Distribución Binomial Negativa

X = Número de intentos antes de lograr el 3er éxito (r = 3)

E F F E F E $x = 6$

X = Número de fracasos antes de lograr el 3er éxito (r = 3)

E F F E F E $x = 3$

E E E $x = 0$

Geométrica como caso particular de la binomial negativa:

X = Número de fracasos antes de lograr el 1er éxito (r = 1)

E $x = 0$

F F E $x = 2$

Ejemplos

- X = Número de intentos fallidos antes de que un estudiante acierte 3 respuestas correctas en una trivia $\rightarrow X \sim BinNeg(r = 3, \pi = 0.25)$ Probabilidad de acertar cada respuesta
- V = Número de ventas fallidas antes de conseguir 5 ventas exitosas $\rightarrow V \sim BinNeg(r = 5, \pi = 0.18)$ Probabilidad de lograr una venta exitosa
- R = Número de fallas que ocurren antes de que un robot logre completar correctamente 6 ensamblajes exitosos $R \sim BinNeg(r = 6, \pi = 0.3)$

Probabilidad de lograr un ensamblaje exitoso

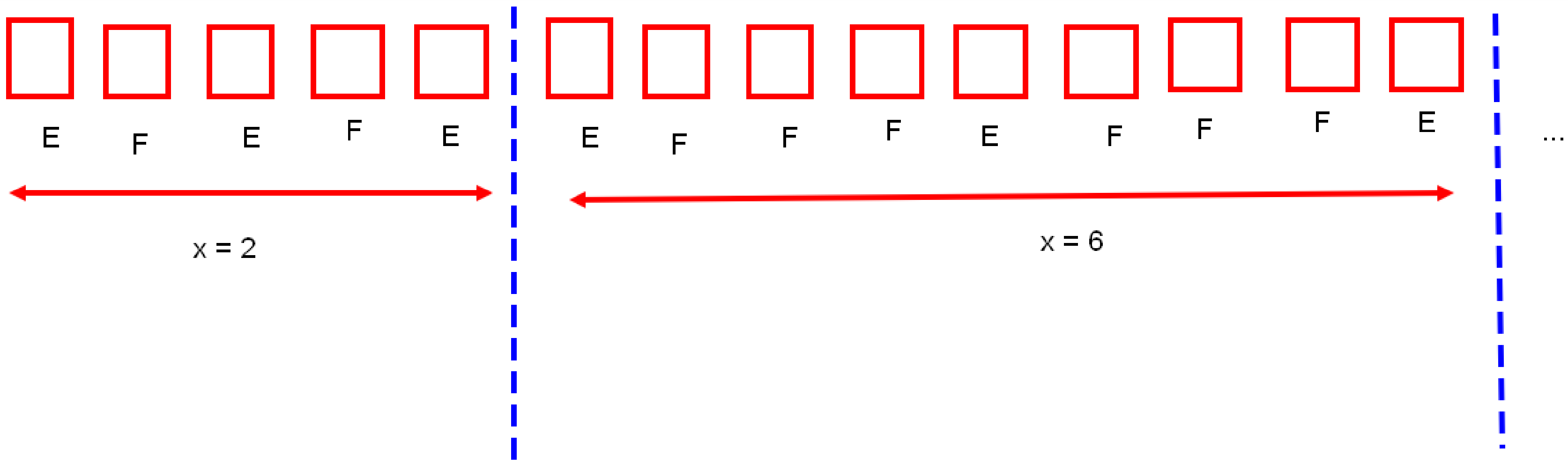
- G = Número de tiros al arco fallidos antes de que un jugador meta su segundo gol, G ~ BinNeg(r = 2, pi = 0.30)

Tercer cuartil de la cantidad de fracasos antes de lograr 3 ventas

```r
qnbinom(p = 0.75, size = 3, prob = 0.2)
```

```
[1] 16
```

En al menos el 75% de los casos, el vendedor fallará como máximo 16 veces antes de cerrar existosamente 3 ventas.



Poisson $\mu = \sigma^2$

Binomial Negativa

$$r \times \frac{1-\pi}{\pi} < r \times \frac{1-\pi}{\pi^2}$$

$$\mu < \sigma^2$$

(sobredispersión)

datos
x1
x2
x3 $\rightsquigarrow$ ¿BN o Poisson?
x4
x5
x6
...

- Distribución Conway–Maxwell–Poisson

Binomial
Cantidad de **éxitos** en **n** ensayos
Cantidad de veces que **gana un equipo** en **8** partidos que juega
Al final, observamos **8 partidos** (no 6, no 5, no 10, no 1)

Binomial negativa
Cantidad de **fracasos** antes de los r éxitos
Cantidad de veces que **pierde un equipo** antes de que gane 3 partidos
Al final, observamos 3 o más partidos

Y = Número de personas vacunadas que **desarrollan la enfermedad por la cual se vacunaron**
Y ~ Bin(n = 30, pi = **0.005**)
Y = f(X1, X2, X3, ..), donde X1, X2, X3, ... son factores clínicos, demográficos, etc que afectan el desarrollo o no desarrollo de la enfermedad → modelo, p. ej. regresión logística.

$$P(Y \leq 4) = P(Y=0) + P(Y=1) + P(Y=2) + P(Y=3) + P(Y=4)$$

DISTRIBUCIONES CONTINUAS

Distribución Uniforme

Función de densidad

Se define la V.A.C. X que toma valores en el intervalo $[a, b]$ donde $a < b$, cuya función de densidad viene dada por:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{de otro modo} \end{cases}$$

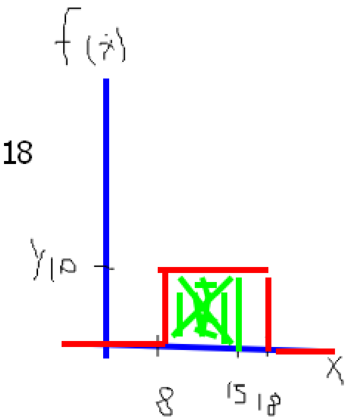
Notación: $X \sim Unif(a, b)$

Función de distribución acumulada

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a, b] \\ 1, & x > b \end{cases}$$

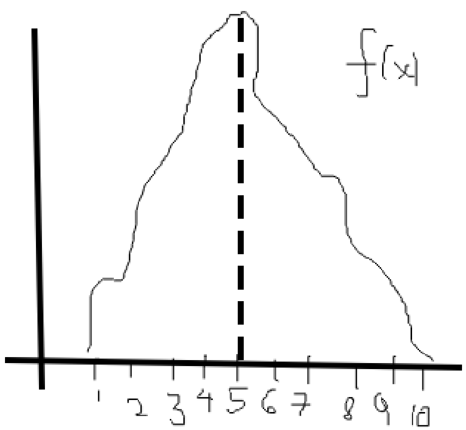
a b

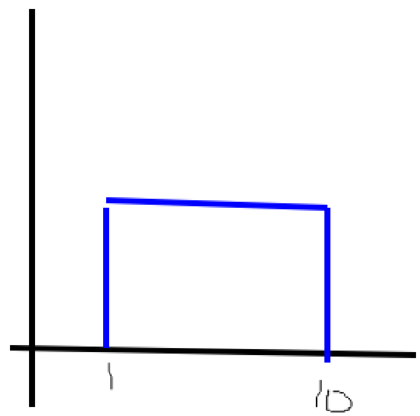
X ~ Unif(8, 18)
f(x) = 1/10, 8<=x<=18
f(10) = 1/10
f(15) = 1/10
f(3) = 0

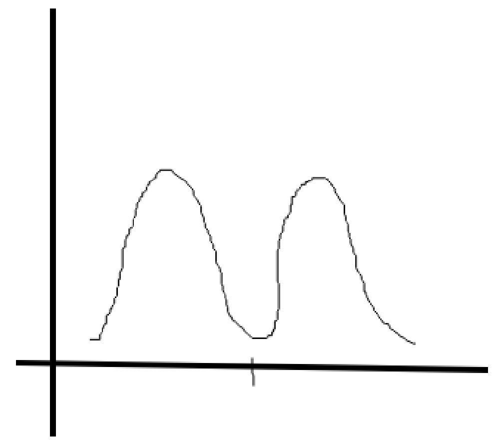


P(X<=15)= F(15)=(15-8)/(18-8)=7/10=0.7
P(X<=11)= F(11)=(11-8)/(18-8)=3/10=0.3
P(X<=20)= F(20)= 1
P(X<=6) = F(6) = 0

En R: `punif`







$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & \text{de otro modo} \end{cases}$$

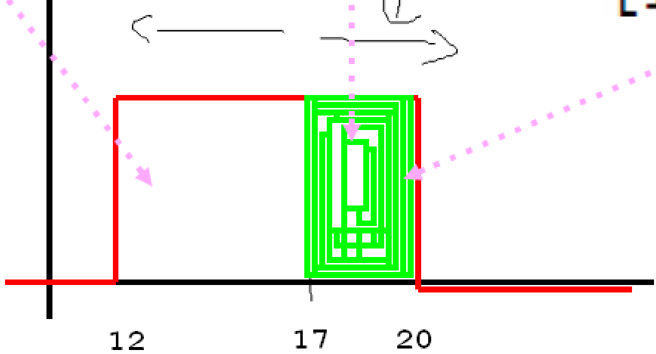
Probabilidad de que una inspección ocurra luego de los 17 segundos

$$P(X > 17) = 1 - P(X \leq 17) = 1 - F(17) = 1 - \frac{17-12}{8} = 0.375$$

```
punif(q = 17, min = 12, max = 20, lower.tail = FALSE)
```

```
[1] 0.375
```

```
> punif(q=17, min=12, max=20)
[1] 0.625
```




```
> 1-punif(q=17, min=12, max=20)
[1] 0.375
```

Percentil 90 del tiempo de inspección:

Se ha acumulado el
90% de probabilidad

$$F(x) = 0.90 \rightarrow \frac{x - 12}{8} = 0.9 \rightarrow x = 8 \times 0.9 + 12 = 19.2$$

```r
qunif(p = 0.9, min = 12, max = 20)
```

```
[1] 19.2
```

Esto significa que el 90% de las inspecciones ocurre hasta los 19.2 segundos.

$$f(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

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

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

$\lambda e^{-\lambda x}$

<u>No es lambda de Poisson</u>, en este caso, lambda será la inversa de la media de Y

$$f(x) = \begin{cases} \frac{1}{\beta} \exp\left(-\frac{1}{\beta} x\right), & x \geq 0 \\ 0, & \text{de otro modo} \end{cases}$$

$$\mu_X = E(X) = \beta$$

$$\sigma_X^2 = V(X) = \beta^2$$

Propiedad de falta de memoria

$$P(X > s + t | X > s) = P(X > t)$$

Es decir, el tiempo restante no depende de cuánto ya se ha esperado.

X = Tiempo que transcurre entre el reporte de un contagio y otro
X = Tiempo que transcurre entre una atención y otra

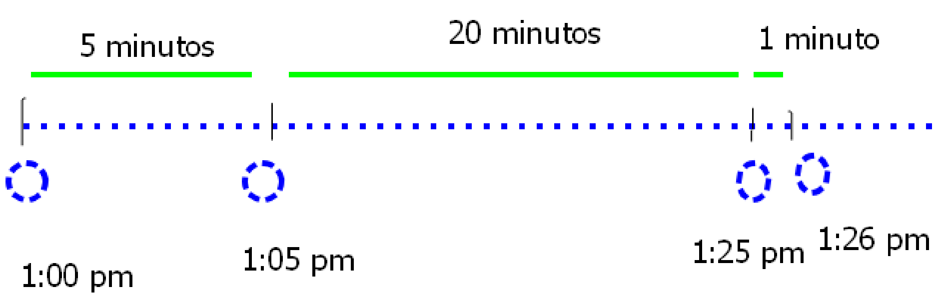
9 min

P(X>39|X>30)=P(X>9)

P(X>15|X>6)=P(X>9)

P(X>18|X>9)=P(X>9)

P(X>10|X>1)=P(X>9)



La probabilidad de que el próximo cliente llegue antes de los 4 minutos:

$$P(X < 4) = F(4) = 1 - \exp\left(-\frac{4}{6}\right) = 0.48658$$

```r
pexp(q = 4, rate = 1/6)
```

λ

```
[1] 0.4865829
```

Distribución Gamma

Función de densidad

Se define la V.A.C. X con distribución Gamma con parámetros $\alpha > 0$ y $\beta > 0$, cuya función de densidad viene dada por:

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Cuando alfa = 1, se llega a una exponencial:

$$f\left(x\right) = \frac{1}{\Gamma(1)\beta^1} x^{1-1} \exp\left(-\frac{x}{\beta}\right) = \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right)$$

El tiempo medio que transcurre hasta 3 eventos es:

$$\mu_X = E(X) = 3 \times 10 = 30$$

Su varianza:

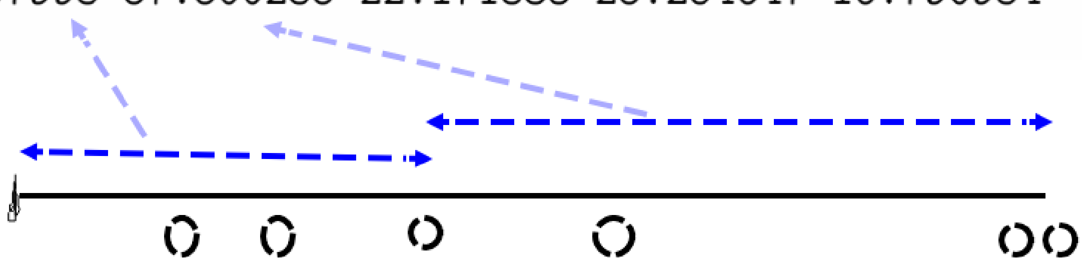
$$\sigma 2_X = V(X) = 3 \times 10^2 = 300$$

Poblacional
(teórico)

Muestra aleatoria de 7 tiempos que transcurren hasta que sucedan 3 eventos:

```r
set.seed(852)
rgamma(n = 7, shape = 3, scale = 10)
```

```
[1] 25.997993 57.500235 22.171333 25.254647 16.790934  9.310327 47.792846
```



Probabilidad de que ocurran los 3 eventos antes de 25 minutos:

$$P(X < 25) = 0.4562$$

```r
pgamma(q = 25, shape = 3, scale = 10)
```

```
[1] 0.4561869
```

Probabilidad de que ocurran los 3 eventos después de 25 minutos:

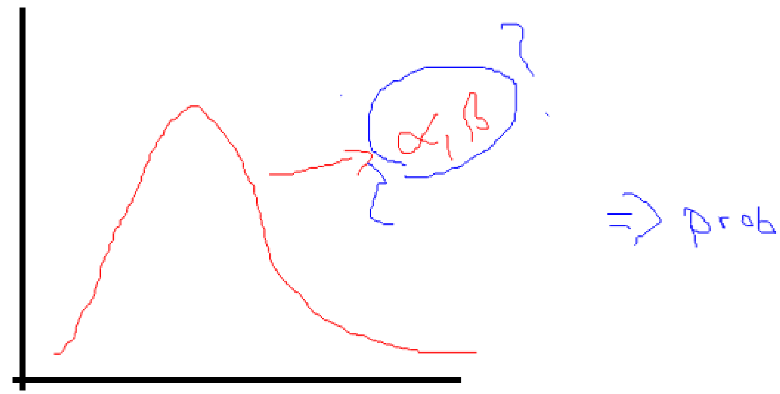
$$P(X > 25) = 0.4562$$

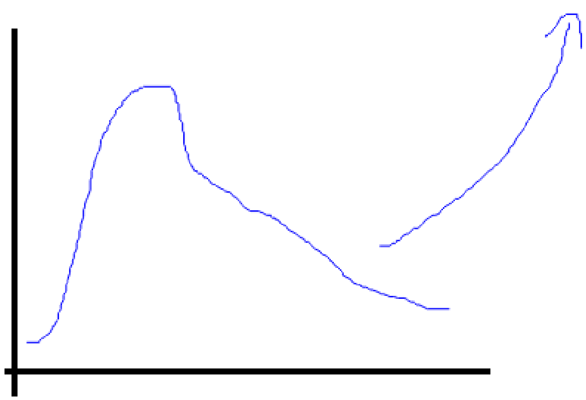
```r
> 1-pgamma(q = 25, shape = 3, scale = 10)
[1] 0.5438131
> pgamma(q = 25, shape = 3, scale = 10, lower.tail = FALSE)
[1] 0.5438131
```

(Método de máxima verosimilitud)

Base de datos

f(x)
"la verdad"





Base de datos

f(x)
"la verdad"

Prueba de bondad de ajuste