PRÁCTICA 8: Distribuciones estadísticas discretas.

DESARROLLO DE LA PRÁCTICA

En esta práctica aprenderemos a trabajar con las principales distribuciones discretas mediante el software estadístico R. Si en RStudio, en la pestaña de ayuda (Help) escribimos "**distributions**". Podemos ver las distribuciones usuales con que trabaja el software R, tanto discretas como continuas. Si disponemos de una distribución discreta xxx, entonces podemos calcular:

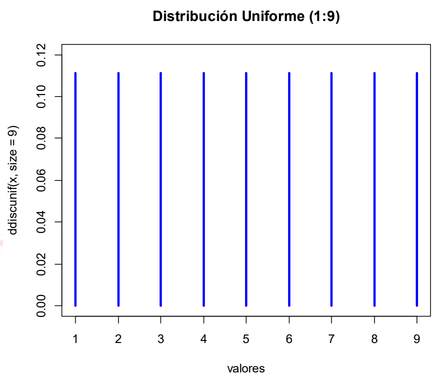
- **a) dxxx** (density functions) distribución de probabilidades de la variable xxxx
- **b) pxxx** (probability distribution functions) la función de distribución de la variable xxxx
- **c) qxxx** (quantile functions) los cuantiles de la distribución de la variable xxxx
- **d) rxxx** (random number generation) la generación de valores aleatorios de la variable xxxx

solo nos resta observar el nombre xxx que R asigna a cada distribución en particular.

EJEMPLO 1. DISTRIBUCIÓN UNIFORME . $X \sim U(1,2,\ldots,n)$.
Si la variable aleatoria X tiene una distribución uniforme sobre n números, entonces las probabilidades de que tome ciertos valores vienen dadas por

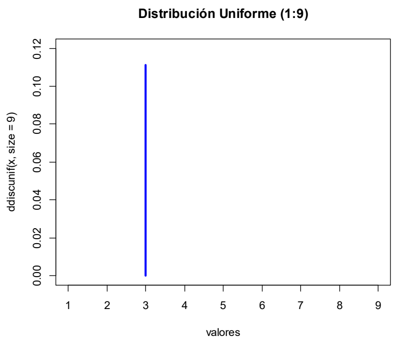
$$P(X=x)=\frac{1}{n},\quad x=1,\ldots,n$$

con los siguientes valores de media $E(X)=\frac{n+1}{2}$ y la varianza $V(X)=\frac{n^2-1}{12}$. Su nombre en R es **"discunif"** exclusivamente para la librería "HH"**.**
Por ejemplo, para la distribución uniforme sobre 9 números, tenemos:



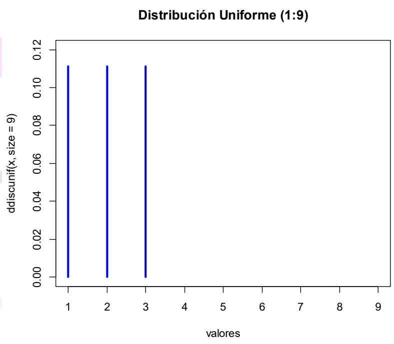
a) ddiscunif(3,size=9) **# valor de la probabilidad en x=3#**

Si $X \sim U(1,2,\ldots,9)$ se tiene $P(X=3)=\frac{1}{9}=0.1111111$



b) pdiscunif(3,size=9) **# función de distribución en 3**

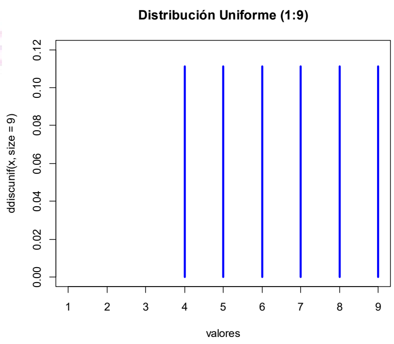
$$F(3) = P(X \le 3) = \sum_{x=1}^{3} \frac{1}{9} = 0.333333$$



Ahora podemos combinar la función de distribución y las probabilidades de una variable para calcular la probabilidad de determinados sucesos.

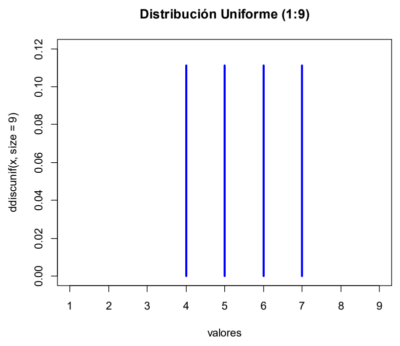
b1) 1-pdiscunif(3,size=9) **# 1 - función de distribución en 3**

$$1 - F(3) = 1 - P(X \le 3) = P(X > 3) = \sum_{x=4}^{9} \frac{1}{9} = 0.6666667 .$$



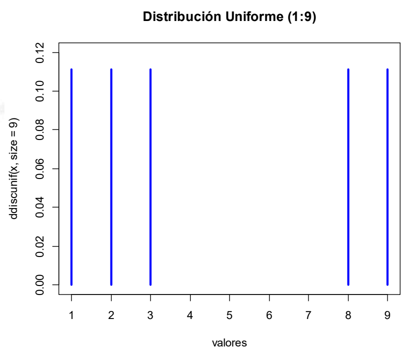
b2) pdiscunif(7,size=9) -pdiscunif(3,size=9) **#combinando dos expresiones**

$$F(7) - F(3) = P(3 < X \leq 7) = \sum_{x=4}^{7} \frac{1}{9} = 0.444444$$



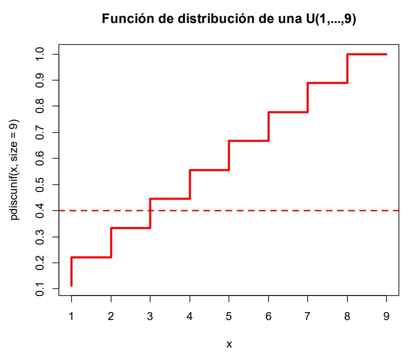
b3) 1-(pdiscunif(7,size=9) -pdiscunif(3,size=9)) **#combinando expresiones**

$$P(1 \leq X \leq 3) + P(X > 7) = \sum_{x=1}^{3} \frac{1}{9} + \sum_{x=8}^{9} \frac{1}{9} = 0.5555555$$



c) qdiscunif(0.4,size=9) **# cuantil de orden 0.4 de la uniforme discreta utilizada**

3



EJEMPLO 2. DISTRIBUCIÓN BINOMIAL . $X \sim Bi(n,p)$.

Si la variable aleatoria X tiene una distribución binomial $Bi(n,p)$ entonces las probabilidades de que tome ciertos valores vienen dadas por

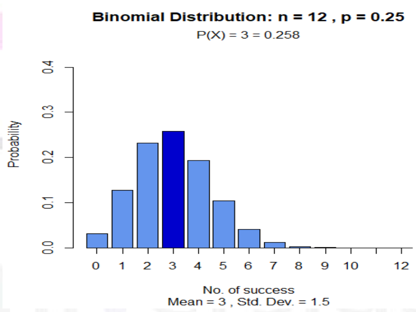
$$P(X=x)=\binom{n}{x}p^x q^{n-x}\ ,\quad x=0,1,\ldots,n$$

con los siguientes valores de media $E(X)=np$ y la varianza $V(X)=np(1-p)$. Su nombre en R es **"binom".**

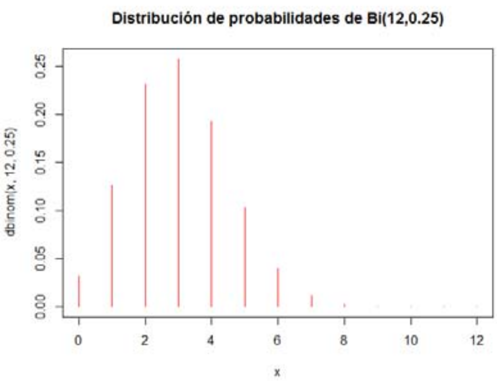
Por ejemplo, para la distribución Binomial $Bi(12,0.25)$ tenemos:

a) dbinom(3,12, 0.25) **# valor de la probabilidad en x=3 para una Bi(12,0.25)**

Si n=12 y p=0.25 se tiene $P(X=3)=\binom{12}{3}0.25^3 0.75^9 = 0.2581036$

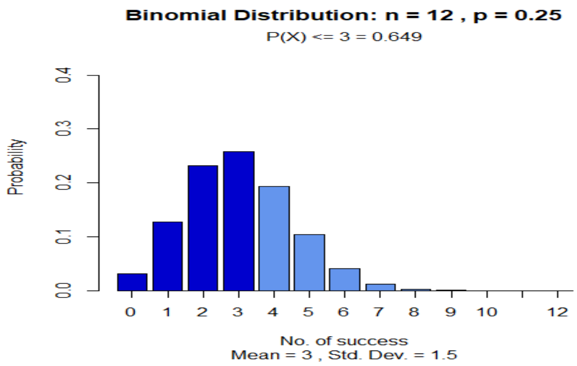


Una gráfica más apropiada, que la proporcionada por "visualize" o "descrptive" para esta distribución binomial es



b) pbinom(3,12, 0.25) **# función de distribución en 3 para una Bin(12,0.25)**

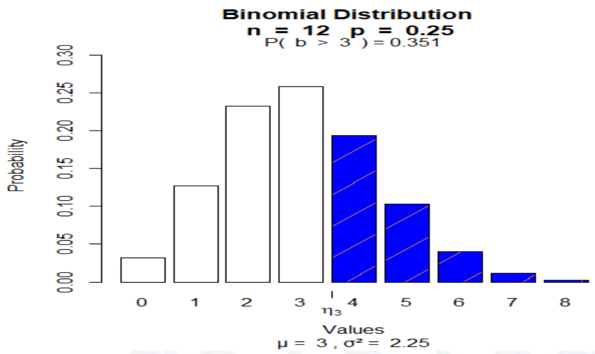
$$F(3)=P(X\le 3)=\sum_{x=0}^{3}\binom{12}{x}0.25^x 0.75^{12-x} = 0.6487786$$



Ahora podemos combinar la función de distribución y las probabilidades de una variable para calcular la probabilidad de determinados sucesos.

b1) pbinom(3,12, 0.25, lower.tail=FALSE) **# 1 - función de distribución en 3**

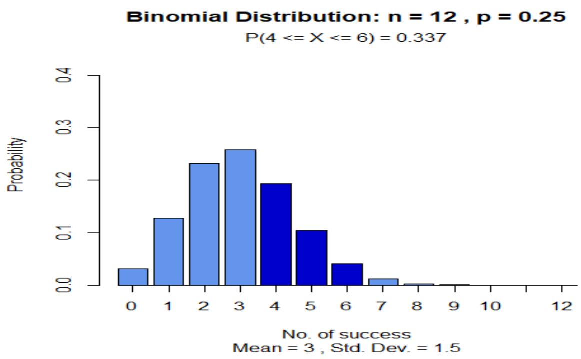
$$1-F(3)=1-P(X\leq 3)=P(X>3)=\sum_{x=4}^{12}\binom{12}{x}0.25^{x}0.75^{12-x}=0.3512214\,.$$



CUIDADO CON LA GRAFICA

b2) pbinom(6,12, 0.25) -pbinom(3,12, 0.25) **#combinando dos expresiones**

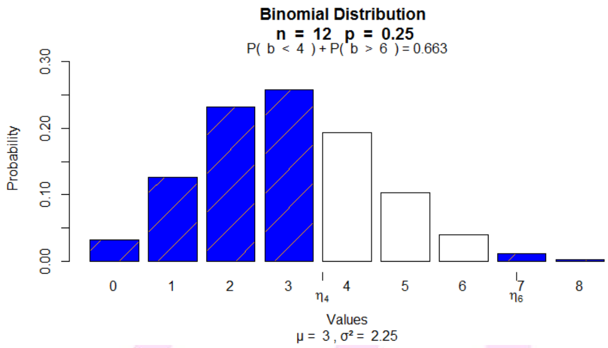
$$F(6)-F(3)=P(3<X\leq 6)=\sum_{x=4}^{6}\binom{12}{x}0.25^{x}0.75^{12-x}=0.3369686$$



---

b3) 1-[pbinom(6,12, 0.25) -pbinom(3,12, 0.25)] **#combinando expresiones**

$$P(0 \le X \le 3) + P(X > 6) = \sum_{x=0}^{3} \binom{12}{x} 0.25^x 0.75^{12-x} + \sum_{x=7}^{12} \binom{12}{x} 0.25^x 0.75^{12-x} = 0.6630314$$



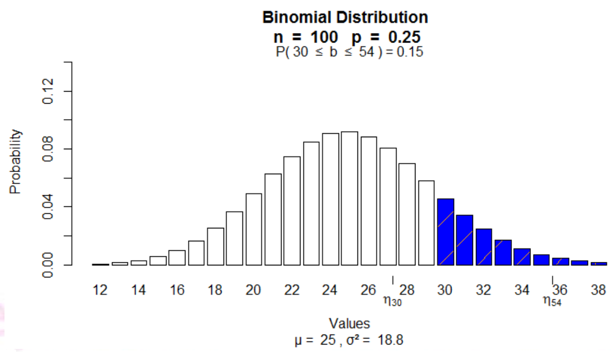
CUIDADO CON LA GRAFICA

b4) Para una binomial de parámetros $Bi(100, 0.25)$ la sentencia

sum(dbinom(30:54, 100, 0.25))

permite evaluar la expresión

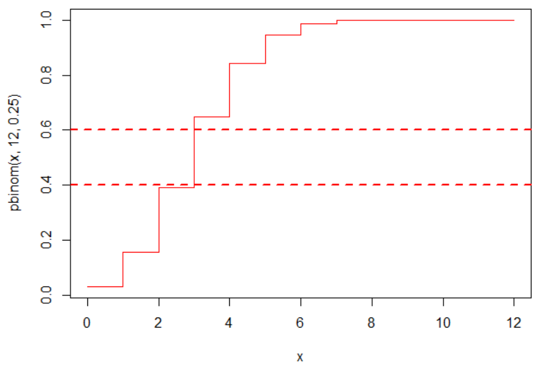
$$P(30 \le X \le 54) = \sum_{x=30}^{54} \binom{100}{x} 0.25^x 0.75^{100-x} = 0.149541$$



CUIDADO CON LA GRAFICA

c) qbinom(0.4,12, 0.25) **# cuantil de orden 0.4 de la Binomial utilizada**

3



c1) qbinom(0.4,12, 0.25, lower.tail=FALSE) # cuantil de orden 0.6 de la Binomial utilizada

3

EJEMPLO 3. DISTRIBUCIÓN GEOMÉTRICA $X \sim Geo(p)$

Si la variable aleatoria X tiene una distribución geométrica $Geo(p)$ entonces las probabilidades de que tome ciertos valores vienen dadas por

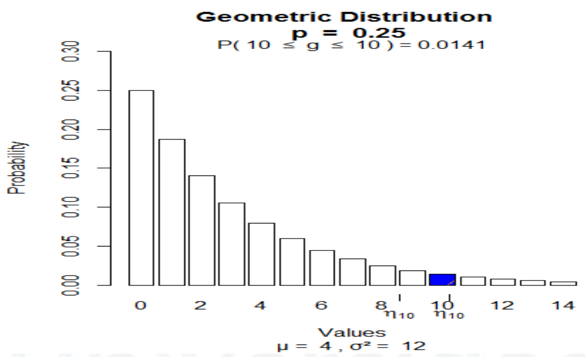
$$P(X=x) = pq^x, \quad x = 0,1,2,\ldots$$

con los siguientes valores de media $E(X) = \frac{1-p}{p}$ y varianza $V(X) = \frac{1-p}{p^2}$. Su nombre en R es **"geom".**

Por ejemplo, tomando una variable Geométrica de parámetro $Geo(0.25)$:

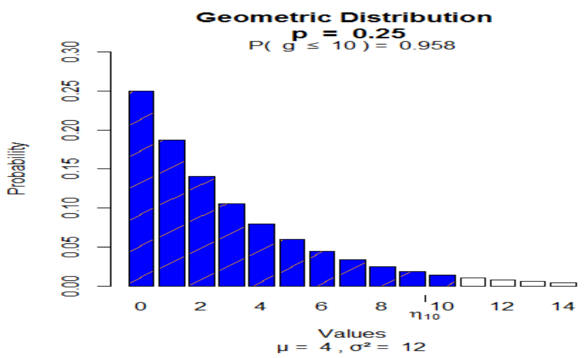
a) dgeom(10,0.25) **# valor de la probabilidad en 10 de una Geo(0.25)**

Si p=0.25 se tiene $P(X=10) = 0.25 * 0.75^{10} = 0.01407838$



b) pgeom(10,0.25) **# función de distribución en 10**

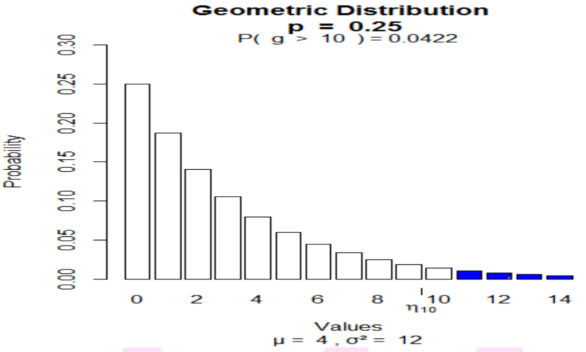
$$F(10) = P(X \le 10) = \sum_{x=0}^{10} 0.25 * 0.75^x = 0.9577649$$



Ahora podemos combinar la función de distribución y las probabilidades de una variable para calcular la probabilidad de determinados sucesos.

b1) pngeom(10,0.25, lower.tail=FALSE) **# 1 - función de distribución en 10**

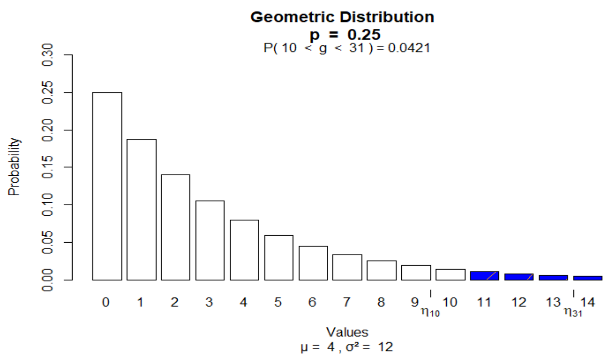
$$1-F(3)=1-P(X\leq 10)=P(X>10)=\sum_{x=11}^{\infty}0.25*0.75^{x}=0.04223514$$



CUIDADO CON LA GRAFICA

b2) pngeom(30,0.25) -pngeom(10,0.25) **#combinando dos expresiones**

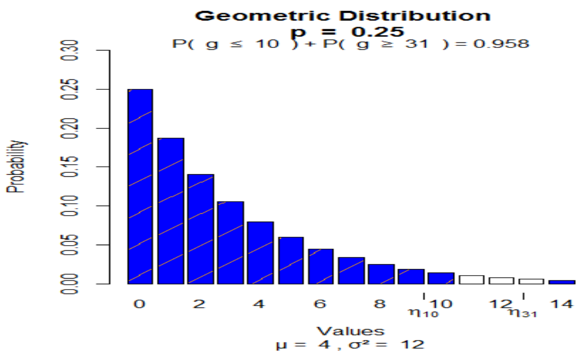
$$F(30)-F(10)=P(11<X\leq 30)=\sum_{x=11}^{30}0.25*0.75^{x}=0.0421012$$



CUIDADO CON LA GRAFICA

b3) 1-(pgeom(30,0.25) -pgeom(10,0.25)) **#combinando expresiones**

$$P(0\leq X\leq 10)+P(X>30)=\sum_{x=0}^{10}0.25*0.75^{x}+\sum_{x=31}^{\infty}0.25*0.75^{x}=0.9578988$$



CUIDADO CON LA GRAFICA

c) qgeom(0.38,0.25) **# cuantil de orden 0.38 de una Geo(0,25)**

1

c1) qgeom(0.38,0.25, lower.tail=FALSE) **# cuantil de orden 0.62 de una Geo(0,25)**

3

EJEMPLO 4. DISTRIBUCIÓN BINOMIAL NEGATIVA $X \sim BiNe(r,p)$

Si la variable aleatoria X tiene una distribución binomial negativa $BiNe(r,p)$ entonces las probabilidades de que tome ciertos valores vienen dadas por

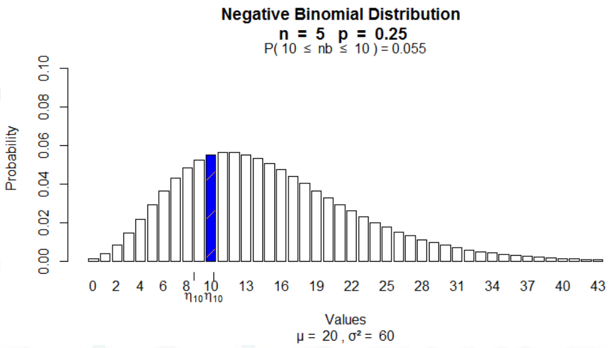
$$P(X = x) = \binom{x+r-1}{r-1} p^r q^x, \quad x = 0,1,2,\dots \; r = 1,2,\dots$$

con los siguientes valores de media $E(X) = \dfrac{r(1-p)}{p}$ y varianza $V(X) = \dfrac{r(1-p)}{p^2}$; y su nombre en R es **"nbinom".**

Por ejemplo, tomando una variable Binomial Negativa de parámetros $BiNe(5, 0.25)$:

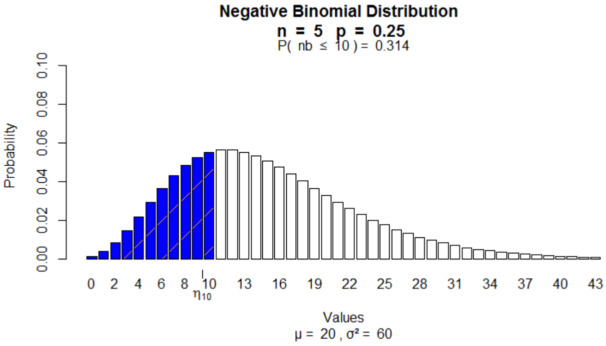
a) dnbinom(10,5, 0.25) **# valor de la probabilidad en 10**

Permite evaluar $P(X = 10) = \binom{10+5-1}{5-1} 0.25^5 0.75^{10} = 0.05504866$.



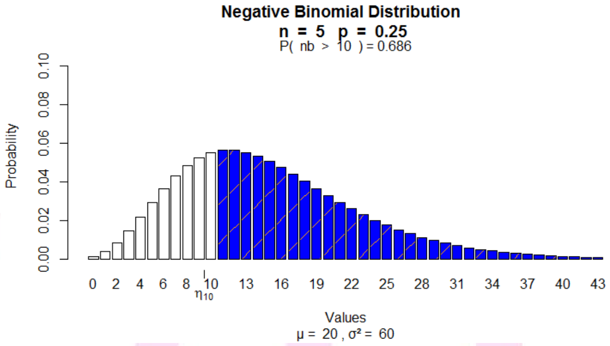
b) pnbinom(10,5, 0.25) **# función de distribución en 10**

$$F(10) = P(X \le 10) = \sum_{x=0}^{10} \binom{x+5-1}{5-1} 0.25^5 0.75^x = 0.3135141$$



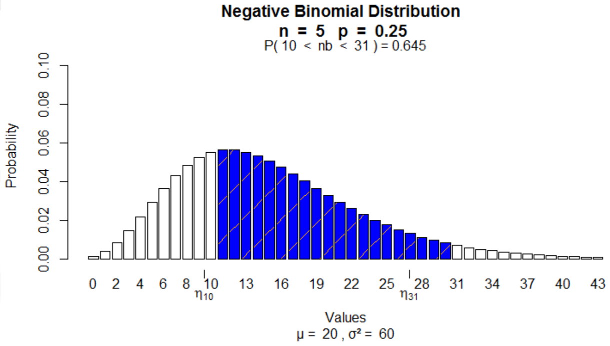
b1) pnbinom(10,5, 0.25, lower.tail=FALSE) **# 1 - función de distribución en 10**

$$1-F(10)=1-P(X\le 10)=P(X>10)=\sum_{x=11}^{\infty}\binom{x+5-1}{5-1}0.25^5 0.75^x=0.6864859$$



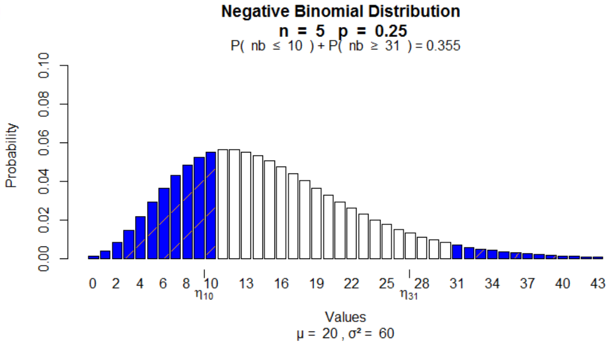
b2) pnbinom(30,5, 0.25) -pnbinom(10,5, 0.25) **#combinando dos expresiones**

$$P(X\le 30)-P(X\le 10)=\sum_{x=11}^{30}\binom{x+5-1}{5-1}0.25^5 0.75^x=0.6454804$$



b3) 1-(pnbinom(30,5, 0.25) -pnbinom(10,5, 0.25)) **#combinando expresiones**

$$1-(P(X\le 30)-P(X\le 10))=P(X\le 10)+P(X>30)=0.3545196$$

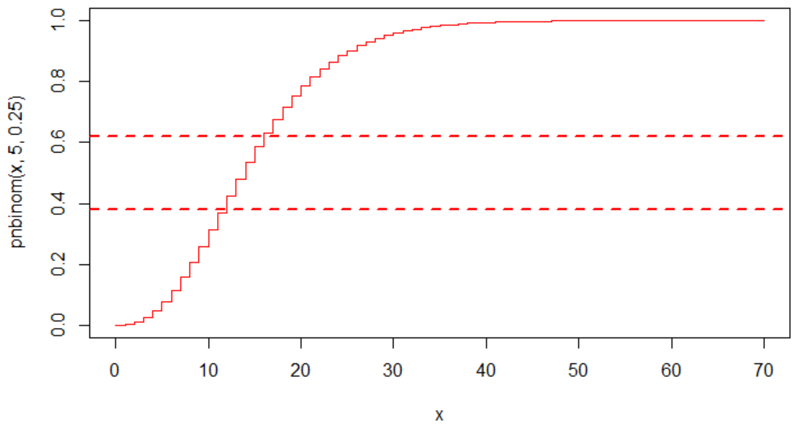


c) qnbinom(0.38,5,0.25) **# cuantil de orden 0.38 de la Binomial Negativa utilizada**

12

c1) qnbinom(0.38,5,0.25,lower.tail=FALSE) **# cuantil de orden 0.62 de la Binomial Negativa**

16



EJEMPLO 5: DISTRIBUCIÓN HIPERGEOMÉTRICA $X \sim H(N,n,r)$

La notación en R es $X \sim H(m,n,k)$ con m=n, n=N-n, m+n=N y k=r para enlazar con nuestra notación. Si la variable aleatoria X tiene una distribución hipergeométrica $H(m,n,k)$ entonces las probabilidades de que tome ciertos valores vienen dadas por

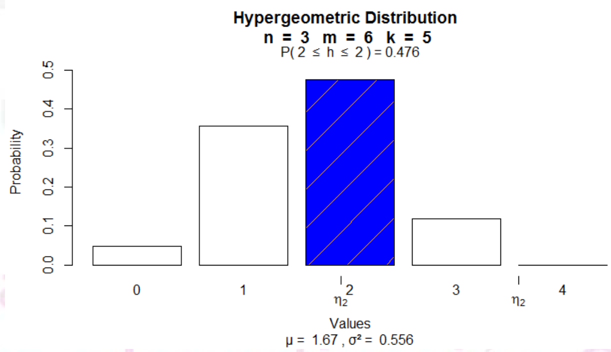
$$P(X=x)=\frac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}, \text{ x=0,1,...,k. } [\max(0,k-n)\le x\le \min(m,k)]$$

con los siguientes valores de media y varianza, $E(X)=\frac{km}{m+n}$ y $V(X)=\frac{(m+n-k)}{m+n-1}\frac{km}{m+n}(1-\frac{m}{m+n})$; y su nombre en R es **"hiper"**

Por ejemplo, tomando una variable Hipergeométrica de parámetros $H(m,n,k)=H(3,6,5)$,tenemos

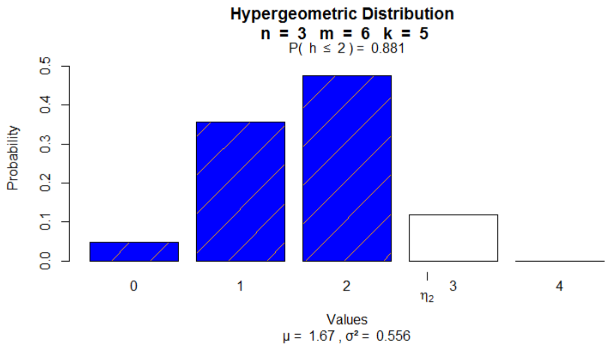
a) dhyper(2,3,6,5) **# valor de la probabilidad en 2**

Permite evaluar $P(X=2)=\frac{\binom{3}{2}\binom{6}{5-2}}{\binom{9}{5}}=0.4761905$.



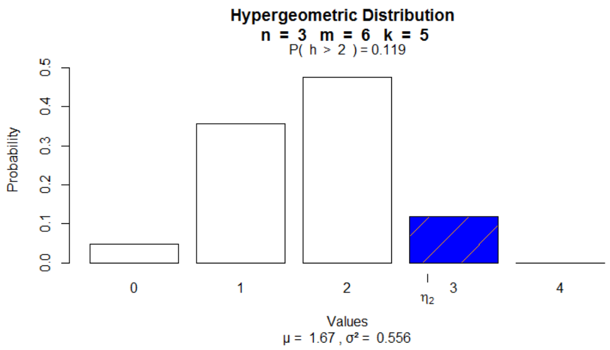
b) phyper(2,3,6,5) **# función de distribución en 2**

$$F(2)=P(X\le 2)=\sum_{x=0}^{2}\frac{\binom{3}{x}\binom{6}{5-x}}{\binom{9}{5}}=0.8809524$$



b1) phyper(2,3,6,5,lower.tail=FALSE) **# 1 - función de distribución en 2**

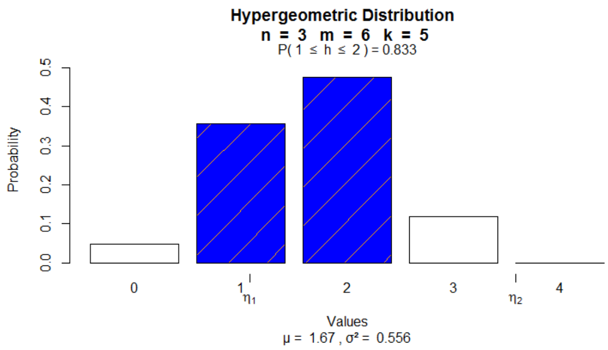
$$1-F(2)=1-P(X\le 2)=P(X>2)=\frac{\binom{3}{3}\binom{6}{5-3}}{\binom{9}{5}}=0.1190476$$



CUIDADO CON LA GRAFICA

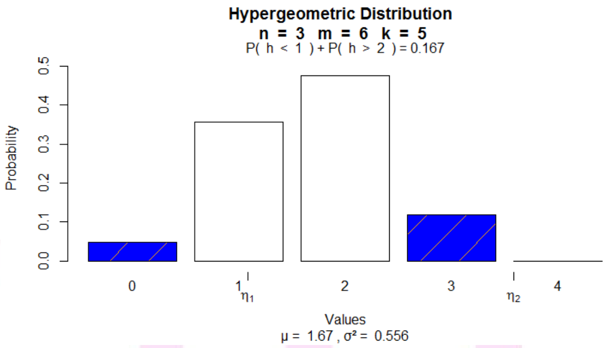
b2) phyper(2,3,6,5) -phyper(0,3,6,5) **#combinando dos expresiones**

$$P(X\le 2)-P(X=0)=\sum_{x=1}^{2}\frac{\binom{3}{x}\binom{6}{5-x}}{\binom{9}{5}}=0.8333333$$



b3) 1-(phyper(2,3,6,5) -phyper(0,3,6,5)) **#combinando expresiones**

$$1-(P(X \leq 2)-P(X=0))=P(X=0)+P(X=3)=0.1666667$$



c) qhyper(0.40,3,6,5) **# cuantil de orden 0.4 de la Hipergeométrica utilizada**

1

c1) qhyper(0.40,3,6,5, lower.tail=FALSE) **# cuantil de orden 0.6 de la Hipergeométrica utilizada**

2

EJEMPLO 6: DISTRIBUCIÓN DE POISSON $X \sim P(\lambda)$

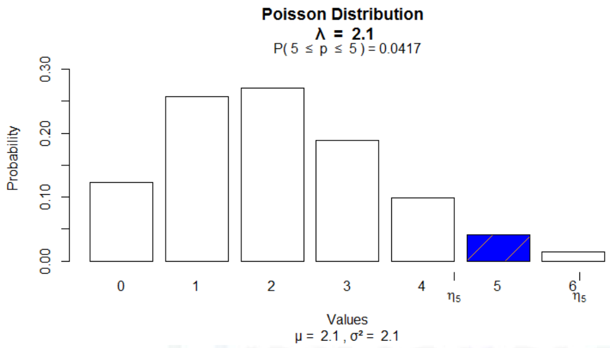
Si la variable aleatoria X tiene una distribución de Poisson $P(\lambda)$ entonces las probabilidades de que tome ciertos valores vienen dadas por

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0,1,\ldots,n$$

con los siguientes valores de media $E(X) = \lambda$ y varianza $V(X) = \lambda$. Su nombre en R es **"pois".**
Por ejemplo, si tomamos una variable de Poisson de parámetro $\lambda = 2.1$, tenemos:

a) dpois(5,2.1) **# valor de la probabilidad en 5 de una P(2.1)**

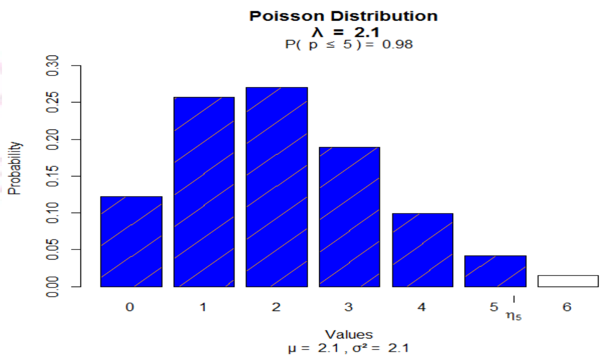
Si $\lambda = 2.1$ se tiene $P(X = 5) = \dfrac{e^{-2.1}2.1^{5}}{5!} = 0.04167704$



CUIDADO CON LA GRAFICA

b) ppois(5,2.1) **# función de distribución en 5**

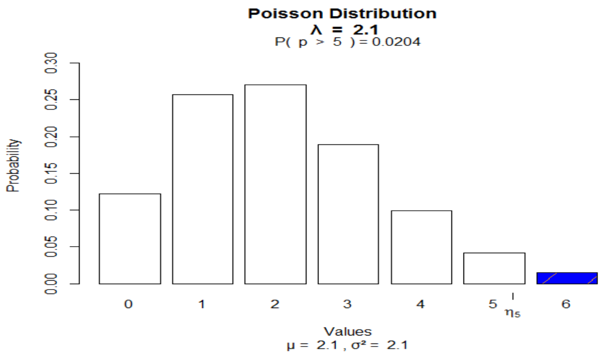
$$F(5) = P(X \le 5) = \sum_{x=0}^{5} \frac{e^{-2.1} * 2.1^{x}}{x!} = 0.9795509$$



CUIDADO CON LA GRAFICA

b1) ppois(5,2.1, lower.tail=FALSE) **# 1 - función de distribución en 5**

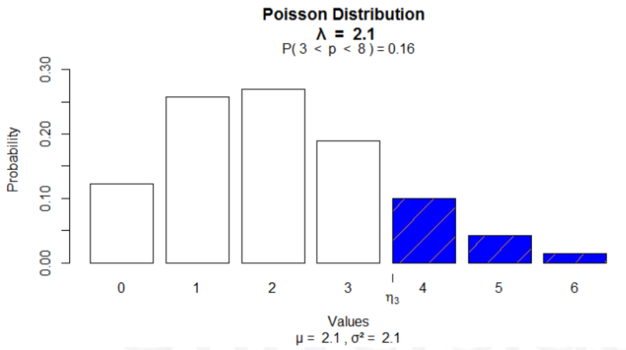
$$1 - F(5) = 1 - P(X \le 5) = P(X > 5) = \sum_{x=6}^{\infty} \frac{e^{-2.1} * 2.1^{x}}{x!} = 0.02044908$$



CUIDADO CON LA GRAFICA

b2) ppois(7,2.1)- ppois(3,2.1) **#combinando dos expresiones**

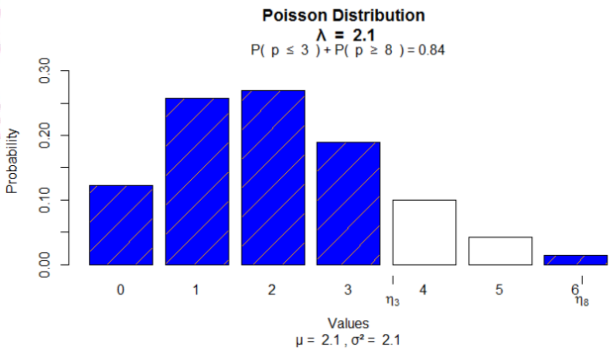
$$P(3 < X \le 7) = \sum_{x=4}^{7} \frac{e^{-2.1} 2.1^{x}}{x!} = 0.1598711$$



CUIDADO CON LA GRAFICA

b3) 1-[ppois(7,2.1)- ppois(3,2.1)] **#combinando expresiones**

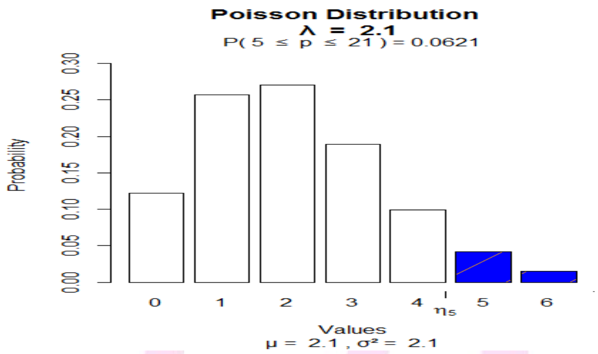
$$P(0 \le X \le 3) + P(X > 7) = \sum_{x=0}^{3} \frac{e^{-2.1} * 2.1^{x}}{x!} + \sum_{x=8}^{\infty} \frac{e^{-2.1} * 2.1^{x}}{x!} = 0.8401289$$



CUIDADO CON LA GRAFICA

b4) sum(dpois(5:21, 2.1)) **#combinando expresiones**

$$P(5 \le X \le 21) = \sum_{x=5}^{21} \frac{e^{-2.1} * 2.1^{x}}{x!} = 0.06212612$$



CUIDADO CON LA GRAFICA

c) qpois(0.35, 2.1) **# cuantil de orden 0.35 de la Poisson utilizada**

1

c1) qpois(0.35,2.1, lower.tail=FALSE) **# cuantil de orden 0.65 de la Poisson utilizada**

3

EJERCICIOS:

1.- En una lista de 1000 correos electrónicos dirigidos a un determinado usuario existe uno con información confidencial muy importante. Un servidor de correo ordena los mensajes de forma aleatoria para remitirlos al usuario. Si este los lee en el orden remitido por el servidor:

a) Hallar la probabilidad de que el mensaje importante sea leído en décimo lugar.
b) Hallar la probabilidad de que el mensaje importante sea leído después de leer k mensajes.
c) ¿Cuál es número esperado de mensajes que debe leer el usuario antes de leer el mensaje importante?

2.- El dos por mil de los DVDs de una determinada marca tienen defectos. Se adquiere una partida de 3500 de estos DVDs. Determinar

a) Número esperado de DVDs sin defectos.
b) Probabilidad de que el número de DVDs defectuosos sea menor o igual que 7.
c) Sabiendo que el número de DVDs defectuosos es mayor o igual que 5, hallar la probabilidad de que dicho número sea mayor o igual que 10.

3.- Un proveedor informático tiene una partida de discos duros de los cuales el 10% tiene algún tipo de error. Un cliente, sin saberlo, compra 20 de esos discos duros. Calcular:

a) Probabilidad de que el cliente adquiera tres discos duros con algún error.
b) Probabilidad de que el cliente adquiera discos duros sin errores.
c) El número esperado de discos defectuosos.

4.-. Sobre una determinada red informática, el intento de conexión es fallido en 5 de cada 30 veces. Hallar

a) La probabilidad de que un usuario no pueda conectarse antes del tercer intento.
b) La probabilidad de que se produzcan dos fallos antes de conseguir realizar diez conexiones con éxito.
c) La probabilidad de que, si realiza ocho intentos, el número de conexiones correctas sea, al menos igual a siete.

5.- El doce por mil de los CDs de una determinada marca resultan defectuosos.

a) Al elegir los CDs aleatoriamente, ¿cuál es la probabilidad de que el primer defectuoso aparezca en sexto lugar? En este caso, ¿cuál es el número esperado de CDs sin defectos antes de que aparezca el primer defectuoso?
b) Al elegir los CDs aleatoriamente, ¿cuál es la probabilidad de que, hasta que aparezca el tercer defectuoso, se hayan elegido diez sin defectos?
c) Si se elige una partida de 600 CDs,¿cuál la probabilidad de que, como mínimo, 590 sean no defectuosos

6.- Cuando lanzamos al aire 100 veces una moneda equilibrada, ¿cuál es la probabilidad de obtener más de 55 cruces?

7.- Sea X una variable que se distribuye según un Binomial $Bi(60, 0.4)$ Calcular:

a) $P(X = 24)$, $P(X > 24)$, $P(X > 30)$
b) $P(20 < X < 30)$, $P(X > 20)$
c) $P(20 < X \text{ o } X > 40)$ y $P(20 < X \text{ y } X > 10)$
e) Los cuantiles de ordenes 0.025, 0.5 y 0.975, respectivamente.

8. Un canal de comunicación recibe 300 señales independientes por microsegundo. y la probabilidad de error en la señal recibida es igual a 0,002, determinar las probabilidades de que, en un microsegundo:

a) No haya ningún error.
b) Haya sólo un error.
c) Haya al menos un error.
d) Haya dos errores