Variables aleatorias discretas (Poisson)

Probabilidad I

25 de noviembre de 2020.

Variable aleatoria Poisson

- Muchos experimentos consisten en observar los tiempos de ocurrencia de llegadas aleatorias. Por ejemplo:
  - El número de marcaciones de teléfono erróneas en un día.
  - El número de personas que entran en un establecimiento cada día
  - El número de personas que sobreviven a cierta edad.
  - El número de accidentes reportadas sobre una avenida en un día.
- La distribución Poisson se usa para modelar el número de arribos que ocurren en un periodo de tiempo **fijo**
- La distribución Poisson también es de utilidad para aproximar a la distribución binomial cuando la probabilidad de éxito es pequeña.

Definición(Variable aleatoria Poisson)

Una v.a. X se dice que tiene distribución Poisson con parámetro $\lambda > 0$, denotada como $X \sim Po(\lambda)$, si su función de masa de probabilidad está dada por la siguiente expresión:

$$f(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & \text{para } x = 0, 1, 2, \ldots \\ 0 & \text{en otro caso} \end{cases}$$

Momentos

Para una variable aleatoria X tal que $X \sim Po(\lambda)$ se tiene que:

$$\mathbb{E}(X) = \lambda$$
$$\mathrm{Var}(X) = \lambda$$
$$m_X(t) = e^{-\lambda(e^t - 1)}$$

Función de masa de probabilidad en R

Se puede encontrar la probabilidad acumulada para una v.a. con distribución Poisson usando la función `dpois()`, por ejemplo:

$\mathbb{P}(X = x)$ para $0 \leq x \leq 15$, $x \in \mathbb{N}$, donde $X \sim Po(4)$

```
round(dpois(0:15,4),3) #Redondeo con tres decimales

##  [1] 0.018 0.073 0.147 0.195 0.195 0.156 0.104 0.060 0.030 0.013
## [13] 0.001 0.000 0.000 0.000
```

$\mathbb{P}(X = 5)$ donde $X \sim Po(4)$

```
dpois(5,lambda=4)

## [1] 0.1562935
```

Ejemplo

El número de erratas en cada página de un libro tiene distribución Poisson con parámetro $\lambda = 4$. Calcule la probabilidad de que al menos haya dos errores en alguna página seleccionada.

Ejemplo

El número de erratas en cada página de un libro tiene distribución Poisson con parámetro $\lambda = 4$. Calcule la probabilidad de que al menos haya dos errores en alguna página seleccionada.

```r
1-(dpois(0,4)+dpois(1,4))

## [1] 0.9084218
```

Función de distribución acumulada en R

Se puede encontrar la probabilidad acumulada para una v.a. con distribución Poisson usando la función `ppois()`, por ejemplo:
$\mathbb{P}(X \leq x)$ para $0 \leq x \leq 15$ donde $X \sim Po(4)$

```
ppois(0:15,4)

##  [1] 0.01831564 0.09157819 0.23810331 0.43347012 0.62883694 0.785
##  [7] 0.88932602 0.94886638 0.97863657 0.99186776 0.99716023 0.999
## [13] 0.99972628 0.99992367 0.99998007 0.99999511
```

$\mathbb{P}(X \leq 6)$ donde $X \sim Po(4)$

```
ppois(6,4)

## [1] 0.889326
```

Ejemplo

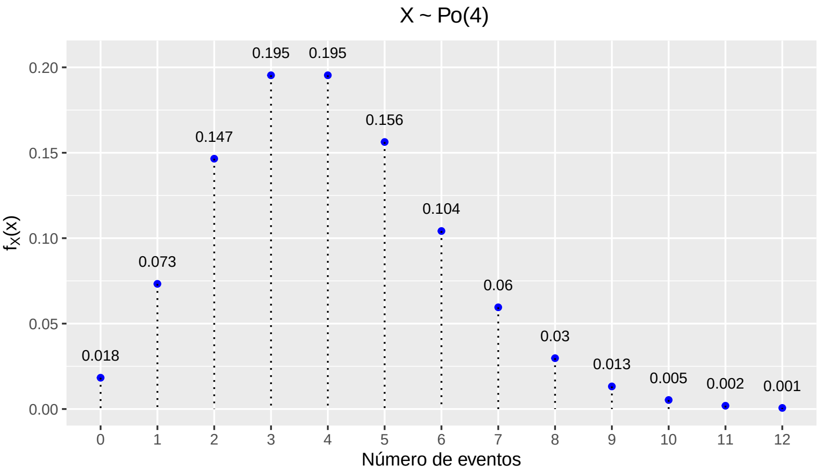
El número de erratas en cada página de un libro tiene distribución Poisson con parámetro $\lambda = 4$. Calcule la probabilidad de que al menos haya dos errores en alguna página seleccionada.

Ejemplo

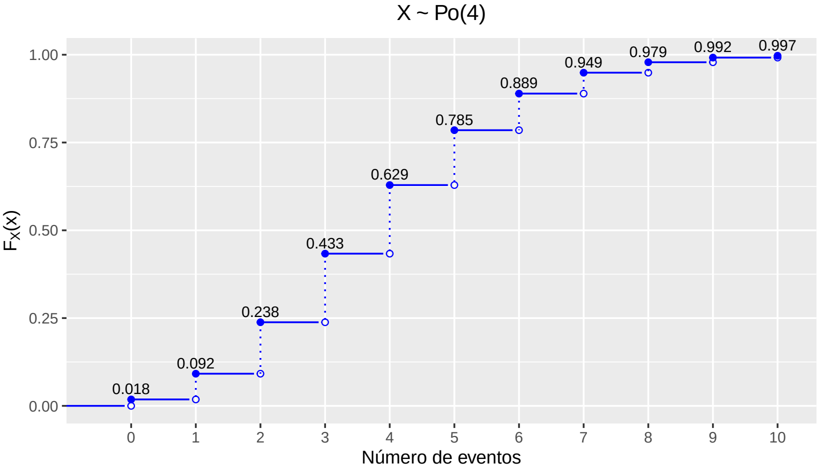
El número de erratas en cada página de un libro tiene distribución Poisson con parámetro $\lambda = 4$. Calcule la probabilidad de que al menos haya dos errores en alguna página seleccionada.

```
1-(ppois(1,4))

## [1] 0.9084218
```

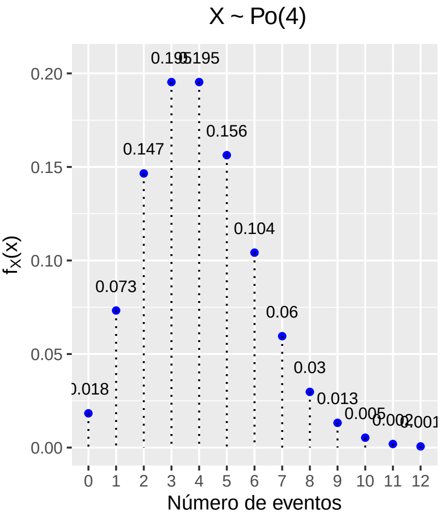
$f(x) = \mathbb{P}(X = x)$

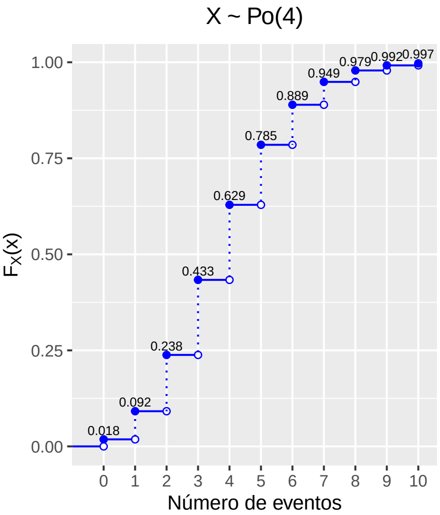


$$F(x) = \mathbb{P}(X \leq x)$$



Gráficas $F_X(x)$ y $f_X(x)$



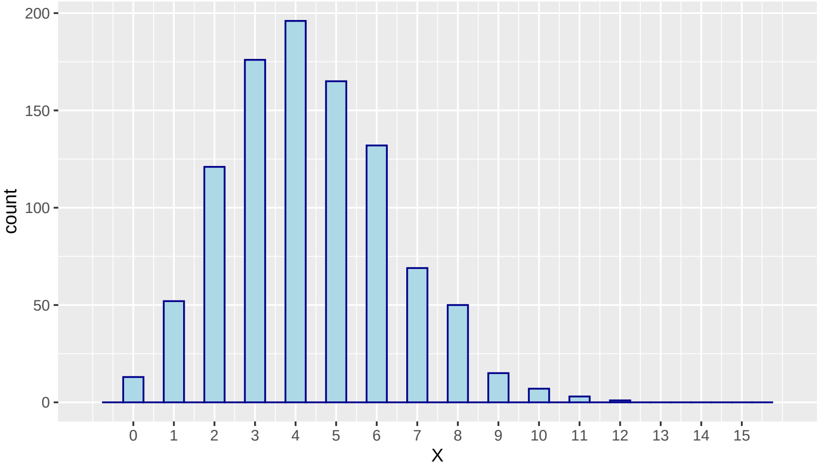


Simulación de valores de la distribución Poisson en R

Para simular valores de la distribución Poisson en R se puede hacer uso de la función `rbinom(valores a simular,`λ` )`, por ejemplo:

```
#Simulación de 150 valores de la distribución Poisson(lambda=10)
rpois(150,10)

##   [1]  6 12  8 10  7 13 10 13 16  8  7  9 17  4  8 12 13 12 12 14  9  9  8  6 15
##  [26]  6  7 12 14 12 14 11  8 11  6 12  9 10  9 10 11  9  7 10  9  6 10 13 11 11
##  [51]  8 14  9  9  6  8  8 14 13  9 12 22 12 12 11  9 13 13 12  8 12 13  7 10  9
##  [76]  9 11 10 11 18 13  9  9 15 11  7  9 11 11  7 13  5 12 11 15  8 10  9  7  8
## [101]  6 11 15  7  9 10 14 13  5 13 11 12  8 11 15  7 10  4 13  9 19  7  9  9 14
## [126] 13 12 11 13 10  7 13 14 12  6 10  8  8 11 10  7 12  7 10 10 11 15 12  7  9
```

```r
simula_pois<-rpois(1000,4.5)
ggplot(mapping = aes(simula_pois))+
  geom_histogram(binwidth = 0.5,color="darkblue",fill="lightblue")+
  scale_x_continuous(name = "X",breaks = seq(0, 15, 1),
                     limits=c(-1,16))
```

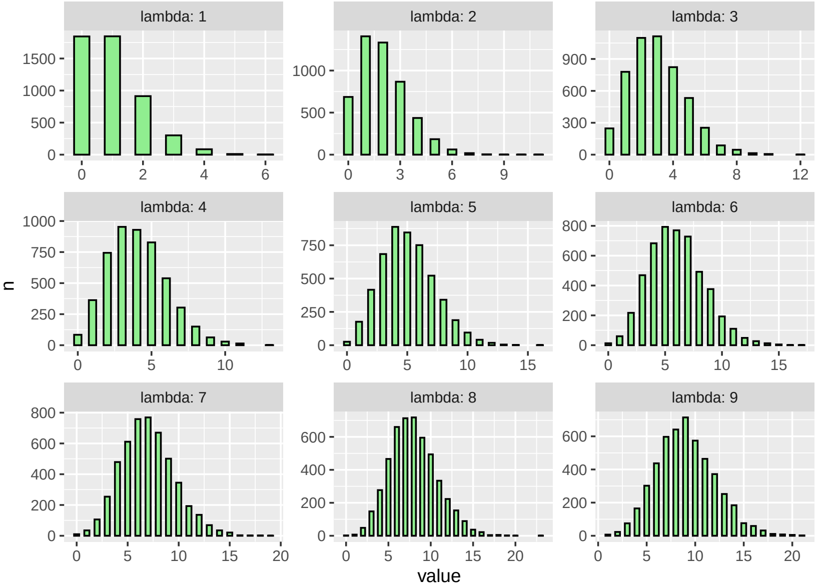


```
table(simula_pois)

## simula_pois
##   0   1   2   3   4   5   6   7   8   9  10  11  12
##  13  52 121 176 196 165 132  69  50  15   7   3   1

cumsum(simula_pois)[1000]/length(simula_pois)

## [1] 4.356
```

```r
library(kableExtra)
library(latex2exp)
kable(data.frame("x"=c(0:10),"y"=dpois(0:10,4),"z"=ppois(0:10,4)),
col.names = c("$x$","$\\mathbb{P}(X)=x$","$\\mathbb{P}(X)\\leq x$"),
escape = F,booktabs = T)%>%
  kable_styling(latex_options = c("striped",'hover'),font_size = 8)
```

x	$\mathbb{P}(X)=x$	$\mathbb{P}(X)\leq x$
0	0.0183156	0.0183156
1	0.0732626	0.0915782
2	0.1465251	0.2381033
3	0.1953668	0.4334701
4	0.1953668	0.6288369
5	0.1562935	0.7851304
6	0.1041956	0.8893260
7	0.0595404	0.9488664
8	0.0297702	0.9786366
9	0.0132312	0.9918678
10	0.0052925	0.9971602

Histograma para diferentes valores de λ



```
library(kableExtra)
library(dplyr)
library(latex2exp)
kable(data_frame("x"=c(0:10),"y"=dpois(0:10,4),"z"=ppois(0:10,4)),
      col.names = c("$x$","$\\mathbb{P}(X)=x$","$\\mathbb{P}(X)\\leq
  kable_styling(latex_options = c("striped"),font_size = 8, bootstra
```

x	$\mathbb{P}(X) = x$	$\mathbb{P}(X) \leq x$
0	0.0183156	0.0183156
1	0.0732626	0.0915782
2	0.1465251	0.2381033
3	0.1953668	0.4334701
4	0.1953668	0.6288369
5	0.1562935	0.7851304
6	0.1041956	0.8893260
7	0.0595404	0.9488664
8	0.0297702	0.9786366
9	0.0132312	0.9918678
10	0.0052925	0.9971602