# Probability – Homework 3.

## Problem 1

You work at a factory that manufactures light bulbs. You have determined that $5\%$ of light bulbs that are produced are defective. For each of the scenarios below:

1. Define an appropriate random variable and distribution.
2. State the values that the random variable can take on.
3. State any assumptions that you need to make.
4. Find the probability that the random variable you defined takes on the value $X = 4$.

**Part a)** ← $\text{Bin}(30, 0.05)$

Out of 30 lightbulbs, $k$ are defective.

**Answer:**

To begin with, we note the following things:

(✓) The sample space $S$ is described as:

$$S = \{(a_1, \ldots, a_{30}) : a_i \in \{0,1\} \text{ for } 1 \leq i \leq 30\}$$

taking into a count that an element $(a_1, \ldots, a_{30}) \in S$ represents the possible states of each lightbulb ( $a_i = 1$ means that the i-th lightbulb is defective and $a_i = 0$ is non defective).

(r) The random variable $X: S \to \mathbb{R}$ is described as:

$$X(a_1, a_2, \ldots, a_{30}) = \sum_{i=1}^{30} a_i = \#\text{ defective lightbulbs}$$

(r) The probability $P: \{T / T \subseteq S\} \to \mathbb{R}$ is defined as:

$$P(\{(a_1, \ldots, a_{30})\}) = (0{,}05)^{L} \cdot (0{,}95)^{30-L}$$

with $L = X(a_1, \ldots, a_{30}) = \boxed{\#\text{ defective lightbulbs}}$

(r) The probability mass function of $X: S \to \mathbb{R}$ relative to the probability $P: \{T / T \subseteq S\} \to \mathbb{R}$ is the function

$f : \mathbb{R} \longrightarrow \mathbb{R}$ described as:

$$f(k) = p(x=k) = \begin{cases} \binom{30}{k} \cdot (0{,}05)^{k} (0{,}95)^{30-k} & \text{if } k \geq 0,\ k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}$$

(r) $f(4) = p(x=4) = \binom{30}{4} \cdot (0{,}05)^{4} \cdot (0{,}95)^{26}$.

(r) The distribution of the random variable $x : S \to \mathbb{R}$ relative to the probability $p : \{T / T \subseteq S\} \to \mathbb{R}$ is the function $g : \mathbb{R} \to \mathbb{R}$ described as:

$$g(k) = p(x \leq k) = \begin{cases} \sum_{i=0}^{m} \binom{30}{i} \cdot (0{,}05)^{i} (0{,}95)^{30-i} & \text{if } k \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

where $m = \lfloor k \rfloor =$ "integer part of $k$".

**Part b)** ← Geo(0.05)

You test each lightbulb as it comes of the line. The $k^{th}$ light bulb is the first defective light bulb you find.

Answer:

(✓) The sample space $S$ is given as:

$$S=\{(1),(0,1),(0,0,1),(0,0,0,1),\dots\}$$

where an element $(0,0,\dots,1)=(a_1,\dots,a_n)\in S$ tells us that the first $n-1$ lightbulbs are non defective and $n$-th lightbulb is defective.

(✓) the random variable $X:S\to\mathbb{R}$ is described as:

$$X(a_1,\dots,a_{n-1},a_n)=X(0,\dots,0,1)=n$$

for $(a_1,\dots,a_n)\in S$.

(v) The probability $p:\{T / T \subseteq S\} \to \mathbb{R}$ is defined as:

$$p(\{(a_1, a_2, \dots, a_{n-1}, a_n)\}) = p(\{(0, \dots, 0, 1)\}) =$$

$$= p \cdot (1-p)^{n-1}.$$

for $(a_1, \dots, a_n) = (0, 0, \dots, 0, 1) \in S$.

(v) The distribution $f: \mathbb{R} \to \mathbb{R}$ of the random variable $x: S \to \mathbb{R}$ relative to the probability $P$ is discribed as:

$$f(k) = p(x=k) = \begin{cases} (0{,}05) \cdot (0{,}95)^{k-1} & \text{if } k \geq 1,\ k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases}$$

(v) $p(x=4) = (0{,}05)(0{,}95)^3$

(v) The distribution of $X : S \to \mathbb{R}$ relative to the probability $P : \{T / T \subseteq S\} \to \mathbb{R}$ is the function $g : \mathbb{R} \to \mathbb{R}$ given by:

$$g(k) = P(X \leq k) = \begin{cases} \sum_{i=1}^{m} (0{,}05) \cdot (0{,}95)^{i-1} & \text{if } k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$g(k) = P(X \leq k) = \begin{cases} 1 - (0{,}95)^{m-1} & \text{if } k \geq 1, \\ 0 & \text{otherwise} \end{cases}$$

where $m = \lfloor k \rfloor =$ "integer part of $k$".

**Part c)** ← NB(2, 0.05)

You find your second defective light bulb after observing $k$ light bulbs in all.

Answer:

(✓) The sample space $\underline{\underline{S}}$ is given as:

$$S=\left\{(a_1,\dots,a_n) \,/\, a_i\in\{0,1\},\ \sum_{i=1}^{n-1}a_i=1,\ a_n=1,\ n\in\mathbb{N}\right\}.$$

where an element $(a_1,\dots,a_n)\in S$ tell us that in the first $n-1$ lightbulbs there is only one defective lightbulb and the n-th lightbulb is defective.

(✓) The random variable $X:S\to\mathbb{R}$ is given by:

$$X(a_1,a_2,\dots,a_n)=n$$

for $(a_1,\dots,a_n)\in S$.

(✓) The probability $P:\{T / T\subseteq S\}\to\mathbb{R}$ is given as:

$$P(\{(a_1,\dots,a_n)\}) = p^2\cdot(1-p)^{n-2}$$

for $(a_1,\dots,a_n)\in S$.

(✓) The probability mass function of the random variable $X:S\to\mathbb{R}$ relative to the probability $P:\{T / T\subseteq S\}\longrightarrow\mathbb{R}$ is the function $f:\mathbb{R}\to\mathbb{R}$ described as:

$$f(k)=P(X=k)=\begin{cases}\binom{k-1}{1}p^2\cdot(1-p)^{k-2} & \text{if } k\geq 2,\ k\in\mathbb{Z},\\ 0 & \text{otherwise.}\end{cases}$$

$$f(k) = p(x=k) = \begin{cases} (k-1)\cdot p^2\cdot(1-p)^{k-2} & \text{if } k \geq 2,\ k \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases}$$

$$\boxed{\begin{array}{l} p = 0{,}05 \\ 1-p = 0{,}95 \end{array}}$$

(v) $f(4) = p(x=4) = 3\cdot(0{,}05)^2\cdot(0{,}95)^2.$

(v) The distribution of the random variable $x: S \to \mathbb{R}$ relative to the probability $p: \{T / T \subseteq S\} \to \mathbb{R}$ is the function $g: \mathbb{R} \to \mathbb{R}$ defined as:

$$g(k) = p(x \leq k) = \begin{cases} \sum_{i=2}^{m} (i-1)\cdot p^2\cdot(1-p)^{i-2} & \text{if } k \geq 2, \\ 0 & \text{otherwise} \end{cases}$$

where $m = \lfloor k \rfloor =$ "integer part of $k$".

## Problem 2

Consider a loaded six-sided die that is twice as likely to roll an even number as an odd number. Let $X$ be random variable for value that is rolled from the die.

### Part a)

What is the Probability Mass Function for $X$. Write this out as a table.

Answer:

To begin with, we shall note that:

($\checkmark$) $S=\{1,2,3,4,5,6\}$ is the random space.

($\checkmark$) The function $\boxed{\begin{aligned} X: S &\longrightarrow \mathbb{R} \\ i &\mapsto X(i)=i \end{aligned}}$ is the random variable of this problem.

($\checkmark$) The probability $P:\{T / T\subseteq S\}\longrightarrow \mathbb{R}$ given in this problem has the following description:

$$P(\{i\})=\begin{cases} \frac{1}{9} & \text{if } i \text{ is an odd number,} \\ \frac{2}{9} & \text{if } i \text{ is an even number} \end{cases}$$

for $i \in \{1,2,3,4,5,6\}$.

(v) The probability mass function $f: \mathbb{R} \to \mathbb{R}$ of $X$ is given as:

$$f(i) = P(X=i) = P(\{i\}) = \begin{cases} \frac{1}{9} & \text{if } i \text{ is an odd number,} \\ \frac{2}{9} & \text{if } i \text{ is an even number} \end{cases}$$

| $X=i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(X=i)$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | $\frac{2}{9}$ |

---

**Part b)**

What is the Cumulative Distribution Function for $X$?

Answer:

The cumulative distribution $g:\mathbb{R}\to\mathbb{R}$ for $x$ is described as:

$$g(i)=P(x\leq i)=\begin{cases} 0 & \text{if } i<1, \\ \frac{1}{9} & \text{if } 1\leq i<2, \\ \frac{3}{9} & \text{if } 2\leq i<3, \\ \frac{4}{9} & \text{if } 3\leq i<4, \\ \frac{6}{9} & \text{if } 4\leq i<5, \\ \frac{7}{9} & \text{if } 5\leq i<6, \\ 1 & \text{if } 6\leq i. \end{cases}$$

---

**Part c)**

What is $E[X]$?

Answer:

$$E(x)=\sum_{x} P(X=x)=\frac{1}{9}(1+3+5)+\frac{2}{9}(2+4+6)$$

$$=1+\frac{24}{9}=\frac{33}{9}=3,666\ldots$$

## Problem 3 ¶

How would we simulate variables from these distributions in R? It'll turn out that the method is fairly similar across all these distributions so, for simplicity, let's just say we want to simulate $X \sim Bin(n, p)$. Take a look at the official documentation for this function [here](). Not extremely clear, is it? Let's go through it one step at a time.

### Part a)

What if we want a random variable from this distribution? That is, we know some underlying distribution and we want to simulate many results from that distribution. Then we would use the "random generation" function `rbinom()`.

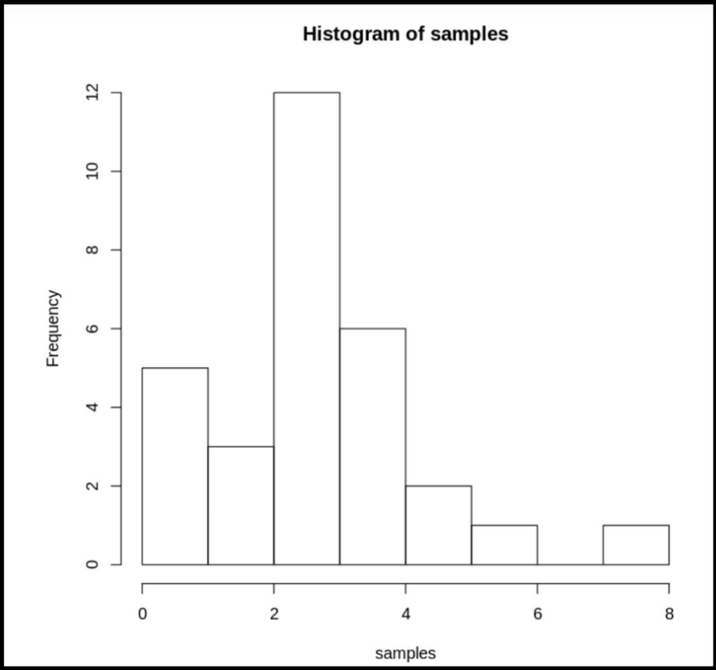
Play around with this function, with different `size` and `prob` parameters to get a feel for how it works. Finally, generate $30$ results from a $Bin(10, 0.3)$ distribution and plot a histogram of the results.

Answer:

In [5]:
```
# YOUR CODE HERE
rbinom(n=30, size=10, prob=0.3)
```

2· 1· 2· 6· 4· 3· 1· 3· 1· 2· 5· 2· 3· 1· 3· 5· 2· 4· 4· 6· 4· 4· 1· 5· 1· 3· 0· 1· 3· 3

```
samples = rbinom(30, size=10, prob=0.3)
hist(samples)
```




---

### Part b)

What if we have some value $k$ and we want to know what's probability of generating $k$? That is, we're solving the Probability Mass Function $P(X = k)$. Then we would use the "density" function `dbinom()`.

Let $X \sim Bin(15, 0.4)$. By hand, solve $P(X = 4)$. Then use the `dbinom()` function to confirm your result.

## Answer:

```r
# YOUR CODE HERE
dbinom(x = 4, size = 15, prob = 0.4)
```

0.12677580324864

$$P(X=4)=\binom{15}{4}(0{,}4)^4\cdot(0{,}6)^{11}=0{,}1267\ldots$$

---

**Part c)**

What if we wanted to solve for some value of the Cumulative Density Function? That is, we know $k$ and want to find $P(X \leq k) = p$. Then we would use the "distribution function" `pbinom()`.

Let $X \sim Bin(15, 0.4)$. By hand, solve $P(X \leq 4)$. Then use the `pbinom()` function to confirm your result.

## Answer:

```r
# YOUR CODE HERE
pbinom(4, size = 15, prob = 0.4)
```

0.217277705650176

$$P(X\leq 4)=\sum_{k=0}^{4}\binom{15}{k}\cdot(0{,}4)^k\cdot(0{,}6)^{15-k}=$$

$$=\binom{15}{0}(0{,}4)^0\cdot(0{,}6)^{15}+\binom{15}{1}\cdot(0{,}4)^1\cdot(0{,}6)^{14}+$$

$$\binom{15}{2}(0{,}4)^2\cdot(0{,}6)^{13}+\binom{15}{3}(0{,}4)^3\cdot(0{,}6)^{12}+$$

$$\binom{15}{4}(0{,}4)^4\cdot(0{,}6)^{11}$$

$\approx 0,0005 + 0.00470 + 0,0219 +$

$0,0634 + 0,12678$

$\approx 0,21728$

---

**Part d)**

Finally, we have the "quantile" function `qbinom()`. This function is the reverse of the `pbinom()` function, in that it takes a probability $p$ as an argument and returns the value $k$ of the CDF that results in that much probability.

Use the `qbinom()` function to confirm your results from **Part c**. That is, plug in the probablity you got from **Part c** and see if you get the same $k$.

Answer:

```
# YOUR CODE HERE
qbinom(p = 0.217277705650176, size = 15, prob = 0.4)
```

```
4
```

$P(x \leq 4) = 0,217277$

$q(0,217277) = 4$