

Phillip Good

Permutation Tests

A Practical Guide to Resampling Methods
for Testing Hypotheses

With 13 Illustrations



Springer

Phillip Good
Box 4216
Huntington Beach CA 92605 USA
72733.434@compuserv.com

Cover illustration: Building a minimal spanning tree.

Library of Congress Cataloging-in-Publication Data
Good, Phillip I.

Permutation tests: a practical guide to resampling methods for
testing hypotheses/Phillip Good.

p. cm.—(Springer series in statistics)

Includes bibliographical references and index.

ISBN 978-1-4757-2348-9 ISBN 978-1-4757-2346-5 (eBook)

DOI 10.1007/978-1-4757-2346-5

1. Statistical hypothesis testing. 2. Resampling (Statistics)

I. Title. II. Series.

QA277.G643 1993

519.5'6—dc20

93-9062

Printed on acid-free paper.

© 1994 Springer Science+Business Media New York
Originally published by Springer-Verlag New York, Inc. in 1994
Softcover reprint of the hardcover 1st edition 1994

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher, Springer Science+Business Media, LLC, except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Production managed by Natalie Johnson; manufacturing supervised by Jacqui Ashri.
Typeset by Asco Trade Typesetting Ltd., Hong Kong.

9 8 7 6 5 4 3 2 (Second corrected printing, 1995)

ISBN 978-1-4757-2348-9

Contents

Preface	v
1. A Wide Range of Applications	1
1.1 Permutation Tests	1
1.2 “I Lost the Labels”	2
1.3 Five Steps to a Permutation Test	4
1.4 What’s in a Name?	6
1.5 Questions	8
2. A Simple Test	10
2.1 Properties of the Test	10
2.2 Fundamental Concepts	11
2.3 Which Test?	18
2.4 World Views	21
2.5 Questions	22
3. Testing Hypotheses	24
3.1 One-Sample Tests	24
3.2 Confidence Intervals	27
3.3 Two-Sample Comparisons	29
3.4 Comparing Variances	31
3.5 k -Sample Comparisons	33
Sidebar: Computing the significance level	
3.6 Blocking	39
3.7 Matched Pairs	41
3.8 Questions	42
4. Experimental Designs	44
4.1 Introduction	44

4.2	Balanced Designs	44
	Sidebars: Computing the significance level	
4.3	Analysis of Covariance	55
4.4	Unbalanced Designs	57
4.5	Clinical Trials	60
4.6	Very Large and Very Small Samples	61
4.7	Questions	62
5.	Multivariate Analysis	64
5.1	Introduction	64
5.2	One- and Two-Sample Comparisons	64
	Sidebar: Computing the significance level	
5.3	Runs Test	71
5.4	Experimental Designs	74
5.5	Repeated Measures	75
5.6	Questions	77
6.	Categorical Data	78
6.1	Contingency Tables	78
6.2	Fisher's Exact Test	78
6.3	Unordered $r \times c$ Contingency Tables	83
6.4	Ordered Contingency Tables	90
6.5	Covariates	91
6.6	Combinations of Tables	93
6.7	Questions	93
7.	Dependence	94
7.1	The Models	94
7.2	Testing for Independence	95
7.3	Testing for Trend	96
7.4	Serial Correlation	97
7.5	Known Models	101
7.6	Questions	103
8.	Clustering in Time and Space	105
8.1	The Generalized Quadratic Form	105
8.2	Applications	107
8.3	Extensions	108
8.4	Questions	109

- 9. Coping with Disaster 110
 - 9.1 Missing Data 110
 - 9.2 Covariates After the Fact 112
 - 9.3 Outliers 113
 - 9.4 Censored Data 117
 - 9.5 Censored Matched Pairs 119
 - 9.6 Adaptive Tests 122
 - 9.7 Questions 122

- 10. Which Statistic? Solving the Insolvable 124
 - 10.1 The Permutation Distribution 124
 - 10.2 New Statistics 124
 - 10.3 Going Beyond 132
 - 10.4 Likelihood Ratio 135
 - 10.5 Questions 138

- 11. Which Test Should You Use? 140
 - 11.1 Sources of Variation 140
 - 11.2 Comparison with the Parametric Test and the Bootstrap ... 142
 - 11.3 A Guide to Selection 142
 - 11.4 Quick Key 146

- 12. Publishing Your Results 149
 - 12.1 Design Methodology 149
 - 12.2 Statistical Software for Exact Distribution-Free Inference .. 151
 - 12.3 Preparing Manuscripts for Publication 151

- 13. Increasing Computational Efficiency 153
 - 13.1 Five Techniques 153
 - 13.2 Monte Carlo 153
 - 13.3 Rapid Enumeration and Selection Algorithms 155
 - 13.4 Focus on the Tails: Branch and Bound Algorithms 156
 - 13.5 Characteristic Functions 160
 - 13.6 Asymptotic Approximations 161
 - 13.7 Sample Size, Power, and Confidence Intervals 162
 - 13.8 Some Conclusions 163
 - 13.9 Questions 164

- 14. Theory of Permutation Tests 166
 - 14.1 Fundamental Concepts 166

14.2 Maximizing the Power	169
14.3 Confidence Intervals	174
14.4 Asymptotic Behavior	175
14.5 Questions	178
 Bibliography	 180
Bibliography Part 1: Randomization	181
Bibliography Part 2: Supporting	212
Bibliography Part 3: Computational Methods	214
Bibliography Part 4: Seminal Articles	219
 Index	 221