Distribución de Poisson

Marina Valdora

Límite de la fpp de una v.a. Binomial :

Proposición: Sea $X \sim Bi(n, p)$ y supongamos que $n \to \infty$ y $p \to 0$, de manera que $np = \lambda$ (fijo), entonces:

$$p_X(k) \longrightarrow \frac{e^{-\lambda}\lambda^k}{k!} \quad \forall k \in N_o$$

Demostración:

$$\begin{aligned}
p_X(k) &= \binom{n}{k} p^k(1-p)^{n-k} = \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)\dots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\frac{\lambda^k}{k!} \\
&= \underbrace{\left[\frac{n}{n}\frac{n-1}{n}\cdot\ldots\cdot\frac{n-k+1}{n}\right]}_{\substack{\downarrow \\ 1}}\underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\substack{\downarrow \\ e^{-\lambda}}}\underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{\substack{\downarrow \\ 1}}\frac{\lambda^k}{k!}
\end{aligned}$$

por lo tanto,

$$p_X(k) \longrightarrow \frac{e^{-\lambda}\lambda^k}{k!},$$

como queríamos demostrar.

Definición

Una v.a. cuya función de probabilidad puntual es la obtenida en la proposición anterior, se dice que tiene distribución de Poisson de parámetro $\lambda (\lambda > 0)$, y se nota $X \sim P(\lambda)$ si su función de probabilidad puntual está dada por:

$$p_x(k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad \forall k \in N_o.$$

Verifiquemos que es una densidad

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1$$

ya que $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ es el desarrollo en serie de Taylor de e^{λ}.

Propiedades de las variables Poisson

Si $X \sim \mathcal{P}(\lambda)$, entonces

1) $E(X) = \lambda$
2) $V(X) = \lambda$

Dem:

1)

$$E(X) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{(k-1)!}$$

$$= \lambda \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)} = \lambda \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!} = \lambda.$$

2) Ejercicio optativo

Una consecuencia del teorema:

- La fpp binomial puede ser aproximada por la fpp de Poisson cuando n es grande y p es pequeño
- Algunos autores sugieren que, para que la aproximación sea buena debe ser $n \geq 100$, $p \leq 0.01$ y $np \leq 20$.

Ejemplo

```
probs_binomiales <- dbinom(0:9, 100, 0.01)
probs_poisson <- exp(-(100*0.01)) *(100*0.01)^(0:9)/factor:
tabla <- round(cbind(probs_binomiales,probs_poisson),4)
colnames(tabla) <- c("binomial","poisson")
tabla
```

```
##       binomial poisson
##  [1,]   0.3660  0.3679
##  [2,]   0.3697  0.3679
##  [3,]   0.1849  0.1839
##  [4,]   0.0610  0.0613
##  [5,]   0.0149  0.0153
##  [6,]   0.0029  0.0031
##  [7,]   0.0005  0.0005
##  [8,]   0.0001  0.0001
##  [9,]   0.0000  0.0000
## [10,]   0.0000  0.0000
```

Proceso de Poisson

Consideremos la variable aleatoria

$X_t =$ cantidad de ocurrencias de un evento en el intervalo de tiempo $[0, t]$

Por ejemplo,

$X_1 =$ cantidad de llamadas que recibe una central en la primera hora de trabajo.

$X_8 =$ cantidad de llamadas que recibe una central en las 8 horas de trabajo.

Proceso de Poisson

Supongamos que

1) La probabilidad de que ocurra un evento en un intervalo de tiempo es aproximadamente proporcional a la longitud del intervalo. Más precisamente, supongamos que existe un número real $\lambda > 0$ tal que

$$P(\text{ocurra un evento en } [t, t+h]) \approx \lambda h$$

2) La probabilidad de que ocurra más de un evento en un intervalo de tiempo muy pequeño es despreciable con respecto a la probabilidad de que ocurra uno.
3) La ocurrencia de un evento en un intervalo de tiempo es independiente de lo que ocurre en otro intervalo disjunto.

Proceso de Poisson

Teorema Sean $(X_t)_{t>0}$ variables aleatorias con rango igual a $\mathbb{N}_0$ definidas como

$X_t =$ cantidad de ocurrencias de un evento en el intervalo de tiempo $[0, t]$

Si se satisfacen las condiciones 1), 2) y 3), entonces

$$X_t \sim P(\lambda t)$$

Proceso de Poisson

Idea de la demostración

¿Qué distribución tiene X_t?

Imaginemos que dividimos el intervalo $(0, t)$ en un número muy grande de pequeños subintervalos de longitud $1/n$ y sea

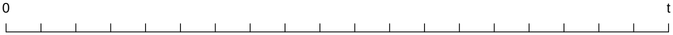
X_t= número de eventos que ocurren en el intervalo $(0, t)$.

Podemos pensar que hay una cantidad muy grande de ensayos Bernoulli siendo éxito la ocurrencia de un evento en cada uno de los subintervalos y

$p = P$(éxito)= probabilidad de que ocurra un evento en un intervalo de longitud $1/n$.

X_t es aproximadamente $B(n, p)$.

Si $n \to \infty$, $p \to 0$ y $np = \lambda t$ se mantiene constante, entonces $X_t \sim \mathcal{P}(\lambda t)$.



Proceso de Poisson

¿Cómo se interpreta la cantidad λ?

Puede interpretarse como la tasa media a la cual ocurren los eventos en la unidad de tiempo. Se la suele llamar tasa media de ocurrencia o intensidad del Proceso de Poisson.

Ejemplos de procesos de Poisson

1) Número de mensajes de correo electrónico que llegan a una casilla de correos.
2) Número de partículas emitidas por una sustancia radioactiva.
3) Número de accidentes que ocurren en un cruce de ruta.
4) Cantidad de meteoritos que caen en la luna
5) Cantidad de clientes que entran a un comercio
6) Cantidad de clicks en una página web
7) Número de larvas de cierto insecto en una laguna. En este caso t no es una unidad de tiempo, sino de volumen.

$$X_t = \text{ número de larvas en } t \text{ cm}^3.$$

Ejemplo: Meteoritos

Todos los años, entre septiembre y noviembre, la Tierra pasa a través de una zona donde se encuentran restos del comete Encke. Las partículas asociadas a este cometa golpean la atmósfera terrestre a aproximadamente 100 km/h y se incendian creando la lluvia de meteoritos de las Tauridas (https://blogs.nasa.gov/Watch_the_Skies/tag/taurid-meteor/).

La lluvia de meteoritos de las Tauridas tiene una tasa de 5 meteoritos por hora. Supongamos que una noche despejada y sin luna, todos los meteoritos que golpean la atmósfera son visibles desde un cierto lugar de la Tierra. ¿Cuál es la probabilidad de que un observador, vea al menos dos meteoritos en 1/2 hora?

Ejemplo: Meteoritos (Resolución)

Supongamos que el número de meteoritos es un proceso de Poisson.

Sea X_t el número de meteoritos observados en t horas. Entonces

$X_{1/2}$ es el número de meteoritos observados en $1/2$ hora.

$X_{1/2} \sim \mathcal{P}(5 \cdot 1/2)$, es decir $X_{1/2} \sim \mathcal{P}(5/2)$.

$$\begin{aligned} P(X_{1/2} \geq 2) &= 1 - P(X_{1/2} = 0) - P(X_{1/2} = 1) \\ &= 1 - \frac{e^{-5/2}(1/2)^0}{0!} - \frac{e^{-5/2}(5/2)^1}{1!} \\ &= 1 - e^{-2.5} - e^{-2.5}0.5 = 0.7127 \end{aligned}$$

Ejemplo: Meteoritos (Resolución en R)

$P(X_{1/2} \geq 2) =$

```r
1-ppois(1,2.5)
```

```
## [1] 0.7127025
```

```r
1-dpois(0,2.5)-dpois(1,2.5)
```

```
## [1] 0.7127025
```

Ejemplo: Errores de tipeo

Supongamos que, una persona comete un promedio de 3 errores de tipeo cada 10 páginas. ¿Cuál es la probabilidad de que tenga al menos un error de tipeo en un documento de 15 páginas?

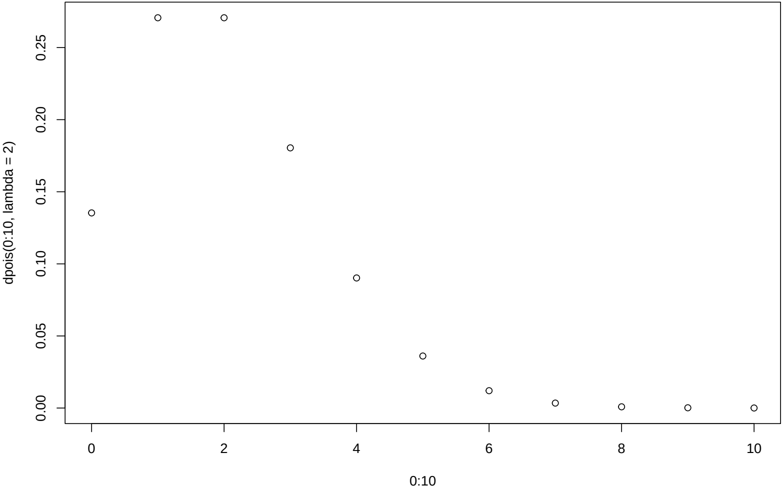
Poisson en R

La fpp

```r
plot(0:10, dpois(0:10, lambda = 2))
```

Poisson en R

La fpp



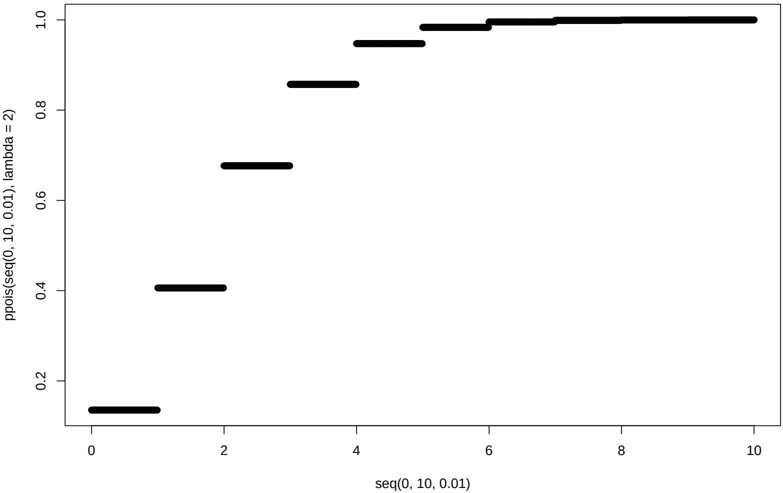
Poisson en R

La fda

```
plot(seq(0,10,0.01), ppois(seq(0,10,0.01), lambda = 2))
```

Poisson en R

La fda



Poisson en R

Simulación

```
rpois(10, lambda = 2)
```

```
##  [1] 1 0 1 2 3 5 5 3 4 3
```

Poisson en R

Calculo aproximado de probabilidades mediante simulaciones

```
realizaciones <- rpois(10000, lambda = 2)
c(mean(realizaciones==0),
mean(realizaciones==1),
mean(realizaciones==2),
mean(realizaciones==3),
mean(realizaciones==4),
mean(realizaciones==5))
```

```
## [1] 0.1390 0.2653 0.2695 0.1823 0.0893 0.0386
```

Comparo con las exactas

```
dpois(0:5, lambda = 2)
```

```
## [1] 0.13533528 0.27067057 0.27067057 0.18044704 0.09022
```

Poisson en R

Cálculo aproximado de la esperanza mediante simulación

```
mean(rpois(10000, lambda = 2))
```

```
## [1] 1.9901
```