Congratulations! You passed!

TO PASS 40% or higher **PRACTICE QUIZ • 20 MIN**

Keep Learning

GRADE
100%

Exploring parameter space

Exploring parameter space

TOTAL POINTS 8

Submit your assignment

Receive grade

TO PASS 40% or higher

Grade
100%

View Feedback

We keep your highest score

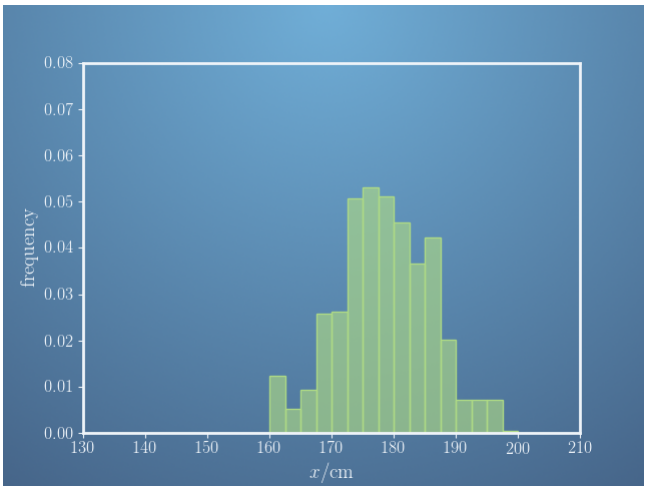
1. In this exercise, we shall see how it is often convenient to use vectors in machine learning. These could be in the form of data itself, or model parameters, and so on. **2 / 2 points** Try again

   The purpose of this exercise is to set the scene for Linear Algebra and the rest of the maths we will cover in the specialization. If this is confusing right now - stick with us! We'll build up your skills throughout the rest of the course. For this reason we've set a low pass mark for this quiz, but even if you don't pass in one go, reading the feedback from a wrong answer can often give more insight than guessing a correct answer!

   * * *

   The problem we shall focus on in this exercise is the distribution of heights in a population.

   If we do a survey of the heights of people in a population, we may get a distribution like this:



   This *histogram* indicates how likely it is for anyone in the survey to be in a particular height *range*. (6 ft is around 183 cm)

   This histogram can also be represented by a vector, i.e. a list of numbers. In this case, we record the frequency of people with heights in little groups at 2.5 cm intervals, i.e. between 150 cm and 152.5 cm, between 152.5 cm and 155 cm, and so on. We can define this as the vector $\mathbf{f}$ with components,

$$\mathbf{f} = \begin{bmatrix} f_{150.0,\ 152.5} \\ f_{152.5,\ 155.0} \\ f_{155.0,\ 157.5} \\ f_{157.5,\ 160.0} \\ f_{160.0,\ 162.5} \\ \vdots \end{bmatrix}$$

   These vector components are then the sizes of each bar in the histogram.

   Of the following statements, select all that you think are true.

   - [ ] None of the other statements.
   - [x] If another sample was taken under the same conditions, the frequencies should be broadly similar.

     Correct
     For a sufficiently large sample, the data will represent the population it is taken from.

   - [ ] No one in the world is less than 160 cm tall.
   - [ ] If another sample was taken under the same conditions, the frequencies would be exactly the same.
   - [x] There are at least 10 elements in the frequency vector, $\mathbf{f}$.

     Correct
     The data has been grouped into around 20 bins each having a width of 2.5 cm, in the range between 150 cm and 210 cm. Around 15 of these have a non-zero frequency.

2. One of the tasks of machine learning is to fit a model to data in order to represent the underlying distribution. **1 / 1 point**

   For the heights of a population, a model we may use to predict frequencies is the Normal (or Gaussian) distribution. This is a model for a bell-shaped curve, which looks like this,