Modelos estocásticos

Soluciones de los ejercicios del tema 2

23 de marzo de 2021

Dr. Víctor Hernández

Universidad Nacional de Educación a distancia
Departamento de Estadística e Investigación operativa

0.1. Ejercicios resueltos del tema 2

Ejercicio 0.1. Muestreo en los intervalos entre llegadas. Este ejercicio toca un tema interesante que ya se ha estudiado en algunos de los ejercicios suplementarios del tema anterior.

Lo plantearemos de esta manera: supongamos que los autobuses llegan a la parada de acuerdo con un proceso de POISSON, cuyos tiempos entre llegadas tienen ua longitud media igual a $1/\lambda$. Una persona llega en un instante t a la parada, ¿cuál es el tiempo medio que tiene que esperar?

Dos respuestas posibles a esta cuestión son las siguientes:

Respuesta A. Puesto que la distribución exponencial de los tiempos entre llegadas no tiene memoria, desde el instate en que llega hasta la llegada del siguiente autobús transcurre un tiempo exponencial de media $1/\lambda$ y el tiempo medio de espera es $1/\lambda$.

Respuesta B. Puesto que llega al azar en un instante entre llegadas, su espera será, en media, la mitad de la longitud media del intervalo y el tiempo medio de espera es $1/2\lambda$.

Para justificar si alguna de las respuestas anteriores es correcta, calcula la distribución de la longitud del intervalo entre llegadas que contiene al punto t y halla el tiempo medio de espera.

Solución. Este ejercicio trata un caso particular de un fenómeno muy importante conocido como *Paradoja de la inspección*, que tiene importancia en el muestreo de procesos. Gracias a las propiedades del proceso de POISSON, o si se quiere de la distribución exponencial, seremos capaces de resolverlo completamente.

Siguiendo mis ideas sobre lo que debe ser una buena explicación, trataré de mostrar primero una explicación intuitiva del fenómeno antes de pasar a los detalle más formales.

Presentemos primero el modelo del problema. Los autobuses llegan en instantes $S_0 = 0 < S_1 < S_2 < S_3 < \cdots$, puesto que el proceso es de POISSON, los tiempos entre llegadas $X_n = S_n - S_{n-1}$, $n = 1, 2, \ldots$, son independientes y tienen distribución común exponencial de parámetro λ, por eso el tiempo medio entre llegadas es $1/\lambda$, y los instantes en que ocurren las llegadas, $S_1 = X_1$, $S_2 = X_1 + X_2$, $\ldots$, $S_n = X_1 + X_2 + \cdots + X_n$, tienen distribuciones gamma con funciones de densidad respectivas dadas por

$$g_n(y) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}, \qquad \text{para } n = 1, 2, \ldots$$

La clave de todo este asunto, allí donde reside la trampa para la intuición poco depurada está en que el intervalo entre llegadas que contiene al instante t no es uno fijo, sino que es aleatorio. En efecto, para que sea X_n el intervalo que contiene a t, debe ser $S_{n-1} < t \le S_n$. Dicho de otra manera, el índice del intervalo que contiene a al instante t es aleatorio y lo designaremos por $N(t)$, de manera que $X_{N(t)}$ es el intervalo que contiene a t y $N(t)$ es una variable aleatoria definida por la condición siguiente

$$\{N(t) = n\} = \{S_{n-1} < t \le S_n\}, \qquad \text{para } n = 1, 2, \dots \tag{1}$$

Ahora podemos intuir en qué consiste el fenómeno; supuesto que el Azar a determinado ya las longitudes de los tiempos entre llegadas, será más probable que el instante t pertenezca a un intervalo largo que a uno corto y la distribución del tiempo entre llegadas que contiene a t no será igual a la común, sino que estará apuntada hacia la derecha, con una mayor probabilidades en la cola, resultando que su media es mayor que $1/\lambda$. Para precisar esta intuición, calcularemos la distribución del intervalo que contiene a t; para ello hallaremos $P(X_{N(t)} \le x)$ para cada $x > 0$.

Por la definición de $N(t)$ que dimos en 1, parece claro que para calcular $P(X_{N(t)} \le x)$ es conveniente condicionar por el valor de S_{n-1} y que hay dos casos posibles.

Caso A. Si $x < t$, necesariamente debe haber alguna llegada antes de t, es decir $n \ge 2$; supongamos que $S_{n-1} = y$, donde $t - x < y < t$, es la última llegada antes del tiempo t, entonces para que $N(t) = n$ debe ser $X_n > t - y$, luego si $x < t$, $n \ge 2$ y $t - x < y < t$, se tiene

$$\begin{aligned} P(X_{N(t)} \le x, N(t) = n \mid S_{n-1} = y) &= P(t - y < X_n \le x) \\ &= e^{-\lambda(t-y)} - e^{-\lambda x} \end{aligned} \tag{2}$$

Caso B. Si $x \ge t$, caben dos posibilidades para tener $X_{N(t)} \le x$. La primera es que $N(t) = 1$ y el primer intervalo entre llegadas X_1 cumpla $t < X_1 \le x$, por lo que si $x \ge t$ y $N(t) = 1$, se tiene

$$\begin{aligned} P(X_{N(t)} \le x, N(t) = 1) &= P(t < X_1 \le x) \\ &= e^{-\lambda t} - e^{-\lambda x} \end{aligned} \tag{3}$$

La otra es que $N(t) = n \ge 2$, entonces, si condicionamos por $S_{n-1} = y$, donde $0 < y < t$, se tendrá $t - y < X_n \le x$ y, si $x \ge t$, $n \ge 2$ y $0 < y < t$, resulta

$$\begin{aligned} P(X_{N(t)} \le x, N(t) = n \mid S_{n-1} = y) &= P(t - y < X_n \le x) \\ &= e^{-\lambda(t-y)} - e^{-\lambda x} \end{aligned} \tag{4}$$

Ahora, de 2 se sigue

$$P(X_{N(t)} \le x, N(t) = n) = \int_{t-x}^{t} (e^{-\lambda(t-y)} - e^{-\lambda x}) g_{n-1}(y)\, dy, \qquad n \ge 2,\ x < t$$

luego

$$\begin{aligned} P(X_{N(t)} \le x) &= \sum_{n=2}^{\infty} \int_{t-x}^{t} (e^{-\lambda(t-y)} - e^{-\lambda x}) g_{n-1}(y)\, dy \\ &= \int_{t-x}^{t} (e^{-\lambda(t-y)} - e^{-\lambda x}) \sum_{n=2}^{\infty} g_{n-1}(y)\, dy, \qquad x < t \end{aligned} \tag{5}$$

pero para hallar la integral anterior no hacen falta muchos cálculos, ya que se tiene

$$\sum_{n=2}^{\infty} g_{n-1}(y) = \sum_{n=2}^{\infty} \frac{\lambda^{n-1} x^{n-2}}{(n-2)!} e^{-\lambda x} = \lambda$$

y resulta

$$\begin{aligned} P(X_{N(t)} \le x) &= \int_{t-x}^{t} \lambda (e^{-\lambda(t-y)} - e^{-\lambda x})\, dy \\ &= 1 - e^{-\lambda x} - \lambda x e^{-\lambda x}, \qquad x < t \end{aligned} \tag{6}$$

De manera similar, de 4 se sigue

$$P(X_{N(t)} \le x, N(t) = n) = \int_0^t (e^{-\lambda(t-y)} - e^{-\lambda x}) g_{n-1}(y)\, dy \tag{7}$$

y de 3 y 7 resulta

$$\begin{aligned} P(X_{N(t)} \le x) &= P(X_{N(t)} \le x, N(t) = 1) + \sum_{n=2}^{\infty} P(X_{N(t)} \le x, N(t) = n) \\ &= e^{-\lambda t} - e^{-\lambda x} + \int_0^t (e^{-\lambda(t-y)} - e^{-\lambda x}) \sum_{n=2}^{\infty} g_{n-1}(y)\, dy \\ &= e^{-\lambda t} - e^{-\lambda x} + \int_0^t \lambda (e^{-\lambda(t-y)} - e^{-\lambda x})\, dy \\ &= 1 - e^{-\lambda x} - \lambda t e^{-\lambda x}, \qquad x \ge t \end{aligned} \tag{8}$$

Derivamos 6 y 8 y obtenemos la función de densidad de $X_{N(t)}$.

$$f_t(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{si } x < t \\ \lambda(1 + \lambda t) e^{-\lambda x} & \text{si } x \ge t \end{cases} \tag{9}$$

El tiempo de espera es igual a la diferencia $T_e = S_{N(t)} - t$. Un razonamiento similar al anterior condicionando por $S_{N(t)-1} = y$ nos permite calcular la función de distribución de T_e.

$$\begin{aligned} P(T \le x) &= e^{-\lambda t} - e^{\lambda(x+t)} + \sum_{n=2}^{\infty} \int_0^t (e^{-\lambda(t-y)} - e^{-\lambda(x+t-y)}) g_{n-1}(y)\, dy \\ &= 1 - e^{-\lambda x} \end{aligned} \tag{10}$$

es decir, T tiene distribución exponencial de parámetro λ, lo que da la razón a la respuesta **A** del enunciado. La consecuencia está contra la intuición ingenua: el tiempo que le resta por esperar es igual que el tiempo de espera entre un autobús y el siguiente.

En general, la Paradoja de la inspección se presenta para cualquier sucesión, $\{X_n\}$, de variables aleatorias independientes, es decir en cualquier *proceso de renovación* como los que estudiamos en el tema 3. El el apartado 7.7 que estudiaremos a continuación, ROSS trata esta paradoja desde un enfoque general.