

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Statistical Methods for Cricket Team Selection

A THESIS PRESENTED IN PARTIAL FULFILMENT
OF THE REQUIREMENT OF THE DEGREE OF
MASTER OF APPLIED STATISTICS
AT MASSEY UNIVERSITY, ALBANY
NEW ZEALAND

Paul J. Bracewell

1999

Abstract

Cricket generates a large amount of data for both batsmen and bowlers. Methods for using this data to select a cricket team are examined. Utilising the assumption that an individual's natural ability is expressed via performance outputs, this thesis seeks to describe and understand the underlying statistical processes of player performance. Randomness is tested for and then the distributional properties of the data are sought.

This information is then used to monitor the estimate of natural ability via widely accepted control methods, such as Shewhart control charts, CUSUM, EWMA and multivariate versions of these procedures. To accommodate the distribution presented by batting scores, a new control chart based on quartiles is also studied.

Further, ranking and selection procedures employ the estimates of individual ability to select the best individuals and note the probability of correct selection.

Major contributions of this study include:

- a) Development of performance measures for cricket
- b) 2 – Dimensional runs test, with further applicability outside cricket.
- c) Statistical interpretation specific to cricket
 - Outliers are very important
 - Form is autocorrelation
 - Zone rules for cricket needed to detect good/poor performance
 - Relatively short nominal ARL's
- d) Control Chart based on quantiles to preserve outlier influences in a non-parametric procedure.
- e) The recommendation of appropriate tools for monitoring batting, bowling and all-rounder performance and also choosing man of the match.
- f) Discriminates between different types of bowlers using the consistency of their performance measures.
- g) Evaluates the members of a team relative to potential contenders.

Acknowledgements

I am fortunate to have had two exceptional people help me through this thesis that I need to thank. The first, my supervisor, Dr Denny Meyer with whom I'm very grateful to have been taught, guided and inspired by in the pursuit of knowledge. I know that I'm not the only student to be appreciative of Denny's assistance.

Secondly, Jenita Higgs who's ongoing support enabled the continuation of full-time study.

I would also like to thank Dr Howard Edwards for his help in the fourth Chapter, on ranking and selection.

Contents

Acknowledgements.....	iii
Table of Contents.....	iv
List of Figures.....	vi
List of Tables.....	vii
Chapter 1 An Overview.....	1
1.1 Introduction.....	1
1.2 General Overview of Cricket Statistics.....	2
1.3 Statistics and Team Selection.....	5
1.4 The Application of Quality Control to Cricket.....	6
1.5 Major Contributions of this Study.....	8
Chapter 2 Analysing Individual Data Characteristics.....	10
2.0 Introduction.....	10
2.1 Literature Review.....	12
2.1.0 Performance Output Measures.....	12
2.1.1 Investigation of Bowling Measures.....	16
2.1.2 Randomness.....	16
2.1.3 Distribution of Performance Measures.....	18
2.2 Data.....	20
2.2.1 Calculation of Batting Measures.....	20
2.2.2 Calculation of Bowling Measures.....	22
2.3 Methods.....	29
2.3.0 Introduction.....	29
2.3.1 Tests for Randomness.....	30
2.3.2 Distribution Fitting.....	34
2.4 Results.....	38
2.4.0 Introduction.....	38
2.4.1 Batting Results.....	38
2.4.2 Bowling Results.....	40

2.5	Discussion.....	41
Chapter 3	Monitoring Player Performance.....	43
3.1	Introduction.....	44
3.2	Previous Studies.....	46
3.2.1	Applying the Methodology of Previous Studies.....	48
3.3	Univariate Quality Control Methodology.....	54
3.3.1	Shewhart Control Charts.....	55
3.3.2	CUSUM.....	63
3.3.3	EWMA.....	66
3.4	Proposed Control Chart Based on Quartiles.....	68
3.5	Comparison of Univariate Control Chart Performance.....	75
3.6	Multivariate Charts.....	82
3.6.1	Hotelling T^2 Chart.....	83
3.6.2	MEWMA.....	90
3.7	Summary.....	92
Chapter 4	Ranking and Selection as Applied to Cricket.....	94
4.1	Ranking the Estimate of an Individual's Ability.....	96
4.2	Homogeneity of Variance.....	97
4.3	Establishing the Probability of Correct Selection.....	101
4.4	New Zealand Statistical XI 94-98.....	103
Chapter 5	Implementation of Statistics in Selection Methodology...	107
Appendix A	Data Description.....	110
Appendix B	Not Out: The Eleventh Form of Dismissal?.....	112
Appendix C	Bowling Results.....	117
Appendix D	Batting Results.....	118
Appendix E	Ducks and Runs Distribution Theoretical Quartiles.....	120
Appendix F	Eligible Players for NZ Statistical XI.....	126
References	127

List of Figures

Figure 1. Relative Effectiveness of the Attack Index	27
Figure 2. Relative Effectiveness of the Economy Index	28
Figure 3. Distributional Comparison of Contribution and Score	42
Figure 4. Non-parametric EWMA for B.R. Hartland	50
Figure 5. Non-parametric EWMA for M.J. Horne	50
Figure 6. Distribution-free CUSUM for B.R. Hartland	52
Figure 7. Distribution-free CUSUM for M.J. Horne	52
Figure 8. Control Chart with Warning Lines	57
Figure 9. Control Chart with Warning Lines for Cricket	58
Figure 10. Shewhart Control Chart of M.J. Horne's Transformed Batting Contribution With Zone Run Rules	61
Figure 11. Shewhart Control Chart of B.R. Hartland's Transformed Batting Contribution With Zone Run Rules	62
Figure 12. \bar{C} USUM Control Chart of M.J. Horne's Transformed Batting Contribution	64
Figure 13. CUSUM Control Chart of B.R. Hartland's Transformed Batting Contribution	65
Figure 14. EWMA Control Chart of M.J. Horne's Transformed Batting Contribution With Zone Run Rules	67
Figure 15. EWMA Control Chart of B.R. Hartland's Transformed Batting Contribution With Zone Run Rules	67
Figure 16. Fitted Line Plot of Mean Contribution Vs Mean Score	69
Figure 17. Quartile Control Chart	71
Figure 18. Establishing Number of Consecutive Increasing Points for Alarm	73
Figure 19. Histograms of simulated Score data with and without Transformation	78
Figure 20. Histograms of simulated Contribution data with and without Transformation	79
Figure 21. Quartile Control Chart for M. J. Horne	80
Figure 22. Quartile Control Chart for B. R. Hartland	81
Figure 23. T2 Control Chart of C.M. Brown for Bowling Indices	85
Figure 24. T2 Control Chart of P.J. Wiseman for Bowling Indices	86
Figure 25. Bivariate Control Chart of C.M. Brown for Bowling Indices	87
Figure 26. Bivariate Control Chart of P.J. Wiseman for Bowling Indices	87
Figure 27. T2 Control Chart of A.C. Barnes	89
Figure 28. MEWMA Control Chart of Bowling Indices for C.M. Brown	90
Figure 29. MEWMA Control Chart of Bowling Indices for P.J. Wiseman.	91
Figure 30. Plot Showing Ranked Nature of Population Batting order.	113

List of Tables

Table 1: Tests for randomness in Batting Performance Measures	38
Table 2: Distribution Fitting for Individual Batting Scores	39
Table 3: Distribution Fitting for Individual Batting Contribution	39
Table 4: Tests for randomness in Batting Performance Measures	40
Table 5: Normality test for Individual Bowling Indices	40
Table 6. Signal Length in Consecutive Points	74
Table 7. Comparative ARL's	76
Table 8. Ratio's Comparing Run Length From In-Control to Out-of-Control.	77
Table 9. Theoretical Quartile Limits for Horne and Hartland	79
Table 10. Comparison of the Number of Points to First Alarm for Horne and Hartland for Different Methods	92
Table 11. P-Values from Test for Homogeneity of Variance for Bowlers	97
Table 12. Probability of Correct Selection for NZ Statistical XI	105