Introducción a distribuciones de probabilidad

Ramon Ceballos

7/2/2021

Distribución Binomial

1. Conceptos teóricos y matemáticos

Si X es variable aleatoria que mide el "número de éxitos" que hay en realizar n ensayos de Bernoulli independientes entre sí, diremos que X se distribuye como una Binomial con parámetros n y p.

$$X \sim \mathbf{B}(n, p)$$

En esta expresión, p es la probabilidad de éxito y $q = 1 - p$ es la probabilidad de fracaso.

Vamos a definir algunas de las propiedades que tendrá una distribución Binomial.

- El **dominio** de X será $D_X = \{0, 1, 2, \ldots, n\}$. Ya que el nº de ensayos a realizar es n.
- La **función de probabilidad** vendrá dada por:

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

En esta expresión k es un nº dentro de D_x. El parámetro $\binom{n}{k}$ es el coeficiente binomial o nº de combinaciones y se calcula:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- La **función de distribución** vendrá dada por:

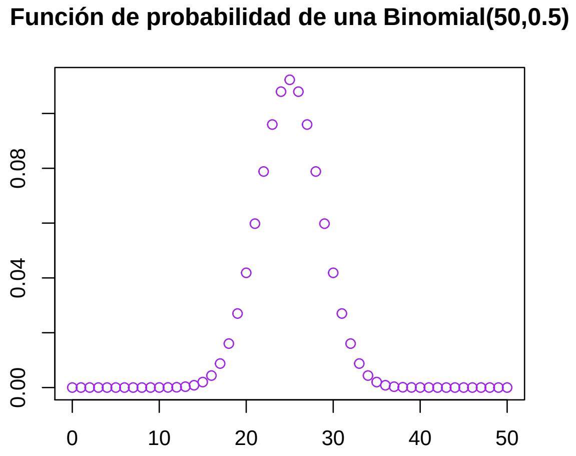
$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^{x} f(k) & \text{si } 0 \leq x < n \\ 1 & \text{si } x \geq n \end{cases}$$

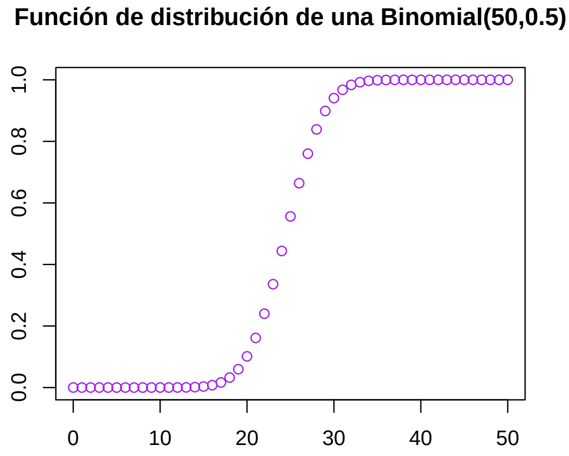
Si x es menor que 0, la distribución será cero. Si está entre 0 (incluido) y n (sin incluir), da lugar a un sumatorio de los elementos del dominio hasta llegar a n. Si x es mayor que n, el resultafo es 1.

- **Esperanza** $E(X) = np$.
- **Varianza** $Var(X) = npq$.

Atención. Fijaos que la distribución de Bernoulli es un caso particular de la Binomial. Basta tomar $n = 1$ y tendremos que $X \sim \text{Be}(p)$ y $X \sim \text{B}(1, p)$ son equivalentes.

```
#Se lleva a cabo el experimento 50 veces (n) con probabilidad de 0.5 (p)
par(mfrow = c(1,2))
plot(0:50,dbinom(0:50,50,0.5),col = "purple", xlab = "", ylab = "", main = "Función de probabilidad de u
plot(0:50, pbinom(0:50,50,0.5),col = "purple", xlab = "", ylab = "", main = "Función de distribución de
```






```
par(mfrow= c(1,1))
```

2. Distribución Binomial en R y Python

El código de la distribución Binomial:

- En **R** tenemos las funciones del paquete básico "stats": **dbinom(x, size, prob)**, **pbinom(q,size, prob)**, **qbinom(p, size, prob)**, **rbinom(n, size, prob)** donde **prob** es la probabilidad de éxito y **size** el número de ensayos del experimento.
- En **Python** tenemos las funciones del paquete **scipy.stats.binom**: **pmf(k,n,p)**, **cdf(k,n,p)**, **ppf(q,n,p)**, **rvs(n, p, size)** donde **p** es la probabilidad de éxito y **n** el número de ensayos del experimento.

3. Ejemplos en código para la distribución Binomial (R y Python)

Sea nuestra Binomial la siguiente: $X = B(n = 30, p = 0.6)$, denotaremos su función de probabilidad y de distribución.

Función de probabilidad:

$$f(k) = \binom{30}{k} 0.6^k (1 - 0.6)^{30-k}$$

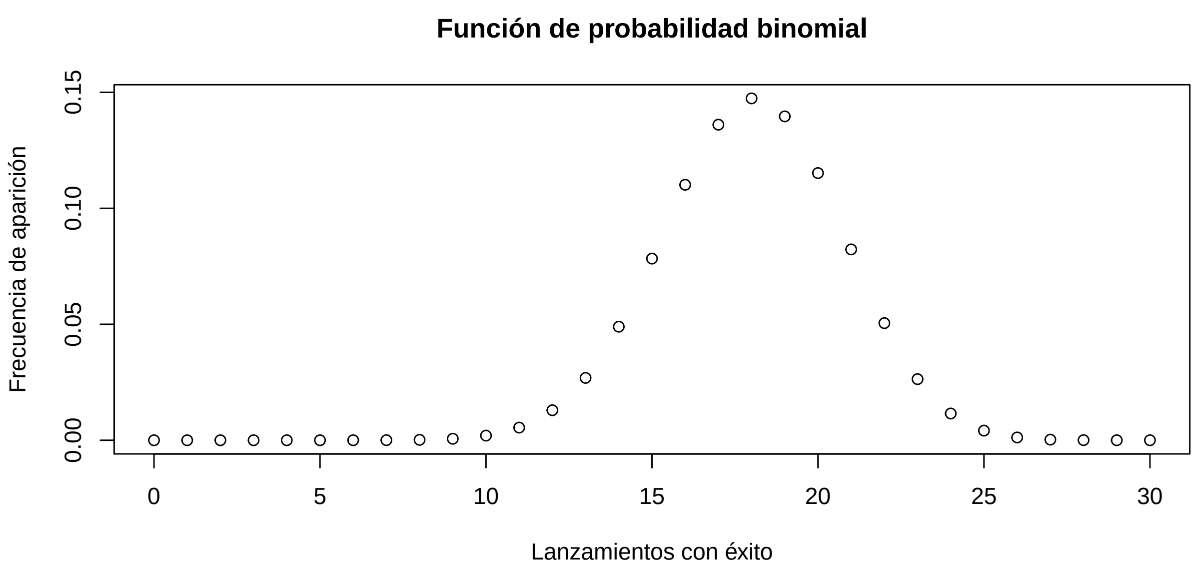
Función de distribución:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^{x} f(k) & \text{si } 0 \leq x < 30 \\ 1 & \text{si } x \geq 30 \end{cases}$$

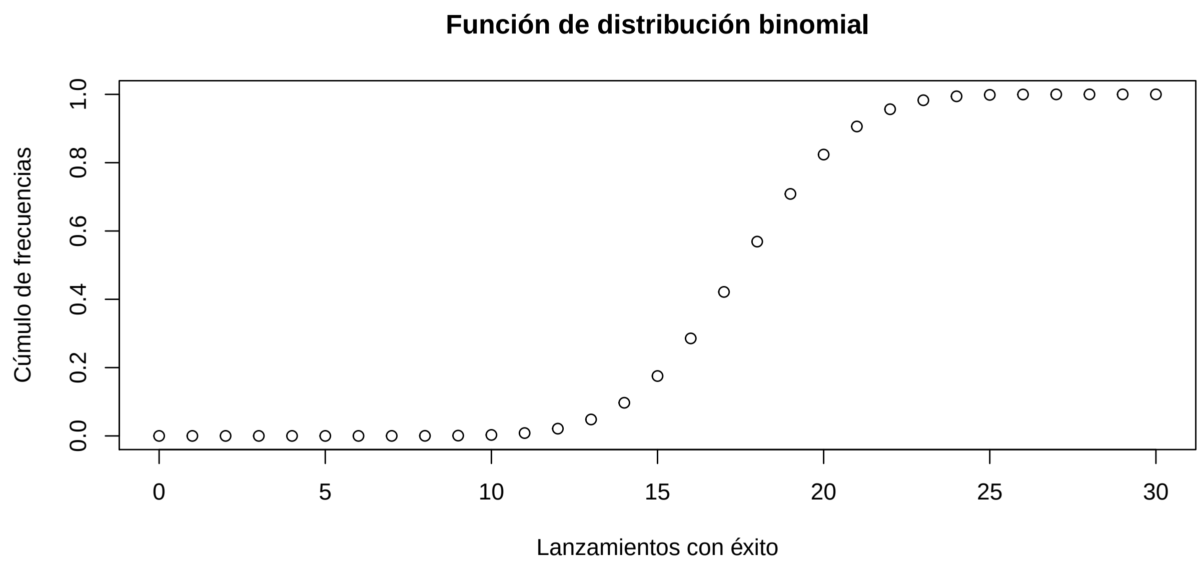
Ejemplo en R Utilizamos las funciones anteriores.

```
#Defino el nº de ensayos y la probabilidad
n = 30
p = 0.6

#Función de probabilidad del dominio, con muesta "n" y probabilidad "p"
plot(0:n, dbinom(0:n, size = n, prob = p),
     title("Función de probabilidad binomial"),
     ylab = "Frecuencia de aparición",
     xlab = "Lanzamientos con éxito")
```




```
#Función de distribución del dominio, con muesta "n" y probabilidad "p"
plot(0:n, pbinom(0:n, size = n, prob = p),
     title("Función de distribución binomial"),
     ylab = "Cúmulo de frecuencias",
     xlab = "Lanzamientos con éxito")
```



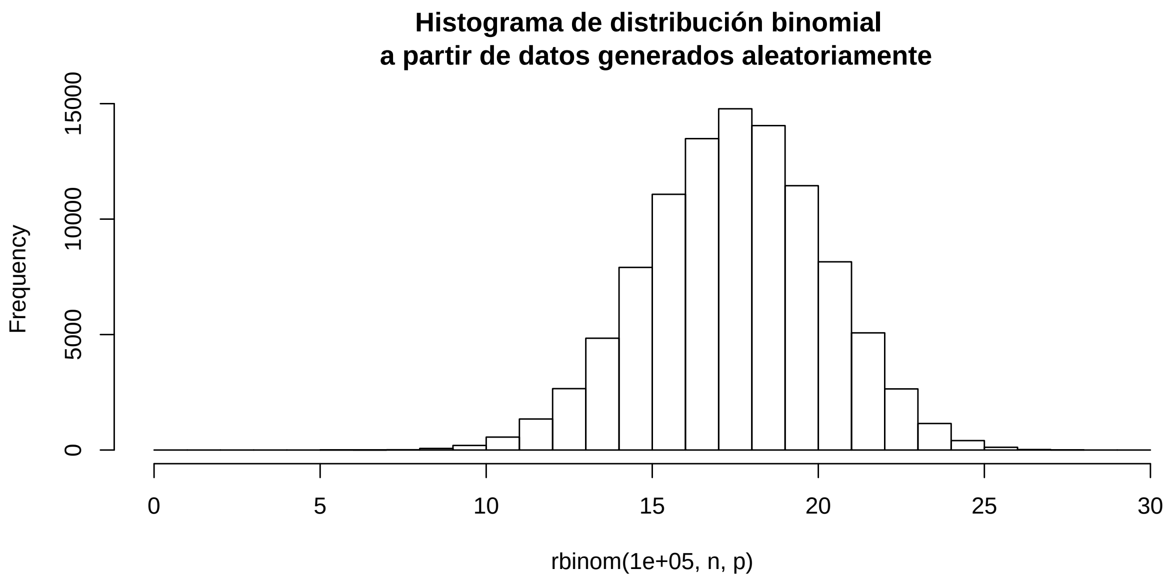

```r
#Sacamos la mediana
qbinom(0.5, n, p)
```

```
## [1] 18
```

```r
#Sacamos el cuantil 0.25
qbinom(0.25, n, p)
```

```
## [1] 16
```

```r
#Generamos datos aleatorios y representamos un histograma
hist(rbinom(100000, n, p),
     breaks = 0:n,
     main = "Histograma de distribución binomial \n a partir de datos generados aleatoriamente")
```



Ejemplo en Python Vete al script de Python del presente tema 6.