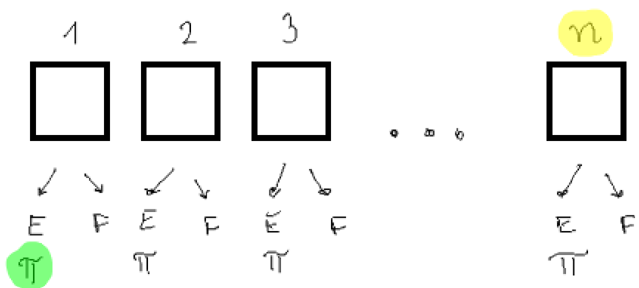
Distribución Binomial



$* \; X = N^o$ de éxitos

$$f(x) = P(X = x) = \begin{cases} \binom{n}{x} \pi^{\overset{\text{éxitos}}{x}} (1-\pi)^{\overbrace{n-x}^{\text{fracasos}}}, & x \in \{0, 1, 2, \ldots, n\} \\ 0, & \text{de otro modo} \end{cases}$$

$$\left.\begin{matrix} \pi \cdot \pi \cdot (1-\pi)(1-\pi)(1-\pi) \\ \pi (1-\pi)(1-\pi)\, \pi\, (1-\pi) \\ (1-\pi)(1-\pi)\, \pi\, (1-\pi)\, \pi \end{matrix}\right\} \binom{n}{x}$$

L L L C C C C C

Éxito: L

P(Éxito) = 3/8 = 0.375 (cte)

constante si el muestreo es con reemplazo

6000 Pre
1000 Pos

Éxito: Pre

P(Éxito) = 6/7 = 0.85714

Elegimos el 1º y es de pregrado ≈ ⇒ cte.

⇒ P(Éxito) = 5999/6999 = 0.85712

Ejm: $X \sim Bin(n = 6, \pi = 0.1)$

$$f(x) = \binom{6}{x} 0.1^{x}\, 0.9^{6-x}$$

Ejemplos

- X = Número de correos promocionales abiertos $\to X \sim Bin(n = 1000, \pi = 0.20)$ → correos → P(abrir el correo)
- Y = Número de piezas defectuosas producidas $\to Y \sim Bin(n = 30, \pi = 0.05)$ → P(pieza def)
- L = Número de respuestas correctas en un examen de opción múltiple (cuando se elige al azar) $\to L \sim Bin(n = 10, \pi = 0.25)$ → P(correcto)
- S = Número de semillas que germinan $\to S \sim Bin(n = 200, \pi = 0.88)$ → semillas → P(germinar)
- C = Cantidad de personas que adquieren el servicio → C ~ Bin(n = 15000, pi = 0.02) → P(adquirir el servicio)
- V = Cantidad de vicuñas hembras preñadas → V ~ Bin(n = 3000, pi = 0.01)
- F = Número de familias con 5 hijos → F ~ Bin(n = 15000, pi = 0.06)

Pruebas de bondad de ajuste: ¿La variable X se distribuye según una (.....)?

dbinom(x, n, p) → $P(X = x)$

pbinom(q, n, p) → $P(X \leq q)$: Probabilidad acumulada

qbinom(pe, n, p) → Percentil pe

rbinom(n, n, p) → Muestra de tamaño n

1)

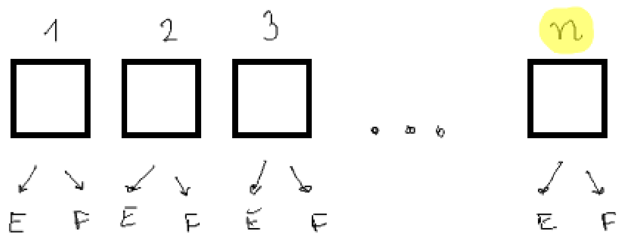
$$X \sim Bin(n = 20, \pi = 0.75)$$

$$P(2 < X \leq 8) = P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7) + P(X=8)$$

$$= P(X \leq 8) - P(X \leq 2)$$

$$\boxed{\boxed{0 \quad 1 \quad 2} \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8}$$

Distribución Hipergeométrica



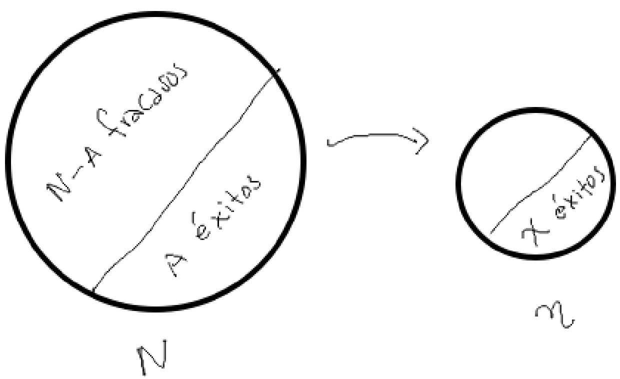
L L L C C C C C

Éxito: L

$p(\text{éxito}) = 3/8$ (1° ensayo) → Sale elegido L

$\Rightarrow p(\text{éxito}) = 2/7$ (2° ensayo) → ...

* X = N° de éxitos en la muestra



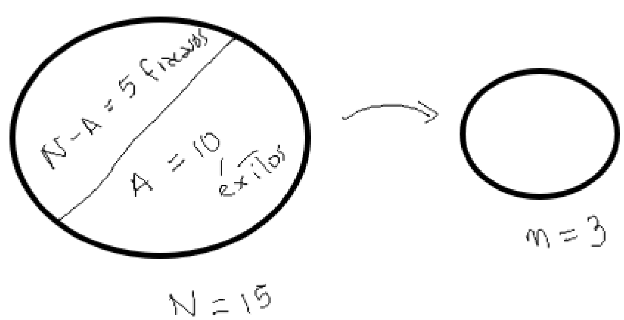
Función de probabilidad

Se define la V.A. Discreta X=Número de éxitos en la muestra de tamaño n. Su distribución de probabilidad es:

$$f(x) = P(X = x) = \begin{cases} \dfrac{\binom{A}{x}\binom{N-A}{n-x}}{\binom{N}{n}}, & \textbf{máx}(0, n + A - N) \leq x \leq \textbf{mín}(n, A) \\ 0, & \text{de otro modo} \end{cases}$$

Notación: $X \sim Hiper(N, n, A)$

$X \sim Hiper(N = 15, n = 3, A = 10)$



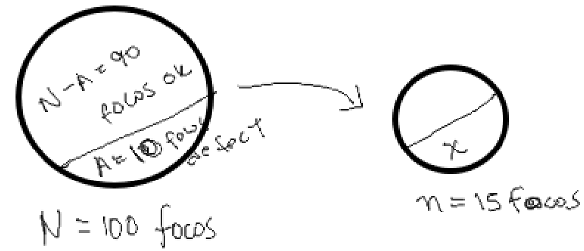
$$\underbrace{\text{máx}(0, -2)}_{0} \leq X \leq \text{mín}(3, 10)$$

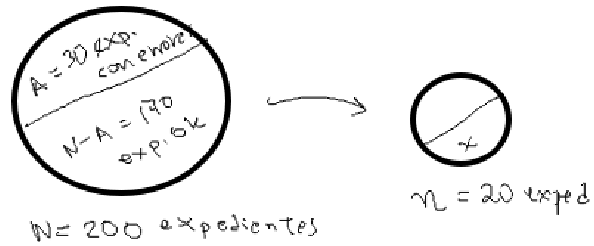
$$0 \leq x \leq 3$$

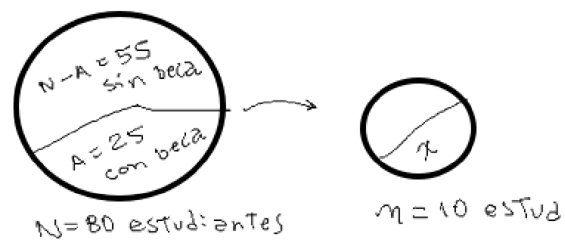
$$x \in \{0, 1, 2, 3\}$$

Ejemplos

- X = Número de focos defectuosos en una muestra
  $X \sim Hiper(N = 100, n = 15, A = 10)$
- Z = Número de expedientes con errores detectados en una auditoría
  $Z \sim Hiper(N = 200, n = 20, A = 30)$
- N = Número de estudiantes que recibieron una beca
  $N \sim Hiper(N = 80, n = 10, A = 25)$

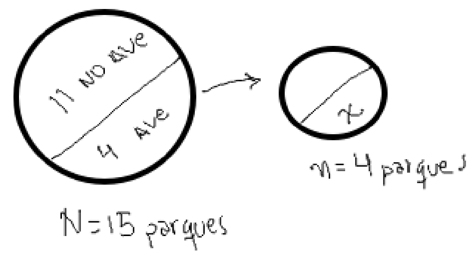






- R = Número de parques nacionales con el ave gallito de las rocas

$R \sim Hiper(N=15, n=4, A=4)$



Hipergeométrica → → → Binomial cuando n/N < 0.05

La probabilidad de encontrar 1 contenedor que no cumpla con los estándares de pureza:

$$P(X = 1) = \frac{\binom{2}{1}\binom{13}{2}}{\binom{15}{3}} = 0.3428571$$

```
dhyper(x = 1, m = 2, n = 13, k = 3) # m = A, n = N-A, k = n
```

A (# éxitos en población)

N − A (# fracasos en población)

n → Tamaño de muestra

```
[1] 0.3428571
```

La probabilidad de encontrar a lo más 1 contenedor que no cumpla con los estándares de pureza:

$$P(X \leq 1) = \frac{\binom{2}{0}\binom{13}{3}}{\binom{15}{3}} + \frac{\binom{2}{1}\binom{13}{2}}{\binom{15}{3}} = 0.9714285$$

```
dhyper(x = 0, m = 2, n = 13, k = 3) + dhyper(x = 1, m = 2, n = 13, k = 3)
```

$p(x=0) + p(x=1)$

```
[1] 0.9714286
```

```
phyper(q = 1, m = 2, n = 13, k = 3)
```

$p(x \leq 1)$

```
[1] 0.9714286
```

Probab. discreta ← dhyper → $P(X = x)$

Probab. Acumul. ← phyper → $P(X \leq x)$

Quantile ← qhyper → percentil

Random ← rhyper → muestra aleatoria

Distribución Poisson

* X = N° de éxitos en un intervalo

tasa de ocurrencia en el intervalo de interés

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x \in \{0, 1, 2, ...\}$$

Ejemplos

- X = Número de llamadas que recibe una central de emergencias en 1 minuto $X \sim Pois(4)$
- R = Número de picaduras en una persona durante una noche en una zona tropical $R \sim Pois(2)$
- E = Número de errores en una página de un libro impreso $E \sim Pois(0.3)$
- A = Cantidad de pacientes que llegan a emergencias en una hora $A \sim Pois(6)$

En promedio, se reciben 4 llamadas por minuto

En promedio, una persona recibe 2 picaduras por noche

En promedio, una página de libro tiene 0.3 errores

X2 = Número de llamadas que recibe una central de emergencias en 5 minutos $\rightarrow X \sim Pois(20)$

$$\mu = \sum x \cdot \frac{e^{-\lambda}\lambda^{x}}{x!} = \bigcirc$$

$$\sigma^2 = \sum x^2 \cdot \frac{e^{-\lambda}\lambda^{x}}{x!} - \mu^2 = \bigcirc$$

$=$

$$\sum_{x=0}^{\infty} \frac{\cancel{x} \cdot e^{-\lambda}\lambda^{x}}{\cancel{x}(x-1)!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x}}{(x-1)!}$$

$$= e^{-\lambda}\left(\frac{\lambda^0}{0!} + \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}\lambda\right)$$

```
1 - ppois(q = 3, lambda = 1.8)
```

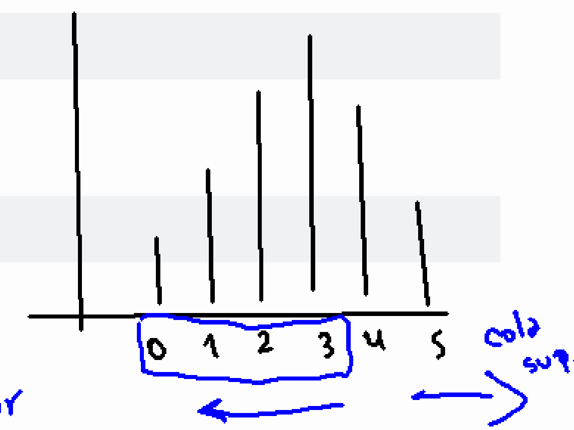
```
[1] 0.1087084
```

NO

```
ppois(q = 3, lambda = 1.8, lower.tail = FALSE)
```

```
[1] 0.1087084
```

$P(X \leq 0)$ cola inferior



$$
\begin{tikzcd}
\text{Hipergeométrica} \arrow[r, "\frac{n}{N} \to 0"] & \text{Binomial} \arrow[d, "\substack{n \to \infty \\ \pi \to 0}"] \\
& \text{Poisson}
\end{tikzcd}
$$

X = Número de personas contagiadas de COVID

$*\quad X \sim Bin(n = 5000, \pi = 0.008) \Rightarrow P(X = 36) =$

```
> dbinom(x = 36, size = 5000, prob = 0.008)
[1] 0.05403984
```

$*\quad Y \sim Pois(\overset{\text{media}}{\lambda} = 40) \Rightarrow P(Y = 36) =$

```
> dpois(x = 36, lambda = 40)
[1] 0.05393184
```

Y = Número de personas contagiadas por COVID en el intervalo .... (intervalo dependiendo de cómo se seleccionaron las 5000 personas, puede ser intervalo temporal, espacial, etc)

Distribución Geométrica

X = Número de $\underbrace{\text{intentos}}$ hasta lograr el 1° éxito

□ E — $X = 1$

□ □ □ □ F F F E — $X = 4$

$$\mu_X = E(X) = \boxed{\frac{1}{\pi}}$$

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

$$\sigma_X^2 = V(X) = \frac{1-\pi}{\pi^2}$$

X = Número de $\overbrace{\text{intentos}}^{\text{fracasos}}$ antes de lograr el 1° éxito (***)

□ E — $X = 0$

□ □ □ □ F F F E — $X = 3$

$$\mu_X = E(X) = \frac{1-\pi}{\pi} = \boxed{\frac{1}{\pi}} - 1$$

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

$$\sigma_X^2 = V(X) = \frac{1-\pi}{\pi^2}$$

Ejemplos

- X = Número de intentos antes de que un cliente realice su primera compra en una tienda online $\rightarrow X \sim Geom(0.2)$ P(compra) = 0.2
- D = Número de inspecciones antes de encontrar el primer producto con algún defecto $\rightarrow D \sim Geom(0.05)$ P(algún defecto) = 0,05
- E = Cantidad de intentos antes de que un estudiante resuelva correctamente un ejercicio sin ayuda $\rightarrow E \sim Geom(0.1)$ P(resolver correct y sin ayuda) = 0,10

- A = Número de veces que desprueba un curso antes de aprobarlo (bajo el supuesto de que no se considere Trika) → A ~ Geom(0.8) ... P(aprobar el curso) = 0.8

- G = Cantidad de tiros al arco antes de meter un gol → G ~ Geom(0.3) ... P(meter un gol) = 0.3

Probabilidad de que se reciba la primera respuesta en el tercer correo enviado:

$$P(X = 2) = 0.85^2 \times 0.15 = 0.108375$$

```
dgeom(x = 2, prob = 0.15)
```

```
[1] 0.108375
```

FRACASO
2 correos sin respuesta

1 correo (3°) con respuesta
ÉXITO

Probabilidad de que se reciba la primera respuesta a más tardar el tercer correo enviado:

$$P(X \leq 2) = 0.85^0 \times 0.15 + 0.85^1 \times 0.15 + 0.85^2 \times 0.15 = 0.108375$$

fracaso fracaso

```
dgeom(x = 0, prob = 0.15) + dgeom(x = 1, prob = 0.15) +
  dgeom(x = 2, prob = 0.15)
```

fracasos

```
[1] 0.385875
```

```
pgeom(q = 2, prob = 0.15)
```

```
[1] 0.385875
```

6 intentos → 5 fracasos

7 intentos → 6 fracasos

6, 7, 8, 9, ...

8 intentos → 7 fracasos

Probabilidad de que se requieran más de 5 intentos para lograr obtener una respuesta:

$$P(X \geq 5) = \sum_{i=5}^{\infty} P(X = i) = P(X = 5) + P(X = 6) + ... = 0.4437$$

5 o más fracasos

```
pgeom(q = 4, prob = 0.15, lower.tail=FALSE)
```

```
[1] 0.4437053
```

```
1-pgeom(q = 4, prob = 0.15)
```

```
[1] 0.4437053
```

$$P(X \geq 5) = 1 - P(X \leq 4)$$

5, 6, 7, 8, ...   0, 1, 2, 3, 4