

Procesos estocásticos

Sesión 7. Procesos de Poisson

Enrique Miranda

Universidad of Oviedo

Máster Universitario en Análisis de Datos
para la Inteligencia de Negocios

Contenidos

1. La distribución exponencial.
2. Definición de un proceso de Poisson.
3. Procesos de Poisson compuestos.
4. Refinamiento y superposición.
5. Procesos condicionados.

Introducción

Comenzamos a continuación el estudio de los procesos a tiempo continuo, que son aquéllos en los que el conjunto de índices de las variables aleatorias es un espacio no numerable (habitualmente $[0, \infty)$).

En esta sesión, vamos a estudiar un tipo de procesos con espacio de estados discreto, los cuales se encuentran en la base de la **teoría de colas**: los procesos de Poisson.

Son un caso particular de los **procesos de conteo**, en los que se considera $\{X_t\}_{t \geq 0}$ y X_t mide el número de apariciones de un suceso hasta el instante t .

La distribución exponencial

Recordemos que una variable aleatoria T posee **distribución exponencial** con parámetro λ cuando tiene función de densidad

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

La distribución exponencial suele utilizarse para representar el tiempo que transcurre hasta que ocurre un suceso de interés.

Propiedades

- ▶ $E(T) = \frac{1}{\lambda}$.
- ▶ $Var(T) = \frac{1}{\lambda^2}$.
- ▶ **Falta de memoria:** $P(T > t + s | T > t) = P(T > s)$.
- ▶ Si $T_i \sim \exp(\lambda_i)$ para $i = 1, \dots, n$ y $T_1, \dots, T_n$ son independientes, entonces $\min\{T_1, \dots, T_n\} \sim \exp(\lambda_1 + \dots + \lambda_n)$.
- ▶ En las condiciones anteriores,
$$P(T_i = \min\{T_1, \dots, T_n\}) = \frac{\lambda_i}{\lambda_1 + \dots + \lambda_n}.$$

Procesos de Poisson: definición

Sea $\{t_n\}_n$ un proceso estocástico compuesto por variables independientes con distribución $\exp(\lambda)$. Definimos

$T_n = t_1 + \cdots + t_n$ para todo n , ($T_0 = 0$), y

$N(s) = \max\{n : T_n \leq s\}$. Entonces, N se dice **proceso de Poisson con parámetro λ** .

Si interpretamos t_n como el tiempo entre dos apariciones consecutivas de un suceso (llegadas de clientes a una cola, etc), $N(s)$ sería el número de apariciones de dicho suceso antes del instante s .

El nombre **Poisson** viene de que este número de apariciones sigue una distribución conocida como distribución de Poisson.

Distribución de Poisson

Se dice que una variable aleatoria discreta X posee **distribución de Poisson** con parámetro λ , y se denota $X \sim \mathcal{P}(\lambda)$, cuando su función de masa de probabilidad es

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

- ▶ En un proceso de Poisson con parámetro λ , $N(s)$ sigue una distribución de Poisson con parámetro λs .
- ▶ **Propiedad de incrementos independientes:** Si $t_0 < t_1 < \dots < t_n$, entonces las variables $N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$ son independientes.

Relación entre la Poisson y la exponencial

- ▶ Si el tiempo entre dos llegadas consecutivas sigue una distribución $\exp(\lambda)$, entonces el número de llegadas en un intervalo de tiempo t sigue una distribución $\mathcal{P}(\lambda t)$.
- ▶ Si el número de llegadas en una hora sigue una distribución $\mathcal{P}(\lambda)$, entonces el número de horas entre dos llegadas consecutivas sigue una $\exp(\lambda)$.

Historia

Si bien el primer estudio de los procesos de Poisson fue realizado por **Michell** en el s. XVIII, la formulación matemática tal y como la conocemos surgió en tres estudios independientes a principios del siglo XX:

- ▶ **Erlang** en teoría de colas.
- ▶ **Rutherford** y **Eiger** en el análisis de partículas alfa.
- ▶ **Campbell** en ruido termiónico.

Caracterización de los procesos de Poisson

$\{N(s)\}_{s \geq 0}$ constituye un proceso de Poisson si y sólo si:

- (i) $N(0) = 0$.
- (ii) $N(t + s) - N(s) \sim \mathcal{P}(\lambda t)$ para todo $t, s \geq 0$.
- (iii) $N(t)$ posee incrementos independientes.

Estas 3 propiedades se utilizan a veces como definición.

Definición alternativa

Equivalente, también se puede definir un proceso de Poisson con parámetro λ como un proceso estocástico $\{N(t)\}_{t \geq 0}$ t.q.:

1. $N(t)$ posee incrementos independientes.
2. $N(t + s) - N(s)$ es independiente de s para todo $s, t \geq 0$.
3. Para h próximo a 0, la probabilidad de que en el intervalo de tiempo $[0, h]$ se dé exactamente una ocurrencia del fenómenos es $\lambda h + o(h)$.
4. Para h próximo a 0, la probabilidad de que en el intervalo de tiempo $[0, h]$ haya 2 o más ocurrencias del fenómeno es $o(h)$.

Interpretación de λ

Bajo la definición anterior, el número medio de ocurrencias por unidad de tiempo sería

$$\lim_{h \rightarrow 0} \frac{\lambda h + o(h)}{h} = \lambda.$$

Es por esto que el parámetro λ del proceso de Poisson se denomina en ocasiones **intensidad** o **tasa** del proceso.

Incrementos estacionarios

Como ya vimos en la segunda sesión, la propiedad 2 en esta definición se conoce como propiedad de **incrementos estacionarios**.

Si un proceso estocástico representa la llegada de individuos a una cola a lo largo del tiempo, es de esperar que posea incrementos independientes (las llegadas en intervalos disjuntos deberían ser independientes), pero no necesariamente estacionarios (podría haber momentos con mayor flujo de llegadas).

Esto último será una de las motivaciones para introducir modelos más generales que los procesos de Poisson.

Procesos de Poisson no homogéneos

Se dice que $\{N(s)\}_{s \geq 0}$ es un proceso de Poisson con parámetro $\lambda(r)$ cuando:

- (i) $N(0) = 0$.
- (ii) $N(t)$ posee incrementos independientes.
- (iii) $N(t + s) - N(s)$ posee una distribución de Poisson con parámetro $\int_s^{t+s} \lambda(r) dr$.

La función λ se conoce también como **función de valor medio del proceso**.

En un proceso no homogéneo los tiempos entre dos llegadas consecutivas ya no siguen una distribución exponencial, ni son independientes.

Ejemplo

En el ejemplo de los tiempos de llegada a un restaurante, utilizaríamos un proceso de Poisson no homogéneo cuando la distribución del tiempo de llegada no es uniforme en todo el intervalo de tiempo.

Un inconveniente de los procesos de Poisson no homogéneos es que los cálculos suelen ser bastante más complicados. Se suelen considerar entonces subintervalos de tiempo en los que podemos considerar el proceso de Poisson como homogéneo.

Procesos de Poisson compuestos: introducción

Sea $\{Y_n\}_n$ una sucesión de variables aleatorias independientes e idénticamente distribuidas. Sea N una variable aleatoria independiente de $(Y_n)_n$ con valores en $\mathbb{N}$, y sea $S = Y_1 + \cdots + Y_N$ (si $N = 0$ entonces $S = 0$). Se cumple:

- (i) Si $E(N) < \infty$, entonces $E(S) = E(N) \cdot E(Y_i)$.
- (ii) Si $E(N^2) < \infty$, entonces
$$\text{var}(S) = E(N)\text{var}(Y_i) + \text{var}(N)(E(Y_i))^2.$$
- (iii) Si $N \sim \mathcal{P}(\lambda)$, entonces $\text{var}(S) = \lambda E(Y_i^2)$.

Procesos de Poisson compuestos: definición

Un proceso estocástico $\{X(t)\}_{t \geq 0}$ es un **proceso de Poisson compuesto** si se puede representar como $X(t) = \sum_{i=1}^{N(t)} Y_i$, donde $\{N(t) : t \geq 0\}$ es un proceso de Poisson e $\{Y_i\}_{i \geq 1}$ es una sucesión de variables aleatorias i.i.d. e independientes de $\{N(t) : t \geq 0\}$.

Ejemplo: Supongamos que el número de coches que llega a un restaurante sigue un proceso de Poisson con parámetro λ . Sea Y_i el número de ocupantes del coche i -ésimo. Entonces, el número total de personas que han llegado hasta el instante N constituiría un proceso de Poisson compuesto.

Si Y_i representa el gasto del cliente i -ésimo, podríamos utilizar el resultado anterior para calcular la media y varianza del gasto.

Refinamiento y superposición

Algunas transformaciones de un proceso de Poisson nos permiten generar nuevos procesos de Poisson:

- ▶ Si dividimos un proceso de Poisson en varias partes utilizando una sucesión de variables independientes e idénticamente distribuidas, hablamos de **refinamiento** o partición.
- ▶ Si sumamos varios procesos de Poisson independientes, el proceso se conoce como **superposición** o mezcla.

Refinamiento: ejemplo

Sea $\{N(t)\}_{t \geq 0}$ un proceso de Poisson con tasa λ . Supongamos que los sucesos se clasifican en dos clases 1 y 2, con probabilidades p y $1-p$, independientemente del resto.

Sean $N_1(t)$, $N_2(t)$, respectivamente, el número de sucesos de la clase 1 y la clase 2 hasta el instante t . Entonces:

- ▶ $N(t) = N_1(t) + N_2(t) \forall t$.
- ▶ $\{N_1(t)\}_{t \geq 0}$ es un proceso de Poisson con tasa λp .
- ▶ $\{N_2(t)\}_{t \geq 0}$ es un proceso de Poisson con tasa $\lambda(1 - p)$.

Superposición: definición

Consideremos $\{N_1(t)\}_{t \geq 0}$, $\{N_2(t)\}_{t \geq 0}$ dos procesos de Poisson independientes con tasas λ_1 , λ_2 , respectivamente.

Entonces, $\{N_1(t) + N_2(t)\}_{t \geq 0}$ es un proceso de Poisson con parámetro $\lambda_1 + \lambda_2$.

Estas dos ejemplos se generalizan al caso en que consideramos c clases independientes, bien para dividir un proceso de Poisson (refinamiento) o para sumar varios (superposición).

Una propiedad importante es que si sumamos un número grande de procesos de conteo, bajo ciertas condiciones podemos aproximar la suma por un proceso de Poisson, incluso aunque los procesos de partida no lo sean.

Propiedades

- ▶ Sea $N(s)$ un proceso de Poisson, $\{Y_n\}_n$ variables aleatorias independientes e idénticamente distribuidas, y $N_j(s) = |\{i \leq N(s) \text{ t.q. } Y_i = j\}|$. Entonces, $N_j(s)$ es un proceso de Poisson con parámetro $\lambda P(Y_i = j)$, y los $N_j(s)$ son independientes.
- ▶ Si $N_1(s), \dots, N_k(s)$ son procesos de Poisson independientes con parámetros $\lambda_1, \dots, \lambda_k$, entonces $N_1(s) + \dots + N_k(s)$ es un proceso de Poisson con parámetro $\lambda_1 + \dots + \lambda_k$.

Ejemplo: carreras de Poisson

Consideremos dos procesos de Poisson independientes con parámetros respectivos λ y μ , representando las llegadas a un servicio de individuos de dos grupos diferentes. ¿Cuál es la probabilidad de que hayan llegado 6 individuos del primer grupo antes de que llegue el cuarto individuo del segundo grupo?

Solución: Mediante un procedimiento de superposición, vemos que el total de llegadas es un proceso de Poisson de parámetro $\lambda + \mu$. Cada llegada tiene probabilidad $\frac{\lambda}{\lambda + \mu}$ de ser del primer grupo. Nos piden la probabilidad de que en las 9 primeras llegadas haya al menos 6 del primer grupo. Usando la distribución binomial, esto es

$$\sum_{k=6}^9 \frac{9!}{k!(9-k)!} \left(\frac{\lambda}{\lambda + \mu} \right)^k \left(\frac{\mu}{\lambda + \mu} \right)^{9-k}.$$

Condicionamiento en procesos de Poisson

Sea $\{T_n\}_n$ la sucesión de tiempos de llegada de un proceso de Poisson con parámetro λ , y sean $\{U_1, \dots, U_n\}$ variables aleatorias independientes con distribución uniforme en $[0, t]$.

- ▶ Si condicionamos a $N(t) = n$, la distribución de $\{T_1, \dots, T_n\}$ es la misma que la de $\{U_1, \dots, U_n\}$.

Es decir, si sabemos que ha habido n llegadas en un tiempo t , la distribución de estos n tiempos de llegada es la misma que la de n observaciones al azar en el intervalo $[0, t]$.

- ▶ Si $s < t$ y $0 \leq m \leq n$, entonces

$$P(N(s) = m | N(t) = n) = \frac{n!}{m!(n-m)!} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}.$$

Generalización a procesos no homogéneos

Sea $\{N(s)\}_{s \geq 0}$ un proceso de Poisson no homogéneo con parámetro $\lambda(r)$. Definimos $\mu(t) = \int_0^t \lambda(r) dr$,

$$g(r) = \begin{cases} \frac{\lambda(r)}{\mu(t)} & \text{si } 0 < r < t \\ 0 & \text{en otro caso .} \end{cases}$$

Dadas $U_1, \dots, U_n$ variables aleatorias independientes con distribución g . Si condicionamos en $N(t) = n$, la distribución de los tiempos de llegada $\{T_1, \dots, T_n\}$ es la misma de $\{U_1, \dots, U_n\}$.