Probability: Random Variables

Bernoulli and Binomial Random Variables

In this document we introduce two special types of discrete random variables. The first one is a 'Bernoulli' random variable, second one is 'Binomial' random variable.

We know that there is a 50% chance that a new born baby will be a girl.
Suppose that a certain hospital has 10 baby deliveries per day.

This means that each baby can be a boy (50% chance) or a girl (also 50% chance). We introduce random variable X such that

$$X = \begin{cases} 1 & \text{Girl} \\ 0 & \text{Boy} \end{cases}$$

Since we have 50% chance that a new born baby will be a girl, we can write:

$$P(X = 1) = 0.5, \quad P(X = 0) = 0.5$$

This random variable X is called **Bernoulli** random variable.

Suppose on Monday, 4 Boys and 6 Girls were born. On Tuesday out of 10 babies, 7 were Boys and 3 were Girls. We see that even though there is 50% chance of Boy or Girl, on any particular day the number of Boys and Girls change, this number is random.

What is the probability that the hospital delivers exactly 7 boys on one day?

To answer this question we need to know:
What is the distribution of the number of boys born on one day?

Consider the random variable

$$Y = \text{number of boys born on a day}$$

Then we say that distribution of Y is **Binomial** with parameters $n = 10$ (because 10 babies are born every day) and $p = 0.5$ (because probability that a new born baby is a boy is 0.5). We can write it mathematically as

$$Y \sim Bin(10, 0.5)$$

The 'Probability Mass Function' (gives probability of the possible values of a discrete random variable) for this random variable is:

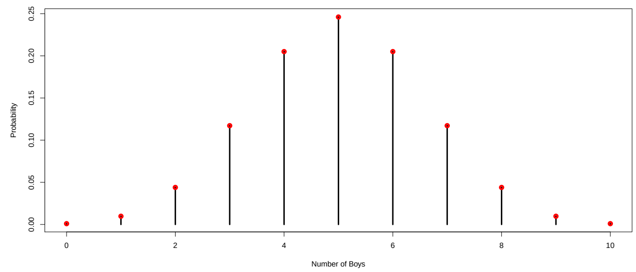

Figure 1: Probability Mass Function for Bin(10,0.5)

Here the number of Boys born is on the x-axis and probabilities of corresponding events are on the vertical axis. Now we can get probability of 7 boys from this plot, it is around 0.12, so

$$P(Y = 7) = 0.12$$

There are some conditions that random variable Y must satisfy in order to be $Bin(n, p)$:

- *There is a fixed number n of independent and identical trials.*
  In our example there are 10 births per day, hence $n = 10$.
- *Each trial can result in one of only two possibilities called success or failure; that is each trial is a Bernoulli trial.*
  In our example the two possibilities are Boy or Girl.
- *The probability p of success is constant from trial to trial.*
  This means that probability that a baby is a Boy on birth one is 0.5 and probability on the second birth is also 0.5. The probability of being a Boy does not change from birth to birth.

EXAMPLE
A survey asked 10 randomly selected people if they approve of the job prime minster is doing while in office. 3 people answered that they approve and 7 people answered that they disapprove.

Here is the bar chart of the result:

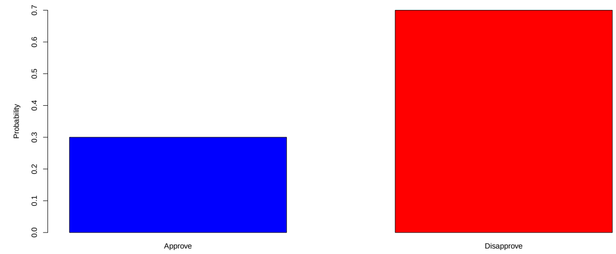

Figure 2: Prime Minister Approval example

We can say that probability of approval is $0.3 = \frac{3}{10}$. Now let X be a random variable such that

$$X = \begin{cases} 1 & \text{if a person approves} \\ 0 & \text{if a person disapproves} \end{cases}$$

Then X has a **Bernoulli** distribution or we can say X is a **Bernoulli** random variable.
Now suppose the probability that randomly selected person approves is 0.3 and probability that he or she disapproves is 0.7.

$$P(X = 1) = 0.3, \quad P(X = 0) = 0.7$$

In this example we can picture these 10 respondents as independent trials. In each trial we have probability 0.3 of success and 0.7 of failure. Also answers of respondents are independent of each other.

Now suppose we conduct another survey, taking a random sample of 10 people. *What is the probability that 4 people would approve of prime minister job?*
Now we construct new discrete random variable Y

$$Y = \text{total number of people that approve of PM job}$$

We want to calculate $P(Y = 4)$, to do that we need to know the distribution of Y.
We have 10 people each one can approve or disapprove, and they are independent of each other. So Y has a **Binomial** distribution with $n = 10$ (number of trials) and $p = 0.3$ (probability of success on each trial).

$$Y \sim Bin(10, 0.3)$$

Let's look at the graph of 'Probability Mass Function' for $Bin(10, 0.3)$

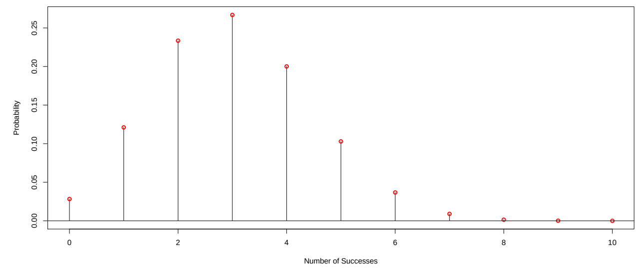

Figure 3: Probability Mass Function for Bin(10,0.3)

Based on the plot the probability of 4 successes is about 0.20, therefore

$$P(Y = 4) \approx 0.20$$

Now suppose we have another question: *What is the probability that at most 2 people approves of prime minister job?*
We need $P(Y \leq 2)$, it is possible to decompose this probability into 3 terms and find each term from the above graph:

$$P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 0.025 + 0.12 + 0.24 = 0.385$$

Thus the probability that at most 2 people approve of prime minister job is about 39%.