TEMA 5. EJERCICIOS RESUELTOS

1. El 8% de los análisis para el control del colesterol tienen resultados erróneos y es necesario repetir dichos análisis.

   a) Si se realiza un análisis, ¿cuál es la probabilidad de que el resultado sea incorrecto?

   b) Si se observan 15 análisis aleatorios, ¿cuál es la probabilidad de que haya que repetir al menos 2 análisis?

   c) ¿Cuál es el número de análisis que se espera repetir de entre 200 análisis realizados?

   d) Si en un día el laboratorio ha realizado 50 análisis, ¿cuál es la probabilidad de que se hayan tenido que repetir a lo sumo 3 análisis?

a) Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Análisis de colesterol erróneos'

$X \sim \text{Binario}(p = 0.08)$

$P(X = 1)$

$> \text{dbinom}(1, 1, 0.08)$

$[1]\ 0.08$

b) En este caso, el experimento binario se repite varias veces por lo que se debe definir una nueva variable aleatoria y la distribución que sigue.

X : 'Número de análisis de colesterol erróneos de entre 15'

$X \sim B(n = 15, p = 0.08)$

$P(X \geq 2) = 1 - P(X \leq 1)$

$> 1 - \text{pbinom}(1, 15, 0.08)$

$[1]\ 0.3402712$

c) Para calcular el número de análisis que se espera repetir, se calcula la media.

$n = 200$

$E(X) = n \cdot p$

```
> n = 200
> p = 0.08
> n*p
[1] 16
```

d) Como en este caso $n = 50 > 30$ y $p = 0.08 < 0,1$ La distribución binomial se podría aproximar por la distribución de Poisson, pero con R Studio no es necesario realizar esta aproximación.

X : 'Número de análisis de colesterol erróneos de entre 50'

$X \sim B(n = 50, p = 0.08)$

Por tanto,

$P(X \leq 3)$

```
> pbinom(3,50,0.08)
[1] 0.4252957
```

2. La probabilidad de que los ensayos realizados con un equipo de ultrasonidos sean efectivos es del 80%. Asumiendo que los ensayos son independientes, calcule:

   a) La probabilidad de que el primer ensayo efectivo ocurra en el quinto intento.

   b) La probabilidad de que tengan que realizarse al menos cuatro ensayos para obtener el primer ensayo efectivo.

   c) La probabilidad de que sean necesarios 12 ensayos para lograr 5 ensayos efectivos.

   d) La probabilidad de tener que realizar un máximo de 10 ensayos y un mínimo de 7 para obtener 3 ensayos efectivos.

a) Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Número de intentos hasta el primer ensayo efectivo'

$X \sim G(p = 0.8)$

Para que el quinto ensayo sea efectivo, los cuatro ensayos anteriores deben ser ineficaces, por lo que:

$P(X = 4)$

$> \ \text{dgeom}(4,\ 0.8)$

$[1]\ 0.00128$

b) Para que el número mínimo de ensayos realizados sea cuatro, al menos tres ensayos deben ser ineficaces.

$P(X \geq 3) = 1 - P(X \leq 2)$

$> \ 1 - \text{pgeom}(2,\ 0.8)$

$[1]\ 0.008$

c) En este caso, la variable aleatoria debe modificarse y la distribución que sigue es diferente, pasando de ser geométrica a ser binomial negativa. Para que 5 intentos sean efectivos, 7 no deben ser efectivos.

X : 'Número de intentos realizados hasta obtener 5 ensayos efectivos'

$X \sim BN(n = 5, p = 0.8)$

$P(X = 7)$

$> \text{ dnbinom}(7,5,0.8)$
$[1]\ 0.00138412$

d) En este caso, para 3 ensayos efectivos, los ensayos ineficaces deben estar entre 4 y 7.

X : 'Número de intentos realizados hasta obtener 3 ensayos efectivos'
$X \sim BN(n=3, p=0.8)$

$P(4 \leq X \leq 7) = P(X \leq 7) - P(X \leq 3)$

$> \text{ pnbinom}(7,3,0.8) - \text{pnbinom}(3,3,0.8)$
$[1]\ 0.01688207$

3. En una empresa eléctrica que fabrica fusibles, la probabilidad de que los fusibles sean defectuosos es de 0,2. Un cliente compra 15 fusibles, pero únicamente utiliza 5 de ellos.

   a) ¿Cuál es la probabilidad de que como mucho 2 de esos 5 fusibles sean defectuosos?
   b) ¿Cuál es el número de fusibles defectuosos que se esperan de entre esos 5 fusibles?
   c) Otro cliente ha adquirido una caja con 200 fusibles para usar 10 de ellos. En este caso, ¿cuál es la probabilidad de que a lo sumo 2 de esos 10 fusibles sea defectuoso?

a) Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Número de fusibles defectuosos'

$X \sim H(N=15, n=5, p=0.2)$

$P(X \leq 2)$

```
> phyper(2,3,12,5)
[1] 0.978022
```

b) Para calcular el número de fusibles que se espera que sean defectuosos, se debe calcular la media.

$n=5;\ p=0.2$

$E(X)=n \cdot p$

```
> n = 5
> p = 0.2
> n*p
[1] 1
```

c) En este caso $N=200$ y $n=10$, so $N>10 \cdot n$; por lo tanto, la distribución hipergeométrica se puede aproximar mediante una distribución binomial, pero en el cálculo con R Studio, no se es necesario realizar dicha aproximación.

X : 'Número de fusibles defectuosos'

$X \sim H(N=200, n=10, p=0.2)$

$P(X \leq 2)$

> 200*0.2

[1] 40

> 200*0.8

[1] 160

> phyper(2, 40,160,10)

[1] 0.6794051

4. En una fábrica de lanas de Edimburgo, por cada 5 metros de tela producida aparece un defecto. Sabiendo que el número de defectos que aparece en la tela sigue una distribución de Poisson, calcule:

   a) Si se compran cinco metros de tela, la probabilidad de que haya más de dos defectos.

   b) Si se compran 50 metros de tela para realizar 15 kilts (típica falda escocesa), la probabilidad de encontrar siete defectos.

a) Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Número de defectos por cada 5 m de tela de lana'

$X \sim P(\lambda = 1)$

$P(X > 2) = 1 - P(X \leq 2)$

```
> 1-ppois(2,1)
[1] 0.0803014
```

b) Es este caso, se debe calcular un nuevo parámetro λ ya que la variable aleatoria ha cambiado. En vez de tener 5 metros de tela hay 50 metros por lo que el parámetro se modifica linealmente

X : 'Número de defectos por cada 50 m de tela de lana'

$X \sim P(\lambda = 1 \cdot 10)$

$P(X = 7)$

```
> dpois(7,10)
[1] 0.09007923
```

5. El tiempo que tarda un estudiante en llegar de su casa a la universidad varía uniformemente entre 35 y 45 minutos. ¿A qué hora debe salir de casa para llegar a tiempo a clase con una probabilidad mínima de 0,8 si las clases empiezan a las 8 a.m.?

Primero, se debe definir la variable aleatoria y la distribución que sigue. $X =$ "Minutos desde casa a la universidad"

La variable sigue una distribución uniforme. $X \sim U[35, 45]$

$$P(X \leq x) \geq 0.8 \Rightarrow \frac{x-35}{45-35} \geq 0.8$$

> qunif $(0.8, 35, 45)$

[1] 43

Por lo tanto, si un estudiante sale de casa a las 7:17 a.m. o antes, llegará a tiempo a clase con una probabilidad de 0.8 o más.

6. En una tienda, el tiempo que esperamos desde la entrada de un cliente hasta la entrada del siguiente cliente se distribuye exponencialmente con una media de 5 minutos. Calcule la probabilidad de que se tenga que esperar entre 8 y 10 minutos hasta la entrada del siguiente cliente.

Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Minutos que debemos esperar hasta la entrada del siguiente cliente'. $X \sim \varepsilon\left(1/\lambda\right)$

El parámetro λ debe ser calculado:

$$E(X) = 5 = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{5}$$

La probabilidad que se pregunta es:

$$P(8 \leq X \leq 10) = P(X \leq 10) - P(X < 8)$$

```
> pexp(10,1/5)-pexp(8,1/5)
[1] 0.06656123
```

7. Un bote de mermelada se clasifica como "sirope" si la cantidad de azúcar está entre 420 y 520 g. El fabricante al analizar las macetas observa que el peso promedio es de 465 g, con una desviación estándar de 30 g. Sabiendo que el peso del azúcar sigue una distribución normal.

   a) ¿Qué porcentaje de la producción del fabricante no se puede etiquetar como "sirope"?

   b) ¿Cuáles son los dos valores centrales que tenemos que fijar para que entre ellos haya el 50% de los botes?

Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Cantidad de azúcar en g dentro del bote'.

$420 \leq X \leq 520 \rightarrow$ Bote que puede ser considerado 'sirope'

$X < 420$ y $X > 520 \rightarrow$ Bote que no puede ser considerado 'sirope'

$E(X) = 465$ g y $\sigma = 30$ g   X ~ N(465;30)

a) Este porcentaje puede calcularse de dos formas:

1)

$$P(420 \leq X \leq 520) = P(X \leq 520) - P(X \leq 420)$$

> pnorm(520, 465, 30) – pnorm(420, 465, 30)

[1] 0.8998163

$$1 - P(420 \leq X \leq 520) = 0.1004$$

Por lo tanto, el 10.02% no puede ser etiquetado como 'sirope'

2)

$$P(X \geq 520) + P(X \leq 420)$$

> pnorm(520, 465, 30, lower.tail = F) + pnorm(420, 465, 30)

[1] 0.1001837

Por lo tanto, el 10.02% no puede ser etiquetado como 'sirope'

2)

Usaremos la distribución normal tipificada

$$P(b \le X \le a) = P(X \le a) - P(X \le b)$$

$$Z = \frac{x-465}{30}$$

Para resolver este ejercicio aplicaremos simetría numerosas veces:

$$P(b \le X \le a) = P(-t \le Z \le t) = P(Z \le \frac{a-465}{30}) - P(Z \le \frac{b-465}{30}) = P(Z \le t) - P(Z \le -t)$$

$$P(b \le X \le a) = P(Z \le t) - P(Z \le -t) = P(Z \le t) - P(Z \ge t) =$$

$$P(Z \le t) - \left(1 - P(Z \le t)\right) = -1 + 2P(Z \le t) = 0.5 \Rightarrow P(Z \le t) = 0.75$$

```
> qnorm(0.75,0,1)
[1] 0.6744898
```

$$P(Z \le t) = 0.75 \Rightarrow t = 0.6744898$$

$$t = \frac{a-465}{30} \Rightarrow a = 485.234694 \text{ g} \quad \text{y}$$

$$-t = \frac{b-465}{30} \Rightarrow b = 444.765306 \text{ g}$$

Por lo tanto los valores centrales a y b son 485.234694 g y 444.765306 g, respectivamente.

8. En una fábrica se producen dos tipos de piezas, una de ellas elaborada con materiales renovables y otra con materias primas derivadas del petróleo.

   a) Si el 2% está defectuoso, ¿cuál es la probabilidad de que al seleccionar 100 piezas no haya más de 3 defectuosos?
   b) El 40% son piezas basadas en materiales renovables. ¿Cuál es la probabilidad de que al seleccionar 600 piezas, 352 o más piezas sean de materias primas derivadas del petróleo?

a) Primero, se debe definir la variable aleatoria y la distribución que sigue.

X : 'Número de piezas defectuosas'

n=100; p=0.02 $\quad$ $X \sim B(100;0.02)$

$P(X \leq 3)$

> pbinom(3,100,0.02)

[1] 0.8589616

b) En este caso, se debe definir una nueva variable aleatoria y, en consecuencia, la distribución que sigue.

X : 'Número de piezas defectuosas realizadas con materias primas derivadas del petróleo'

n=600; p=0.6 $\quad$ $X \sim B(600;0.6)$

$P(X \geq 352) = 1 - P(X \leq 351)$

> pbinom(351,600,0.6,lower.tail = F)

[1] 0.7610461

En este caso, la distribución binomial no debe ser aproximada a una distribución normal ya que R Studio puede hacer el cálculo muy fácilmente.

Nota: el separador decimal corresponde con un punto sólo es tema dado que el software R es el delimitador que utiliza.