

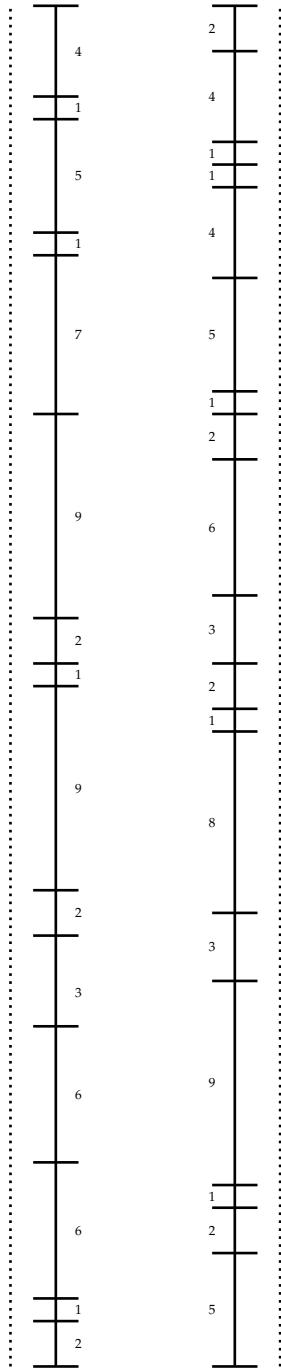
Activity 2.13

[f2008/f2008-10]

This is an exercise about sampling. You will need to follow the directions carefully.

The drawing consists of a pair of dotted lines, each of which is next to a cross-ticked scale marked with numbers. The meaning of these numbers will be explained later. For now, do the following:

1. Use a loose piece of paper to cover up all the drawing except the left dotted line.
2. Mark the dotted line at 10 random places with a pencil. Make sure your marks are clearly visible.
3. Now do the same thing to the dotted line on the right, marking another 10 random places on the dotted line. Make sure to cover up the cross-ticked scale with the loose piece of paper.
4. Remove the loose piece of paper so that you can see the scales.
5. Each of your 20 random marks falls into one of the intervals marked on the axis. For each of the marks, write down the number that labels the interval. (If more than one of your marks falls into the same interval, record the same number for each of the marks.)
6. Write down your 20 numbers here:



The number written next to each interval give the length of that interval.

Assuming that you made your marks at random, then you have made a fair, representative sample of locations on the dotted lines. The question now is whether you also made a fair, representative sample of the intervals. Remember, you had covered up the intervals so you couldn't be influenced by them when you marked the dotted lines.

As a simple way to see if your sampling of the intervals was representative, calculate the mean of the 20 intervals you recorded. Then calculate the mean of the actual intervals, counting each interval once: 4, 1, 5, 1, 7, 9 and so on.

Check whether the mean from your random sampled of intervals is a close match to the mean of the actual intervals. (Later in the book, you'll see how to do a formal comparison of your random sample to the actual value. But for now use your intuition.)

If your random sample isn't representative of the actual intervals, then your sample is **biased**. Can you see any reason for there to be a bias in the sampling procedure? (Hint: The individual intervals are the cases in the sampling frame. A simple random sample will be genuinely representative only when it's equally likely to pick any of the cases in the sampling frame.)