



**Parkland College**  
**Arba Minch**  
**Short Term Trainings Programs**

---

**Introduction to Scan Statistics and Methods**

**About the Course**

In statistics, a Scan statistic or window statistic is a problem relating to the clustering of randomly positioned points. Scan statistics are the most important tool for detecting clusters. A scan statistic has the objective of detecting and evaluating the statistical significance of clusters that cannot be explained by the assumption of randomness. This is done by moving a window over the study region and identifying a region, if there is one, with a higher concentration of points than should occur by chance. Detection of such regions is traditionally performed by maximizing a likelihood ratio as will be shown.

Researchers encounter at least nine types of datasets where they might wish to identify spatial clusters using scan statistics: point data, case-control data, aggregated data, spatial-temporal data, spatial survival data, event data, multinomial data, ordinal data, and ime series data.

This training course introduce Scan-statistics and summarizes useful formulae for approximate and exact probabilities and moments of scan statistics with retrospective and prospective scanning of events over time, and references available tables and computer programs.

**Course Duration**

- 7 Days at 8 hours a day

**Course Learning Outcomes**

After studying this course, you should be able to:

- Understand the types of Scan Statistics
- Differentiate a Retrospective and Prospective Scanning of Events over Time
- Understand Success Scans in a Sequence of Trials
- Understand Higher-Dimensional Scans

**Course Outline**

**1. Introduction**

**1.1** Definition of Scan Statistics

**1.2** Types of Scan Statistics

1.2.1 Spatial, Temporal and Space-Time Scan Statistics

**1.3** Scan Statistic Models

1.3.1 Bernoulli versus Ordinal Mode

1.3.2 Normal versus Exponential Model

1.3.3 Normal versus Ordinal Model

1.3.4 Discrete versus Homogeneous Poisson Model



- 1.4 The Discrete Scan Statistics
  - 1.4.1 One-Dimensional Case
  - 1.4.2 Multidimensional Case
- 1.5 The Continuous Scan Statistics
  - 1.5.1 One-Dimensional Case
  - 1.5.2 Multidimensional Case
- 1.6 Power of the Scan Statistic
- 1.7 Clusters and Intuition
- 1.8 Applications of Scan Statistic for Medical and Health Sciences
2. **Retrospective Scanning of Events Over Time**
  - 2.1 Conditional Case: Uniform Distribution of Events
  - 2.2 The Scan Statistic on the Circle
  - 2.3 The Ratchet Scan Statistic
  - 2.4 Moments of Scan Statistics
    - 2.4.1 Exact Values for Moments of the Scan Statistic
3. **Prospective Scanning of Events over Time**
  - 3.1 Poisson Distribution of Events
  - 3.2 Handling Trends or Seasonality in Data
  - 3.3 Moments of Scan Statistics
    - 3.3.1 The Expected Waiting Time until a Cluster
  - 3.4 The Distribution of the Number of Clusters
  - 3.5 The Scan Statistic on the Circle
4. **Success Scans in a Sequence of Trials**
  - 4.1 Binomial Distribution of Events: Discrete Time, Unconditional Case
  - 4.2 A Null Model for the Unconditional Case: The Bernoulli Process
  - 4.3 The Charge Problem
  - 4.4 Binomial Distributed Events: Discrete Time, Conditional Case
  - 4.5 Moments of Scan Statistics
5. **Higher-Dimensional Scans**
  - 5.1 Introduction
  - 5.2 The Conditional Problem
  - 5.3 The Unconditional Problem
  - 5.4 Clustering on the Lattice
6. **Scan Statistics in DNA and Protein Sequence Analysis**
  - 6.1 Introduction
  - 6.2 Scanning for Clusters of Patterns
  - 6.3 Matching in DNA Sequences
  - 6.4 Matching in Multiple Random Letter Sequences
  - 6.5 Sequencing Fragments to Reconstruct a Genome
  - 6.6 Using Double Scans for More Effective Searches for Homologies
  - 6.7 Correlated Descendant Sequences Scan Statistics

### **Target Participants**

The intended audience for this training activity is Statisticians, Biostatisticians and Researchers including Epidemiologist, Biologists, Clinicians, Laboratory and Health professionals with the following prerequisite.

### **Prerequisites**



No prior working knowledge of SaTScan software is required for this course. Understanding of the basic statistical/ bio-statistical concepts, basic computer operational skills and data intuition are required.

### **Training Approach**

This training on Scan Statistics and Methods is delivered by our seasoned trainers who have vast experience as expert professionals using SaTScan Software. The course is mainly focused on theoretical parts of Scan Statistics and Methods and taught through a mix of practical activities, theory and group works.

Training manuals and additional reference materials are provided to the participants.

### **Certification**

Upon successful completion of this course, participants will be issued with a certificate.

### **Tailor-Made Course**

We can also do this as a tailor-made course to meet organization-wide needs. A training needs assessment will be done on the training participants to collect data on the existing skills, knowledge gaps, training expectations, and tailor-made needs.