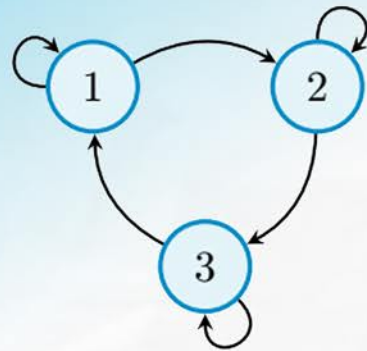
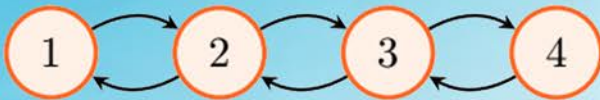
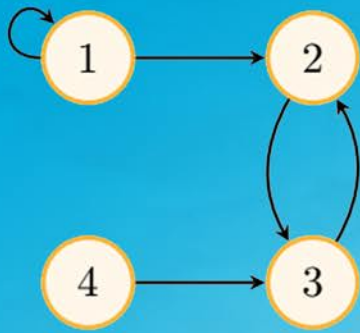


Procesos Estocásticos



Proceso Poisson

Se dice que el proceso a tiempo continuo $\{N_t : t \geq 0\}$ es un proceso Poisson de parámetro λ tal que cumple

- 1) $N_0 = 0$
- 2) Tiene incrementos independientes
- 3) Tiene incrementos estacionarios
- 4) Si $s < t$, entonces $N_t - N_s \sim \text{Poisson}(\lambda \cdot (t - s))$

Considerando $s = 0$ en el último inciso, se tiene que $N_t \sim \text{Poisson}(\lambda \cdot t)$.

Además, recordemos que una variable $X \sim \text{Poisson}(\lambda)$ tiene como función de densidad

$$f_X(x) = \mathbb{P}(X = x) = \exp(-\lambda) \cdot \frac{\lambda^x}{x!}$$

Ejercicio 1

Los arribos de la línea 1 del camión forman un proceso de Poisson de tasa 1 por hora y los de la línea 7 forman un proceso de Poisson independiente de tasa 7 camiones por hora.

- 1) ¿Cuál es la probabilidad que exactamente pasen 3 camiones en una hora?
- 2) ¿Cuál es la probabilidad de que pasen exactamente 3 camiones de la línea 7 mientras espero al de la línea 1 en una hora?
- 3) Cuando el equipo de mantenimiento realiza una huelga, la mitad de los camiones no llegan a mi parada.
¿Cuál es entonces la probabilidad de no ver ningún camión durante 30 minutos?

Solución

Consideremos a X como los arribos de la línea 1 y Y los arribos de la línea 7.

Por lo tanto, tenemos que $X \sim \text{Poisson}(1)$ y $Y \sim \text{Poisson}(7)$.

Además, por Probabilidad 2, sabemos que si $X \perp Y$ entonces $X + Y = Z \sim \text{Poisson}(8)$ cada hora.

La probabilidad de que pasen exactamente 3 camiones en una hora, sin importar la línea, es

$$\mathbb{P}[Z = 3] = e^{-8} \cdot \frac{8^3}{3!} \approx 0.02862614$$

La probabilidad de que pasen exactamente 3 camiones de la línea 7 mientras se espera al de la línea 1 en una hora, significa que no ha llegado algún camión de la línea 1 en una hora. Nos piden que ambos eventos ocurran simultáneamente

$$\mathbb{P}[X = 0, Y = 3] = \mathbb{P}[X = 0] \cdot \mathbb{P}[Y = 3] \approx 0.01917728$$

El que la mitad de los camiones no llegan a la parada, implica que la tasa de arribos se ha dividido entre 2.

Lo que significa que ahora $X \sim \text{Poisson}(1/2)$ y $Y \sim \text{Poisson}(7/2)$ cada hora.

Más aún, el arribo de todos los camiones sin importar la línea, se tiene que $Z \sim \text{Poisson}(4)$ cada hora.

Sin embargo, el único detalle es que estamos pensando aún en que los parámetros (λ 's) están en horas.

El ejercicio nos pide intervalos de tiempo de 30 minutos.

Por las propiedades del Proceso Poisson, se tiene que $Z \sim \text{Poisson}(2)$ cada 30 minutos.

Por lo tanto, la probabilidad de no ver ningún camión durante 30 minutos es

$$\mathbb{P}[Z = 0] = e^{-2} \cdot \frac{2^0}{0!} \approx 0.1353353$$

Alguien podría interpretar la pregunta anterior de la siguiente manera,

¿Cuál es la probabilidad de que el intervalo de tiempo entre un camión y otro supere los 30 minutos?

Sin embargo, es una pregunta que se relaciona con el tiempo transcurrido entre un evento y otro, no con el conteo.

Sabemos que el tiempo interarribo tiene una distribución $T \sim \text{Exponencial}(2)$

Por lo que la respuesta a la pregunta anterior podría pensarse como

$$\mathbb{P}(T \geq 30) = 1 - \mathbb{P}(T < 30) = 1 - F_T(30) = 1 - (1 - e^{-2 \cdot 30}) \approx 0$$

Ejercicio 2

Un peatón desea cruzar una calle de un sólo sentido. Suponga que la cantidad de vehículos que han pasado hasta el tiempo t forman un proceso de Poisson de intensidad λ y que toma a unidades de tiempo cruzar la calle.

Asumiendo que el peatón no cruza hasta estar seguro de que no lo atropellan,

¿Cuanto tarda en promedio en cruzar la calle?

Responda la misma pregunta pero ahora considerando si debe cruzar dos calles sucesivas cuando hay un camellón en medio y el caso cuando no hay ningún camellón.

Solución

Consideremos a X el tiempo transcurrido entre el paso de un vehículo y otro, esto es el tiempo inter-arribo.

Recordemos que si $X \sim \text{Exponencial}(\lambda)$, entonces tenemos que $\mathbb{E}[X] = 1/\lambda$

Finalmente, el tiempo promedio que tarda en cruzar la calle es justamente $\mathbb{E}[X|X > a]$ donde $X \sim \text{Exponencial}(\lambda)$

No es muy complicado de demostrar que $\mathbb{E}[X|X > a] = a + 1/\lambda$

En el caso de que exista un camellón en medio de dos calles consecutivas, es $2 \cdot \mathbb{E}[X|X > a] = 2(a + 1/\lambda)$.

Lo anterior ocurre debido a que cuando llega al camellón, el tiempo en el que pasa un auto y otro sigue siendo una distribución exponencial, además de que sabemos que la distribución tiene pérdida de memoria, por lo que $P(X > t + s | X > s) = P(X > t)$.

El proceso se reinicia sin importar el tiempo que le haya tomado llegar al camellón. Por lo que el tiempo promedio para cruzar 2 calles con un camellón es sumar el tiempo promedio de cruzar dos calles de manera individual.

En el caso de que no exista un camellón, el problema es muy similar, porque ahora consideremos que es una sola calle pero el doble de ancha. Esto implica que el tiempo que toma en cruzar la calle se duplica, esto es $2a$.

El tiempo promedio que tarda en cruzar dos calles sin camellón es $\mathbb{E}[X|X > 2a]$ donde $X \sim \text{Exponencial}(\lambda)$

Ejercicio 3

Suponga que las familias de México emigran a Estados Unidos a una tasa de $\lambda = 2$ por día y que las llegadas a EU forman un proceso de Poisson de tasa λ . Si el número de miembros de cada familia puede tomar los valores 1, 2, 3 y 4 con probabilidades $1/6, 1/3, 1/3, 1/6$ respectivamente y además supongamos que las familias son independientes, entonces ¿Cuál es el valor esperado y la varianza de la cantidad de individuos que llegan a EU en semanas?

Realice un programa en R que permita observar la cantidad de individuos que llegan a EU durante 2 semanas.

Solución

El valor esperado y varianza (en semanas) corresponden a la esperanza y varianza de una distribución *Poisson*(14). La esperanza y varianza de una distribución *Poisson*(λ) es su parámetro λ , en este caso 14 sería la respuesta.

A continuación se muestra el código de R donde también se añade una simulación de esta primer parte.

```
#####
# Parámetros Globales
set.seed(1)          # Semilla de números pseudo-aleatorios
lambda  <- 2         # Num. Promedio de Eventos por Unidad de Tiempo (Días en este caso)
tiempo_t <- 7         # Tiempo que "observaremos" el Proceso Poisson en Días (7 días = 1 Semana)
num_sims <- 50000     # Número de simulaciones para estimar la media y varianza

# ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
# Primera Parte :: Estimación de la Esperanza y Varianza
# Se simulan la cantidad de "eventos" en el total de tiempo y se estima la media y varianza
vector_simulaciones <- vector(mode = "numeric", length = num_sims)
for( sim in seq(num_sims) ){
  vector_simulaciones[ sim ] <- rpois(n = 1, lambda = lambda*tiempo_t)
}
media_estimada  <- mean(vector_simulaciones) # 14.02228
varianza_estimada <- var(vector_simulaciones) # 14.04742

# ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
# Segunda Parte :: Modelación la cantidad de individuos en EEUU en 2 semanas
set.seed(1)          # Semilla de números pseudo-aleatorios
lambda  <- 2         # Num. Promedio de Eventos por Unidad de Tiempo (Días en este caso)
tiempo_t <- 7*2      # Tiempo que "observaremos" el Proceso Poisson en Días (14 días = 2 Semanas)

# Se simulan la cantidad de "eventos" en el total de tiempo
num_simulaciones <- rpois(n = 1, lambda = lambda*tiempo_t)

# Se simulan los tiempos "inter-arribo" de cada evento
# Los cuales se suman para obtener los tiempos en los que se presentan los eventos
simulaciones_exp <- rexp(n = num_simulaciones, rate = lambda)
tiempos_llegada  <- cumsum(simulaciones_exp)

# Se simulan los tamaños de las familias de cada evento
# Los cuales se suman para obtener la altura de la función
simulaciones_fam <- sample(x = c(1,2,3,4), size = num_simulaciones, replace = TRUE,
                           prob = c(1/6,1/3,1/3,1/6) )
cantidad_individuos <- cumsum(simulaciones_fam)

# Generamos los vectores que representen las coordenadas de los puntos y graficamos
valores_x <- tiempos_llegada
valores_y <- c( 0 , cantidad_individuos ) # (se añade la altura 'inicial')
function_a_trozos <- stepfun(x = valores_x, y = valores_y, f = 0)
plot(function_a_trozos, main="Inmigrantes a EU", verticals = FALSE, pch=19, col.points="red",
      lwd=1.5, cex.points=1.1, xlim=c(0, tiempo_t), xlab = "Días", ylab = "Individuos")
#####
```

Ejercicio 4

Un cable submarino tiene defectos de acuerdo a un proceso de Poisson de parámetro $\lambda = 0.1$ por km.

- 1) ¿Cuál es la probabilidad de que no haya defectos en los primeros dos kilómetros de cable?
- 2) Si no hay defectos en los primeros dos kilómetros, ¿cuál es la probabilidad de que tampoco los haya en el tercer kilómetro?

Solución

Sea entonces nuestro proceso Poisson $N_1 \sim \text{Poisson}(0.1)$

Como el parámetro están en km y la pregunta está hecha para 2 km, debemos transformar nuestra tasa λ . Por lo tanto, nuestro proceso Poisson para 2 km es $N_2 \sim \text{Poisson}(0.2)$

Con lo que la probabilidad de que no haya defectos en los primeros 2 kilómetros es

$$\mathbb{P}(N = 0) = e^{-0.2} \approx 0.8187$$

Nos piden calcular

$$\mathbb{P}(N_3 - N_2 = 0 | N_2 = 0)$$

Sabemos que el proceso tiene incrementos independientes, por lo tanto el número de defectos entre el segundo y tercer kilómetro es independiente del número de defectos que hay entre el kilómetro cero y el segundo. Esto es que $N_3 - N_2 \perp N_2 - N_0$

De modo que nos queda

$$\mathbb{P}(N_3 - N_2 = 0 | N_2 = 0) = \mathbb{P}(N_3 - N_2 = 0) = \mathbb{P}(N_1 = 0) = e^{-0.1} \approx 0.9048$$

Ejercicio 5

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.?

Solución

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

$$\begin{aligned} \mathbb{P}(N_{1/2} = 1, N_{5/2} = 5) &= \mathbb{P}(N_{1/2} = 1, N_{5/2} - N_{1/2} = 4) \\ &= \mathbb{P}(N_{1/2} = 1) \cdot \mathbb{P}(N_{5/2} - N_{1/2} = 4) \\ &= \mathbb{P}(N_{1/2} = 1) \cdot \mathbb{P}(N_2 = 4) \end{aligned}$$

Ahora, tenemos que $N_{1/2} \sim \text{Poisson}(2)$ y $N_2 \sim \text{Poisson}(8)$. Por lo que

$$\mathbb{P}(N_{1/2} = 1, N_{5/2} = 5) = \left(e^{-2} \cdot \frac{2^1}{1!} \right) \left(e^{-8} \cdot \frac{8^4}{4!} \right) \approx 0.0154965155$$

Proceso de Poisson Compuesto

Asociamos ahora una variable aleatoria Y_k a cada evento de un proceso de Poisson

Suponemos que las variables Y_k con $k \in \{1, 2, 3, \dots\}$ son i.i.d y también son independientes del proceso Poisson.

Algunos ejemplos que se pueden modelar con estos procesos son los carros que llegan a un centro comercial y las variables asociadas Y_k son el número de pasajeros que hay en cada auto; o modelar los mensajes que llegan a un servidor para ser transmitidos via internet y las variables Y_k representan el tamaño de los mensajes.

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

$$S(t) = Y_1 + \dots + Y_{N(t)} = \sum_{k=1}^{N(t)} Y_k$$

Además, si $N(t) = 0$ entonces $S(t) = 0$ debido a que no hay nada que sumar.

Si consideramos $Y_k = 1$ para cualquier k , obtenemos el Proceso Poisson que hemos analizado anteriormente.

Realizando algunas operaciones se pueden demostrar las siguientes igualdades

$$\mathbb{E}[S(t)] = \mathbb{E}[N(t)] \cdot \mathbb{E}[Y] \quad \text{Var}[S(t)] = \mathbb{E}[N(t)] \cdot \text{Var}[Y] + \text{Var}[N(t)] \cdot (\mathbb{E}[Y])^2$$

En nuestro caso tenemos $N(t) \sim \text{Poisson}(\lambda t)$. Por lo tanto $\mathbb{E}[N(t)] = \text{Var}[N(t)] = \lambda t$

Demostración 1

Para resolver esto, debemos utilizar las propiedades de probabilidad condicional

$$\mathbb{E}[S(t)] = \mathbb{E}[\mathbb{E}[S(t)|N(t)]]$$

Resolvamos primero el caso cuando tenemos una suma fija de variables aleatorias $\mathbb{E}[S(t)|N(t) = n]$

$$\mathbb{E}[S(t)|N(t) = n] = \mathbb{E}[Y_1 + Y_2 + Y_3 + \dots + Y_n] = \mathbb{E}[Y_1] + \mathbb{E}[Y_2] + \mathbb{E}[Y_3] + \dots + \mathbb{E}[Y_n] = n \cdot \mathbb{E}[Y]$$

Por lo tanto, $\mathbb{E}[S(t)|N(t)] = N(t) \cdot \mathbb{E}[Y]$. Finalmente tenemos que

$$\mathbb{E}[S(t)] = \mathbb{E}[\mathbb{E}[S(t)|N(t)]] = \mathbb{E}[N(t) \cdot \mathbb{E}[Y]] = \mathbb{E}[N(t)] \cdot \mathbb{E}[Y]$$

Demostración 2

De manera similar, utilizaremos las propiedades de probabilidad condicional para resolverlo

$$\text{Var}[S(t)] = \mathbb{E}[\text{Var}[S(t)|N(t)]] + \text{Var}[\mathbb{E}[S(t)|N(t)]]$$

Resolvamos la varianza de $S(t)$ para una suma fija de variables aleatorias $\text{Var}[S(t)|N(t) = n]$

$$\text{Var}[S(t)|N(t) = n] = \text{Var}[Y_1 + Y_2 + Y_3 + \dots + Y_n] = \text{Var}[Y_1] + \text{Var}[Y_2] + \text{Var}[Y_3] + \dots + \text{Var}[Y_n] = n \cdot \text{Var}[Y]$$

Por lo tanto, $\text{Var}[S(t)|N(t)] = N(t) \cdot \text{Var}[Y]$. Análogamente $\mathbb{E}[S(t)|N(t)] = N(t) \cdot \mathbb{E}[Y]$

$$\begin{aligned} \text{Var}[S(t)] &= \mathbb{E}[\text{Var}[S(t)|N(t)]] + \text{Var}[\mathbb{E}[S(t)|N(t)]] \\ &= \mathbb{E}[N(t) \cdot \text{Var}[Y]] + \text{Var}[N(t) \cdot \mathbb{E}[Y]] \\ &= \mathbb{E}[N(t)] \cdot \text{Var}[Y] + \text{Var}[N(t)] \cdot (\mathbb{E}[Y])^2 \end{aligned}$$

Ejercicio 1

El número de clientes de una tienda durante el día tiene distribución de Poisson de media 30 y el gasto que cada cliente realiza sigue una distribución Gamma con media \$150 y desviación típica de \$50.

Indice el gasto esperado de lo que se gasta en dicha tienda en un día. También proporcione su desviación estándar.

Solución

Por los cálculos anteriores sabemos que el ingreso medio por día es $30 \cdot \$150 = \$4,500$ (pesos)

La varianza del ingreso total es (en unidades cuadradas) es la siguiente

$$30 \cdot (\$50)^2 + 30 \cdot (\$150)^2 = \$^2 750.000 \quad (\text{pesos}^2)$$

Lo que significa una desviación típica que es de \$866,02.

A continuación se presenta una código en R que permite verificar los resultados encontrados

```
#####
set.seed(1)
num_simulaciones <- 500000
poisson_compuesto <- function() sum( rgamma(n = rpois(n = 1, lambda = 30), shape = 9, rate = 3/50 ) )
simulaciones <- replicate(n = num_simulaciones, poisson_compuesto() )
mean(simulaciones) # 4499.316
var(simulaciones) # 752060.5
#####
```

Ejercicio 2

Sea N el número de reclamaciones en un mes de una aseguradora con distribución $Poisson(\lambda = 1200)$ y sea Y el monto de cada reclamación con distribución en común $Weibull(\alpha = 5, \beta = 15,000)$

Determina el monto esperado que la aseguradora debe pagar al mes.

Solución

$$\mathbb{E}[Y] = \beta \Gamma \left(1 + \frac{1}{\alpha} \right) \quad \text{Var}[Y] = \beta^2 \left[\Gamma \left(1 + \frac{2}{\alpha} \right) - \Gamma^2 \left(1 + \frac{1}{\alpha} \right) \right]$$

$$\mathbb{E}[N] = 1200 \quad \text{Var}[N] = 1200 \quad \mathbb{E}[Y] \approx 13772.5 \quad \text{Var}[Y] = 15000^2 \left[\Gamma \left(\frac{7}{5} \right) - \Gamma^2 \left(\frac{6}{5} \right) \right] \approx 9951750$$

De tal manera que el monto esperado sería $\mathbb{E}[N] \cdot \mathbb{E}[Y] \approx \$16,527,000$

Ejercicio moral

¿Podría usted proporcionar un intervalo de confianza para el gasto esperado?

Simulando en R, podemos tener con un 95% de confianza un intervalo de \$15,576,508 a \$17,494,507.

```
#####
set.seed(1)
num_simulaciones <- 50000
poisson_compuesto <- function() sum( rweibull(n = rpois(n = 1, lambda = 1200), shape = 5, scale = 15000 ) )
simulaciones <- replicate(n = num_simulaciones, poisson_compuesto() )
mean(simulaciones) # 16,526,916
quantile(x = simulaciones, probs = c(0.025, 0.975)) # 2.5% = 15,576,508 97.5% = 17,494,507
#####
```


Ejercicio 3

Sea N el resultado al lanzar un dado equilibrado, es decir $Uniforme(1, 2, 3, 4, 5, 6)$.

Posteriormente lanzaremos nuevamente el dado N ocasiones, sea Y_1 los puntos del dado en la 1ra ocasión, Y_2 los puntos del dado en la 2da ocasión y así sucesivamente.

Y finalmente sea S el resultado de sumar los puntos obtenidos en las variables Y_1, Y_2, \dots, Y_N

- 1) ¿Cuáles son los posibles valores que puede tomar S ?
- 2) Calcula el valor esperado y la varianza de la variable S

Solución

Para la cuota inferior, consideremos que todos los dados caen 1. Significaría que $N = 1$ y luego en cada lanzamiento posterior se obtiene 1 punto, por lo cual S podría tomar como mínimo 1 punto.

Para la cuota superior, consideremos que todos los dados caen 6. Significaría que $N = 6$ y luego en cada lanzamiento posterior se obtiene 6 puntos, por lo cual S podría tomar como máximo 36 puntos

Sabemos que si existe $Z \sim Uniforme(1, 2, \dots, n)$ entonces

$$\mathbb{E}[Z] = \frac{n+1}{2} \quad \text{Var}[Z] = \frac{(n-1)(n+1)}{12}$$

En nuestro caso, sustituyendo obtenemos

$$\mathbb{E}[N] = \frac{6+1}{2} = \frac{7}{2} \quad \text{Var}[N] = \frac{(6-1)(6+1)}{12} = \frac{5 \cdot 7}{12} = \frac{35}{12}$$

Para la variable Y simplemente usaríamos las siguientes ecuaciones

$$\mathbb{E}[S] = \mathbb{E}[N] \cdot \mathbb{E}[Y] \quad \text{Var}[S] = \mathbb{E}[N] \cdot \text{Var}[Y] + \text{Var}[N] \cdot (\mathbb{E}[Y])^2$$

Además, las variables Y_k con $k \in \{1, 2, 3, \dots, X\}$ tienen la misma distribución que N , simplificando

$$\begin{aligned} \mathbb{E}[S] &= \mathbb{E}[N]^2 & \text{Var}[S] &= (\mathbb{E}[N] \cdot \text{Var}[N]) (1 + \mathbb{E}[N]) \\ \mathbb{E}[S] &= \frac{49}{4} = 12.25 & \text{Var}[S] &= \frac{245}{24} \cdot \frac{9}{2} = \frac{2205}{48} = 45.9375 \end{aligned}$$

Ejercicio moral

¿Podría dar usted una función de densidad para S ?

A continuación se presenta un código para comprobar los resultados anteriormente obtenidos

```
#####
set.seed(1)
num_simulaciones <- 50000
poisson_compuesto <- function() sum( sample(x = 1:6, size = sample(x = 1:6, size = 1), replace = TRUE) )
simulaciones <- replicate(n = num_simulaciones, poisson_compuesto() )
mean(simulaciones) # 12.25868
var(simulaciones) # 45.91024
quantile(x = simulaciones, probs = c(0.025, 0.975)) # 2.5% = 1 97.5% = 26
#####
```


Si continuamos con el ejercicio anterior

- 1) ¿Cuál sería la probabilidad de que $S = 36$?
- 2) ¿Cuál sería la probabilidad de que $S = 35$?
- 3) ¿Cuál sería la probabilidad de que $S = 34$?

Para resolver el primer inciso es necesario que el primer dado resulte 6, y también que se obtenga 6 en cada uno de los 6 dados que lancemos. En otras palabras, necesitamos calcular que 7 dados independientes se obtenga 6.

Como $\mathbb{P}[N = 6] = \frac{1}{6}$ y son independientes, la probabilidad que buscamos es $\left(\frac{1}{6}\right)^7 = \frac{1}{279,936} \approx 0.000003572245084$

Para el segundo inciso las únicas maneras de obtener 35 puntos es que el primer dado nos pida que lancemos 6 dados y que en uno de esos dados obtengamos 5 puntos. Esto último tiene $\binom{6}{1}$ combinaciones posibles.

La probabilidad que buscamos es $\frac{1}{6} \left[\binom{6}{1} \left(\frac{1}{6}\right)^6 \right] = \frac{1}{46,656} \approx 0.00002143347$

Para el tercer inciso, se complica la situación. También es necesario que el primer dado nos pida que lancemos 6 dados, pero podemos tener varias situaciones.

- + Un dado resulte 6 y otro dado resulte 4
- + Dos dados resulten 5

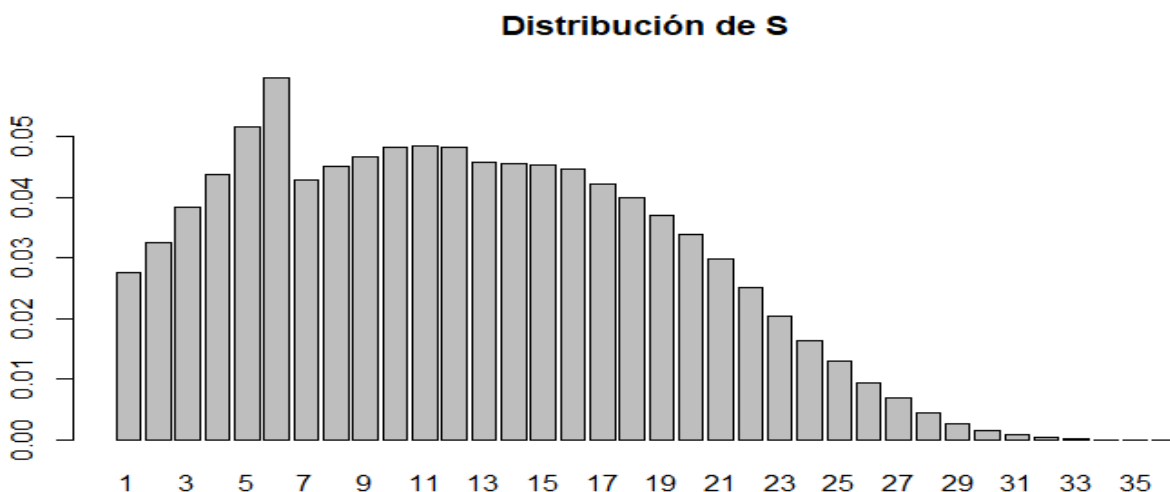
La primera situación tiene $\binom{6}{2}$ combinaciones posibles mientras que la segunda tiene $\binom{6}{2}$ combinaciones posibles. La probabilidad que buscamos es $\frac{1}{6} \left[\binom{6}{2} \left(\frac{1}{6}\right)^6 + \binom{6}{2} \left(\frac{1}{6}\right)^6 \right] = \frac{5}{46,656} \approx 0.000107167$

Como pueden observar, hacer esto empieza ser complicado al momento de generar estas probabilidades.

Sin embargo, podemos utilizar R para estimar la densidad de la variable S con solo aumentar el número de simulaciones, lo anterior con el fin de tener valores en todo el soporte de S , desde 1 hasta 36.

El siguiente código muestra la densidad de S estimada. A diferencia de lo que podríamos intuir, no es simétrica.

```
#####
set.seed(1)
num_simulaciones <- 500000
poisson_compuesto <- function() sum( sample(x = 1:6, size = sample(x = 1:6, size = 1), replace = TRUE) )
simulaciones <- replicate(n = num_simulaciones, poisson_compuesto() )
barplot(prop.table(table(simulaciones)),main="Distribución de S")
#####
```



Proceso Poisson No Homogéneo

A menudo son más realistas los modelos basados en procesos de Poisson no Homogéneos, en los que la tasa de llegadas es una función del parámetro de tiempo, $\Lambda(t)$.

Formalmente esto significa que un proceso de Poisson no homogéneo es un proceso de contar que satisface:

- 1) $N_0 = 0$
- 2) Tiene incrementos independientes (en intervalos ajenos).
- 3) $\mathbb{P}(X_{t+h} - X_t \geq 1) = \lambda(t) \cdot h + o(h)$ para $t \geq 0, h \rightarrow 0$
- 4) $\mathbb{P}(X_{t+h} - X_t \geq 2) = o(h)$ para $t \geq 0, h \rightarrow 0$

Notemos que no necesariamente tiene incrementos estacionarios, además $N_t \sim \text{Poisson}(\Lambda(t))$ donde

$$\Lambda(t) = \int_0^t \lambda(x) dx \quad t \geq 0$$

Más generalmente, si consideramos dos tiempos $0 \leq s \leq t$ del proceso, entonces $N_t - N_s \sim \text{Poisson}(\Lambda(t) - \Lambda(s))$

$$\Lambda(t) - \Lambda(s) = \int_0^t \lambda(x) dx - \int_0^s \lambda(x) dx = \int_s^t \lambda(x) dx$$

Además, recordemos que una variable $X \sim \text{Poisson}(\lambda)$ tiene como función de densidad

$$f_X(x) = \mathbb{P}(X = x) = \exp(-\lambda) \cdot \frac{\lambda^x}{x!}$$

Observemos que si $\lambda(x) = \lambda$, se obtiene el proceso Poisson Homogéneo con tasa constante en todo momento.

Ejercicio 1

Consideremos a una farmacia que abre las 24 horas del día, además supongamos que la tasa con la que llegan los clientes tiene un comportamiento Poisson no Homogéneo tal que $\lambda(x) = -\frac{x(x-24)}{10} = \frac{1}{10}(24x - x^2)$ $x \in [0, 24]$. Considere el parámetro t como la hora del día (entre 00:00 horas y 24:00 horas). Responda lo siguiente

- 1) ¿Cuál es la probabilidad de que lleguen exactamente 3 personas entre las 9:00 y 10:00 horas?
- 2) ¿Cuál es la probabilidad de que lleguen exactamente 3 personas entre las 14:00 y 15:00 horas?
- 3) ¿Cuál es la probabilidad de que lleguen exactamente 3 personas entre las 23:00 y 24:00 horas?
- 4) De los 3 eventos anteriores, ¿Cuál es el más probable?
- 5) ¿Cuál es la probabilidad de que lleguen a lo más 20 personas antes de las 6:00 horas?
- 6) ¿Cuál es la probabilidad de que lleguen al menos 100 personas entre las 9:00 y 18:00 horas?
- 7) ¿Cuál es la cantidad de arribos esperados en un día completo (entre 00:00 horas y 24:00 horas)?

Solución

Lo primero que necesitamos es determinar la tasa de arriba, para la cual usaremos $\Lambda(b) - \Lambda(a)$.

Lo calcularemos de manera general, esto nos ayudará a determinar la tasa de arribo de cualquier intervalo.

$$\begin{aligned} \Lambda(b) - \Lambda(a) &= \int_a^b \lambda(x) dx = \frac{1}{10} \int_a^b (24x - x^2) dx = \frac{1}{10} \left(12x^2 - \frac{x^3}{3} \right) \Big|_{x=a}^{x=b} \\ &= \frac{1}{30} (36x^2 - x^3) \Big|_{x=a}^{x=b} = \frac{1}{30} \cdot (36b^3 - 36a^2 + a^3 - b^3) \end{aligned}$$

Finalmente tenemos que $\Lambda(b) - \Lambda(a) = \frac{1}{30} \cdot (36b^3 - 36a^2 + a^3 - b^3)$

Evaluemos lo anterior para determinar la tasa de arriba en cada intervalo de los incisos.

$$\Lambda(10) - \Lambda(9) = \frac{413}{30} \approx 13.767 \quad \Lambda(15) - \Lambda(14) = \frac{413}{30} \approx 13.767 \quad \Lambda(24) - \Lambda(23) = \frac{7}{6} \approx 1.167$$

Las variables aleatorias asociadas a los 3 primeros incisos son los siguientes

$$X_a \sim \text{Poisson}(413/30) \quad X_b \sim \text{Poisson}(413/30) \quad X_c \sim \text{Poisson}(7/6)$$

Por lo que nos queda que

$$\mathbb{P}(X_a = 3) = \exp(-413/30) \cdot \frac{(413/30)^3}{3!} = 0.0004566125$$

Notemos que la tasa con la que llegan las personas entre las 9:00 y 10:00 es la misma que para las 14:00 y 15:00

$$\mathbb{P}(X_b = 3) = \exp(-413/30) \cdot \frac{(413/30)^3}{3!} = 0.0004566125$$

También tenemos que

$$\mathbb{P}(X_c = 3) = \exp(-7/6) \cdot \frac{(7/6)^3}{3!} = 0.08241613$$

De los 3 eventos anteriores, es más probable que lleguen exactamente 3 personas de 23:00 a 24:00.

Para los ejercicios siguientes, tenemos que

$$\Lambda(6) - \Lambda(0) = \frac{155}{6} = 25.833 \quad \Lambda(18) - \Lambda(9) = \frac{243}{2} = 121.5 \quad \Lambda(24) - \Lambda(0) = \frac{1152}{4} = 230.5$$

Las variables aleatorias asociadas a los 3 últimos incisos son los siguientes

$$X_e \sim \text{Poisson}(155/6) \quad X_f \sim \text{Poisson}(243/2) \quad X_g \sim \text{Poisson}(1152/4)$$

Lo que nos lleva a

$$\mathbb{P}(X_e \leq 20) = \sum_{k=0}^{20} \exp(-155/6) \cdot \frac{(155/6)^k}{k!} = 0.1457796$$

Notemos que la tasa con la que llegan las personas entre las 9:00 y 10:00 es la misma que para las 14:00 y 15:00

$$\mathbb{P}(X_f \geq 100) = \sum_{k=100}^{\infty} \exp(-243/2) \cdot \frac{(243/2)^k}{k!} = 0.979599$$

Por último, tenemos que

$$\mathbb{E}(X_f) = 1152/4 = 230.5$$

Ejercicio moral

El gerente desea que la tienda tenga un horario de 9 horas (sin interrumpir) como las demás farmacias.

¿En qué horario le conviene abrir y cerrar la farmacia de tal manera que maximice la cantidad de personas esperadas? ¿Cuánto valdría dicha esperanza?

A continuación se presenta un ejemplo de simulación del proceso Poisson no Homogéneo

```
#####
# EJEMPLO DE PROCESO POISSON NO HOMOGENEO
#####

# Parámetros Globales
set.seed(1)
lambda <- function(t) 1 * ( sin( t*2*pi/48 - pi/2 ) + 1 ) / 2
tiempos <- seq(from=0, to=24*3, by = 1)
plot(tiempos, lambda(tiempos), type="l", col="red", main="lambda(x)", xlab="x", ylab="Lambda(x)" )
plot(tiempos, cumsum(lambda(tiempos)), type="l", col="red", main="Lambda(t)", xlab="t", ylab="Tasa" )

# Se simulan los eventos en cada sub intervalo de tiempo
num_periodos <- length(tiempos)-1
num_simulaciones <- vector(mode = "numeric",length = num_periodos)
simulaciones_exp <- NULL
for( t in seq( num_periodos ) ){
  tasa_t <- integrate( f = lambda, lower = tiempos[t], upper = tiempos[t+1] )["value"]
  # Se simulan la cantidad de "eventos" en el total de tiempo
  num_simulaciones[t] <- rpois(n = 1, lambda = tasa_t )
  # Se simulan los tiempos "inter-arribo" de cada evento
  simulaciones_exp <- c(simulaciones_exp, rexp(n = num_simulaciones[t], rate = tasa_t ) + (t-1) )
}
simulaciones_exp <- sort( simulaciones_exp )

# Generamos los vectores que representen las coordenadas de los puntos
valores_x <- simulaciones_exp # Tiempos en los que se presentan los eventos
valores_y <- c( 0 , seq( sum(num_simulaciones) ) ) # Altura de la función
function_a_trozos <- stepfun(x = valores_x, y = valores_y, f = 0 )
# Gráfica del proceso
plot(function_a_trozos, main="Proceso Poisson No Homogéneo", verticals = FALSE,
      pch=19, col.points ="red", xlim=c(0, max(tiempos)), xlab = "Tiempo", ylab = "Eventos")
lines(tiempos, cumsum(lambda(tiempos)), type="l", col="purple", lwd=2)
#####
# FIN DEL ARCHIVO
#####
```