Repaso de la distribución Exponencial de probabilidad y de los procesos estocásticos tipo Poisson

Se resumen aquí los principales conceptos vinculados a la distribución exponencial de probabilidad y a los procesos Poisson, ambos muy utilizados en teoría de colas.

1) Distribución Exponencial de Probabilidad

La distribución exponencial es una distribución de probabilidad continua que empleamos frecuentemente en teoría de colas para modelizar el tiempo que transcurre entre eventos. Estos eventos pueden ser, por ejemplo, la llegada de un cliente, la finalización de un servicio, la interrupción (o la reactivación) del funcionamiento de un centro de atención.

Una variable aleatoria exponencial es una variable aleatoria continua cuya función de densidad de probabilidad es de la forma:

$$f(t) = \lambda e^{-\lambda t} \qquad \text{... (1)}$$

$con: t > 0;\ \ \lambda > 0 \ y\ constante$

Se puede demostrar que:

- La función de distribución acumulada es:

$$F(t) \equiv Pr\{T \leq t\} = 1 - e^{-\lambda t} \qquad \text{... (2)}$$

- La función complementaria de la distribución acumulada es:

$$G(t) \equiv Pr\{T > t\} = e^{-\lambda t} \qquad \text{... (3)}$$

- La media es:

$$E\,[\,T\,] \;=\; 1\,/\lambda \qquad \text{… (4)}$$

- La varianza es:

$$Var\,[\,T\,] \;=\; 1\,/\lambda^2 \qquad \text{… (5)}$$

El parámetro λ representa la cantidad de eventos por unidad de tiempo; es la tasa o la frecuencia a la cual ocurren los eventos. Su inversa $1/\lambda$ tiene unidades de tiempo.

Variables aleatorias sin memoria

Decimos que una variable aleatoria no tiene memoria si:

$$Pr\,\{\,T\;>(t\;+\;s)\,|\,T\;>s\,\}\;=\;Pr\,\{\,T\;>t\,\},(\,s,t\;\geq\;0) \qquad \text{… (6)}$$

En otros términos, si ya ha transcurrido el tiempo s, la probabilidad de que el evento demore un tiempo t adicional, es independiente de que ya transcurrió el tiempo s. Por ejemplo, si los colectivos llegan a la parada de acuerdo con una distribución aleatoria sin memoria y he esperado un tiempo s en el cual el colectivo no llegó, eso no significa que el colectivo estará por llegar. La probabilidad de que demore un tiempo t es independiente del tiempo ya transcurrido.

Teorema 1:

Una variable aleatoria exponencial tiene la propiedad de no tener memoria

El teorema de Bayes permite calcular la probabilidad de un suceso A, condicionado a la ocurrencia de otro suceso B:

$$Pr(A\,|B) = \frac{Pr(B|A)\,.\,Pr(A)}{Pr(B)}$$

En este caso el suceso $A = T > (t+s)$ y el suceso $B = T > s$

$$\Pr\{\,T > (t\ +\ s)\mid T\ > s\,\} =$$
$$= \frac{Pr\ \{T > s|T\ > (t+s)\}.Pr\{\,T > (t\ +\ s)\,\}}{Pr\{\,T\ > s\,\}}$$

Si $T\ > (t\ +\ s)$ entonces necesariamente $T > s$ y

$Pr(T\ > s\,|\,T > t+s) = 1$

$$Pr\{\,T\ > (t\ +\ s)\mid T\ > s\,\} = \frac{Pr\{\,T\ > (t\ +\ s\,)\}}{Pr\ \{T\ > s\,\}} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$$

$$Pr\{\,T\ > (t\ +\ s\,)|\,T\ > s\,\} = e^{-\lambda t} = Pr\{\,T >\ t\,\}$$

Cuando T mide los tiempos entre llegadas, el tiempo hasta la siguiente llegada es independiente del tiempo desde la última llegada.

Si T representa los tiempos de los servicios, entonces el tiempo necesario para concluir un servicio es independiente del tiempo que el cliente ya ha estado en servicio.

Teorema 2:

La distribución exponencial de probabilidad es la única distribución continua que tiene la propiedad de no tener memoria.

No demostraremos aquí el teorema, pero veremos con un ejemplo que la distribución uniforme de probabilidad sí tiene memoria: El tiempo hasta que arriba un colectivo a la parada sigue una distribución uniforme [0;10] minutos. La probabilidad de que el arribo demore más de 1 minuto es

$P(T > 1) = 9/10$. Ahora supongamos que ya transcurrió un tiempo $s = 4$ minutos y calculemos la probabilidad de que el arribo demore más de un minuto.

$$Pr\{T > (t + s) \mid T > s\} = \frac{Pr\{T > (t + s)\}}{Pr\{T > s\}} =$$

$$Pr\{T > (1 + 4) \mid T > 4\} = Pr\{\frac{T > 5}{T > 4}\} =$$

$$= (5 / 10) / (6 / 10) = 5/6$$

$$Pr\{T > (1 + 4) \mid T > 4\} = 5/6 \neq P(T > 1) = 9/10.$$

En este ejemplo, el evento se vuelve cada vez más probable conforme transcurre el tiempo. Se trata entonces de un proceso con memoria ya que la probabilidad de arribo no es independiente del tiempo transcurrido.

Teorema 3:

Sean $T_1, T_2 \dots\dots T_n$ variables aleatorias exponenciales independientes con parámetros $\lambda_1, \lambda_2, \dots\dots\ \lambda_n$.

Entonces:

$min(T_1, T_2 \dots\dots T_n)$ *distribuye exponencialmente con parámetro* $\lambda_1 + \lambda_2 + \dots\dots + \lambda_n$

<u>Actividad 1</u>: El tiempo hasta que arriba un cliente tipo “1” distribuye exponencialmente con media 40 minutos. El tiempo hasta que arriba un cliente tipo “2” distribuye exponencialmente con media 10 minutos. ¿Cuál es la probabilidad de esperar al menos 10 minutos por un cliente tipo “1” o tipo “2”?

Respuesta: 0.29

Teorema 4:

Sean $T_1, T_2 \ldots \ldots T_n$ *variables aleatorias exponenciales independientes con parámetros* $\lambda_1, \lambda_2, \ldots .. \ \lambda_n$.

Entonces:

$$Pr\{T_i = min(T_1, T_2 \ldots \ldots T_n)\} = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \ldots . + \lambda_n}$$

<u>Actividad 2</u>: El tiempo hasta que se arriba un cliente tipo "1" distribuye exponencialmente con media 40 minutos. El tiempo hasta que arriba un cliente tipo "2" distribuye exponencialmente con media 10 minutos. a) ¿Cuál es la probabilidad que el primer cliente en arribar sea tipo "1"? b) ¿Cuál es la probabilidad que el primer cliente en arribar sea tipo "2"?

Respuesta: a) 0.2; b) 0.8

Valor medio y distribución de la variable

La mayor parte de los valores de una variable aleatoria con distribución exponencial de probabilidad son menores que el valor de la media. Efectivamente, si T distribuye exponencialmente con parámetro λ, la media de T es $1/\lambda$ y por la ecuación (2):

$$\Pr\left(T \leq \frac{1}{\lambda}\right) = 1 - e^{-\lambda t} = 1 - e^{-1} = 0.632$$

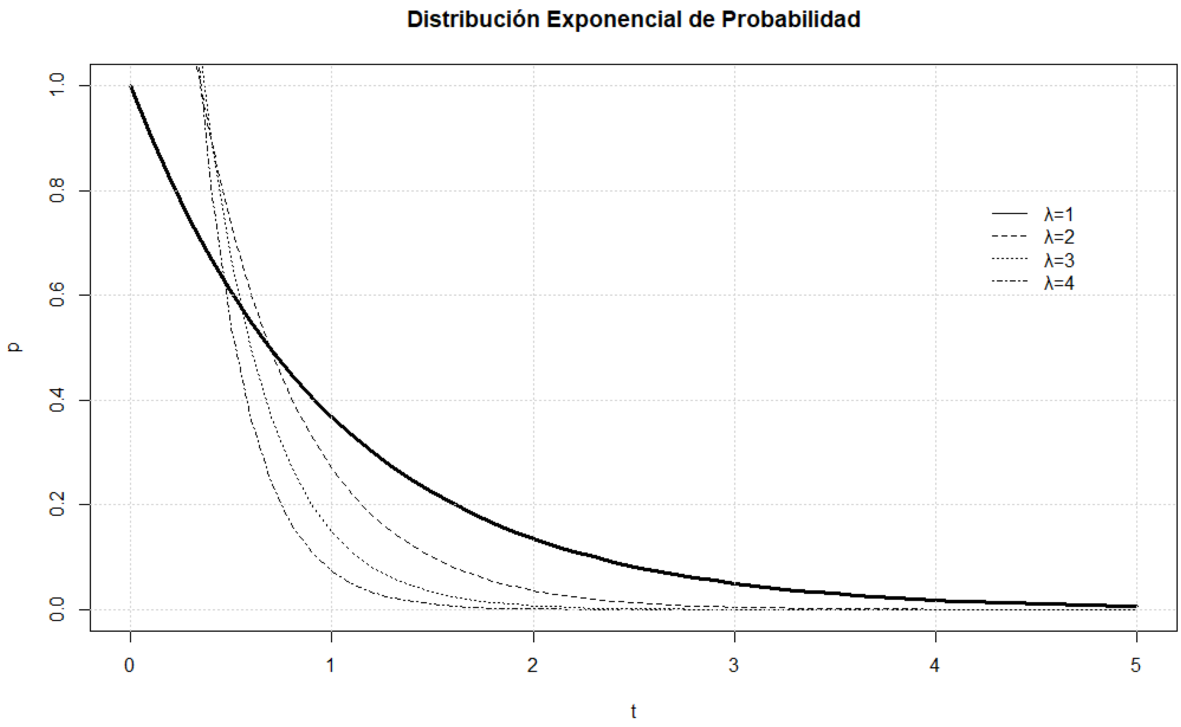
Por tanto, el 63.2% de los valores son menores que la media de la distribución.

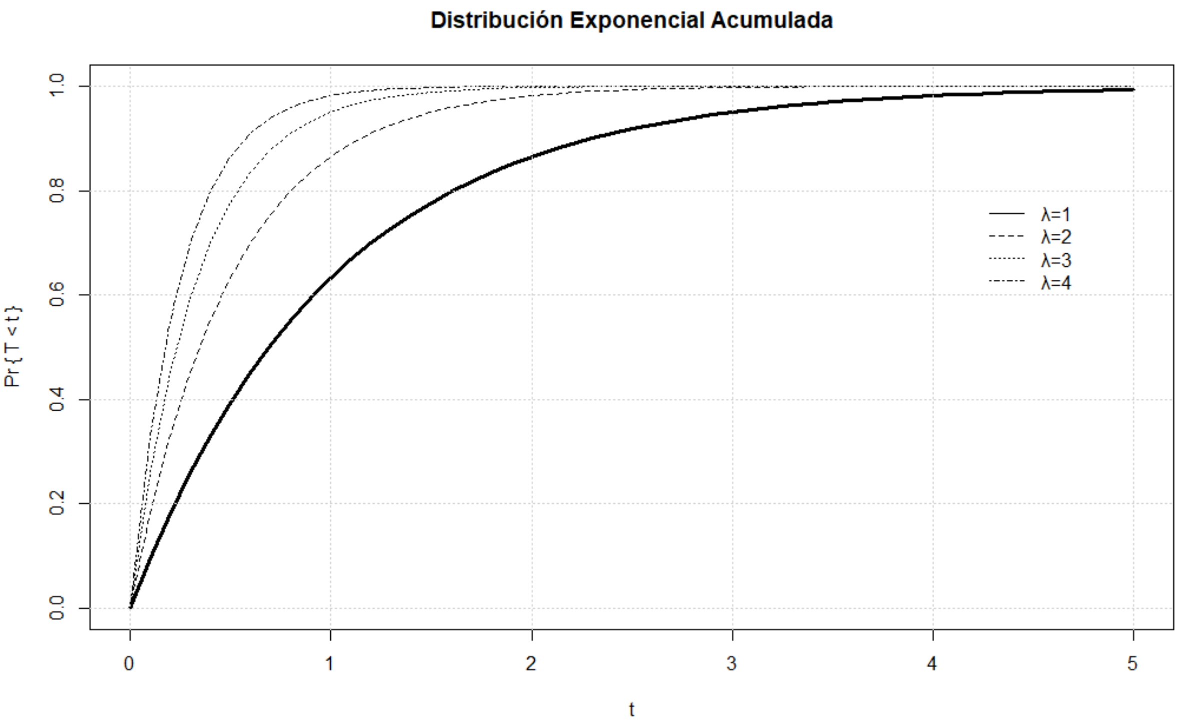
Cuando aplicamos la distribución exponencial para modelizar los tiempos de un servicio en un sistema de atención, estamos suponiendo la preponderancia de servicios con duración menor a la media combinados con algunos servicios de mayor duración.

Visto de otro modo existe una mayor proporción de servicios de tratamiento simple (duración menor a la media) y una menor proporción de servicios complejos (duración mayor a la media).

Una interpretación similar podemos hacer respecto de los arribos a un sistema de atención. Si los tiempos entre arribos son modelizados mediante una distribución exponencial esto implica el arribo de una mayoría de clientes espaciados un tiempo menor a la media seguidos de otros clientes espaciados con tiempos más largos. Durante el arribo de los clientes poco espaciados en el tiempo se formará una fila de espera y durante el arribo de los clientes más espaciados el sistema se recuperará.

La distribución exponencial no resulta entonces muy adecuada para modelizar aquellos casos en los que el servicio es el mismo para todos los clientes como por ejemplo una línea de montaje.


Distribución Exponencial de Probabilidad
p
t
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
λ=1
λ=2
λ=3
λ=4


Distribución Exponencial Acumulada
Pr{ T < t }
t
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
λ=1
λ=2
λ=3
λ=4

2) Procesos estocásticos tipo Poisson

Los procesos Poisson son comúnmente utilizados para modelizar los arribos a un sistema de atención. Los arribos (en general los eventos) ocurren en forma aleatoria en el tiempo.

Consideremos una secuencia de acontecimientos o eventos E que resultan de la repetición de una misma experiencia y que se suceden en el tiempo. El número n de sucesos que se producen en el intervalo de tiempo t es una variable aleatoria N. A la probabilidad de que $N = n$ en el intervalo de tiempo t la llamaremos $p_n(t)$.

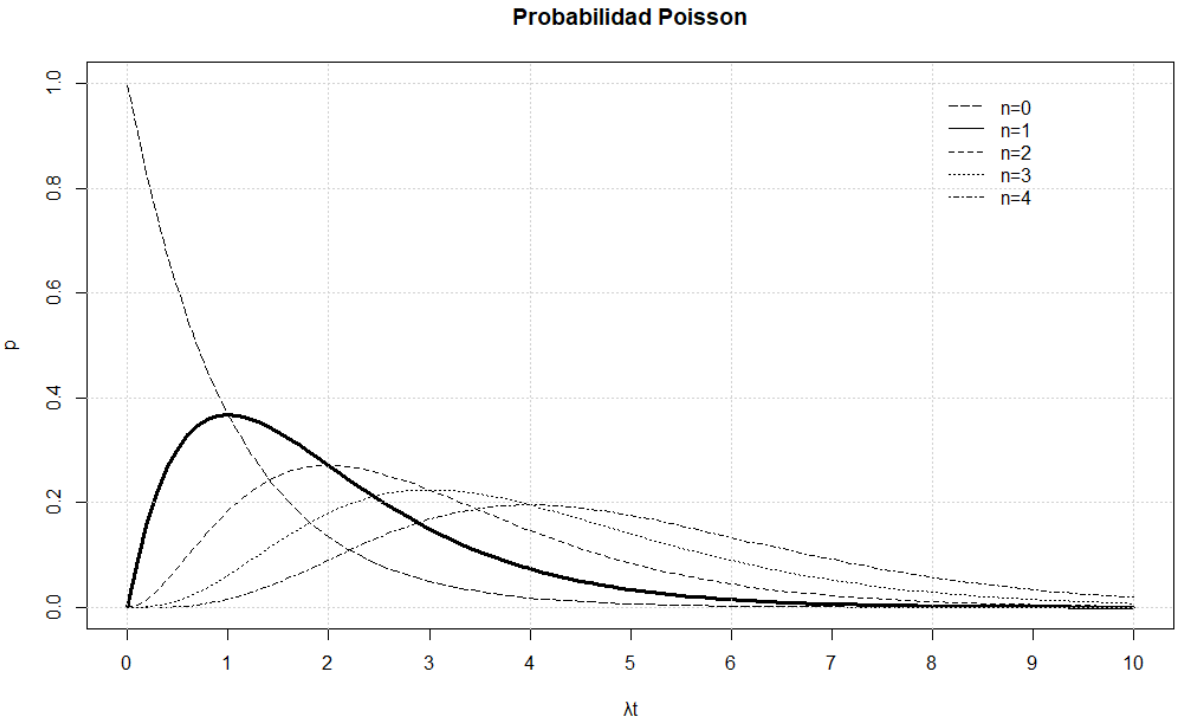
Hacemos las siguientes hipótesis:

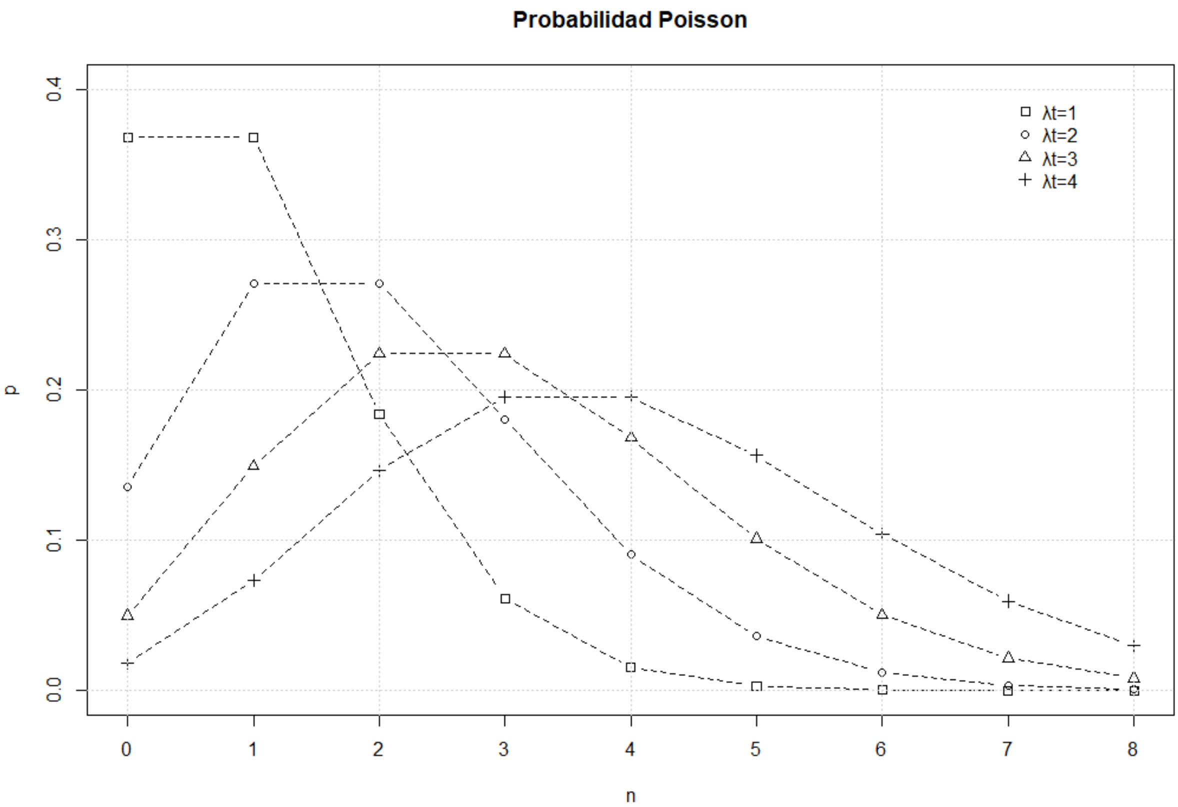
1. La probabilidad de que se verifiquen n de sucesos en el intervalo de tiempo t, que hemos llamado $p_n(t)$, depende solo del intervalo de tiempo y no depende del instante inicial (homogeneidad en el tiempo).
2. La probabilidad de que el suceso E se produzca más de una vez en el intervalo de tiempo dt, es infinitamente pequeña con respecto a dt.
3. La probabilidad de que E se produzca una sola vez en el intervalo de tiempo infinitesimal dt, es proporcional a dt; es decir λdt.
4. La variable aleatoria N es una función de t que puede tomar los valores: 0, 1, 2, …..i en los instantes aleatorios $t_1, t_2, \ldots\ldots t_i$. El incremento de N en un intervalo de tiempo es igual al número de sucesos n que acontecen en ese intervalo (su probabilidad es $p_n(t)$ y es independiente de los valores previos de $N(t)$).

La función aleatoria $N(t)$ así descripta es un proceso de *Markov* (sin memoria) que está definido completamente por la probabilidad $p_n(t)$. Se puede demostrar que $p_n(t)$ sigue la ley Poisson:

$$p_n(t) = \Pr\{N(t) = n\} = \frac{(\lambda t)^n \, e^{-\lambda t}}{n!} \quad (n = 0, 1, 2, \ldots)$$

La media de la distribución es $E(N) = \lambda t$ y la varianza es $Var = \lambda t$.





Nota: Las líneas de puntos se agregaron para facilitar la lectura del gráfico; la variable "n" es discreta.

Teorema 5: *Un proceso Poisson tiene intervalos estacionarios. Esto es* $para\ t > s,\ N(t) - N(s)\ está\ distribuido\ en\ forma\ idéntica\ a\ N(t + h) - N(s + h), para\ cualquier\ h > 0.$

Esto significa que la distribución del número de eventos en un intervalo depende de la longitud, pero no de la posición o localización del intervalo en el tiempo.

Teorema 6: *Sea* $N(t)$ $un\ proceso\ Poisson\ con\ tasa$ $\lambda > 0$. *Entonces los tiempos entre eventos sucesivos son independientes y exponencialmente distribuidos con tasa* λ.

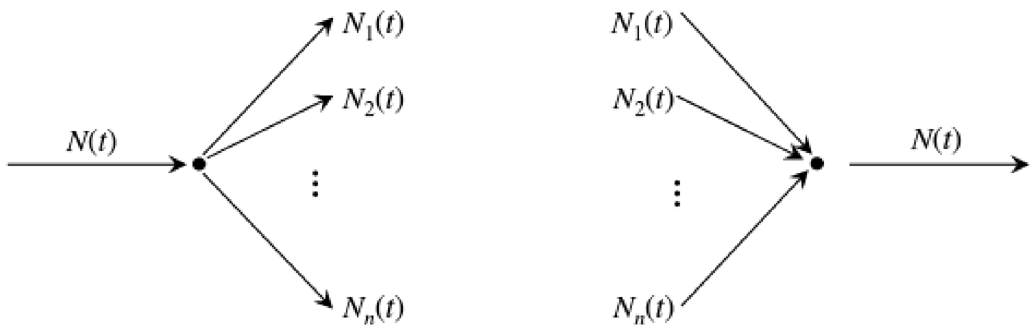
La distribución exponencial y la distribución Poisson resultan así asociadas: La distribución Poisson da la probabilidad de que ocurra un número n de sucesos en un intervalo t de tiempo. La distribución exponencial da la probabilidad de que dos eventos consecutivos estén separados por un tiempo t.

Teorema 7: *Sea* $N(t)$ $un\ proceso\ Poisson\ con\ tasa$ $\lambda > 0$. *Supongamos que cada evento es identificado como tipo(i) con probabilidad* p_i, *todas ellas independientes entre sí. Entonces,* $N_i(t)$ *es un proceso Poisson con tasa* λp_i. *Además* $N_i(t)$ *y* $N_j(t)$ *son independientes para todo* $i \neq j$.

Teorema 8: *Sean* $N_1(t), N_2(t), \ldots\ldots N_n(t)$ *procesos Poisson independientes con tasas* $\lambda_1(t), \lambda_2(t), \ldots\ldots \lambda_n(t)$.

Entonces $N(t) = N_1(t) + + N_2(t), \ldots\ldots + N_n(t)$ *es un proceso Poisson con tasa* $\lambda = \lambda_1(t) + \lambda_2(t) + \cdots + \lambda_n(t)$

Los teoremas 7 y 8 refieren a la división y combinación de procesos Poisson. El primero nos indica que, bajo condiciones de independencia, la división de un proceso Poisson, genera a su vez procesos Poisson independientes. El segundo de los teoremas indica que la combinación de procesos Poisson independientes en un único proceso conlleva que el proceso originado de la combinación es también un proceso Poisson.



Problemas complementarios:

1) El tiempo que tarda en llegar un autobús a una parada determinada sigue una distribución exponencial con una media de 10 minutos. ¿Cuál es la probabilidad de que el autobús llegue en los próximos 5 minutos?

Respuesta: 0.3935

2) El número de errores de impresión en una impresora láser sigue un proceso de Poisson con una tasa media de 0.1 errores por página. Si un trabajo de impresión tiene 50 páginas, ¿cuál es la probabilidad de que haya al menos 5 errores de impresión en el trabajo completo?

Respuesta: 0.5595

3) El número de llegadas de clientes por hora en una tienda sigue un proceso de Poisson con una tasa media de 4 llegadas por hora. ¿Cuál es la probabilidad de que lleguen más de 6 clientes en los próximos 30 minutos?

Respuesta: 0.0045

4) La vida útil de una bombilla sigue una distribución exponencial con una media de 1000 horas. ¿Cuál es la probabilidad de que la bombilla dure al menos 1500 horas?

Respuesta: 0.2231

5) El número de errores de codificación en un proyecto de software sigue un proceso de Poisson con una tasa media de 2 errores por hora. Si un programador trabaja en el proyecto durante 8 horas, ¿cuál es la probabilidad de que haya al menos 5 errores de codificación?

Respuesta: 0.9996

6) La cantidad de solicitudes de crédito por día en una institución financiera sigue un proceso de Poisson con una tasa media de 10 solicitudes por día. Si la institución financiera está abierta de lunes a viernes, ¿cuál es la probabilidad de que reciban al menos 40 solicitudes de crédito en una semana?

Respuesta: 0.9354

7) El tiempo que tarda un paquete en llegar por correo sigue una distribución exponencial con una media de 3 días. ¿Cuál es la probabilidad de que el paquete llegue en un plazo de 5 días o menos?

Respuesta: 0.8111

8) Calcular la derivada respecto del tiempo de $p_{n=1}(t) = \lambda t\, e^{-\lambda t}$ valuada en $t = 0$.

Respuesta: λ

BIBLIOGRAFÍA

Bronson, Richard. **Investigación de Operaciones**. Mc Graw Hill (1998).

García, Roberto Mariano. **Inferencia Estadística y Diseño de Experimentos**. EUDEBA (2004).

Kaufmann, Arnold. **Métodos y Modelos de la Investigación de Operaciones** (Tomo 1). C.E.C.S.A (1984).

Shortle, John F; Thompson, James M; Gross, Donald; Harris, Carl M. **Fundamentals of Queueing Theory**. Wiley (2018).