Variables discretas famosas

Distribución binomial

Ensayos de Bernoulli

Definición: Un ensayo de Bernoulli es un experimento que tiene dos posibles resultados que llamaremos 'éxito' y 'fracaso'. La probabilidad de éxito la notaremos p y por lo tanto la probabilidad de fracaso será $1 - p$

Ejemplos de ensayos de Bernoulli

1) Tirar un dado y considerar como éxito al evento 'sale 1'
2) Tirar un dardo en un blanco de radio 8cm y considerar éxito al evento 'el dardo cae a menos de 5cm del centro'
3) Elegir un votante de la ciudad de Buenos Aires al azar y preguntarle si está a favor de cierta medida de gobierno. Consideramos 'éxito' que el votante esté a favor y 'fracaso' que no. p en este caso es la verdadera proporción de votantes que están a favor de la medida. Suele ser desconocida pero de interés!

Distribución de Bernoulli

Consideremos un experimento de Bernoulli y sea $X =$ cantidad de éxitos

Hallar R_X, p_X, $E(X)$ y $V(X)$.

Proceso de Bernoulli

Un proceso de Bernoulli es una sucesión (finita o infinita) de ensayos de Bernoulli idénticos e independientes.

Ejemplos:

1) Tirar 8 veces un dado, considerarndo como éxito al evento 'sale 1'.

2) Tirar 8 veces a un blanco de radio 8cm, considerando como éxito al evento 'el dardo cae a menos de 5cm del centro'. ¿Es un ensayo de Bernoulli?

3) Sacar 8 bolitas con reposición de una urna que contiene 4 rojas y 5 verdes, considerando como éxito al evento 'la bolilla extraída es verde'. ¿Y si es sin reposición?

4) Elegir 80 personas al azar de la ciudad de Buenos Aires y preguntarles si están a favor de cierta medida de gobierno, considerando como éxito al evento 'la persona está de acuerdo con la medida'. ¿Es un proceso de Bernoulli? No, pero se parece mucho.

Definición de variable aleatoria binomial

Consideremos una sucesión de n ensayos de Bernoulli con probabilidad de éxito p. La variable aleatoria

$$X = \text{ número de éxitos en los } n \text{ ensayos}$$

se llama binomial y se nota $X \sim B(n, p)$.

La función de probabilidad puntual de una v.a. binomial

Teorema: Si $X \sim B(n, p)$ entonces su f.p.p es

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Demostración de la fpp binomial para n=3 (idea)

Calculemos la fpp de una v.a. binomial, con un ejemplo

Ensayo de Bernoulli: tirar dado y considerar como éxito que salga 1.

Experimento: Repetir 3 veces el ensayo de bernoulli

X = cantidad de unos (o de éxitos)

$$\Omega = \{(EFF), (EFE), (EEF), (EEE), (FEE), (FEF), (FFE), (FFF)\}$$

$$\#\Omega = 2^3 = 8$$

$$\Omega = \{(EFF), (EFE), (EEF), (EEE), (FEE), (FEF), (FFE), (FFF)\}$$

¿Es equiprobable?

$P(\{(EEE)\}) = 1/6 \cdot 1/6 \cdot 1/6, \quad P(\{(EEF)\}) = 1/6 \cdot 1/6 \cdot 5/6$

$P(\{(EFF)\}) = 1/6 \cdot 5/6 \cdot 5/6, \quad P(\{(FFF)\}) = 5/6 \cdot 5/6 \cdot 5/6$

No es equiprobable

$R_X = \{0, 1, 2, 3\}$

$p_X(0) = P(X = 0) = 5/6 \cdot 5/6 \cdot 5/6$

$p_X(1) = P(X = 1) = 3 \cdot 1/6 \cdot 5/6 \cdot 5/6$

El 3 aparece porque son todas las posibles posiciones del 'éxito' Puede ser que el éxito aparezca en el 1er ensayo, en el 2do en el 3ro.

$p_X(2) = P(X = 2) = 3 \cdot 1/6 \cdot 1/6 \cdot 5/6$

El 3 aparece porque son todas las posibles posiciones de los 2 'éxitos' Puede ser que los éxitos aparezca en el 1er y 2do ensayo, en el 2do 3er ensayo o en el 1er y 3er ensayo.

$p_X(3) = P(X = 3) = 3 \cdot 1/6 \cdot 1/6 \cdot 1/6$

Observar:

$$3 = \binom{3}{2}$$

Demostración de la fpp binomial para n=4 (idea)

$X =$ cantidad de unos al tirar 4 veces un dado

$R_X = \{0, 1, 2, 3, 4\}$

$p_X(0) = (5/6)^4$

$p_X(4) = (1/6)^4$

$p_X(1) = 4 \cdot 1/6 \cdot (5/6)^3$

$p_X(3) = 4 \cdot 5/6 \cdot (1/6)^3$

$p_X(2) =$

Demostración de la fpp binomial para n=4 (idea)

Una posibilidad es que salga $(EEFF)$,

$P(\{(EEFF)\}) = (1/6)^2 \cdot (5/6)^2$

otra es $(EFEF)$,

$P(\{(EFEF)\}) = (1/6)^2 \cdot (5/6)^2$

$p_X(2) = (1/6)^2 \cdot (5/6)^2 \cdot ($ cantidad de veces que hay 2 éxitos$)$

¿cuántas son? ¿Cuantas formas hay de distribuir dos éxitos en 4 ensayos? ¿Cuántas formas hay de elegir 2 lugares en un vector de longitud 4? ¿Cuántas formas hay de elegir 2 números entre 1 y 4?

La cantidad de formas de distibuir 2 éxitos en 4 ensayos es:

$$\binom{4}{2} = \frac{4!}{2!2!} = 6$$

```
choose(4,2)
```

```
## [1] 6
```

Por lo tanto,

$$p_X(2) = P(X = 2) = \binom{4}{2} \cdot (1/6)^2 \cdot (5/6)^2 = 6 \cdot (1/6)^2 \cdot (5/6)^2$$

Ejemplo: Dado

La fpp y la fda de la binomial están implementadas en R. *dbinom* es la fpp y *pbinom* es la pda.

Ejemplo: Se tira un dado 8 veces.

a) Calcular la probabilidad de que salgan exactamente 3 unos.
b) Calcular la probabilidad de que salgan entre 3 y 8 unos.

Ejemplo: Dado

Resolución:

$X =$ Cantidad de unos al tirar el dado 8 veces

a) $P(X = 3) = \binom{8}{3}(1/6)^3(5/6)^7$

```
choose(8,3)*(1/6)^3*(5/6)^7
```

```
## [1] 0.0723545
```

b) $P(3 \leq X \leq 8)$

$= \binom{8}{3}(1/6)^3(5/6)^5 + \binom{8}{4}(1/6)^4(5/6)^4 + \binom{8}{5}(1/6)^5(5/6)^3 +$
$\binom{8}{6}(1/6)^6(5/6)^2 + \binom{8}{7}(1/6)^7(5/6)^1 + \binom{8}{8}(1/6)^8$

```
choose(8,3:8)*(1/6)^(3:8)*(5/6)^(8-(3:8))
```

```
## [1] 1.041905e-01 2.604762e-02 4.167619e-03 4.167619e-04
## [6] 5.953742e-07
```

```
sum(choose(8,3:8)*(1/6)^(3:8)*(5/6)^(8-(3:8)))
```

```
## [1] 0.1348469
```

Ejemplo: Dado. Reslolución en R

a)

```
dbinom(3, size=8, prob = 1/6)
```

```
## [1] 0.1041905
```

b)

```
sum(dbinom(3:8, size=8, prob = 1/6))
```

```
## [1] 0.1348469
```

que es lo mismo que

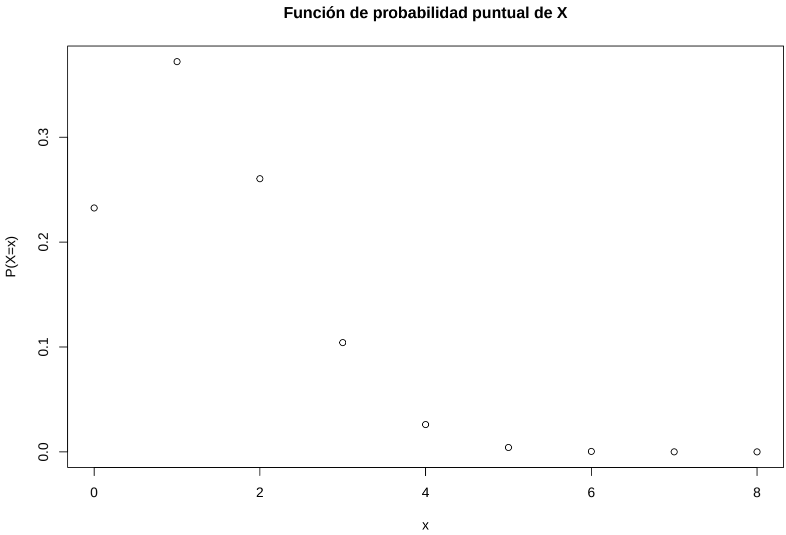
```
pbinom(8, size = 8, prob = 1/6)-pbinom(2, size = 8, prob =
```

```
## [1] 0.1348469
```

Ejemplo: Dado.

c) Graficar la fpp de la la v.a X

```
r_x <- 0:8
p_x <- dbinom(0:8, prob = 1/6, size = 8)
plot(r_x, p_x, main = "Función de probabilidad puntual de X
```


Función de probabilidad puntual de X
P(X=x)
0.0
0.1
0.2
0.3
0
2
4
6
8
x

Verifiquemos que la fpp hallada cumple las propiedades.

1) $p_X(x) \geq 0$
2) $\sum_{x=0}^{n} p_X(x) = 1$

Verificación de que la fpp de la binomial cumple las propiedades

1) es obvio
2) Para la demostración, recordar la fórmula del **binomio de Newton**:

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = (p+1-p)^n = 1$$

Función de distribución acumulada de una v.a. binomial

Recordar $F_X(x) = P(X \leq x)$

Si $X \sim B(n, p)$,

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sum_{k=0}^{[x]} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} & \text{si } 0 \leq x \leq n \\ 1 & \text{si } x > n \end{cases}$$

donde $[x]$ denota la parte entera de x.

En R: **pbinom(n, p)**

La esperanza una v.a. binomial

Teorema: Si $X \sim B(n,p)$ entonces $E(X) = np$.

Recordar

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

$$(a+b)^{n-1} = \sum_{i=0}^{n-1} \binom{n-1}{i} a^i b^{n-i-1}$$

En particular

$$\sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-i-1} = (p+1-p)^{n-1} = 1$$

La esperanza una v.a. binomial

Teorema: Si $X \sim B(n,p)$ entonces $E(X) = np$.

Demostración: $E(X) =$

$$= \sum_{k=0}^{n} k \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{j=0}^{n-1} \frac{n!}{j!(n-(j+1))!} p^{j+1} (1-p)^{n-(j+1)}$$

$$= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} = np(p + (1-p))^{n-1} = np$$

La varianza una v.a. binomial

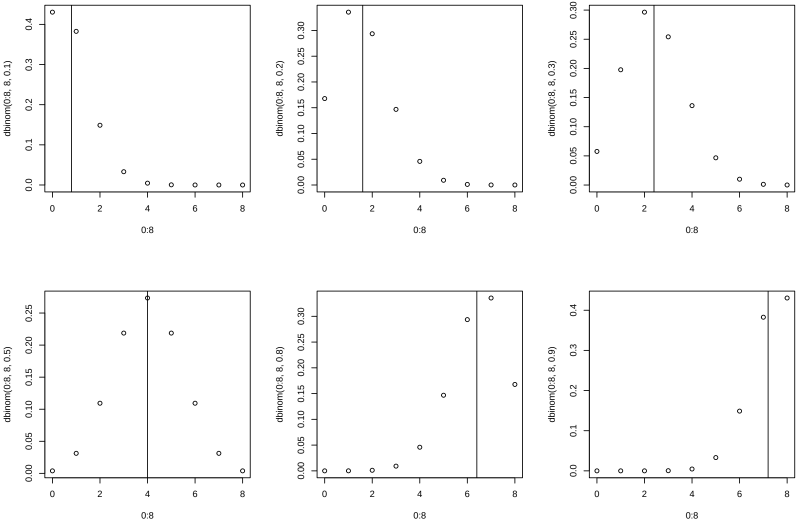
Teorema: Si $X \sim B(n, p)$ entonces $V(X) = np(1-p)$.

Demostración: Ver, por ejemplo, el apunte de Ana Bianco y Elena Martinez

Función de probabilidad puntual y esperanza de la binomial

```
par(mfrow=c(2,3))
plot(0:8, dbinom(0:8, 8, 0.1));abline(v=0.8)
plot(0:8, dbinom(0:8, 8, 0.2));abline(v=1.6)
plot(0:8, dbinom(0:8, 8, 0.3));abline(v=2.4)
plot(0:8, dbinom(0:8, 8, 0.5));abline(v=4)
plot(0:8, dbinom(0:8, 8, 0.8));abline(v=6.4)
plot(0:8, dbinom(0:8, 8, 0.9));abline(v=7.2)
```

Función de probabilidad puntual y esperanza de la binomial



Esperanza y varianza de la binomial

```
##   n   p E(X) V(X)
## 1 8 0.1  0.8 0.72
## 2 8 0.2  1.6 1.28
## 3 8 0.3  2.4 1.68
## 4 8 0.5  4.0 2.00
## 5 8 0.8  6.4 1.28
## 6 8 0.9  7.2 0.72
```

Resultados de experimentos simulados

```
rbinom(20, 8, 0.1)
```

```
##  [1] 0 0 0 1 0 0 0 0 0 1 0 3 0 0 1 1 0 0 0 1
```

```
rbinom(20, 8, 0.2)
```

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

```
rbinom(20, 8, 0.3)
```

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

```
rbinom(20, 8, 0.5)
```

```
##  [1] 3 3 6 4 5 3 4 4 2 3 5 3 4 5 2 6 5 4 5 3
```

```
rbinom(20, 8, 0.8)
```

```
##  [1] 5 7 4 7 7 6 6 7 6 6 8 7 7 5 7 7 5 8 8 6
```

```
rbinom(20, 8, 0.9)
```

```
##  [1] 8 7 8 7 7 6 8 8 7 8 7 7 4 8 6 7 6 6 8 8
```

Definición de variable aleatoria geométrica

Se repiten ensayos de Bernoulli independientes e idénticos hasta que ocurre el primer éxito. Sea $X =$ cantidad de repeticiones hasta el primer éxito

Se dice que X sigue una distrución geométrica y se nota
$X \sim Ge(p)$

Función de probabilidad puntual de una geométrica

$$R_X = \mathbb{N}$$

$$P(X = 1) = p$$

$$P(X = 2) = (1 - p)p$$

$$P(X = 3) = (1 - p)^2 p$$

$$P(X = 4) = (1 - p)^3 p$$

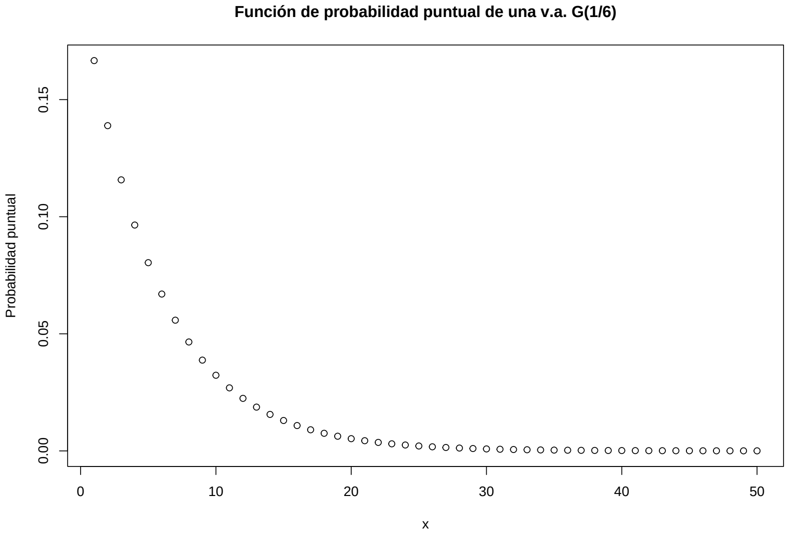
$$\ldots$$

$$P(X = k) = (1 - p)^{k-1} p$$

Gráfico de la fpp de la geométrica

```
equis <- 1:50
p <- 1/6
imequis1 <- p*(1-p)^(equis-1)
plot(equis,imequis1,xlab = "x",
     ylab="Probabilidad puntual",
     main = "Función de probabilidad puntual de una
     v.a. G(1/6)")
```

Gráfico de la fpp de la geométrica



Verifiquemos que es una fpp

1) $p_X(x) \geq 0$
2) $\sum_{x=1}^{\infty} p(1-p)^{x-1} = 1$

Demostración

$$\sum_{x=1}^{\infty} p(1-p)^{x-1} = \sum_{x=0}^{\infty} p(1-p)^{x} = p\frac{1}{1-(1-p)} = 1$$

Hallar la fda de una v.a. geométrica

Si $x \geq 0$ y x es entero,

$$F_X(x) = P(X \leq x) = \sum_{t=1}^{x} p(1-p)^{t-1} = \sum_{t=0}^{x-1} p(1-p)^{t}$$
$$= p\left(\frac{(1-(1-p))^x}{1-(1-p)}\right)$$

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1-(1-p)^{[x]} & \text{si } x \geq 1 \end{cases}$$

Esperanza y varianza de una variable aleatoria geométrica:

Si $X \sim \mathcal{G}(p)$ entonces

$$E(X) = \frac{1}{p} \text{ y } V(X) = \frac{1}{p^2}$$

—

Esperanza de una variable aleatoria geométrica:

Si $X \sim \mathcal{G}(p)$

$E(X) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=1}^{\infty} k(1-p)^{k-1}$

$= -p\sum_{k=1}^{\infty} \frac{\partial}{\partial p}(1-p)^k$

$= -p\frac{\partial}{\partial p}\left(\sum_{k=1}^{\infty}(1-p)^k\right) = -p\frac{\partial}{\partial p}\left(\frac{1}{1-(1-p)} - 1\right)$

$= -p\frac{\partial}{\partial p}\left(\frac{1}{p} - 1\right) = -p\left(-\frac{1}{p^2}\right) = \frac{1}{p}$

Propiedad de falta de memoria

Sea $X \sim \mathcal{G}(p)$ y sean n y m números naturales cualesquiera,

$$P(X > n + m | X > n) = P(X > m)$$

Demostración

$$P(X > n+m|X > n) = \frac{P(X > n + m \text{ y } X > n)}{P(X > n)} = \frac{P(X > n + m)}{P(X > n)}$$

$$\frac{1 - F_X(n + m)}{1 - F_X(n)} = \frac{1 - (1 - (1 - p)^{n+m})}{1 - (1 - (1 - p)^n)}$$

$$\frac{(1 - p)^{n+m}}{(1 - p)^n} = (1 - p)^m = 1 - F_X(m) = P(X > m)$$

Ejemplo: Dados

Se tira un dado hasta que salga un 6. Sabiendo que no salió un 6 en las primeras 5 tiradas, cuál es la probabilidad de sea necesario tirarlo al menos 3 veces más.

X=cantidad de tiros hasta que salga un 6.

La probabilidad pedida es

$$P(X > 2) = 1 - P(X = 1) - P(X = 2) = 1 - 1/6 - 1/6 \cdot (5/6) = 25/36.$$

Como ejercicio, calcularla usando la probabilidad condicional y verificar que da lo mismo.

Variables discretas famosas

Binomial negativa

Marina Valdora

Binomial negativa

Supongamos que se repite en forma independiente un ensayo de Bernoulli con probabilidad de Éxito p constante en todas las pruebas. Se define la v.a.

X : número de repeticiones hasta obtener el r-ésimo Éxito ($r \geq 1$).

Decimos que X tiene distribución binomial negativa

Notación: $X \sim BN(r, p)$.

Binomial negativa

$R_X = \{r, r+1, r+2, \ldots\}$

Sea k un número natural, $k \geq r$. Para que sean necesarias k repeticiones para obtener el primer éxito, el r-ésimo Éxito debe ocurrir en la repetición k y en las $(k-1)$ repeticiones previas debe haber exactamente $(r-1)$ Éxitos. Como las repeticiones son independientes la probabilidad de una configuración de ese tipo es $p^r(1-p)^{k-r}$, pero hay varias configuraciones de esta forma. ¿Cuántas? Tantas como formas de elegir entre las $(k-1)$ primeras repeticiones, aquellas donde ocurrirán los $(r-1)$ Éxitos, o sea $\left(\begin{array}{c} k-1 \\ r-1 \end{array} \right)$.

Por lo tanto la función de probabilidad puntual será:

$$P(X=k) = \left(\begin{array}{c} k-1 \\ r-1 \end{array} \right) p^r(1-p)^{k-r} \quad \forall k \in \{r, r+1, r+2, \ldots .\}$$