LES FACULTÉS
DE L'UNIVERSITÉ
CATHOLIQUE DE LILLE

Inference & Estimation theory

STATISTICAL SAMPLING METHODS
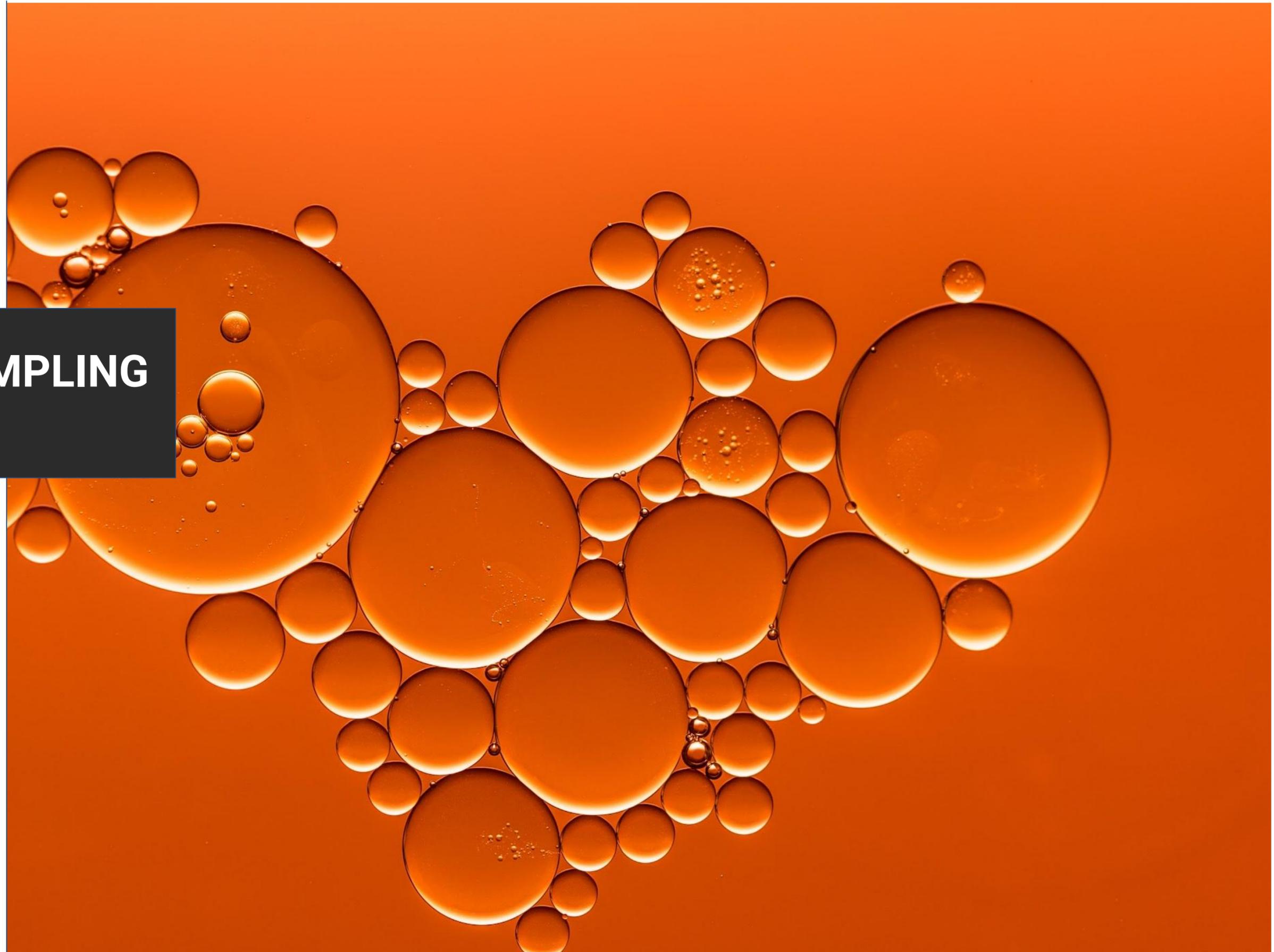
Baptiste Mokas

baptiste.mokas@gmail.com

weeki.io/dynamical-système

linktr.ee/baptistemokas

+33 7 69 08 54 19

weeki
scientific workflow system

KNOWLEDGE TREE

Part 1: Introduction to Inference and Estimation Theory

1.1 Basics of Statistical Inference

- What Is Statistical Inference?
- The Role of Probability in Inference
- Types of Statistical Inference

1.2 Point Estimation

- Estimators and Estimates
- Properties of Good Estimators (Bias, Consistency, Efficiency)
- Maximum Likelihood Estimation (MLE)

1.3 Interval Estimation

- Confidence Intervals
- Construction of Confidence Intervals
- Interpreting Confidence Intervals

Part 2: Hypothesis Testing and Model Evaluation

2.1 Foundations of Hypothesis Testing

- Null and Alternative Hypotheses
- Significance Level (Alpha)
- p-Values and Their Interpretation

2.2 Parametric Hypothesis Tests

- Z-Tests and T-Tests
- One-Sample and Two-Sample Tests
- Analysis of Variance (ANOVA)

2.3 Nonparametric Hypothesis Tests

- Wilcoxon Signed-Rank Test
- Mann-Whitney U Test
- Kruskal-Wallis Test

2.4 Model Evaluation Metrics

- Accuracy, Precision, and Recall
- F1-Score and ROC Curve
- Confusion Matrix

Part 3: Statistical Sampling Methods

3.1 Simple Random Sampling

- Principles of Simple Random Sampling
- Methods for Generating Random Samples
- Sampling Error and Bias

3.2 Systematic Sampling

- Systematic Sampling Process
- Advantages and Disadvantages
- Sample Size Considerations

3.3 Stratified Sampling

- Stratification and Strata Selection
- Stratified Sampling Techniques
- Improved Precision and Efficiency

3.4 Cluster Sampling

- Cluster Sampling Methodology
- Advantages and Challenges
- Cluster Sampling vs. Simple Random Sampling

Part 4: Advanced Sampling Techniques

4.1 Multistage Sampling

- Multistage Sampling Process
- Sampling in Multiple Stages
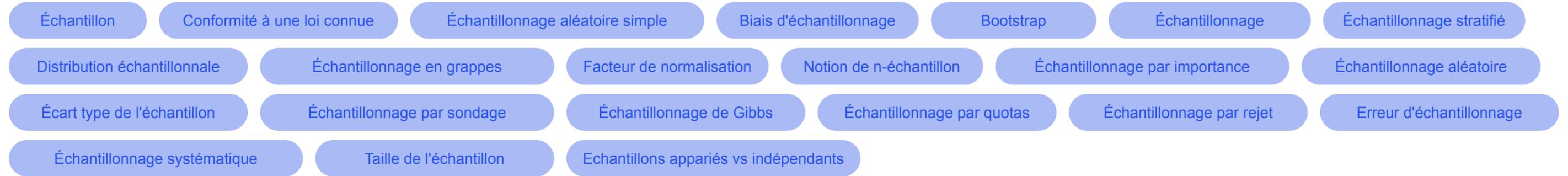- Complex Survey Designs

4.2 Sampling for Surveys and Polls

- Polling Methods and Techniques
- Survey Sampling Strategies
- Margin of Error and Confidence Intervals in Surveys

4.3 Nonprobability Sampling

- Convenience Sampling
- Purposive Sampling
- Snowball Sampling

Échantillon

Conformité à une loi connue

Échantillonnage aléatoire simple

Biais d'échantillonnage

Bootstrap

Échantillonnage

Échantillonnage stratifié

Distribution échantillonnale

Échantillonnage en grappes

Facteur de normalisation

Notion de n-échantillon

Échantillonnage par importance

Échantillonnage aléatoire

Écart type de l'échantillon

Échantillonnage par sondage

Échantillonnage de Gibbs

Échantillonnage par quotas

Échantillonnage par rejet

Erreur d'échantillonnage

Échantillonnage systématique

Taille de l'échantillon

Echantillons appariés vs indépendants

KEYWORDS

- Inference and Estimation Theory
- Statistical Inference
- Probability in Inference
- Types of Statistical Inference
- Point Estimation
- Estimators
- Bias in Estimation
- Consistency in Estimation
- Efficiency in Estimation
- Maximum Likelihood Estimation (MLE)
- Interval Estimation
- Confidence Intervals
- Hypothesis Testing
- Null Hypothesis
- Alternative Hypothesis
- Significance Levels
- Alpha Level
- P-Values
- Parametric Hypothesis Tests
- Z-Tests
- T-Tests
- ANOVA
- Nonparametric Hypothesis Tests
- Wilcoxon Signed-Rank Test
- Mann-Whitney U Test
- Kruskal-Wallis Test
- Model Evaluation Metrics
- Accuracy
- Precision
- Recall
- F1-Score
- ROC Curve
- Confusion Matrix
- Statistical Sampling Methods
- Simple Random Sampling
- Systematic Sampling
- Stratified Sampling
- Cluster Sampling
- Advanced Sampling Techniques
- Multistage Sampling
- Sampling for Surveys and Polls
- Margin of Error
- Nonprobability Sampling
- Convenience Sampling
- Purposive Sampling
- Snowball Sampling
- Data-Driven Decisions
- Statistical Analysis
- Research Methods
- Sampling Methodology
- Statistical Survey Design
- Statistical Polling Methods
- Confidence Intervals in Surveys

In the context of the course on Inference and Estimation Theory, covering topics related to statistical inference, point estimation, interval estimation, hypothesis testing, model evaluation metrics, and statistical sampling methods, let's explore a use case related to survey sampling and polling.

Description:

In this use case, we will focus on applying statistical sampling methods to conduct surveys and polls effectively. Survey sampling is crucial in various fields, including politics, market research, and public opinion analysis, where making accurate inferences about a large population based on a smaller sample is essential.

Key Components:

Introduction to Inference and Estimation Theory: Understanding the fundamentals of statistical inference, including point estimation, interval estimation, hypothesis testing, and model evaluation metrics.

Statistical Sampling Methods: Utilizing techniques such as simple random sampling, systematic sampling, stratified sampling, cluster sampling, multistage sampling, and nonprobability sampling to gather representative samples.

Survey Sampling and Polling: Applying sampling methods to design surveys and polls. Calculating margin of error and confidence intervals for survey results.

Model Evaluation Metrics: Using metrics like accuracy, precision, recall, F1-score, ROC curve, and the confusion matrix to assess the quality of survey and polling data.

Python Code Example (Survey Sampling and Polling):

```python
import numpy as np
import random

# Simulate a population of voters (0 for 'A', 1 for 'B')
population = np.random.randint(2, size=100000)

# Survey sampling using simple random sampling
sample_size = 1000
sample = random.sample(list(population), sample_size)

# Calculate the proportion of voters for candidate 'A' in the sample
proportion_a = np.mean(sample)

# Calculate the margin of error and confidence interval
z_score = 1.96  # 95% confidence level
margin_of_error = z_score * np.sqrt((proportion_a * (1 - proportion_a)) / sample_size)
confidence_interval = (proportion_a - margin_of_error, proportion_a + margin_of_error)

print(f"Sample Proportion for Candidate 'A': {proportion_a:.2f}")
print(f"Margin of Error: {margin_of_error:.4f}")
print(f"95% Confidence Interval: {confidence_interval}")
```

In this code, we simulate a population of voters with two candidates ('A' and 'B'). We then perform simple random sampling to select a sample of voters from the population. We calculate the proportion of voters for candidate 'A' in the sample and determine the margin of error and confidence interval for the survey results.

This use case demonstrates how statistical sampling methods and concepts from the course on Inference and Estimation Theory can be applied in real-world scenarios, such as conducting surveys and polls to make accurate inferences about a larger population.

- Cochran, W. G. (2007). Sampling Techniques. John Wiley & Sons.
- Särndal, C.-E., Swensson, B., & Wretman, J. (2003). Model Assisted Survey Sampling. Springer.
- Lohr, S. L. (2019). Sampling: Design and Analysis. CRC Press.
- Kish, L. (1965). Survey Sampling. John Wiley & Sons.
- Yates, F. (1949). Sampling Methods for Censuses and Surveys. Charles Griffin & Company.
- Thompson, S. K. (2012). Sampling. John Wiley & Sons.
- Kalton, G. (1983). Introduction to Survey Sampling. Sage Publications.
- Lee, E. S., & Forthofer, R. N. (2006). Analyzing Complex Survey Data. Sage Publications.
- Bethlehem, J. (2009). Applied Survey Methods: A Statistical Perspective. Wiley.
- Groves, R. M., et al. (2009). Survey Methodology. John Wiley & Sons.

- Casella, G., & Berger, R. L. (2002). Statistical Inference. Duxbury.
- Lehmann, E. L., & Casella, G. (1998). Theory of Point Estimation. Springer.
- Efron, B., & Tibshirani, R. J. (1994). An Introduction to the Bootstrap. CRC press.
- Wasserman, L. (2004). All of Statistics: A Concise Course in Statistical Inference. Springer.
- Lehmann, E. L., & Romano, J. P. (2005). Testing Statistical Hypotheses. Springer.
- DeGroot, M. H., & Schervish, M. J. (2012). Probability and Statistics. Pearson.
- Fawcett, T. (2006). An Introduction to ROC Analysis. Pattern Recognition Letters.
- Jolliffe, I. T. (2011). Principal Component Analysis. Springer.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). Nonparametric Statistical Methods. John Wiley & Sons.
- Davis, J., & Goadrich, M. (2006). The Relationship Between Precision-Recall and ROC Curves. Proceedings of the 23rd International Conference on Machine Learning.

Embark on a comprehensive exploration of "Inference and Estimation Theory" with this course, covering the core concepts and techniques essential for drawing meaningful conclusions from data.

In the initial section, "Basics of Statistical Inference" introduces you to the foundations of statistical inference, including its definition, the pivotal role of probability in inference, and the various types of statistical inference. You'll then delve into "Point Estimation," where you'll gain an understanding of estimators, the properties that make them good (such as bias, consistency, and efficiency), and the powerful concept of Maximum Likelihood Estimation (MLE). Additionally, "Interval Estimation" comes into focus, covering confidence intervals, their construction, and effective interpretation.

Transitioning to "Hypothesis Testing and Model Evaluation" in Part 2, you'll start with the "Foundations of Hypothesis Testing." Here, you'll learn about null and alternative hypotheses, significance levels (alpha), and how to interpret p-values. The course demystifies "Parametric Hypothesis Tests," including Z-tests, T-tests, and ANOVA, along with their applications. It also explores "Nonparametric Hypothesis Tests," unveiling tests like the Wilcoxon Signed-Rank Test, Mann-Whitney U Test, and Kruskal-Wallis Test, expanding your analytical toolkit. Furthermore, essential model evaluation metrics such as accuracy, precision, recall, F1-score, ROC curve, and the confusion matrix will be covered.

In Part 3, "Statistical Sampling Methods" takes center stage. You'll explore principles like Simple Random Sampling, Systematic Sampling, Stratified Sampling, and Cluster Sampling, each with its methodology, advantages, and considerations. These techniques will equip you to make informed decisions when selecting a sampling method for your research or analysis.

The course culminates in Part 4 with "Advanced Sampling Techniques." You'll delve into Multistage Sampling, a complex process involving multiple stages, and examine its role in complex survey designs. Additionally, you'll learn about Sampling for Surveys and Polls, exploring polling methods, survey sampling strategies, and how to calculate the margin of error and confidence intervals in surveys. Lastly, "Nonprobability Sampling" is discussed, covering techniques such as Convenience Sampling, Purposive Sampling, and Snowball Sampling, shedding light on their applications and limitations.
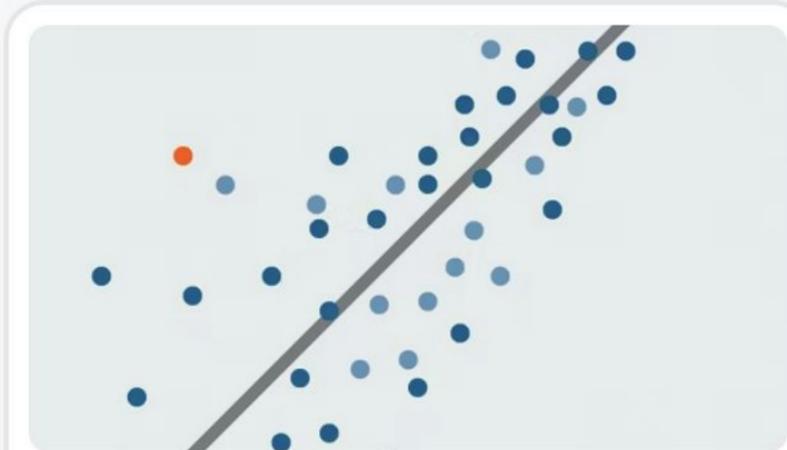
By the end of this course, you'll have a solid grasp of inference and estimation theory, enabling you to make data-driven decisions and conduct rigorous statistical analysis.
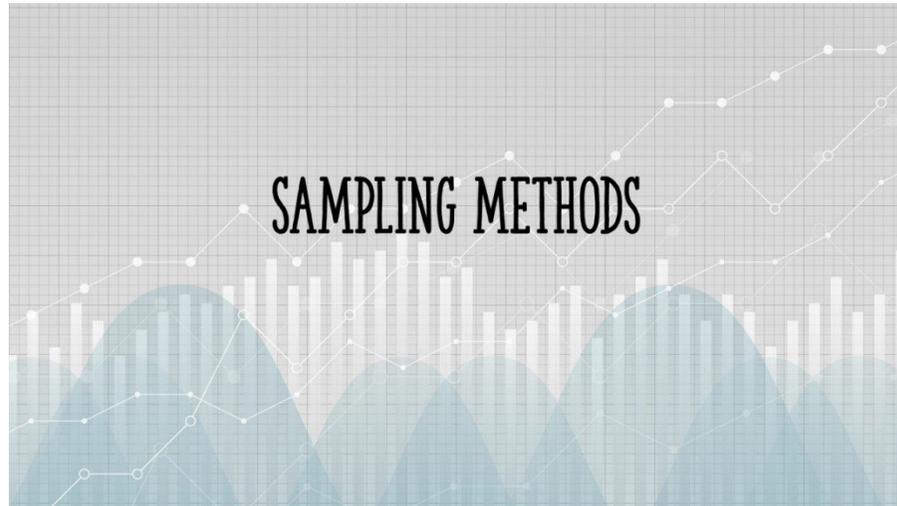
SAMPLING METHODS

Author: Baptiste Mokas, Weeki

Course Name: Simple Linear Regression

#InferenceAndEstimationTheory
#HypothesisTesting
#StatisticalSamplingMethods

Duke University

Linear Regression and Modeling

Compétences que vous acquerrez: Probability & Statistics, Regression, Business Analysis, Data Analysis, General Statistics, Statistical Analysis,…

⭐ **4.8** (1.7k avis)

Débutant · Course · 1 à 4 semaines