

# Short Course

## Modern Methods in Sport Statistics

### Session 1

## Overview of Modern Statistical Methods in Sport Science Research



Chapman & Hall/CRC  
Handbooks of Modern  
Statistical Methods

## Handbook of Statistical Methods and Analyses in Sports

*Edited by*  
Jim Albert  
Mark E. Glickman  
Tim B. Swartz  
Ruud H. Koning

 CRC Press  
Taylor & Francis Group  
A CHAPMAN & HALL BOOK

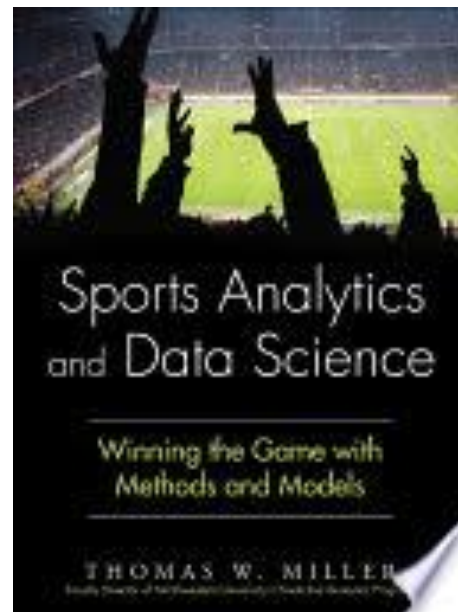
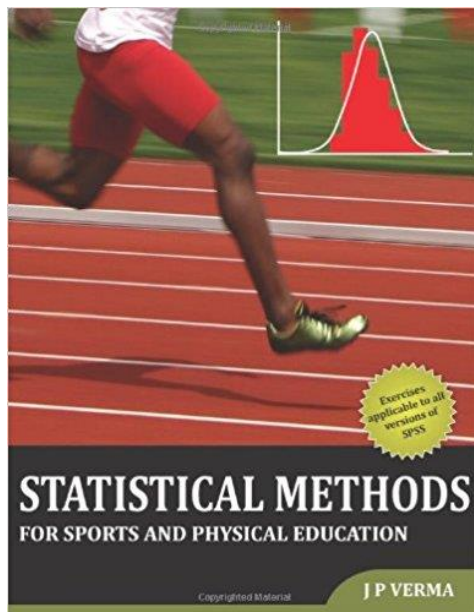
J. P. VERMA

## SPORTS RESEARCH

WITH ANALYTICAL  
SOLUTION USING SPSS®



WILEY

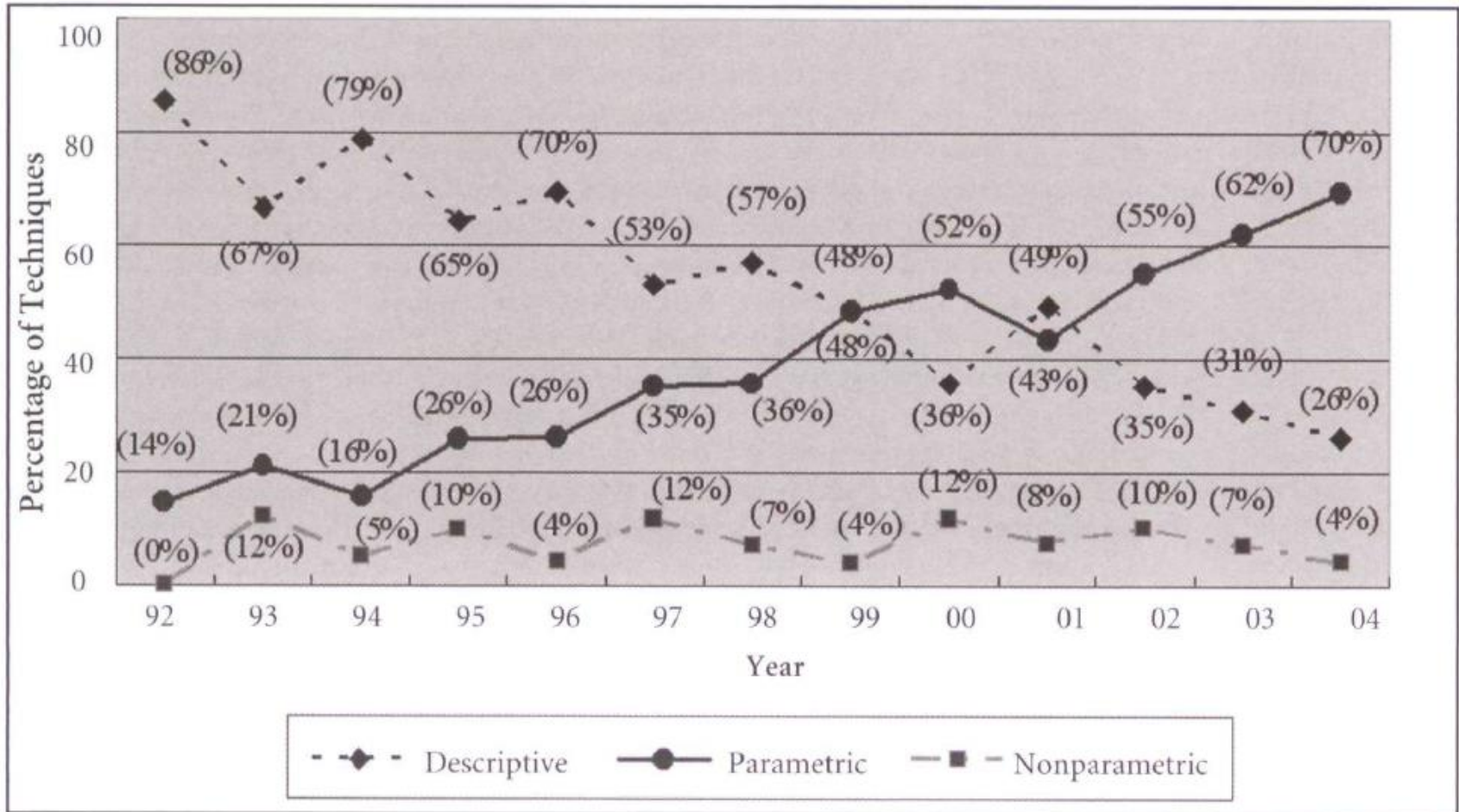


# What are the current methods?

Technique	Definition	Example
Descriptive statistics	Describe, organise, summarise numerical data	mean, mode, median, variance, standard deviation, range, percentage, frequency, rank
Parametric statistics	Make inferences about a population when data have been measured on an interval or ratio scale	t-test, ANOVA, MANOVA, ANCOVA, MANCOVA, correlation, linear and logistic regression, trend analysis, factor analysis, discriminant analysis, cluster analysis, multidimensional scaling, path analysis, structural equation model, conjoint analysis, Z score
Nonparametric statistics	Make inferences about a population when data have been measured on a nominal or ordinal scale	chi-square, Mann-Whitney, Kruskal Wallis, Spearman rank correlation

Quarterman et al. (2005) Statistical Data Analysis Techniques Employed in Sports Marketing Quarterly: 1992 to 2004. Scholarworks@Georgia State University

# What are trends?



Quartermann et al. (2005) Statistical Data Analysis Techniques Employed in Sports Marketing Quarterly: 1992 to 2004. Scholarworks@Georgia State University

# What's the problem?

## 1. “Playing the stats”

- p-hacking
- Hypothesis after results are known (HARKing)
- File-drawer effect
- Null hypothesis significance testing (ignoring magnitude of effect)

## 2. Simplistic statistical models

## 3. Complex data

# What's new?

- Statistical Machine Learning approaches
- Bayesian approaches
- Algorithms, software, consultancies

Bernards, J.R. et al. (2017) Current Research and Statistical Practices in Sport Science and a Need for Change. Sports 5, 87.

Mengersen et al. (2016) PLoS ONE.

Claire Gormley, Blog post: The science of statistics: its application in sports performance analytics, 09 August 2016



# Statistical Machine Learning

## Classification

- Logistic and Multinomial Regression (Generalised Linear Models)
- Support Vector Machines (SVM)
- Neural Networks
- Classification Trees (CART)
- K-Nearest Neighbour (KNN)
- Bayesian Networks

## Clustering

- Mixture Models
- K-means
- Agglomerative Clustering

## Regression

- Linear Models
- Regression Trees
- State Space Models
- Neural Networks
- Functional analysis
- Bayesian Networks

## Dimension reduction

- Functional Data Analysis
- Principal Components Analysis



# What's new?

## Statistical Machine Learning

### Classification

- Regression (Generalised linear models)
- Classification Trees (CART)

### Regression

- Regression
  - Linear models
  - Hierarchical Models
  - State Space Models
- Regression Trees