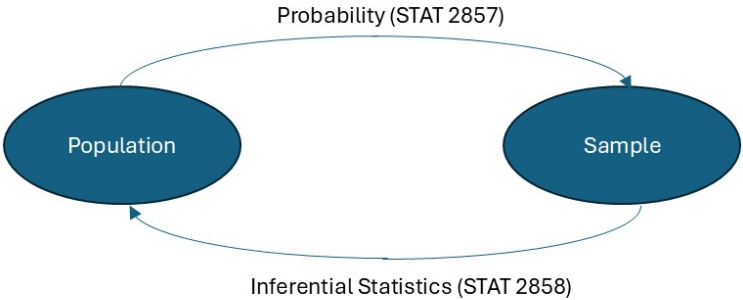
Probability and Statistics I

24. Statistics and their Distributions

6.1 Statistics and their Distributions


Probability (STAT 2857)
Population
Sample
Inferential Statistics (STAT 2858)

Statistic

A statistic is any quantity whose value can be calculated from sample data.

Statistic

A statistic is any quantity whose value can be calculated from sample data.

Statistics are random variables.

Sampling Distribution

The sampling distribution is the probability distribution of a statistic.

We use this term to highlight the fact that it is the distribution we expect to see if we repeatedly sampled many, many, many times.

Sampling

There is an entire field within statistics that considers how to sample from a population to ensure that the observations are representative (SS3843A: Introduction to Study Design).

Random Sample

The random variables $X_1, \ldots, X_n$ form a random sample of size n if

1. The X_is are mutually independent.
2. Every X_i has the same probability distribution.

We also say that the X_is are independent and identically distributed or *iid*.

Example 24.1

The website

www.roulettesimulator.net/

provides a free roulette simulator. Playing in free mode we can make bets in the fictional FUN currency starting with a balance of FUN 5000.

I believe that the simulator may be rigged in free mode so that players win more often than they should – possibly prompting them to "Play for Real Money".

Example 24.1 ctd

Spin the wheel at www.roulettesimulator.net/ 5 times placing FUN 1000 bets on black each time. Enter your final balance in the spreadsheet here.

a) Describe the simulated distribution for the balance.

b) What are some statistics you could compute from our sample?

Example 24.2

For student i denote:

- Number of wins: $X_i \sim \text{Binomial}(5, 18/37)$
- Final balance: $W_i = 5000 + 1000X - 1000(5 - X) = 2000X_i$.

The average balance over all students is

$$\bar{W} = \left(\sum_{i=1}^{n} W_i \right) / n.$$

a) What is the sampling distribution of $\bar{W}$ if the simulator is realistic?

b) What are the mean and variance of $\bar{W}$?

c) Do you think the simulator is realistic?

Questions?

Exercise 24.1

Suppose that each student spun the wheel on the online roulette simulator betting on black repeatedly until they had won 5 times. Let X_i be the number of times that the i-th student played and

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

be the average number of times played per student.

a) What is the sampling distribution of $\bar{X}$ if the simulator is realistic?

b) What are the mean and variance of $\bar{X}$?

c) How could you use the sample values to determine if the simulator is realistic?