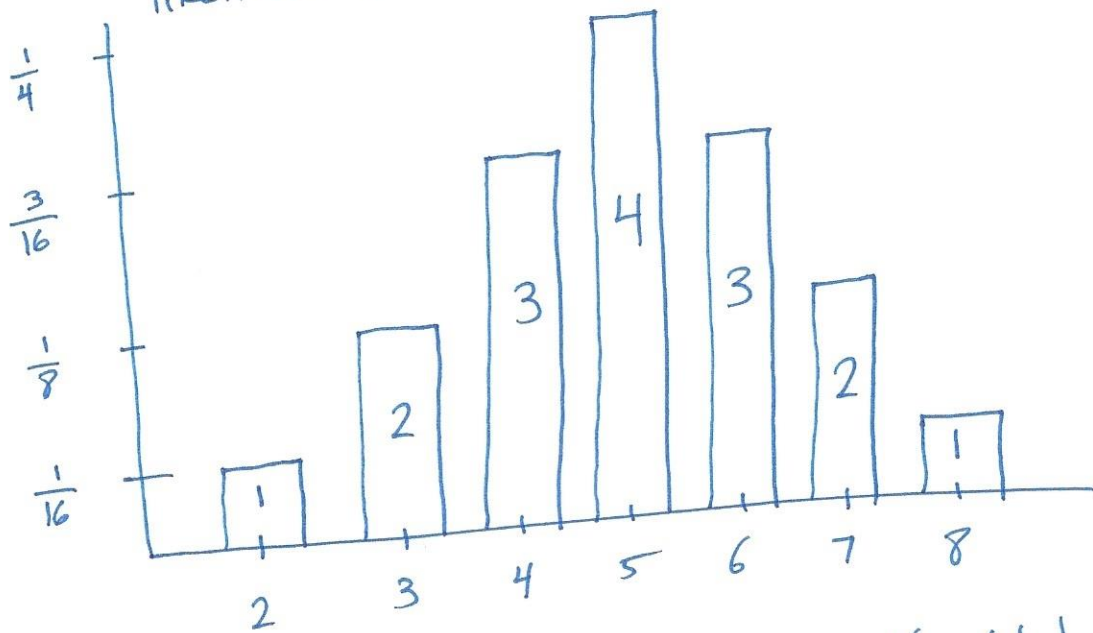


## Section 4.5: Probability

Elementary probability typically concerns itself with the study of discrete random variables, in which the possible outcomes of a scenario are discontinuous.

eg Suppose we roll two 4-sided dice and add together the results. The possible outcomes are 2, 3, 4, 5, 6, 7 or 8. We can depict the likelihood of each outcome as follows:



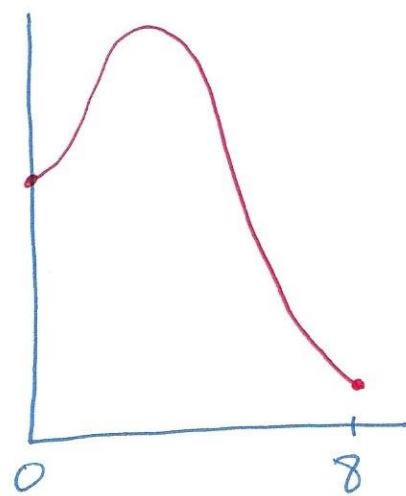
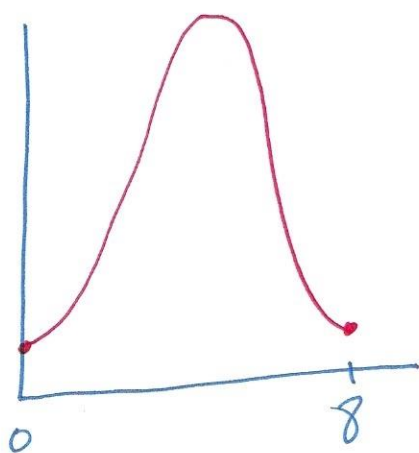
We normally scale probabilities such that the total probability of all possible outcomes is exactly 1.

For discrete random variables, we can consider the probability of a single outcome, or of multiple outcomes, such as all the outcomes which lie on a given interval.

When studying continuous random variables, the possible outcomes will be defined by a continuous function.

eg Suppose you choose two videos from your YouTube search history and add together your viewing time ~~from~~ from the first 4 minutes of each. The possible outcomes include any real number between 0 and 8.

Depending on one's viewing habits, many different functions might depict the relative likelihood of the possible outcomes, such as:



The function that describes the possible outcomes of a scenario is a probability density function  $f(x)$ .

For continuous random variables, we are only interested in the probability of outcomes which lie on the interval  $[a, b]$ , denoted by  $P(a \leq X \leq b)$ .

Given a probability density function  $f(x)$ ,

$$P(a \leq X \leq b) = \int_a^b f(x) dx.$$

Note that a probability density function must obey the following properties:

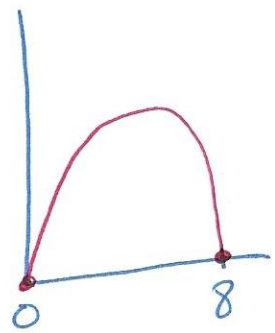
①  $f(x) \geq 0$  for all real numbers  $x$

②  $\int_{-\infty}^{\infty} f(x) dx = 1$

eg Suppose we want to model our YouTube search history viewing habits with a parabolic function. Then

$$f(x) = \begin{cases} kx(8-x), & \text{for } 0 \leq x \leq 8 \\ 0, & \text{otherwise} \end{cases}$$

Determine the value of the constant  $k$  which ensures that  $f(x)$  is a probability density function.



First, we note that  $f(x)$  is non-negative as long as  $k \geq 0$ .

Next, we have  $\int_{-\infty}^{\infty} f(x) dx = \int_0^8 kx(8-x) dx$

$$= k \int_0^8 (8x - x^2) dx$$

$$= k \left[ 8 \cdot \frac{x^2}{2} - \frac{x^3}{3} \right]_0^8$$

$$= k \left[ 4x^2 - \frac{1}{3}x^3 \right]_0^8$$

$$= \frac{256}{3} k$$

We must have  $\frac{256}{3} k = 1$ , so  $k = \frac{3}{256}$ .

eg Determine the probability that the total viewing time of the two videos is 6 minutes or longer.

$$\begin{aligned}P(X \geq 6) &= P(6 \leq X \leq 8) \\&= \int_6^8 f(x) dx \\&= \int_6^8 \frac{3}{256} x (8-x) dx \\&= \frac{3}{256} \int_6^8 (8x - x^2) dx \\&= \frac{3}{256} \left[ 4x^2 - \frac{x^3}{3} \right]_6^8 \quad \boxed{= \frac{5}{32}}\end{aligned}$$

In general, the probability of an outcome falling in the interval  $[x_{i-1}, x_i]$  is approximately  $f(x_i^*) \Delta x_i$  where  $x_i^*$  is a sample point on  $[x_{i-1}, x_i]$  and  $\Delta x_i = x_i - x_{i-1}$ .

Let  $N$  be the total number of events under consideration. Then the number of events falling into  $[x_{i-1}, x_i]$  is given by  $N f(x_i^*) \Delta x_i$ .

The total viewing time of these events is about  $x_i^* N f(x_i^*) \Delta x_i$

and the total viewing time of all the events is

$$\sum_{i=1}^n x_i^* N f(x_i^*) \Delta x_i = N \sum_{i=1}^n x_i^* f(x_i^*) \Delta x_i.$$

The average viewing time is given

$$\frac{N \sum_{i=1}^n x_i^* f(x_i^*) \Delta x_i}{N} = \sum_{i=1}^n x_i^* f(x_i^*) \Delta x_i.$$

This is a Riemann sum corresponding to the definite integral

$$\mu = \int_a^b x f(x) dx$$

which gives the mean value  $\mu$  of the events.