Parameters vs. Statistics

Learning Objective: Identify and distinguish between a parameter and a statistic.

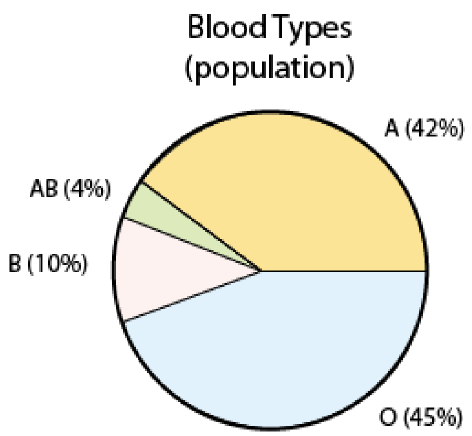
Learning Objective: Explain the concepts of sampling variability and sampling distribution.

Parameters vs. Statistics

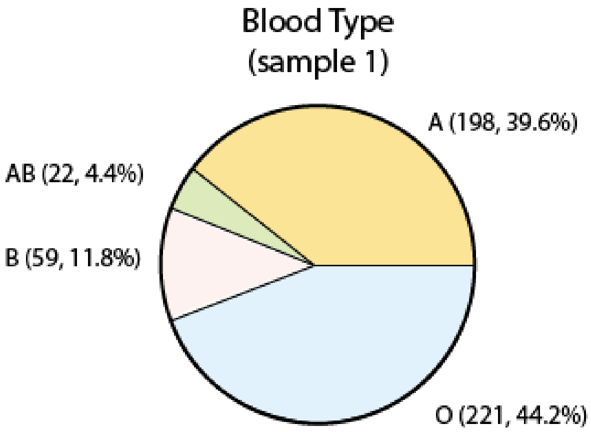
To better understand the relationship between sample and population, let's consider the two examples that were mentioned in the introduction.

Example: Example #1: Blood Type

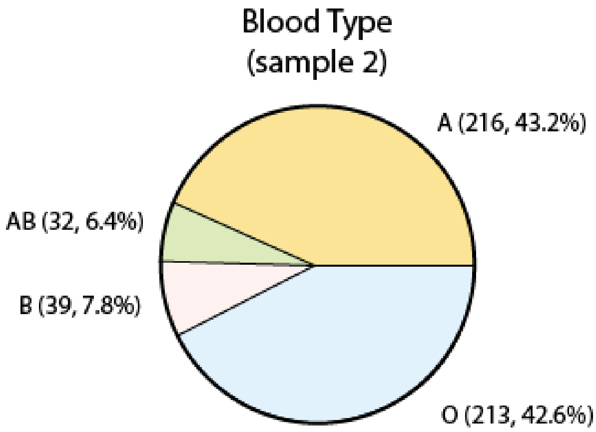
In the probability section, we presented the distribution of blood types in the entire U.S. **population**:



Assume now that we take a **sample** of 500 people in the United States, record their blood type, and display the sample results:



Note that the percentages (or proportions) that we got in our sample are slightly different than the population percentages. This is really not surprising. Since we took a sample of just 500, we cannot expect that our sample will behave exactly like the population, but if the sample is random (as it was), we expect to get results which are not that far from the population (as we did). If we took yet another sample of size 500:



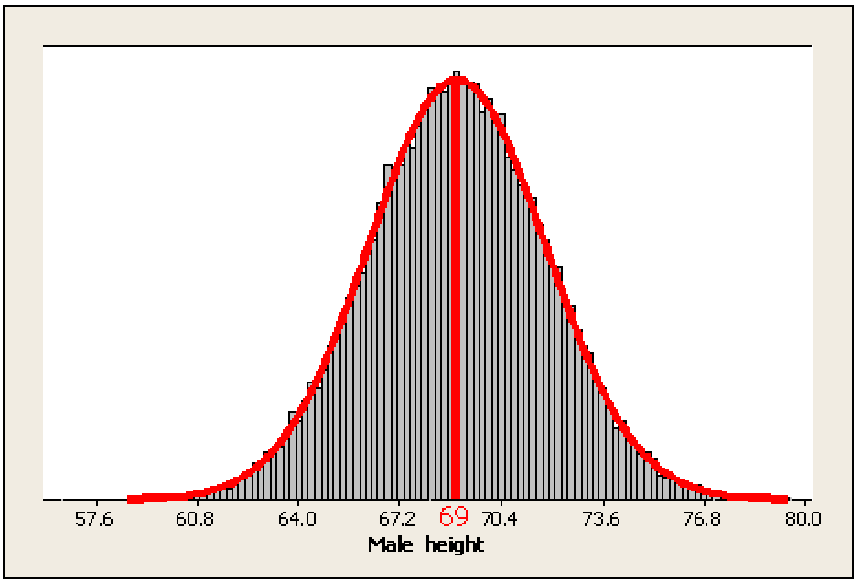
we again get sample results that are slightly different from the population figures, and also different from what we got in the first sample. This very intuitive idea, that sample results change from sample to sample, is called **sampling variability.**

Let's look at another example:

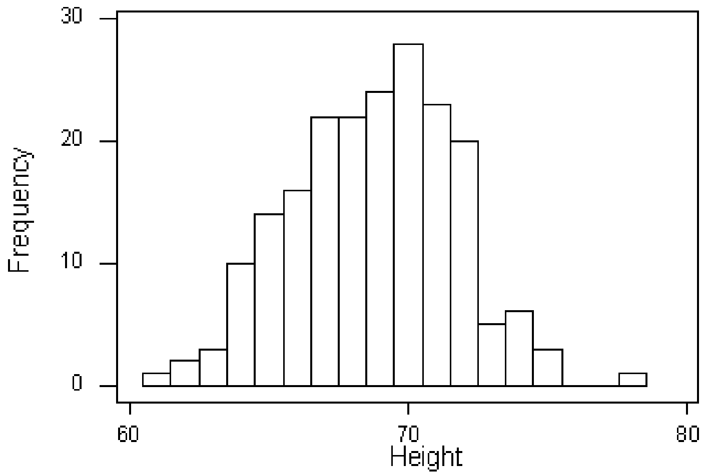
Example: Example #2: Heights of Adult Males

Heights among the population of all adult males follow a normal distribution with a mean $\mu = 69$ inches and a standard deviation

$\sigma = 2.8$ inches. Here is a probability display of this population distribution:

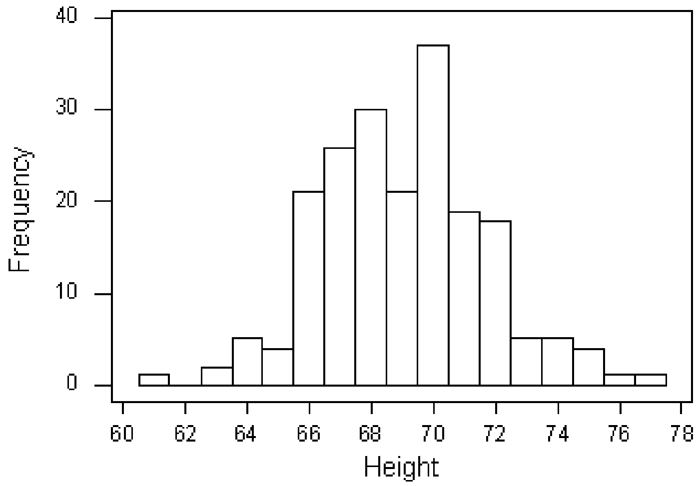


A sample of 200 males was chosen, and their heights were recorded. Here are the sample results:



The sample mean is $\overline{x} = 68.7$ inches and the sample standard deviation is s = 2.95 inches.

Again, note that the sample results are slightly different from the population. The histogram we got resembles the normal distribution, but is not as fine, and also the sample mean and standard deviation are slightly different from the population mean and standard deviation. Let's take another sample of 200 males:



The sample mean is $\overline{x} = 69.065$ inches and the sample standard deviation is s = 2.659 inches.

Again, as in Example 1 we see the idea of **sampling variability.** Again, the sample results are pretty close to the population, and different from the results we got in the first sample.

In both the examples, we have numbers that describe the population, and numbers that describe the sample. In Example 1, the number 42% is the population proportion of blood type A, and 39.6% is the sample proportion (in sample 1) of blood type A. In Example 2, 69 and 2.8 are the population mean and standard deviation, and (in sample 1) 68.7 and 2.95 are the sample mean and standard deviation.

parameter and statistic

(definition) A **parameter** is a number that describes the population; a **statistic** is a number that is computed from the sample.

In Example 1: 42% is the parameter and 39.6% is a statistic.

In Example 2: 69 and 2.8 are the parameters and 68.7 and 2.95 are the statistics.

In this course, as in the examples above, we focus on the following parameters and statistics:

- population proportion and sample proportion
- population mean and sample mean
- population standard deviation and sample standard deviation

The following table summarizes the three pairs, and gives the notation

	(Population) Parameter	(Sample) Statistic
Proportion	p	$\hat{p}$
Mean	μ	$\overline{x}$
Standard Deviation	σ	s

The only new notation here is p for population proportion (p = 0.42 for type A in Example 1), and $\hat{p}$ for sample proportion

($\hat{p}$ = 0.396 for type A in Example 1).

Comments

1. Parameters are usually unknown, because it is impractical or impossible to know exactly what values a variable takes for every member of the population.

2. Statistics are computed from the sample, and vary from sample to sample due to **sampling variability**.

In the last part of the course, statistical inference, we will learn how to use a statistic to draw conclusions about an unknown parameter, either by estimating it or by deciding whether it is reasonable to conclude that the parameter equals a proposed value. In this module, we'll learn about the behavior of the statistics assuming that we know the parameters. So, for example, if we know that the population proportion of blood type A in the population is 0.42, and we take a random sample of size 500, what do we expect the sample proportion ($\hat{p}$) to be?

Here are some more examples:

Example

If students picked numbers completely at random from the numbers 1 to 20, the proportion of times that the number 7 would be picked is .05. When 15 students picked a number "at random" from 1 to 20, 3 of them picked the number 7. Identify the parameter and accompanying statistic in this situation.

The parameter is the population proportion of random selections resulting in the number 7, which is p = 0.05. The accompanying statistic is the sample proportion of selections resulting in the number 7, which is $\hat{p} = 3/15 = 0.20$.

Example

The length of human pregnancies has a mean of 266 days and a standard deviation of 16 days. A random sample of 9 pregnant women was observed to have a mean pregnancy length of 270 days, with a standard deviation of 14 days. Identify the parameters and accompanying statistics in this situation.

The parameters are population mean $\mu = 266$ and population standard deviation $\sigma = 16$. The accompanying statistics are sample mean $\overline{x} = 270$ and sample standard deviation $s = 14$.

Scenario: SAT Verbal Scores

The SAT-Verbal scores of a sample of 300 students at a particular university had a mean of 592 and standard deviation of 73.

According to the university's reports, the SAT-Verbal scores of all its students had a mean of 580 and a standard deviation of 110.

Did I Get This

1/1 point (graded)
Which of the following is a statistic?

- 300
- 592 ✔
- 580
- 110

Answer
Correct: Indeed, 592 is the sample mean, which is a statistic.

Submit

Did I Get This

1/1 point (graded)
Which of the following is a parameter?

- 300
- 592
- 110 ✔
- 73

Answer
Correct: Indeed, 110 is the population standard deviation, which is a parameter.

Submit

Did I Get This

1/1 point (graded)
What is the appropriate value for $\bar{x}$?

592 ✔

Answer
Correct: $\bar{x}$ is the sample mean which is 592.

Submit

Did I Get This

1/1 point (graded)
What is the appropriate value for μ?

580 ✔

Answer
Correct: μ is the population mean which is 580.

Submit

Did I Get This

1/1 point (graded)
What is the appropriate value for s?

73 ✔

Answer

Correct: s is the sample standard deviation which is 73.

Submit

Did I Get This

1/1 point (graded)
What is the appropriate value for σ?

110

Answer

Correct: σ is the population standard deviation which is 110.

Submit