

Populations, Samples, and Random Variables

An experiment's **population** is a complete collection of objects that we would like to study. These objects can be people, machines, processes, or anything we would like to understand experimentally.

Of course, we can't measure **all** of the population, usually. Instead, we take a subset of the population—called a **sample**—and infer the characteristics of the entire population from this sample.

However, this inference that the population is somehow representative of the population assumes that the sample size is sufficiently large and that the sampling is **random**. This means that selection of the sample should be such that no one group within a population are systematically overrepresented in the sample.

Example Suppose we wanted to see if the diameters of adults' heads were being studied in relation to their intelligence. What is the population? What would be a way to get a random sample of the measurements?