento en t viene del proceso N con probabilidad $\alpha(t)/(\alpha(t) + \beta(t))$.

Demostración. Ver S.M. Ross, *Introduction to Probability Models* 10th. Ed. p. 340.

Supongamos ahora que $\{N(t), t \geq 0\}$ es un proceso de Poisson no-homogéneo con función de intensidad acotada $\lambda(t)$ tal que $\lambda(t) \leq \lambda$ para todo t . Sea $\{M(t), t \geq 0\}$ otro proceso de Poisson no-homogéneo con intensidad $\mu(t) = \lambda - \lambda(t)$ e independiente de $N(t)$. Por la proposición anterior tenemos que $\{N(t), t \geq 0\}$ se puede considerar como el proceso que se obtiene a partir del proceso homogéneo $\{N(t) + M(t), t \geq 0\}$, donde un evento que ocurre en el tiempo t es observado con probabilidad $p(t) = \lambda(t)/\lambda$.

La función $\mu(t)$ definida por

$$\mu(t) = \int_0^t \lambda(s) ds$$

es continua y no decreciente y representa el valor esperado del número de eventos que ocurren en el intervalo $[0, t]$, $E(N(t)) = \mu(t)$. Definimos su inversa generalizada $\nu(t)$ por

$$\nu(t) = \inf\{s : \mu(s) > t\}, \quad t \geq 0.$$

Usando estas funciones tenemos el siguiente resultado

Teorema 4.8 *Sea N un proceso de Poisson no homogéneo y sea $M(t) = N(\nu(t))$, $t \geq 0$. Entonces M es un proceso de Poisson homogéneo con intensidad 1.*

Demostración. Fijamos s, t y ponemos $t' = \nu(t)$, $t' + s' = \nu(t + s)$, $s' = \nu(t + s) - \nu(t)$. Entonces

$$\begin{aligned} E(M(t + s) - M(t) | M(u), u \leq t) &= E(N(t' + s') - N(t') | N(u), u \leq t) \\ &= E(N(t' + s') - N(t')) = \mu(t' + s') - \mu(t') \\ &= t + s - t = s. \end{aligned}$$

Por el teorema 4.4 obtenemos el resultado. ■

Denotemos por τ_n el instante en el cual ocurre el n -ésimo evento de un proceso no-homogéneo $\{N(t), t \geq 0\}$. Entonces

$$\begin{aligned} P(t < \tau_n < t + h) &= P(N(t) = n - 1, \text{ y al menos un evento ocurre en } (t, t + h)) \\ &= P(N(t) = n - 1, \text{ y un evento ocurre en } (t, t + h)) + o(h) \\ &= P(N(t) = n - 1)P(\text{un evento ocurre en } (t, t + h)) + o(h) \\ &= e^{-\mu(t)} \frac{\mu(t)^{n-1}}{(n-1)!} [\lambda(t)h + o(h)] + o(h) \\ &= \lambda(t) e^{-\mu(t)} \frac{(\mu(t))^{n-1}}{(n-1)!} h + o(h). \end{aligned}$$

Dividiendo por h y haciendo $h \rightarrow 0$ obtenemos que la densidad de esta variable es

$$f_{\tau_n}(t) = \lambda(t) e^{-\mu(t)} \frac{(\mu(t))^{n-1}}{(n-1)!}.$$

4.9.2. Procesos de Cox

Un proceso de Cox es un proceso de Poisson no-homogéneo en el cual la intensidad ($\lambda(t)$, $t \geq 0$) es a su vez un proceso aleatorio. En general, los incrementos sobre intervalos disjuntos para un proceso de Cox no son independientes.

Sea $\{N(t), t \geq 0\}$ un proceso de Poisson con intensidad constante $\lambda = 1$. El proceso de Cox más simple requiere seleccionar el valor de una v.a. Θ y luego observar el proceso $M(t) = N(\Theta t)$. Dado el

valor de Θ , M es, condicionalmente, un proceso de Poisson con intensidad constante $\lambda = \Theta$. Θ es aleatoria y, típicamente, no es observable. Si Θ tiene distribución continua con densidad $f(\theta)$ entonces, por la ley de probabilidad total obtenemos la distribución marginal

$$P(M(t) = k) = \int_0^\infty \frac{(\theta t)^k e^{-\theta t}}{k!} f(\theta) d\theta.$$

4.10. La Distribución Uniforme

Consideremos un segmento de longitud t y escojamos sobre él n puntos al azar, de manera independiente y con distribución uniforme, es decir, consideramos una muestra aleatoria simple de tamaño n de la distribución uniforme sobre $[0, t]$. Llamemos U_1, \dots, U_n a estas variables. La densidad de probabilidad de c/u de ellas es

$$f_U(u) = \frac{1}{t}, \quad \text{para } 0 \leq u \leq t.$$

Consideremos ahora esta misma muestra pero ordenada y llamemos $U_{(i)}$, $1 \leq i \leq n$, a sus valores, de modo que

$$U_{(1)} \leq U_{(2)} \leq \dots \leq U_{(n)}.$$

Como veremos a continuación, la densidad conjunta de $U_{(1)}, U_{(2)}, \dots, U_{(n)}$ es

$$f_{U_{(1)}, \dots, U_{(n)}}(u_1, \dots, u_n) = \frac{n!}{t^n} \quad \text{para } 0 < u_1 < \dots < u_n \leq t \quad (4.6)$$

Dada cualquier colección X_1, \dots, X_n de v.a.i.i.d. las variables $X_{(1)}, \dots, X_{(n)}$, que corresponden a los valores de las variables originales pero ordenadas

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)},$$

se conocen como los *estadísticos de orden*. El próximo teorema nos muestra cómo se obtiene su distribución en el caso general.

Teorema 4.9 Sean $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ los estadísticos de orden para una muestra aleatoria simple de variables aleatorias continuas con densidad $f(x)$. La densidad conjunta de los estadísticos de orden es

$$g_n(x_{(1)}, x_{(2)}, \dots, x_{(n)}) = \begin{cases} n! \prod_{i=1}^n f(x_{(i)}), & x_{(1)} < \dots < x_{(n)}, \\ 0, & \text{en otro caso.} \end{cases} \quad (4.7)$$

Demostración. Haremos la prueba para el caso general n y la ilustraremos detalladamente cuando $n = 2$. Definimos los conjuntos $A = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}, x_i \neq x_j \text{ para } i \neq j\}$ y $B = \{(x_{(1)}, \dots, x_{(n)}) : -\infty < x_{(1)} < \dots < x_{(n)} < \infty\}$. La transformación que define los estadísticos de orden es una función de A a B pero no es 1-1, ya que cualquiera de las $n!$ permutaciones de los valores observados produce los mismos estadísticos de orden. Por ejemplo, cuando $n = 2$, $(x_1, x_2) = (1.6, 3.4)$ y $(x_1, x_2) = (3.4, 1.6)$ ambos producen $(x_{(1)}, x_{(2)}) = (1.6, 3.4)$

Si dividimos A en $n!$ subconjuntos de modo que c/u corresponda a un orden particular de la muestra observada, vemos que ahora la transformación que define los estadísticos de orden define una biyección de cada uno de estos conjuntos al conjunto B . Como ilustración, cuando $n = 2$, dividimos A en $A_1 = \{(x_1, x_2) : -\infty < x_1 < x_2 < \infty\}$ y $A_2 = \{(x_1, x_2) : -\infty < x_2 < x_1 < \infty\}$. En el primero de estos conjuntos la transformación es $x_1 = x_{(1)}$ y $x_2 = x_{(2)}$. Por lo tanto el valor absoluto del Jacobiano de la transformación es

$$|J_1| = \left| \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \right| = 1$$

mientras que en A_2 la transformación es $x_1 = x_{(2)}$ y $x_2 = x_{(1)}$ y el valor absoluto del Jacobiano de la transformación es

$$|J_2| = \left| \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \right| = 1.$$

En el caso general se ve similarmente que el Jacobiano de cada una de las biyecciones de una de las $n!$ regiones en que dividimos a A sobre B , tiene valor absoluto igual a 1. La densidad conjunta de los estadísticos de orden es, en consecuencia, la suma de las contribuciones de cada conjunto de la partición. En el caso particular $n = 2$ tenemos para $-\infty < x_{(1)} < x_{(2)} < \infty$,

$$\begin{aligned} g(x_{(1)}, x_{(2)}) &= f(x_{(1)})f(x_{(2)})|J_1| + f(x_{(2)})f(x_{(1)})|J_2| \\ &= 2f(x_{(1)})f(x_{(2)}). \end{aligned}$$

Para n cualquiera, la densidad conjunta (4.7) se obtiene tomando en cuenta que la contribución de cada una de las $n!$ particiones es $\prod_{i=1}^n f(x_{(i)})$. ■

En el caso particular de la distribución uniforme obtenemos la ecuación (4.6) para la densidad de los estadísticos de orden.

Teorema 4.10 Sean τ_1, τ_2, \dots los instantes en los cuales ocurren los sucesivos eventos de un proceso de Poisson de parámetro λ . Dado que $N(t) = n$, las variables τ_1, τ_2, \dots tienen la misma distribución conjunta que los estadísticos de orden de n v.a.i. con distribución uniforme en $[0, t]$.

Demostración. Consideremos un proceso de Poisson $\{N(t), t \geq 0\}$ y supongamos que en el intervalo $[0, t]$ han ocurrido n eventos. Sea $[t_i, t_i + h_i]$, $1 \leq i \leq n$, una sucesión de intervalos disjuntos en $[0, t]$. Dado que han ocurrido n eventos hasta t , la probabilidad de que ocurra exactamente un evento en cada uno de los intervalos que listamos, y ningún evento fuera de ellos es

$$\begin{aligned} P(t_1 \leq \tau_1 \leq t_1 + h_1, \dots, t_n \leq \tau_n \leq t_n + h_n | N(t) = n) \\ &= \frac{\lambda h_1 e^{-\lambda h_1} \dots \lambda h_n e^{-\lambda h_n} e^{-\lambda(t-h_1-\dots-h_n)}}{e^{-\lambda t} (\lambda t)^n / n!} \\ &= \frac{n!}{t^n} h_1 \dots h_n. \end{aligned} \quad (4.8)$$

Pero, por definición de la función de densidad, el lado izquierdo de (4.8) es, para valores pequeños de h_i , $1 \leq i \leq n$, aproximadamente igual a

$$f_{\tau_1, \dots, \tau_n | N(t)=n}(t_1, \dots, t_n) h_1 \dots h_n.$$

Esto es suficiente para demostrar que la densidad condicional de los tiempos τ_i dado que han ocurrido n eventos en el intervalo $[0, t]$ es igual a $n!/t^n$. ■

Ejemplo 4.10

El teorema anterior nos da una manera de probar si un conjunto de observaciones es de Poisson. Supongamos que hemos observado el proceso por un período de tiempo t durante el cual han ocurrido n eventos. Sea τ_1, \dots, τ_n los instantes en los cuales han ocurrido los eventos y sea W_1, \dots, W_n una permutación de los instantes escogida al azar. Si los eventos ocurrieron de acuerdo a un proceso de Poisson, las variables W_i son independientes y tienen distribución uniforme sobre el intervalo $[0, t]$. Por lo tanto podemos hacer un test sobre estas variables para ver si cumplen esta hipótesis, para lo cual podemos hacer una prueba de Kolmogorov-Smirnov o de Cramér-von Mises. También es posible usar el TCL ya que para valores moderados o grandes de n , la suma $S_n = \sum_1^n U_i$ es aproximadamente normal con media $E(S_n) = n E(U_1) = nt/2$ y varianza $\text{Var}(S_n) = n \text{Var}(U_1) = nt^2/12$.

Por ejemplo, si en $t = 10$ minutos de observación, $n = 12$ eventos ocurren, entonces la suma S_{12} de los instantes en los cuales ocurren los eventos es aproximadamente normal con media 60 y desviación estándar 10. En consecuencia, si S_{12} satisface las desigualdades

$$60 - (1.96)10 \leq S_{12} \leq 60 + (1.96)10,$$

aceptaríamos la hipótesis de que los eventos provienen de un proceso de Poisson con un nivel de significación de 95%. ▲

Ejemplo 4.11

Consideremos una masa de material radioactivo que emite partículas alfa de acuerdo a un proceso de Poisson de intensidad λ . Cada partícula existe por un período aleatorio de tiempo y luego desaparece. Supongamos que los tiempos de vida sucesivos Y_1, Y_2, \dots de las diferentes partículas son v.a.i. con distribución común $G(y) = P(Y_k \leq y)$. Sea $M(t)$ el número de partículas que existen en el instante t . Queremos hallar la distribución de probabilidad de $M(t)$ bajo la condición de que $M(0) = 0$.

Sea $N(t)$ el número de partículas creadas hasta el tiempo t . Observamos que $M(t) \leq N(t)$. Dado que $N(t) = n$ sean $\tau_1, \dots, \tau_n \leq t$ los instantes en los cuales se crean las partículas. La partícula k existe en el instante t si y sólo si $\tau_k + Y_k \geq t$. Por lo tanto

$$P(M(t) = m | N(t) = n) = P\left(\sum_{k=1}^n \mathbf{1}_{\{\tau_k + Y_k \geq t\}} = m | N(t) = n\right).$$

Usando el teorema 4.10 y la simetría entre las partículas tenemos

$$P\left(\sum_{k=1}^n \mathbf{1}_{\{\tau_k + Y_k \geq t\}} = m | N(t) = n\right) = P\left(\sum_{k=1}^n \mathbf{1}_{\{U_k + Y_k \geq t\}} = m\right) \quad (4.9)$$

donde U_1, U_2, \dots, U_n son v.a.i. con distribución uniforme en $[0, t]$. El lado derecho de (4.9) es una distribución binomial con parámetros n y

$$\begin{aligned} p &= P(U_k + Y_k \geq t) = \frac{1}{t} \int_0^t P(Y_k \geq t - u) du \\ &= \frac{1}{t} \int_0^t [1 - G(t - u)] du = \frac{1}{t} \int_0^t [1 - G(z)] dz. \end{aligned} \quad (4.10)$$

Escribiendo explícitamente la distribución binomial tenemos

$$P(M(t) = m | N(t) = n) = \binom{n}{m} p^m (1 - p)^{n-m}$$

con p dado por la ecuación (4.10). Finalmente

$$\begin{aligned} P(M(t) = m) &= \sum_{n=m}^{\infty} P(M(t) = m | N(t) = n) P(N(t) = n) \\ &= \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \frac{(\lambda t)^n e^{-\lambda t}}{n!} \\ &= e^{-\lambda t} \frac{(\lambda p t)^m}{m!} \sum_{n=m}^{\infty} \frac{(1-p)^{n-m} (\lambda t)^{n-m}}{(n-m)!}. \end{aligned} \quad (4.11)$$

La suma es una serie exponencial que se reduce a

$$\sum_{n=m}^{\infty} \frac{(1-p)^{n-m} (\lambda t)^{n-m}}{(n-m)!} = \sum_{j=0}^{\infty} \frac{[\lambda t(1-p)]^j}{j!} = e^{\lambda t(1-p)}$$

y usando esto (4.11) se reduce a

$$P(M(t) = m) = \frac{e^{-\lambda pt} (\lambda pt)^m}{m!} \quad \text{para } m \geq 0,$$

es decir, el número de partículas que existen en el instante t tiene distribución de Poisson de media

$$\lambda pt = \lambda \int_0^t (1 - G(y)) dy. \quad (4.12)$$

Veamos que ocurre cuando $t \rightarrow \infty$. Sea $\mu = E[Y_k] = \int_0^\infty (1 - G(y)) dy$ la vida media de una partícula alfa. Vemos a partir de (4.12) que cuando $t \rightarrow \infty$, la distribución de $M(t)$ converge a una distribución de Poisson con parámetro $\lambda\mu$. Por lo tanto, asintóticamente, la distribución de probabilidad para el número de partículas que existen depende únicamente de la vida media μ . ▲

Ejemplo 4.12

Un procedimiento común en estadística es observar un número fijo n de v.a.i.i.d. X_1, \dots, X_n y usar su media muestral

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

como estimador de la media de la población $E[X_1]$. Consideremos en cambio la siguiente situación: Una compañía nos pide estimar el tiempo medio de vida en servicio de cierto componente de una máquina. La máquina ha estado funcionando por dos años y se observó que el componente original duró 7 meses, el siguiente duró 5 meses y el tercero 9. No se observaron fallas en los tres meses restantes del período de observación. La pregunta es si es correcto estimar la vida media en servicio por el promedio observado $(7 + 9 + 5)/3 = 7$ meses.

Este ejemplo presenta una situación en la cual el tamaño de la muestra no está fijo de antemano sino que se determina a través de una 'cuota' prefijada $t > 0$: Observamos una sucesión de variables i.i.d. X_1, X_2, \dots y continuamos el muestreo mientras la suma de observaciones sea menor que la cuota t . Llamemos $N(t)$ al tamaño de la muestra,

$$N(t) = \text{máx}\{n \geq 0 : X_1 + \dots + X_n < t\}.$$

La media muestral es

$$\bar{X}_{N(t)} = \frac{X_1 + \dots + X_{N(t)}}{N(t)}.$$

Puede suceder que $X_1 \geq t$, en este caso $N(t) = 0$ y no podemos definir la media muestral. Por lo tanto tenemos que suponer que $N(t) \geq 1$. Una pregunta importante en estadística matemática es si este estimador es insesgado. Es decir, ¿cómo se relaciona el valor esperado de este estimador con el valor esperado de $E[X_1]$?

En general, determinar el valor esperado de la media muestral en esta situación es muy difícil. Es posible hacerlo, sin embargo, en el caso en el cual los sumandos tienen distribución exponencial de parámetro común λ , de modo que $N(t)$ es un proceso de Poisson. La clave es usar el teorema 4.10 para evaluar la esperanza condicional

$$E[\tau_{N(t)} | N(t) = n] = E[\text{máx}(U_1, \dots, U_n)] = t \left(\frac{n}{n+1} \right),$$

donde U_1, \dots, U_n son independientes y tienen distribución uniforme sobre el intervalo $[0, t]$. Observamos además que

$$P(N(t) = n | N(t) > 0) = \frac{(\lambda t)^n e^{-\lambda t}}{n!(1 - e^{-\lambda t})}.$$

Entonces,

$$\begin{aligned}
 E\left[\frac{\tau_{N(t)}}{N(t)} \mid N(t) > 0\right] &= \sum_{n=1}^{\infty} E\left[\frac{\tau_{N(t)}}{n} \mid N(t) = n\right] P(N(t) = n \mid N(t) > 0) \\
 &= \sum_{n=1}^{\infty} t \binom{n}{n+1} \left(\frac{1}{n}\right) \left(\frac{(\lambda t)^n e^{-\lambda t}}{n!(1 - e^{-\lambda t})}\right) \\
 &= \frac{1}{\lambda} \left(\frac{1}{e^{\lambda t} - 1}\right) \sum_{n=1}^{\infty} \frac{(\lambda t)^{n+1}}{(n+1)!} \\
 &= \frac{1}{\lambda} \left(\frac{1}{e^{\lambda t} - 1}\right) (e^{\lambda t} - 1 - \lambda t) \\
 &= \frac{1}{\lambda} \left(1 - \frac{\lambda t}{e^{\lambda t} - 1}\right).
 \end{aligned}$$

Podemos ver el efecto de este tipo de muestreo si expresamos el resultado anterior en términos del cociente del sesgo entre el verdadero valor de la esperanza $E(X_1) = 1/\lambda$. Tenemos

$$\frac{E[X_1] - E[\bar{X}_{N(t)}]}{E[X_1]} = \frac{\lambda t}{e^{\lambda t} - 1} = \frac{E[N(t)]}{e^{E[N(t)]} - 1}.$$

El lado izquierdo representa la fracción del sesgo y el lado derecho expresa esta fracción como función del tamaño esperado de la muestra para este tipo de muestreo. La siguiente tabla presenta algunos valores:

$E(N(t))$	Fracción
1	0.58
2	0.31
3	0.16
4	0.07
5	0.03
6	0.015
10	0.0005

En el ejemplo inicial, observamos $N(t) = 3$ fallas en un período de un año y en la tabla anterior vemos que la fracción del sesgo es del orden de 16%. Como mencionamos $\bar{X}_{N(t)} = 7$, una estimación más adecuada podría ser $7/0.84 = 8.33$, que intenta corregir, en promedio, el sesgo debido al método de muestreo. ▲

Los resultados anteriores se pueden generalizar al caso de procesos no-homogéneos con intensidad $\lambda(r)$. Sea $\mu(t) = \int_0^t \lambda(r) dr$ y $g(r) = \lambda(r)/\mu(t)$ para $0 < r < t$.

Teorema 4.11 Sean U_1, U_2, \dots, U_n v.a.i. con densidad g . Dado que $N(t) = n$, los tiempos de llegadas τ_1, \dots, τ_n tienen la misma distribución que los estadísticos de orden correspondientes a las variables U_1, \dots, U_n .

La demostración es similar a la del caso homogéneo y queda como ejercicio.

4.11. Procesos Espaciales de Poisson

Sea S un conjunto en un espacio n -dimensional y sea \mathcal{A} una familia de subconjuntos de S . Un proceso puntual en S es un proceso estocástico $N(A)$ indexado por los subconjuntos A en \mathcal{A} que tiene como valores posibles los elementos del conjunto $\{0, 1, 2, \dots\}$. La idea es que los puntos se encuentran dispersos en S de manera aleatoria y $N(A)$ cuenta los puntos en el conjunto A . Como $N(A)$ es una función que cuenta, hay varias condiciones obvias que debe satisfacer. Por ejemplo, si A y B son disjuntos, están en \mathcal{A} y su unión $A \cup B$ también está en \mathcal{A} entonces necesariamente $N(A \cup B) = N(A) + N(B)$.

El caso unidimensional, en el cual S es la semirecta positiva y \mathcal{A} es la colección de los intervalos de la forma $A = (s, t]$ para $0 \leq s < t$, lo consideramos al estudiar el proceso puntual de Poisson. La generalización al plano o al espacio tridimensional tiene interés cuando consideramos la distribución espacial de estrellas o galaxias en Astronomía, de plantas o animales en Ecología, de bacterias sobre una placa de laboratorio en Biología o de defectos sobre una superficie en Ingeniería.

Definición 4.5 Sea S un subconjunto de \mathbb{R}, \mathbb{R}^2 o \mathbb{R}^3 . Sea \mathcal{A} una familia de subconjuntos de S y para cualquier $A \in \mathcal{A}$ sea $|A|$ el tamaño (longitud, área o volumen) de A . Entonces $\{N(A) : A \in \mathcal{A}\}$ es un proceso puntual homogéneo de Poisson de intensidad $\lambda > 0$ si

1. Para todo $A \in \mathcal{A}$, la variable $N(A)$ tiene distribución de Poisson con parámetro $\lambda|A|$.
2. Para toda colección finita A_1, \dots, A_n de conjuntos disjuntos de \mathcal{A} , las variables $N(A_1), \dots, N(A_n)$ son independientes.

Ejemplo 4.13

Una zona de Londres se dividió en $N = 576 = 24 \times 24$ áreas de $1/4$ de kilómetro cuadrado. Esta área recibió 535 impactos de bomba durante la II Guerra Mundial, un promedio de $535/576 = 0.9288$ por cuadrado. La siguiente tabla presenta N_k , el número de cuadrados que recibieron el impacto de exactamente k bombas y lo compara con el valor esperado si los impactos tuviesen una distribución de Poisson con esta media

k	0	1	2	3	4	≥ 5
N_k	229	211	93	35	7	1
Poisson	226.74	211.39	98.54	30.62	7.14	1.57

El ajuste es muy bueno y puede ser verificado haciendo una prueba χ^2 . ▲

Muchas de las propiedades que hemos estudiado para el caso unidimensional tienen una extensión natural para el caso de dimensiones mayores. Veamos, como ejemplo, la propiedad de uniformidad de la distribución de la ubicación de los puntos en una región dado que conocemos el número de puntos. Consideremos inicialmente una región A de tamaño positivo $|A| > 0$ y supongamos que sabemos que A contiene exactamente un punto: $N(A) = 1$. Entonces, la distribución de este punto es uniforme en el siguiente sentido:

$$P(N(B) = 1 | N(A) = 1) = \frac{|B|}{|A|}$$

para cualquier $B \subset A$. Para ver esto escribimos $A = B \cup C$ donde $C = A \setminus B$ y en consecuencia $N(B)$ y $N(C)$ son v.a.i. de Poisson con medias respectivas $\lambda|B|$ y $\lambda|C|$. Entonces

$$\begin{aligned} P(N(B) = 1 | N(A) = 1) &= \frac{P(N(B) = 1, N(C) = 0)}{P(N(A) = 1)} \\ &= \frac{\lambda|B|e^{-\lambda|B|}e^{-\lambda|C|}}{\lambda|A|e^{-\lambda|A|}} \\ &= \frac{|B|}{|A|}. \end{aligned}$$

Para generalizar este resultado consideremos una región A de tamaño positivo $|A| > 0$ que contiene $N(A) = n \geq 1$ puntos. Entonces estos puntos son independientes y están distribuidos uniformemente en A en el sentido de que para cualquier partición disjunta A_1, \dots, A_m de A , donde $A = A_1 \cup \dots \cup A_m$ y para cualesquiera enteros positivos k_1, \dots, k_m con $k_1 + \dots + k_m = n$, tenemos

$$P(N(A_1) = k_1, \dots, N(A_m) = k_m | N(A) = n) = \frac{n!}{k_1! \dots k_m!} \left(\frac{|A_1|}{|A|} \right)^{k_1} \dots \left(\frac{|A_m|}{|A|} \right)^{k_m}$$

es decir, dado que $N(A) = n$, la distribución de $N(A_1), \dots, N(A_m)$ es multinomial.

Ejemplo 4.14

Consideremos un proceso de Poisson compuesto sobre la recta y supongamos que las variables asociadas U_1, U_2, \dots tienen distribución uniforme $\mathcal{U}(0, 1)$. Esto nos da una sucesión de puntos sobre la banda $\{(t, u) : 0 \leq t < \infty, 0 < u < 1\}$. En este caso, como las U_i tienen distribución continua, el número de puntos sobre una recta fija $\{(t, u) : u = x\}$ es 0 con probabilidad uno.

Si en lugar de rectas consideramos bandas

$$\{(t, u) : 0 \leq t < \infty, a_{m-1} < u \leq a_m\}$$

y si $0 \leq a_0 < a_1 < \dots < a_n \leq 1$, entonces los puntos en las bandas $1 \leq m \leq n$ son independientes.

Usando esta propiedad y la propiedad de incrementos independientes de los procesos de Poisson unidimensionales obtenemos el siguiente resultado: Sean $R_m = \{(t, u) : a_m < t \leq b_m, c_m < u \leq d_m\}$ rectángulos con $a_m \geq 0$ y $0 \leq c_m < d_m \leq 1$ y sea $N(R_m)$ el número de puntos (T_i, U_i) que están en el rectángulo R_m . Si los rectángulos R_m son disjuntos entonces las variables $N(R_m)$ son independientes y tienen distribución de Poisson con media

$$\lambda_m = \lambda(b_m - a_m)(d_m - c_m) = \lambda|R_m|.$$

▲

4.11.1. Procesos no homogéneos en el plano

Decimos que N es un proceso no homogéneo de Poisson con intensidad $\lambda(x, y)$ si cuando R_m , $1 \leq m \leq n$ son conjuntos disjuntos, las variables $N(R_m)$ son independientes con distribución de Poisson de media

$$\mu(R_m) = \int_{(x,y) \in R_m} \lambda(x, y) dy dx.$$

Si la integral vale ∞ , el número de puntos es ∞ .

Ejemplo 4.15

Sean τ_1, τ_2, \dots los instantes en los que ocurren eventos de un proceso de Poisson con intensidad λ . Supongamos que si un evento ocurre en el instante s , lo registramos con probabilidad $p(s)$. El proceso de eventos registrados es un proceso de Poisson no-homogéneo de intensidad $\lambda p(s)$.

Para ver esto, asociamos a cada τ_i una v.a.i. con distribución uniforme sobre $(0, 1)$, y aceptamos el punto si $U_i < p(\tau_i)$. El número de puntos τ_i aceptados en un intervalo (a, b) es igual al número de puntos (τ_i, U_i) que caen en la región

$$\{(t, u) : a < t < b, 0 < u < p(t)\}.$$

En consecuencia, este número es Poisson con media igual al producto de λ por el área del conjunto, es decir $\lambda \int_a^b p(s) ds$. Es claro que el número de eventos en intervalos disjuntos son independientes, de modo que tenemos un proceso de Poisson no-homogéneo. Este resultado nos da una manera de construir un proceso de Poisson no-homogéneo de intensidad dada. ▲