

## 2.2 Example: Empirical distributions of oyster sizes

To demonstrate the empirical distribution tool in PRIMER, we shall examine a dataset consisting of length measurements (in mm) of the Sydney rock oyster (*Saccostrea commercialis*) settling on various surfaces in Quibray Bay, New South Wales, Australia ( [Anderson \(1992\)](#) , [Anderson & Underwood \(1994\)](#) ). Settlement panels (measuring 10 cm x 10 cm) of four different substrata commonly introduced by humans into marine environments (concrete, marine plywood, fibreglass and aluminium) were placed in intertidal estuarine habitats (an oyster farm) in the bay. The greatest length (the longest distance from the umbo to the tip of the furthest growing edge) of all oysters settling on these four different types of surfaces were recorded after a period of 4 months (January - May, 1992).

A subset of the data (i.e., from just one of the sticks deployed in field, see [Anderson & Underwood \(1994\)](#) for logistic details of the experiment) are contained in a file called 'Quibray\_oyster\_sizes\_subset.pri' (found in the 'Quibray\_oysters' folder in 'Examples\_P8'). In this file, the lengths of oysters from the four different substrata are provided as 4 different levels of a factor called 'Substratum'. Note that there were different numbers of oysters on each of these different types of surfaces (hence, different sample sizes for different levels of the factor), but this is not of any concern here. We are comparing only the *shape* of the *distribution of sizes* of oysters that have settled among these four different types of surfaces; we are not comparing the total number of individuals that have settled on them.

1. Start running **PRIMER 8**, then click **File > Open...** and open the data file named 'Quibray\_oyster\_sizes\_subset.pri' (in 'Examples\_P8 > Quibray\_oysters').

PRIMER

File Edit Select View Wizards Pre-treatment Analyse Plots PERMANOVA+ Tools Window Help

Workspace

Quibray\_oyster\_sizes\_subset

Quibray\_oyster\_sizes\_subset

*Sizes of oysters from Quibray Bay, NSW (subset)*

Other

| Variables |             |
|-----------|-------------|
|           | Length (mm) |
| 1         | 2.5         |
| 2         | 3           |
| 3         | 4           |
| 4         | 12.5        |
| 5         | 5           |
| 6         | 5           |
| 7         | 5           |
| 8         | 5           |
| 9         | 7           |
| 10        | 24          |
| 11        | 25          |
| 12        | 6.5         |
| 13        | 16          |
| 14        | 24          |
| 15        | 15          |
| 16        | 9           |
| 17        | 2           |
| 18        | 2           |

Samples - Individual oysters

- It would be useful to see the empirical distributions for the four different surfaces side by side on a single plot. From the data sheet, click **Edit > Factors..** and you can see the factor of 'Substratum' that shows the type of surface on which each individual measured oyster had settled.

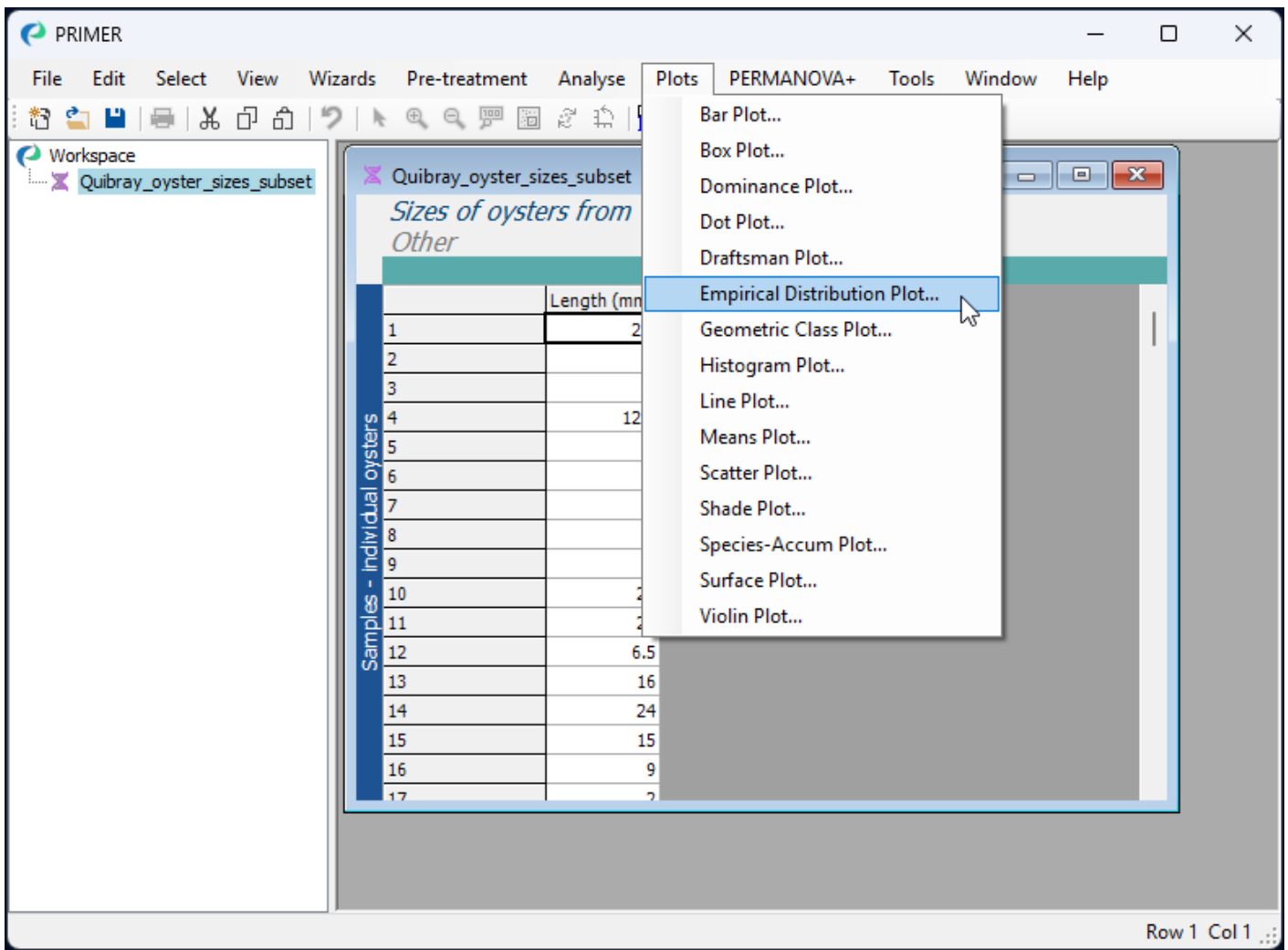
Factors

Edit Fill

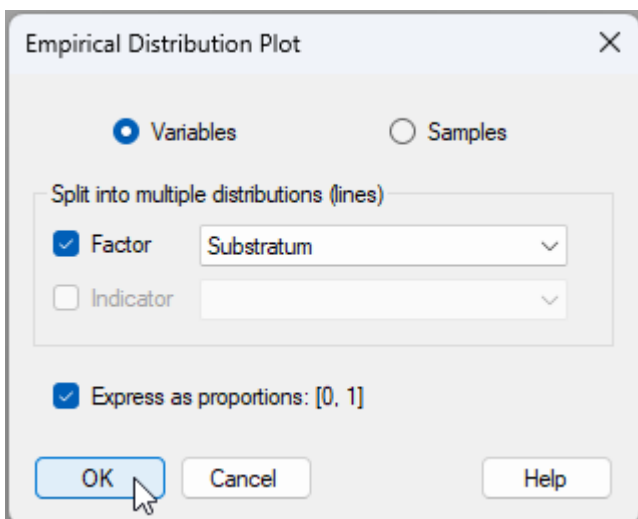
Add... Duplicate Combine... Rename... Rename Levels... Reorder... Delete... Key... Import... OK Cancel Help

| Label | Substratum | Block | Stick |
|-------|------------|-------|-------|
| 1     | Concrete   | 1     | 1A    |
| 2     | Concrete   | 1     | 1A    |
| 3     | Concrete   | 1     | 1A    |
| 4     | Concrete   | 1     | 1A    |
| 5     | Concrete   | 1     | 1A    |
| 6     | Concrete   | 1     | 1A    |
| 7     | Concrete   | 1     | 1A    |
| 8     | Concrete   | 1     | 1A    |
| 9     | Concrete   | 1     | 1A    |
| 10    | Concrete   | 1     | 1A    |
| 11    | Concrete   | 1     | 1A    |
| 12    | Concrete   | 1     | 1A    |
| 13    | Concrete   | 1     | 1A    |
| 14    | Concrete   | 1     | 1A    |
| 15    | Concrete   | 1     | 1A    |
| 16    | Concrete   | 1     | 1A    |
| 17    | Concrete   | 1     | 1A    |
| 18    | Concrete   | 1     | 1A    |
| 19    | Concrete   | 1     | 1A    |

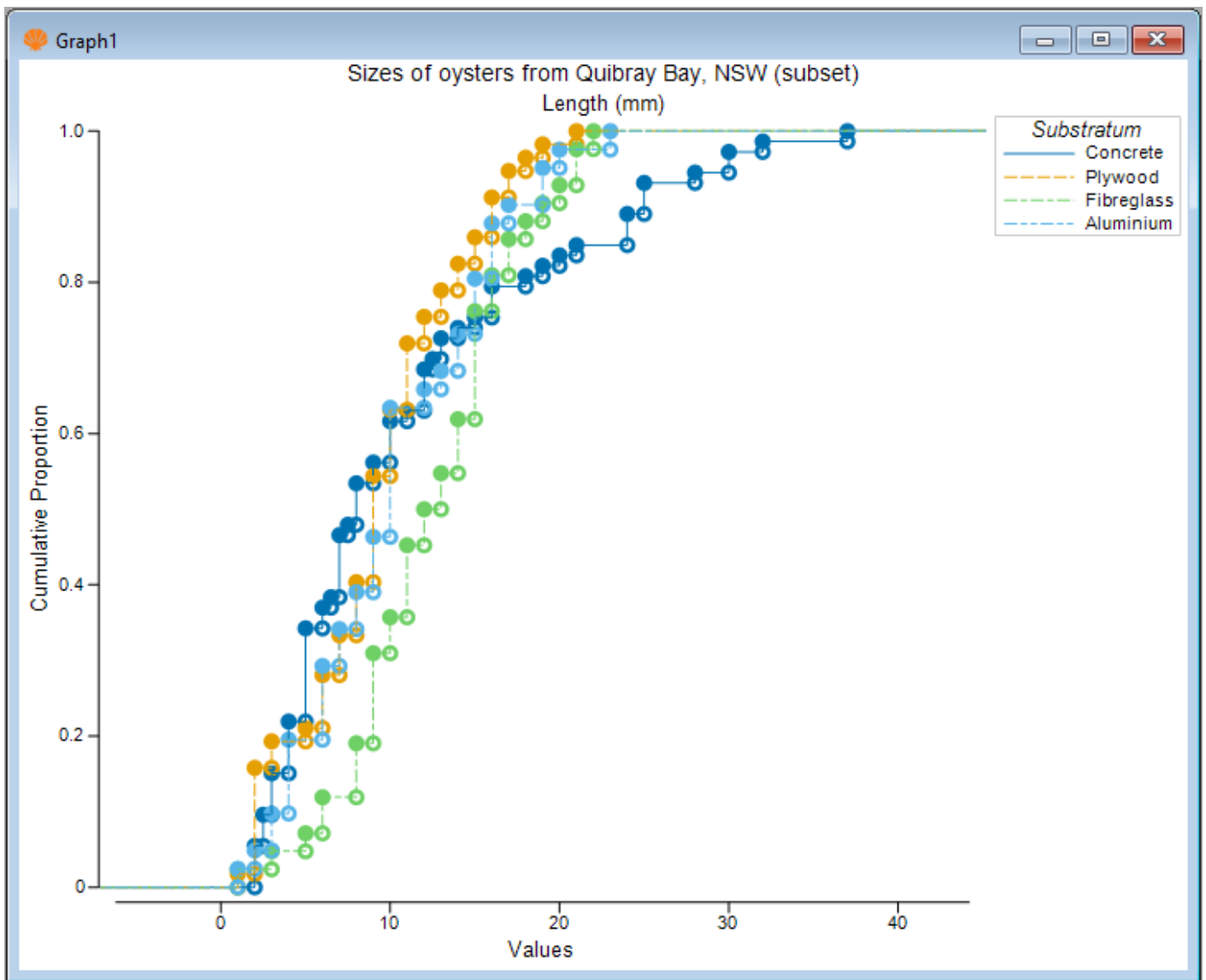
3. Now we are ready to create the plot. From the data sheet, click **Plots > Empirical Distribution Plot...**



In the resulting dialog box, choose to draw the lines for 'Variables' (there is only one variable here, so just one plot will be given in the output) and choose to 'Split into multiple distributions (lines)' > ☒ Factor > **Substratum**. Also, tick the box that says ☒ Express as proportions: [0, 1]. The resulting dialog will look like this:



The resulting graphic (called '**Graph1**' in the Explorer tree) will look like this:



From this graphic, we can see that there were, proportionately, quite a lot more large oysters measured on concrete surfaces (dark blue line) compared to the other substrata. In addition, fibreglass surfaces (green line) had proportionately fewer smaller-sized oysters than the other substrata. We can also consider looking at these distributions using [dot plots](#), [violin plots](#), or [histograms](#). It is also possible to formally test the null hypothesis of 'no difference' in the underlying distributions for any pair of these groups, using the [Kolmogorov-Smirnov test](#).

Revision #50

Created 26 March 2025 23:06:00 by Marti

Updated 1 December 2025 21:49:38 by Marti