Table Number:\_\_\_\_\_\_\_\_\_\_ Group Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Group Members:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Sampling Distribution of Sample Proportions

Purpose: to help you understand what a **sampling distribution** is.

Consider a small population of 5 students. Of these five students, two are female and three are male. Let’s refer to them as F1, F2, M1, M2, and M3. Note that the proportion of females in this population is 0.40.

1. **SAMPLE SIZE 2**. Suppose now we want to consider samples of this population with sample size 2, i.e. . List ALL the possible samples below. (You should have 10 different samples with ).

 In the table below, list the sample and the proportion of females in the sample;

|  |  |  |
| --- | --- | --- |
| SAMPLE | PROPORTION OF FEMALES in the sample(***statistic***),  | PROPORTION OF FEMALES *in the population**(****parameter****),p.* |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Which can change from sample to sample, the statistic or parameter? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which always stays the same from sample to sample, the statistic or parameter? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Value of | Probability of  |
| 0 |  |
| 0.50 |  |
| 1.00 |  |

Write the probability of each  in the table below

Find the mean, or expected value, of this probability distribution and write it here:\_\_\_\_\_\_\_\_\_\_\_\_\_

This probability distribution of the sample proportions,, (a statistic) is called a **sampling distribution.**

 Now sketch a probability histogram of , the proportion of females you obtained in each sample.

1. **SAMPLE SIZE 3**. Suppose now we randomly select a sample of 3 students. List ALL the possible outcomes below. (You should have 10 different samples with).

In the table below, list the sample and the proportion of females in the sample;

|  |  |  |
| --- | --- | --- |
| SAMPLE | PROPORTION OF FEMALES in the sample(***statistic***),  | PROPORTION OF FEMALES in the population*(****parameter****),p.* |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Which can change from sample to sample, the statistic or parameter?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which always stays the same from sample to sample, the statistic or parameter?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Value of | Probability of  |
| 0 |  |
|  |  |
|  |  |

Write the probability of each  in the table below.

Find the mean, or expected value, of this probability distribution and write it here:\_\_\_\_\_\_\_\_\_\_\_\_\_

This probability distribution of the sample proportions (a statistic) is called a **sampling distribution.**

Now sketch a probability histogram of , the proportion of females you obtained in each sample.

1. **SAMPLE SIZE 4**. Suppose now we randomly select a sample of 4 students. List ALL the possible outcomes below. (You should have 5 different samples with).

In the table below, list the sample and the proportion of females in the sample;

|  |  |  |
| --- | --- | --- |
| SAMPLE | PROPORTION OF FEMALES in the sample(***statistic***),  | PROPORTION OF FEMALES in the population*(****parameter****),p.* |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Which can change from sample to sample, the statistic or parameter? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Which always stays the same from sample to sample, the statistic or parameter?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Value of | Probability of  |
| 0.25 |  |
| 0.5 |  |

Write the probability of each  in the table.

Now sketch a probability histogram of , the proportion of females you obtained in each sample.

Complete: a **sampling distribution** is **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

No matter how many different samples we take, the value of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is always the same, but the value of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ can change from sample to sample.

The expected value, or mean, of the sampling distribution is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. This means that *on average*, the estimator, , is the same as the population parameter. **Bias** of an estimator is measured as the distance between the mean (expected value) of the sampling distribution and the population parameter. Thus, our estimator has **no bias.** (This is a good thing!)

The **precision** of an estimate is reflected in the spread of the sampling distribution. The standard deviation of the sampling distribution measures this spread and thus is a measure of the estimator’s (’s) precision. The standard deviation of a sampling distribution is called the **standard error.**

Take another look at the sampling distributions above. Comment about the change in the spread of the distributions as the sample size increases.