STA 3180 Statistical Modelling: Sampling

Sampling for STA 3180 Statistical Modelling

Definition

Sampling is the process of selecting a subset of data from a larger population. This subset of data is used to make inferences about the population as a whole.

Key Concepts

- Sampling is an important tool for making inferences about a population.

- Sampling techniques can be divided into two categories: probability sampling and non-probability sampling.

- Probability sampling techniques involve selecting a sample from a population using random selection, while non-probability sampling techniques involve selecting a sample from a population without using random selection.

Coding Examples

Example 1 - Simple Random Sampling
Start of Code
```
#import the necessary libraries
import numpy as np
#define the population
population = [1,2,3,4,5,6,7,8,9,10]
#select a sample of size 5 from the population
sample = np.random.choice(population, size=5, replace=False)
#print the sample
print(sample)
```
End of Code

Example 2 - Stratified Sampling
Start of Code
```
#import the necessary libraries
import numpy as np
#define the population
population = [1,2,3,4,5,6,7,8,9,10]
#divide the population into two strata
```

```
stratum_1 = population[0:5]
stratum_2 = population[5:]
#select a sample of size 5 from each stratum
sample_1 = np.random.choice(stratum_1, size=5, replace=False)
sample_2 = np.random.choice(stratum_2, size=5, replace=False)
#combine the samples
sample = sample_1 + sample_2
#print the sample
print(sample)
```
End of Code

Practice Multiple Choice Questions

Q1. Which of the following is an example of a probability sampling technique?

A. Simple random sampling

B. Stratified sampling

C. Convenience sampling

D. Cluster sampling

Answer: A. Simple random sampling