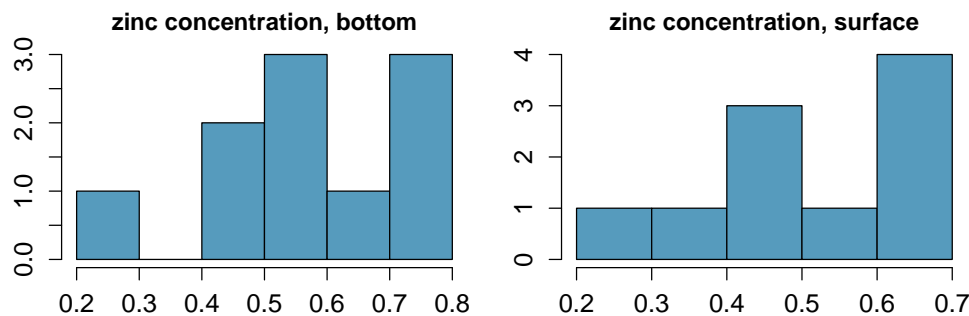# Application exercise 4.2:
Comparing two means, Part 1

Submit your responses on Sakai, under the appropriate assignment. Only one submission per team is required. One team will be randomly selected and their responses will be discussed.

Trace metals in drinking water affect the flavor and an unusually high concentration can pose a health hazard. Ten pairs of data were taken measuring zinc concentration in bottom water and surface water at 10 randomly sampled locations. The distributions are shown below. We want to evaluate whether the true average concentration in the bottom water *exceeds* that of surface water? Note that water samples collected at the same location, on the surface and in the bottom, cannot be assumed to be independent of each other.

To complete this application exercise you will need various sample statistics. You can calculate these in R using the raw data. The dataset can be loaded using the following command:

```
download("http://stat.duke.edu/~mc301/data/zinc.csv", destfile = "zinc.csv")
zinc = read.csv("zinc.csv")
```

1. Define the parameter of interest and the point estimate and calculate the point estimate.

2. Conduct a hypothesis test answering the research question. Don't forget to check conditions first. Use $\alpha = 0.05$. Make sure to frame your conclusion in context of the data and the research question.

3. Calculate a confidence interval for the parameter of interest at the confidence level equivalent to the previous hypothesis test. Make sure to interpret the interval in context of the research question.

4. Describe how you would construct this interval using bootstrapping and the standard error method.

5. Construct the bootstrap interval and compare it to the theoretical interval you calculated earlier.