Probabilidad y Análisis de Datos

Daniel Fraiman

Maestría en Ciencia de Datos, Universidad de San Andrés

VARIABLES ALEATORIAS

Variables aleatorias

Muchas veces uno está interesado en una función de los resultados de un experimento ($\omega' s$), sin importar específicamente qué ω salió.

Ejemplo

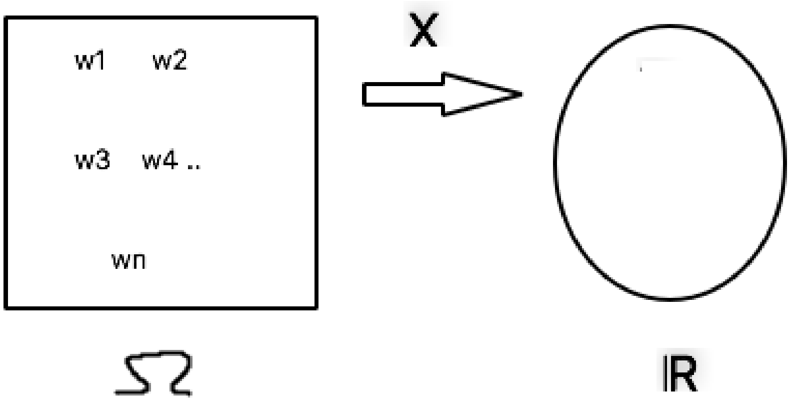
- Si tiramos dos dados podemos estar interesados en la suma de los 2 dados. Nos puede interesar saber si la suma dió 4, sin importar si salió el $\omega_3 = (1,3)$, $\omega_{13} = (3,1)$, ó el $\omega_8 = (2,2)$.
- Tiro 3 monedas y nos interesa el número de caras.

Variables aleatorias

Variables aleatorias

Una variable aleatoria, X, es una función que le asigna a cada uno de los resultados del experimento aleatorio un valor numérico (pero puedo no serlo).

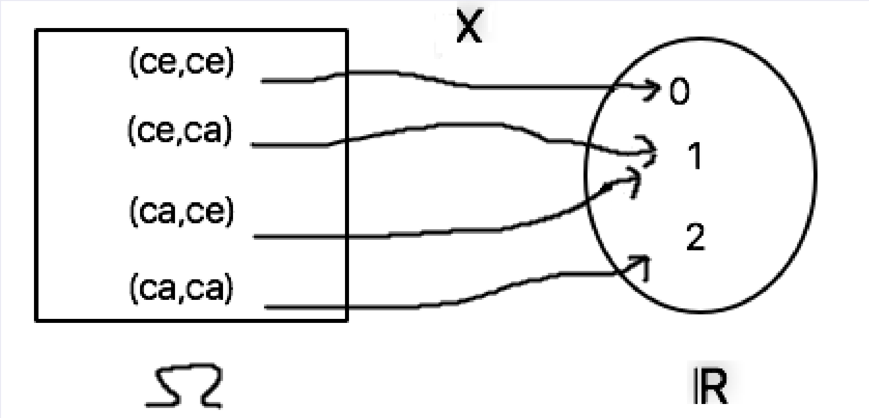
$$X : \Omega \to \mathbb{R}$$



Variables aleatorias

Exp.: Tiro dos monedas

X = cantidad de caras.



- ¿Por qué X es aleatorio?

Variables aleatorias

Una variable aleatoria (v.a.) está bien determinada al conocer:

- Los valores que toma la v.a.
- La probabilidad de cada uno de los valores.

Variables aleatorias

Existen

- variables aleatorias discretas.
- variables aleatorias continuas.
- variables aleatorias mixtas.

Si compro una bombita de luz, y apenas la conecto se rompe, tiene valor 0, es discreta, Si no se rompe, tiene un tiempo de duracion continuo

VARIABLES ALEATORIAS DISCRETAS

Variables aleatorias discretas

Variables aleatorias discretas:

Son v.a. que pueden tomar un conjunto finito de valores, o un conjunto infinito de valores pero numerable (ej. $\mathbb{N}$)

Ejemplo

$X = \{0, 2, 6\}$ con $\mathbb{P}(X = 0) = 0{,}1$, $\mathbb{P}(X = 2) = 0{,}5$ y $\mathbb{P}(X = 6) = 0{,}4$.

Variables aleatorias discretas

Ejemplo a partir de un exp.: Tiro una moneda dos veces.

$\Omega = \{(ca, ca), (ca, ce), (ce, ce), (ce, ca)\}$. Definimos la variable aleatoria: $X =$ número de caras.

¿Qué necesitamos saber de X para realmente comprenderla?

- ¿Qué valores toma?
  - $X = \{0, 1, 2\}$
- ¿Probabilidades puntuales $(\mathbb{P}(X = 0), \mathbb{P}(X = 1), \mathbb{P}(X = 2))$?
  - $\mathbb{P}(X = 0) = \mathbb{P}(\{(ce, ce)\}) \stackrel{equip.}{=} \frac{1}{4}$
  - $\mathbb{P}(X = 1) = \mathbb{P}(\{(ca, ce), (ce, ca)\}) \stackrel{equip.}{=} \frac{2}{4} = \frac{1}{2}$
  - $\mathbb{P}(X = 2) = \mathbb{P}(\{(ca, ca)\}) \stackrel{equip.}{=} \frac{1}{4}$
- Notar: $\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 1$

Ejemplo

En un bolillero hay 6 bolillas numeradas del 1 al 6. Se toman 3 al azar sin reposición, y definimos la v.a.

Z=número más alto obtenido

¿$Z = \{a, b, c, d\}$, $\mathbb{P}(Z = a)$, $\mathbb{P}(Z = b)$, $\mathbb{P}(Z = c)$, $\mathbb{P}(Z = d)$?
¿$\mathbb{P}(Z = a)$, $\mathbb{P}(Z = b)$, $\mathbb{P}(Z = c)$, $\mathbb{P}(Z = d)$?

Z={3,4,5,6}

P(z=3) = (1 1)(2 2)/(6 3)
P(z=4) = (1 1)(3 2)/(6 3)
P(z=5) = (1 1) (4 2)/(6 3)

generalizando:
Sumatoria (i=k —> n-1) (i k) = (n k+1)

Variables aleatorias discretas

Ejemplo

En un bolillero hay 6 bolillas numeradas del 1 al 6. Se toman 3 al azar sin reposición, y definimos la v.a.

Z=número más alto obtenido

$Z = \{3, 4, 5, 6\}$

- $\mathbb{P}(Z = 3) = \dfrac{\binom{3}{3}}{\binom{6}{3}}$
- $\mathbb{P}(Z = 4) = \dfrac{\binom{3}{2}\binom{1}{1}}{\binom{6}{3}}$
- $\mathbb{P}(Z = 5) = \dfrac{\binom{4}{2}\binom{1}{1}}{\binom{6}{3}}$
- $\mathbb{P}(Z = 6) = \dfrac{\binom{5}{2}\binom{1}{1}}{\binom{6}{3}}$
- Notar: $\mathbb{P}(Z = 3) + \mathbb{P}(Z = 4) + \mathbb{P}(Z = 5) + \mathbb{P}(Z = 6) = 1$

Variables aleatorias discretas

Definición:

Función de probabilidad puntual o función de frecuencia

$$p_X(x_i) = \mathbb{P}(X = x_i) = \mathbb{P}(\{w : X(w) = x_i\})$$

Propiedad

$\sum_i p_X(x_i) = 1$

Demostración.

$\sum_i p_X(x_i) = \sum_i \mathbb{P}(\{w : X(w) = x_i\}) = \mathbb{P}\left(\cup_i \{w : X(w) = x_i\}\right) =$
$\mathbb{P}(\Omega) = 1$ □

Variables aleatorias discretas

Definición:

Función de distribución (acumulada) Va acumulando valores de probabilidades a medida que avanza

$$F_X(x) = \mathbb{P}(X \leq x) \qquad \forall x \in \mathbb{R}$$

Variable aleatoria

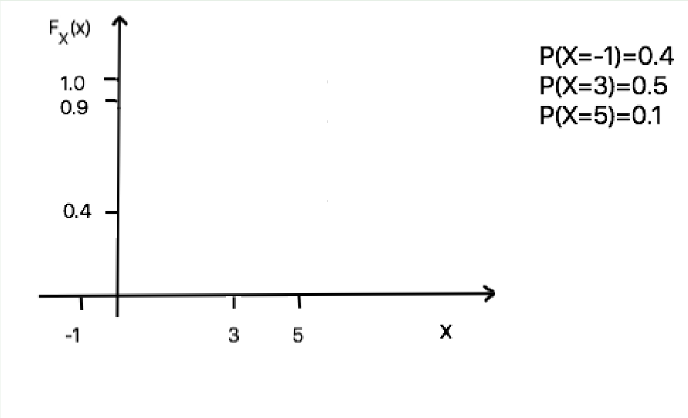
- Está definida para todos los reales.

Probabilidad puntual —> valor de probabilidad en un punto exacto
Probabilidad acumulada —> Valor de probabilidad menor o igual a un punto

Variables aleatorias discretas

Ejemplo

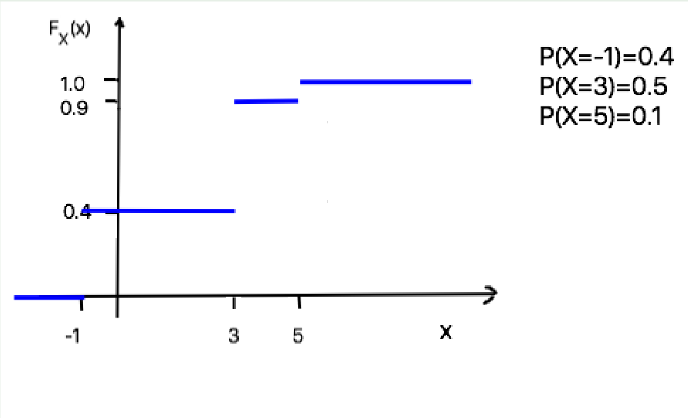
$X = \{-1, 3, 5\}$,
$\mathbb{P}(X = -1) = 0{,}4, \mathbb{P}(X = 3) = 0{,}5, \mathbb{P}(X = 5) = 0{,}1$



Variables aleatorias discretas

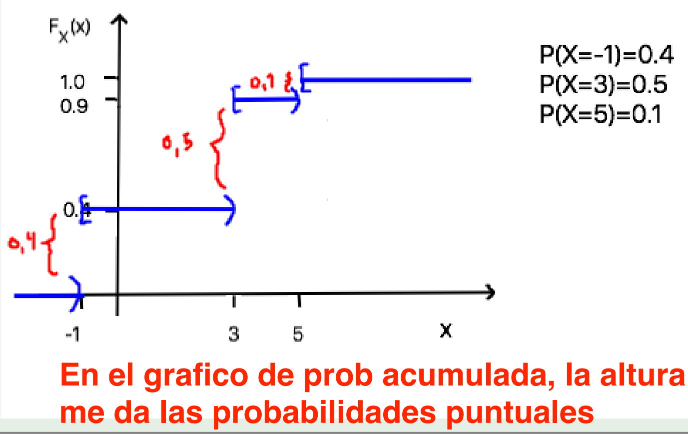
Ejemplo

$X = \{-1, 3, 5\}$,
$\mathbb{P}(X = -1) = 0{,}4, \mathbb{P}(X = 3) = 0{,}5, \mathbb{P}(X = 5) = 0{,}1$



- $F_X(-1{,}2) = \mathbb{P}(X \leq -1{,}2) = 0$
- $F_X(4) = \mathbb{P}(X \leq 4) = \mathbb{P}(X = -1) + \mathbb{P}(X = 3) = 0{,}4 + 0{,}5 = 0{,}9$

Ejemplo



Si X es un v.a. discreta $F_X(x)$ es una función escalera con saltos en los puntos donde toma valores X, y la altura de los escalones es la probabilidad puntual.

Propiedades

- $\lim_{x \to \infty} F_X(x) = 1$
- $\lim_{x \to -\infty} F_X(x) = 0$
- $F_X(x)$ es monótona no decreciente (creciente, pero no estrictamente creciente).
- $F_X(x)$ es continua a derecha ($F_X(b) = \underset{x \to b^+}{lim} F_X(x)$)
- $\mathbb{P}\,(a < X \leq b) = F_X(b) - F_X(a)$
- $\mathbb{P}\,(a \leq X \leq b) = F_X(b) - F_X(a - \epsilon)$ con $0 < \epsilon << 1$

Variables aleatorias discretas

Variables muy utilizadas

- Bernoulli
- Binomial
- Geométrica
- Binomial Negativa
- Poisson

Bernoulli

La distribucion de bernoulli tiene un solo parametro "p"

Variable de Bernoulli

$X \sim$ Bernoulli(p) si $X = \{0, 1\}$ y $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$
El 1 uno lo asocia al "éxito"
El 0 al "fracaso"

Ejemplos

- Tiro una moneda y asocio el éxito con la cara. En este caso $p = 1/2$
- Tiro una dado y asocio el éxito a que el dado ≤ 2. En este caso $p = 1/3$

Binomial

Variable Binomial

Z =número de "éxitos" (o unos) en n ensayos de Bernoulli independientes.

- $Z = X_1 + X_2 + \cdots + X_n$ con $X_i \sim$ Bernoulli(p) independientes.
- $Z \sim$ Binom(n,p) si $Z = \{0, 1, \ldots, n\}$ y ¿$\mathbb{P}(Z = k)$?

Una variable binomial, se construye a partir de varios experimentos de bernoulli independientes

Recibe 2 parametros, hereda el p de bernoulli y el numero de experimentos

Binomial

Ej tirar 2 monedas, y sale cara o seca, la primera moneda no le dice a la moneda posterior, sali seca, asi q te toca cara, no, por lo tanto son independientes

Variable Binomial

Z =número de "éxitos" (o unos) en n ensayos de Bernoulli independientes.

$Z = \{0, 1, \ldots, n\}$

$\mathbb{P}(Z = 0) = \mathbb{P}(X_1 = 0 \cap X_2 = 0 \cap X_3 = 0 \cap \cdots \cap X_n = 0) \overset{indep.}{=}$
$\mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 0) \ldots \mathbb{P}(X_n = 0) = (1 - p)(1 - p) \ldots (1 - p)$

$$\mathbb{P}(Z = 0) = (1 - p)^n$$

ej : El 20% de los articulos que fabrica una empresa salen defectuosos. Si se seleccionan 6 articulos ¿Cual es la probabilidad que salgan 3 articulos defectuosos?

tengo definido n=6, a diferencia de la Distribucion Geometrica

Binomial

Variable Binomial

Z =número de "éxitos" (o unos) en n ensayos de Bernoulli independientes.

$Z = \{0, 1, \ldots, n\}$

$$\mathbb{P}\left(Z = 0\right) = (1 - p)^n$$

$$\mathbb{P}\left(Z = 1\right) = \binom{n}{1} p(1 - p)^{n-1}$$

Probabilidad de 1 exito=de todos los n que conforman los experimentos, dame 1 exito, y multiplicalo por la probabilidad de fracaso de los demas experimentos

Binomial

Variable Binomial

Z =número de "éxitos" (o unos) en n ensayos de Bernoulli independientes.

$Z = \{0, 1, \ldots, n\}$

$$\mathbb{P}\left(Z = 0\right) = (1 - p)^n \qquad \mathbb{P}\left(Z = 1\right) = \binom{n}{1} p(1 - p)^{n-1}$$

$$\mathbb{P}\left(Z = 2\right) = \binom{n}{2} p^2(1 - p)^{n-2}$$

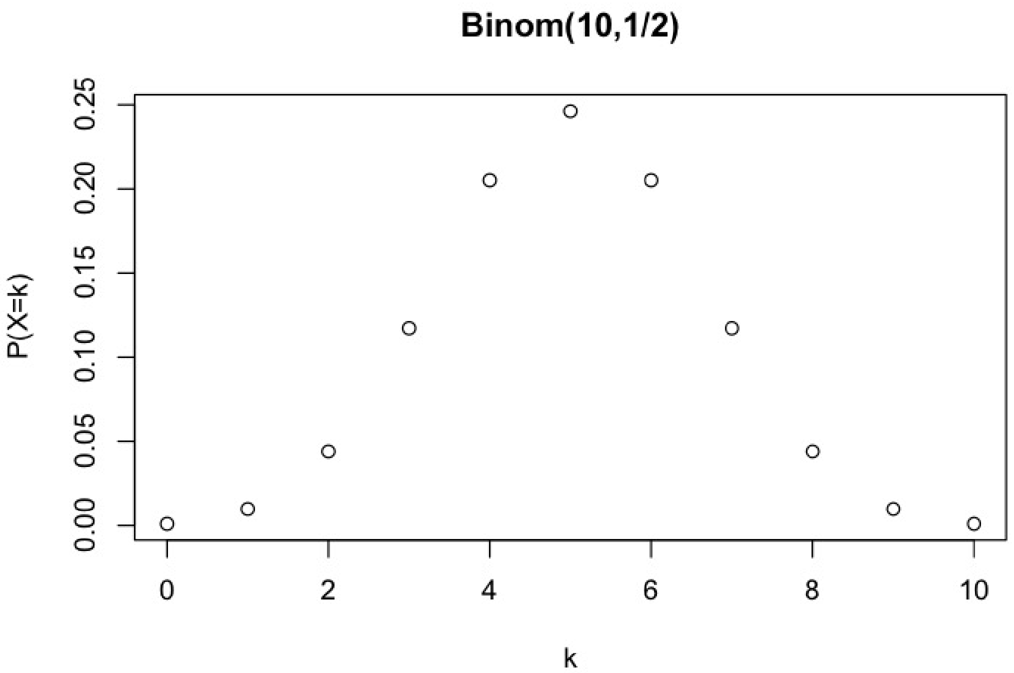
Binomial

Variable Binomial

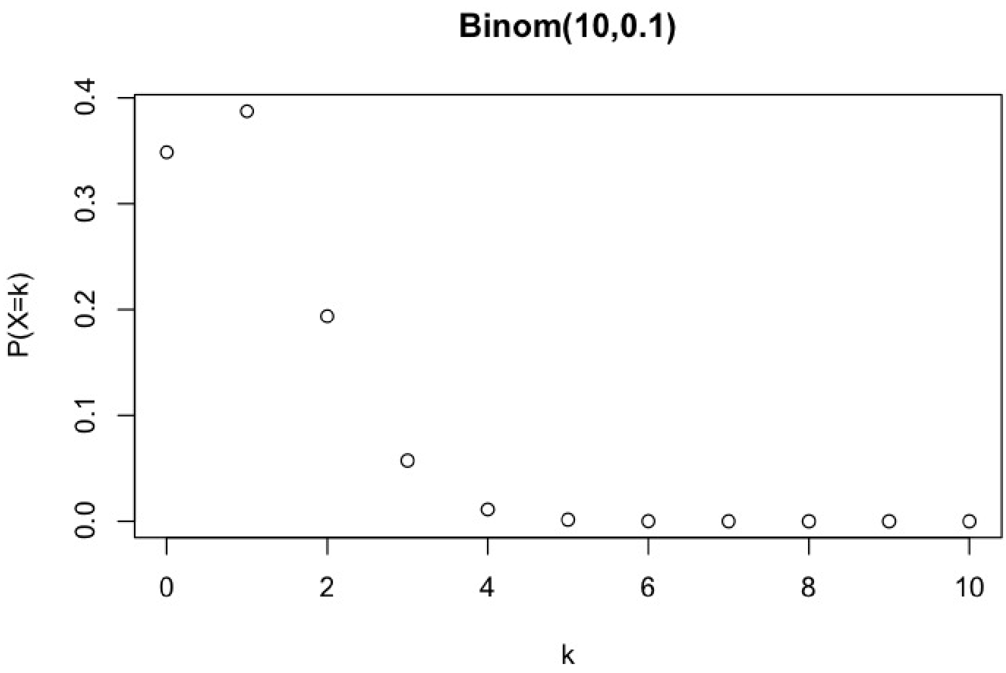
Z =número de “éxitos” (o unos) en n ensayos de Bernoulli independientes.

- $Z \sim Binom(n, p)$
- $Z = \{0, 1, \ldots, n\}$
- $\mathbb{P}(Z = k) = \binom{n}{k} p^k (1-p)^{n-k}$ con $k = \{0, 1, 2, \ldots, n\}$
- Notar: $\sum_{k=0}^{n} \mathbb{P}(Z = k) = 1$

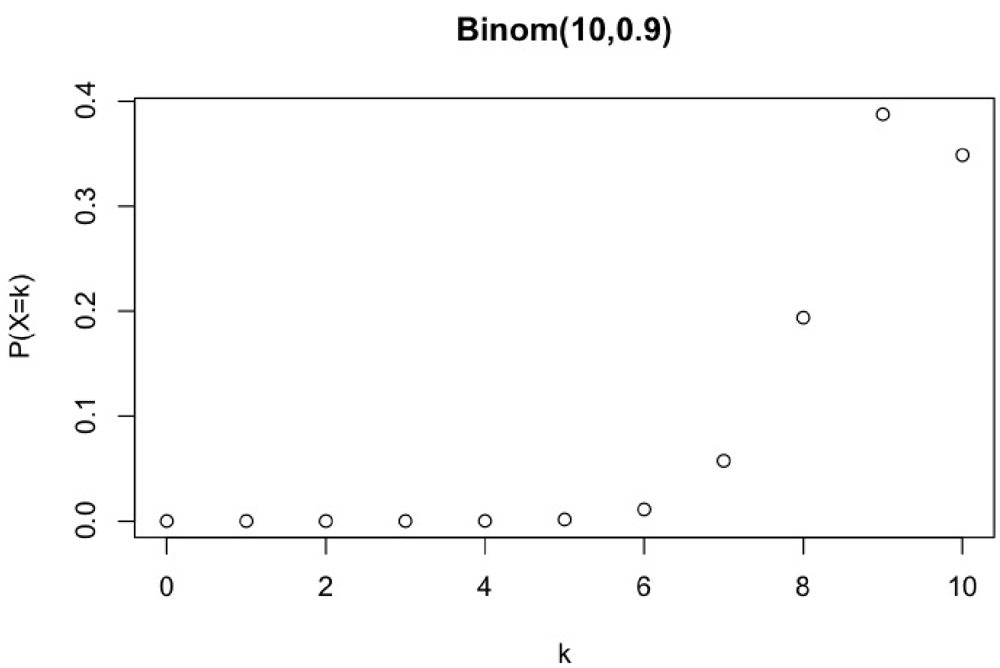
Distribución Binomial



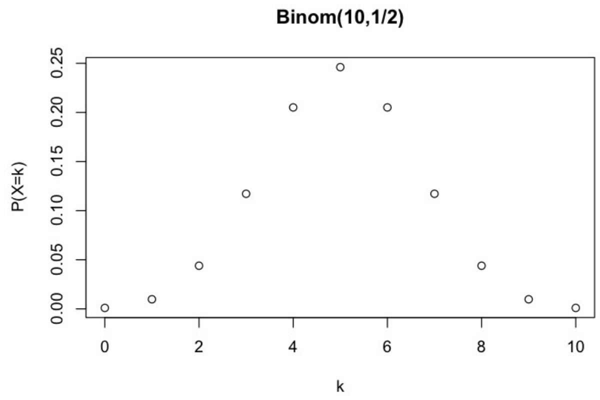
Distribución Binomial

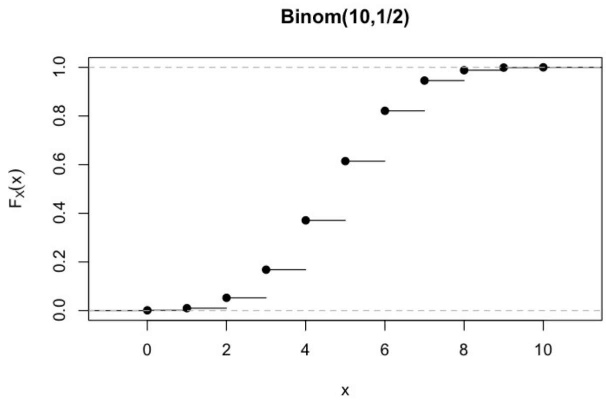


Distribución Binomial

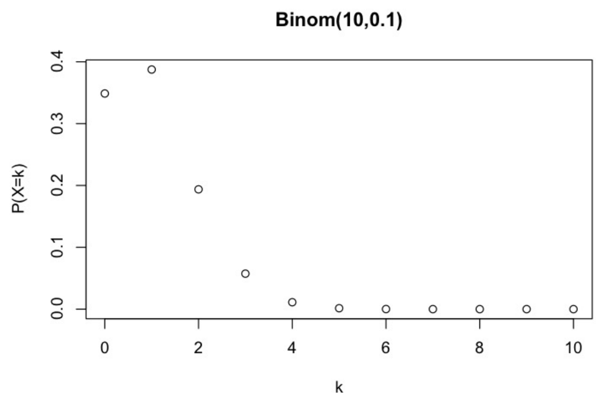


Distribución Binomial

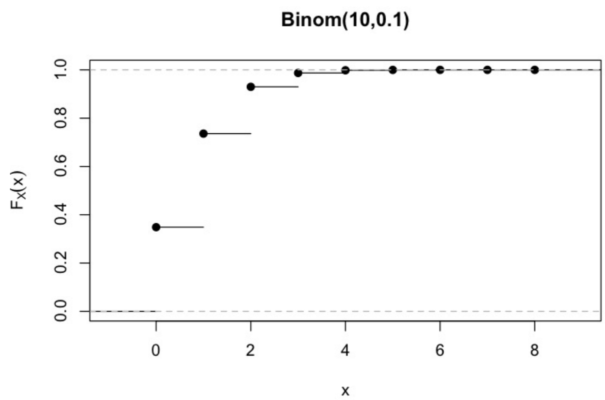




Distribución Binomial

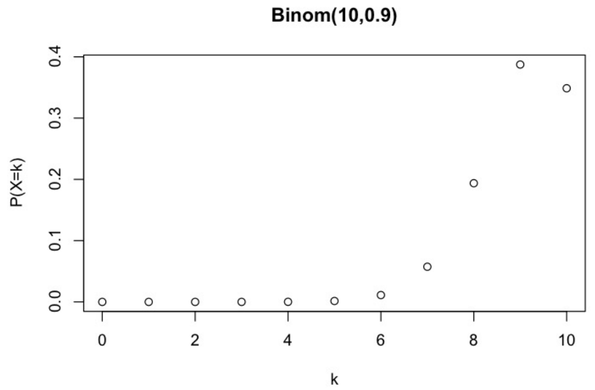


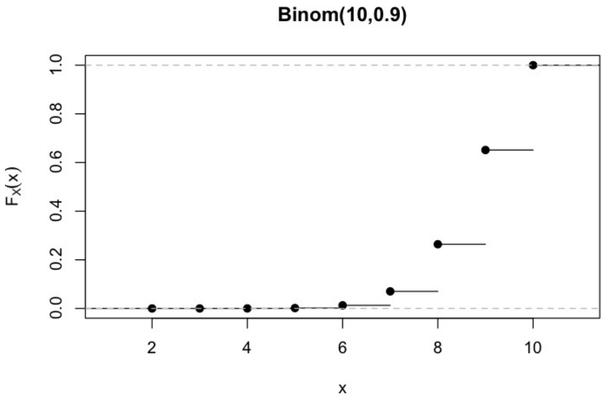
Distribucion de probabilidad



Acumulada

Distribución Binomial





Variables muy utilizadas

Ejemplo

¿Cuál es la probabilidad de tener menos de 3 caras al tirar 10 veces la moneda?

$Z =$número de "éxitos" (o unos) en n ensayos de Bernoulli independientes.

- $Z =$número de caras al tirar 10 veces la moneda.
- $Z \sim Binom(10, 1/2)$ $\quad\quad p = \mathbb{P}\,(\text{éxito}) = \mathbb{P}\,(\text{cara}) = 1/2$
- ¿$\mathbb{P}\,(Z < 3)$?
- $\mathbb{P}\,(Z < 3) = \mathbb{P}\,(Z = 0) + \mathbb{P}\,(Z = 1) + \mathbb{P}\,(Z = 2)$
- $\mathbb{P}\,(Z < 3) = \binom{10}{0}(1/2)^0(1-1/2)^{10} + \binom{10}{1}(1/2)^1(1-1/2)^9 + \binom{10}{2}(1/2)^2(1-1/2)^8$
- $\mathbb{P}\,(Z < 3) = \binom{10}{0}(1/2)^{10} + \binom{10}{1}(1/2)^{10} + \binom{10}{2}(1/2)^{10}$
- $\mathbb{P}\,(Z < 3) = 1(1/2)^{10} + 10(1/2)^{10} + 45(1/2)^{10} = 56(1/2)^{10}$

Involucramos la binomial, ya que sabemos como resolver el problema, sumamos las probabilidades y no pienso mas

Geométrica

Variable Geométrica

W =número de ensayos de Bernoulli independientes hasta obtener el primer “éxito”.

- $W \sim Geom(p)$
- $W = \{1, 2, 3, \ldots\}$ y ¿$\mathbb{P}(W = k)$?

Geométrica

Variable Geométrica

W =número de ensayos de Bernoulli independientes hasta obtener el primer “éxito”.

- $W \in \mathbb{N}$, ¿$\mathbb{P}(W = k)$?
- $\mathbb{P}(W = 1) = \mathbb{P}(E) = p$
- $\mathbb{P}(W = 2) = \mathbb{P}(FE) \overset{indep}{=} (1 - p)p$
- $\mathbb{P}(W = 5) = \mathbb{P}(FFFFE) \overset{indep}{=} (1 - p)^4 p$
- $\mathbb{P}(W = k) = (1 - p)^{k-1} p$

Geométrica

Variable Geométrica

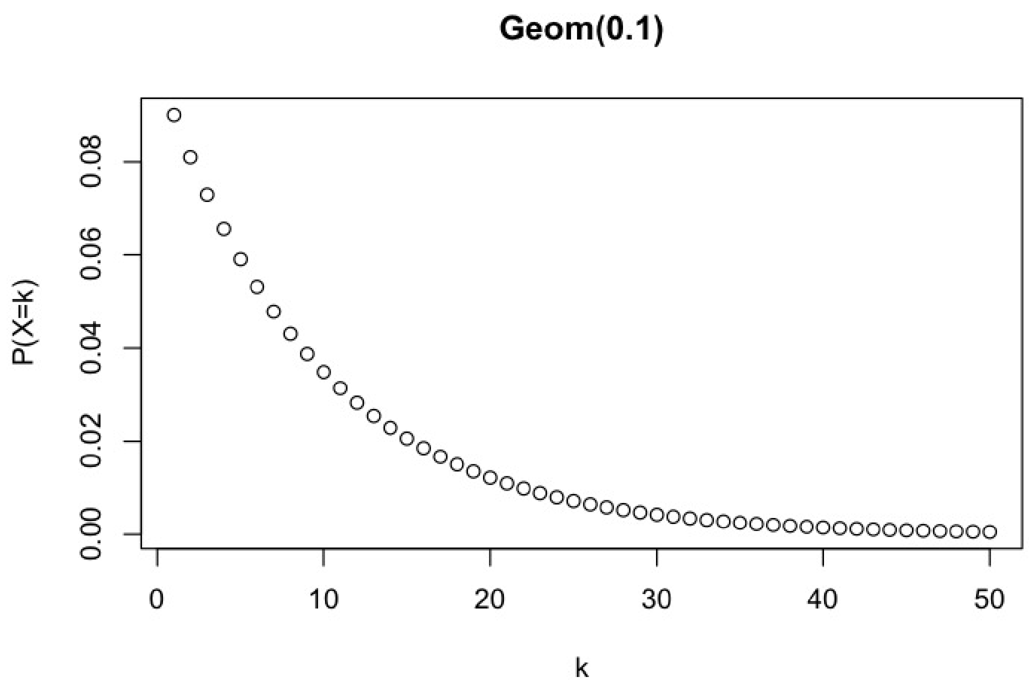
W =número de ensayos de Bernoulli independientes hasta obtener el primer “éxito”.

- $W \sim Geom(p)$
- $W = \{1, 2, 3, \dots\}$
- $\mathbb{P}(W = k) = (1-p)^{k-1}p \quad$ con $k = \{1, 2, 3, \dots\}$
- Notar: $\sum_{k=1}^{\infty} \mathbb{P}(W = k) = 1$

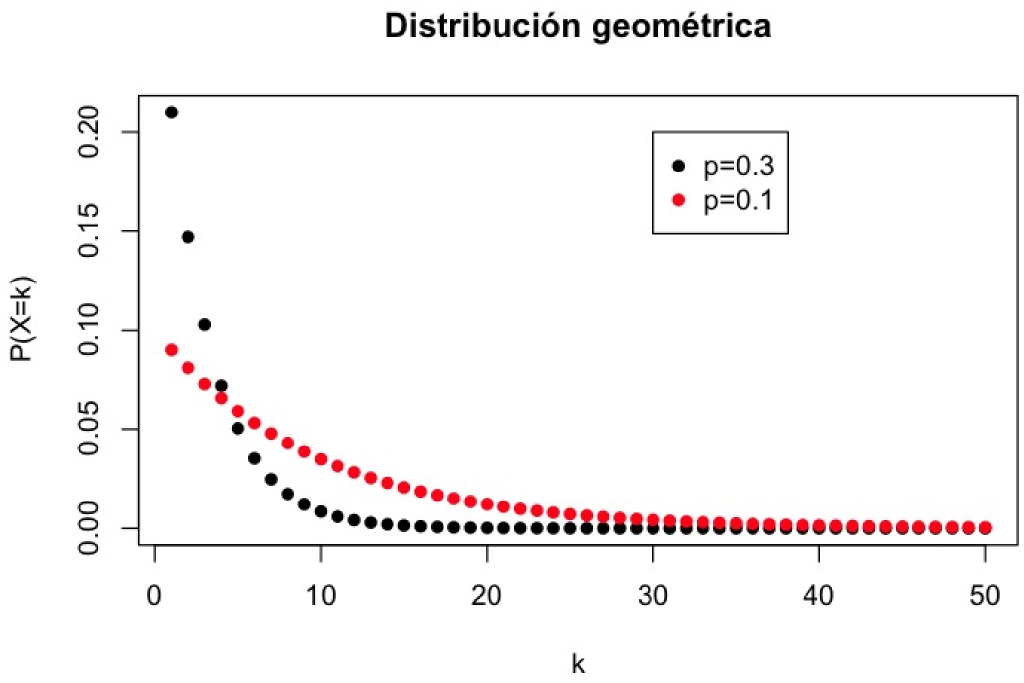
En la geometrica no hay un “n” de experimentos, porque vos haces experimentos HASTA que se logra un exito

En la geometrica, yo busco el “n”, de numero de experimentos independientes hasta obtener el PRIMER “exito”

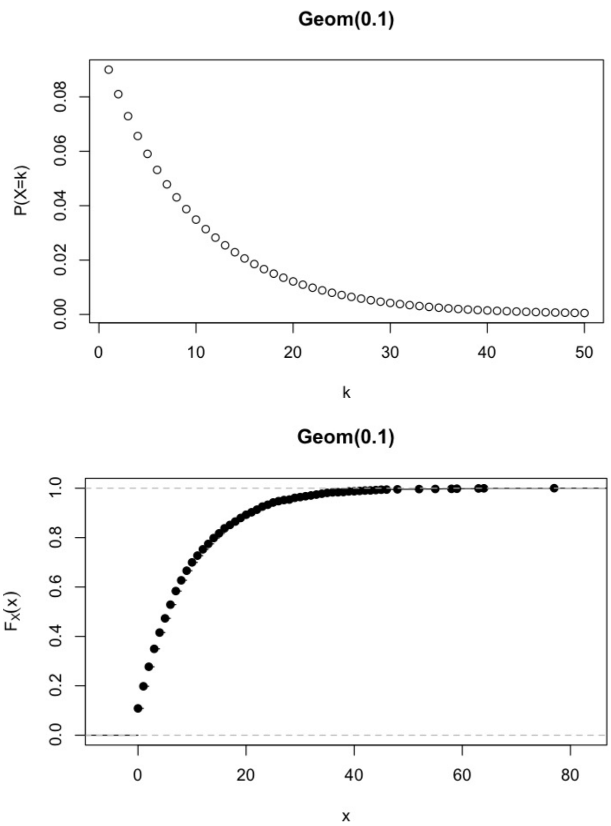
Distribución Geométrica



Distribución Geométrica



Distribución Geométrica



Geométrica

Propiedades de una v.a. Geométrica

W =número de ensayos de Bernoulli independientes hasta obtener el primer "éxito".

1. $F_W(j) = \mathbb{P}(W \leq j) = 1 - (1-p)^j$ para $j \in \mathbb{N} \cup \{0\}$

Demostración.

$F_W(j) = \mathbb{P}(W \leq j) = 1 - \mathbb{P}(W > j) =$
$1 - \mathbb{P}(\text{el primer éxito aparece luego de } j \text{ intentos}) =$
$1 - \mathbb{P}(\text{los primeros } j \text{ ensayos son fracasos}) = 1 - \mathbb{P}(F_1 F_2 \ldots F_j) =$
$1 - (1-p)^j$ $\square$

2. $\mathbb{P}(W > j) = (1-p)^j$ para $j \in \mathbb{N} \cup \{0\}$
3. Recordemos: $\mathbb{P}(a < W \leq b) = F_W(b) - F_W(a)$ vale para toda v.a.

Geométrica

Propiedades de una v.a. Geométrica

W =número de ensayos de Bernoulli independientes hasta obtener el primer "éxito".

1. $F_W(j) = \mathbb{P}(W \leq j) = 1 - (1-p)^j$ para $j \in \mathbb{N} \cup \{0\}$
2. $\mathbb{P}(W > j) = (1-p)^j$ para $j \in \mathbb{N} \cup \{0\}$
3. $\mathbb{P}(W > n + k | \underset{W}{\cancel{X}} > n) = \mathbb{P}(\underset{W}{\cancel{X}} > k)$ (Falta de memoria)
4. Recordemos: $\mathbb{P}(a < W \leq b) = F_W(b) - F_W(a)$ vale para toda v.a.

- $\mathbb{P}(2 < W \leq 10) = F_W(10) - F_W(2) = (1-p)^2 - (1-p)^{10}$
- $\mathbb{P}(2 \leq W \leq 10) = \mathbb{P}(1 < W \leq 10) = F_W(10) - F_W(1) = (1-p)^1 - (1-p)^{10}$

Geométrica

Ejemplo

Si le apuesto a un número en la ruleta (pleno) la probabilidad de ganar es 1/37. Voy a apostar muchas veces hasta ganar, (a) ¿cuál es la probabilidad de gane antes de la apuesta 20? , (b) ¿y que gane después de la apuesta 30?

W =número de veces que apuesto hasta ganar.

- $W \sim Geom(1/37)$
- (a) ¿$\mathbb{P}(W < 20)$?,   (b) ¿$\mathbb{P}(W > 30)$?
- $\mathbb{P}(W < 20) = \mathbb{P}(W \leq 19) = F_W(19) = 1 - (1 - 1/37)^{19} = 1 - (36/37)^{19} \approx 0{,}406$
- $\mathbb{P}(W > 30) = 1 - \mathbb{P}(W \leq 30) = 1 - F_W(30) = 1 - (1 - (1 - 1/37)^{30}) = (36/37)^{30} \approx 0{,}439$

La ventaja q me da utilizar la geometrica, es que directamente me da la acumulada a un valor, por eso la utilizo

Binomial Negativa

Es la Geometrica, pero si quiero MAS DE UN EXITO

Variable Binomial Negativa

X =número de ensayos de Bernoulli independientes hasta obtener el r-ésimo "éxito".

- ¿Qué valores toma X? ¿$\mathbb{P}(X = k)$?
- $X = \{r, r+1, r+2, \dots\}$
- $\mathbb{P}(X = r) = \mathbb{P}(E_1 E_2 E_3 \dots E_r) = p^r$
- $\mathbb{P}(X = r+1) = \mathbb{P}(F_1 E_2 E_3 \dots E_r E_{r+1}) + \cdots = \binom{r}{1}(1-p)p^r = \binom{r}{r-1}(1-p)p^r$
- $\mathbb{P}(X = r+2) = \mathbb{P}(F_1 F_2 E_3 \dots E_r E_{r+1} E_{r+2}) + \cdots = \binom{r+1}{2}(1-p)^2 p^r = \binom{r+1}{r-1}(1-p)^2 p^r$
- $\mathbb{P}(X = k) = \binom{k-1}{r-1}(1-p)^{k-r} p^r$

Aca no tenemos 1 solo exito, buscamos mas exitos

Binomial Negativa

Variable Binomial Negativa

X =número de ensayos de Bernoulli independientes hasta obtener el r-ésimo "éxito".

- $X \sim BN(r, p)$
- $X = \{r, r+1, r+2, \dots\}$
- $\mathbb{P}(X = k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r \quad$ con $k = \{r, r+1, r+2, \dots\}$
- Notar: $\sum_{k=r}^{\infty} \mathbb{P}(X = k) = 1$
- Notar: $BN(r = 1, p) =$ Geométrica(p).
- $BN(r, p)$ es una generalización de la Geométrica(p)

Binomial Negativa

Ejemplo

Todos los días uso internet en mi casa, y puede funcionar bien o mal. La cuarta vez que funcione mal llamo al servicio técnico. Suponiendo que la probabilidad de que funcione bien es 0.9 y que cada día funciona en forma independiente, ¿cuál es la probabilidad de que llame dentro de 10 días al servicio técnico.

X =número de días hasta que sea la cuarta vez que funcione mal internet.

- $X \sim BN(4, 0{,}1)$
- ¿$\mathbb{P}(X = 10)$?
- $\mathbb{P}(X = 10) = \binom{10-1}{4-1}(1-0{,}1)^{10-4}0{,}1^4 = \binom{9}{3}(0{,}9)^6 0{,}1^4$

Hipergeométrica

Variable Hipergeométrica

Tenemos N bolitas en una urna, de las cuales r son rojas y N-r son blancas.

$Y =$ número de bolitas rojas extraídas al extraer n sin reposición.

- ¿Qué valores toma Y? Supongamos que $N = 5, r = 4, n = 3$
- $Y = \{2, 3\}$ P(y=2) = (4 2) (1 1)/(5 3) P(y=3) = (4 3)(1 0)/(5 3)
- En el caso gral, $Y = \{max\{0, n - N + r\}, \dots, min\{n, r\}\}$
- ¿$\mathbb{P}(Y = k)$? con k uno de los valores posibles.
- $\mathbb{P}(Y = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$

Hipergeométrica

Variable Hipergeométrica

Tenemos N bolitas en una urna, de las cuales r son rojas y N-r son blancas.

$Y =$ número de bolitas rojas extraídas al extraer n sin reposición.

- $Y \sim H(N, r, n)$
- $Y = \{max\{0, n - N + r\}, \dots, min\{n, r\}\}$
- $\mathbb{P}(Y = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$ con k valores compatibles.
- Notar: $\sum_k \mathbb{P}(Y = k) = 1$

Hipergeométrica

Ejemplo

En una cajón hay 40 tomates de los cuales 5 están podridos. Se sacan 8 tomates al azar sin reposición.

(a) Hallar la probabilidad de haber sacado exactamente 2 podridos.

(b) Hallar la probabilidad de haber sacado al menos 2 podridos.

Definimos $Y =$ número de tomates podridos al comprar 8 (extracción sin reposición).
Entonces $Y \sim H(40, 5, 8)$.

(a) $\mathbb{P}(Y = 2) = \frac{\binom{5}{2}\binom{35}{6}}{\binom{40}{8}} \approx 0{,}211.$

(b) $\mathbb{P}(Y \geq 2) = 1 - \mathbb{P}(Y < 2) = 1 - \mathbb{P}(Y \leq 1)$
$= 1 - [\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1)]$
$= 1 - \left[\frac{\binom{5}{0}\binom{35}{8}}{\binom{40}{8}} + \frac{\binom{5}{1}\binom{35}{7}}{\binom{40}{8}}\right] \approx 0{,}25676.$

La diferencia con la binomial, es que en este tipo de problemas, es sin reposicion el experimento.

Con reposición vs. sin reposición

Ejercicio

Tenemos N bolitas en una urna, de las cuales r son rojas y N-r son blancas.

$Y =$ número de bolitas rojas extraídas al extraer n sin reposición.
$X =$ número de bolitas rojas extraídas al extraer n con reposición.

- ¿Qué distribución tiene Y? **Hipergeometrica**
- ¿Qué distribución tiene X? **Binomial(p,n)**
- ¿Cuál es la diferencia entre con y sin reposición? ¿Hay independencia en cada una de las bolitas extraídas en algún caso?
- $Y \sim H(N, r, n), \qquad X \sim Binomial(n, r/N)$

Hay independencia cuando se utiliza con reposicion, ya que las bolitas que fueron saliendo, no afectan la probabilidad de las que aun no han salido

Poisson

En poisson tengo la cantidad de "eventos", en un tiempo determinado

Variables Poisson

Y = número de "eventos" en dado período de tiempo.
Y = número de terremotos en un año en la ciudad de Santiago.
Y = número de autos que pasan por un peaje por minuto.
Y = número de llamadas entradas en un call center. **en 5 minutos**
Muchas veces este tipo de v.a. está bien descripta por una v.a. de Poisson.

- $Y \sim \text{Poisson}(\lambda)$ Lambda, me dice, en promedio tuviste "tantos" eventos en ese intervalo de tiempo
- $Y = \{0, 1, 2, 3, \dots\}$ y $\mathbb{P}(Y = k) = \frac{e^{-\lambda}\lambda^k}{k!}$
- λ es el número promedio de ocurrencias en el período de tiempo.
- $\lambda > 0$ y depende de la escala de tiempo utilizada (días, horas, minutos).
- $\sum_{k=0}^{\infty} \mathbb{P}(Y = k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$

Poisson

Variable Poisson

Supongamos que Y = número de terremotos en un año en la ciudad de Santiago, tiene distribución de Poisson con $\lambda = 9{,}3$

9.3 terremotos al año

$$Y \sim Poisson(9{,}3), \quad \mathbb{P}(Y = k) = \frac{e^{-9{,}3}(9{,}3)^k}{k!} \text{ con } k \in \mathbb{N} \cup \{0\}$$

W = número de terremotos en un año en la ciudad de Mendoza, tiene distribución de Poisson con $\lambda = 2{,}6$

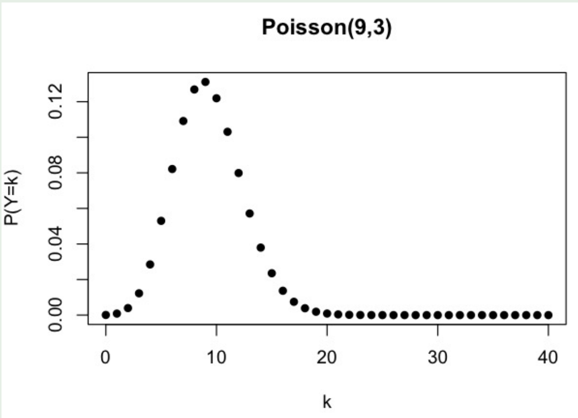
$$W \sim Poisson(2{,}6), \quad \mathbb{P}(W = k) = \frac{e^{-2{,}6}(2{,}6)^k}{k!} \text{ con } k \in \mathbb{N} \cup \{0\}$$

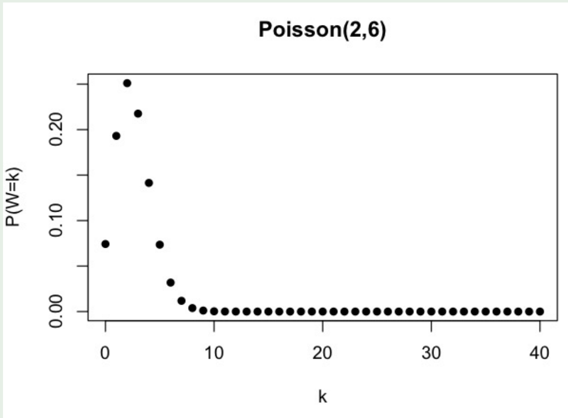
Distribución de Poisson

Ejemplo

Y = número de terremotos en un año en la ciudad de Santiago.
W = número de terremotos en un año en la ciudad de Mendoza.

$$Y \sim Poisson(9{,}3), \qquad W \sim Poisson(2{,}6)$$





Poisson

Ejemplo

Supongamos que Y = número de terremotos en un año en la ciudad de Santiago. tiene distribución de Poisson con $\lambda = 9{,}3$

$$Y \sim Poisson(9{,}3), \quad \mathbb{P}\left(Y = k\right) = \frac{e^{-9{,}3}(9{,}3)^k}{k!} \text{ con} k \in \mathbb{N} \cup \{0\}$$

- ¿Cuál es la probabilidad de no tener un solo terremoto el año que viene?
- ¿Cuál es la probabilidad de no tener menos de 2 terremotos el año que viene?

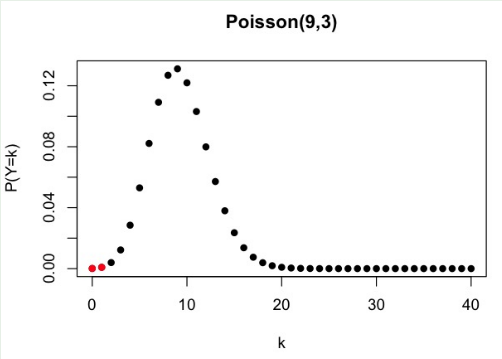
Poisson

Ejemplo

Supongamos que Y = número de terremotos en un año en la ciudad de Santiago. tiene distribución de Poisson con $\lambda = 9{,}3$

$$Y \sim Poisson(9{,}3), \quad \mathbb{P}(Y = k) = \frac{e^{-9{,}3}(9{,}3)^k}{k!} \text{ con} k \in \mathbb{N} \cup \{0\}$$

- ¿$\mathbb{P}(Y = 1)$? ¿$\mathbb{P}(Y < 2)$?



Poisson

Ejemplo

$$Y \sim Poisson(9{,}3), \quad \mathbb{P}(Y = k) = \frac{e^{-9{,}3}(9{,}3)^k}{k!} \text{ con} k \in \mathbb{N} \cup \{0\}$$

- $\mathbb{P}(Y = 1) = \frac{e^{-9{,}3}(9{,}3)^1}{1!} \approx 0{,}0009$
- $\mathbb{P}(Y < 2) = \mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) = \frac{e^{-9{,}3}(9{,}3)^0}{0!} + \frac{e^{-9{,}3}(9{,}3)^1}{1!} = e^{-9{,}3} + e^{-9{,}3}9{,}3 = 10{,}3e^{-9{,}3}$

Poisson

Ejemplo cambio de escala temporal

Supongamo que $Y =$ número de terremotos en un año en la ciudad de Santiago. tiene distribución de Poisson con $\lambda = 9{,}3$

¿Cuál es la probabilidad de que el mes que viene haya 2 terremotos?

- $Z =$ número de terremotos en un mes en la ciudad de Santiago.
- ¿Qué distribución tiene Z?
- $Z \sim Poisson(\lambda_{mes})$
- ¿λ_{mes}?
- $\lambda_{mes} = \frac{\lambda}{12} = \frac{9{,}3}{12} = 0{,}775$
- $Z \sim Poisson(0{,}775)$, ¿$\mathbb{P}(Z = 2)$?

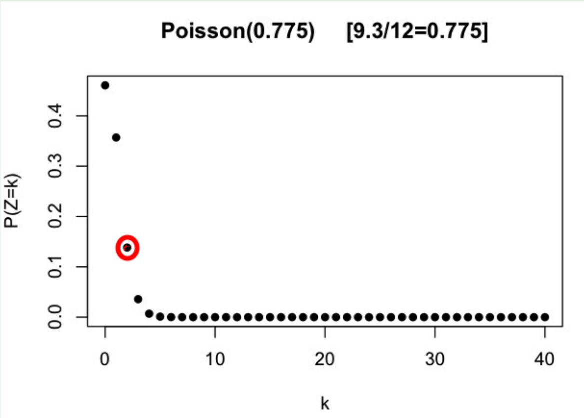
Poisson

Ejemplo

$Z =$ número de terremotos en un mes en la ciudad de Santiago.

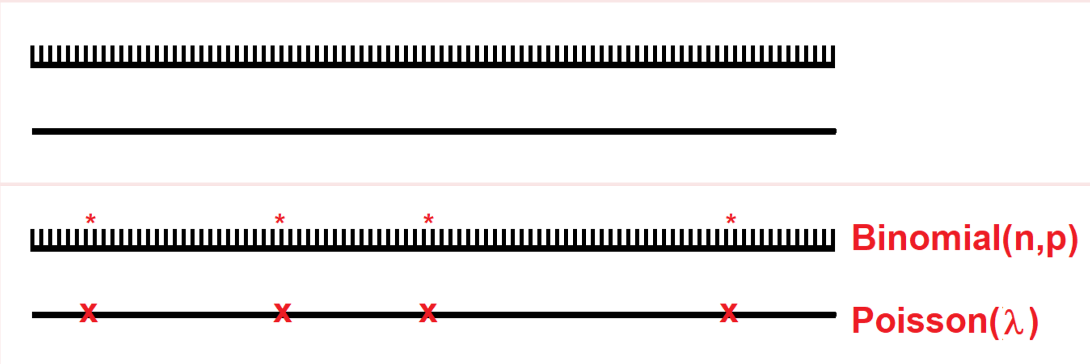
$$Z \sim Poisson(0{,}775)$$

$$\mathbb{P}(Z = 2) = \frac{e^{-0{,}775}(0{,}775)^2}{2!}$$



Poisson como una aproximación a una Binomial(n,p) con $n \to \infty, p \to 0, np \to c > 0$

Pensemos en una Binomial(n, p) con $n >> 1$, $p << 1$ y $np = \lambda$



Poisson como una aproximación a una Binomial(n,p) con $n \to \infty, p \to 0, np \to c > 0$

Pensemos en una Binomial(n, p) con $n >> 1$, $p << 1$ y $np = \lambda$

$$X \sim Binom(n, p = \frac{\lambda}{n}) \leftrightarrow \mathbb{P}\left(X = k\right) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mathbb{P}\left(X = k\right) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k}$$

$$\mathbb{P}\left(X = k\right) = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k}$$

$$\mathbb{P}\left(X = k\right) = \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} \frac{\lambda^k}{k!} (1 - \frac{\lambda}{n})^n (1 - \frac{\lambda}{n})^{-k}$$

Poisson como una aproximación a una Binomial(n,p) con $n \to \infty, p \to 0, np \to c > 0$

Pensemos en una Binomial(n, p) con $n >> 1$, $p << 1$ y $np = \lambda$

$$\mathbb{P}(X = k) = \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} \frac{\lambda^k}{k!} (1 - \frac{\lambda}{n})^n (1 - \frac{\lambda}{n})^{-k}$$

- $\lim_{n \to \infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} = 1$
- $\lim_{n \to \infty} (1 - \frac{\lambda}{n})^{-k} = 1$
- $\lim_{n \to \infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda}$

$$\bigstar \quad \mathbb{P}(X = k) \approx \frac{\lambda^k e^{-\lambda}}{k!} \quad \bigstar$$