Case-Control Studies

The case-control approach is a powerful method for investigating factors that may explain a particular event. It is extensively used in epidemiology to study disease incidence, one of the best-known examples being the investigation by Bradford Hill and Doll of the possible connection between cigarette smoking and lung cancer. More recently, case-control studies have been increasingly used in other fields, including sociology and econometrics.

With a particular focus on statistical analysis, this book is ideal for applied and theoretical statisticians wanting an up-to-date introduction to the field. It covers the fundamentals of case-control study design and analysis as well as more recent developments, including two-stage studies, case-only studies, and methods for case-control sampling in time. The latter have important applications in large prospective cohorts that require case-control sampling designs to make efficient use of resources. More theoretical background is provided in an appendix for those new to the field.

RUTH H. KEOGH is a Lecturer in the Department of Medical Statistics at the London School of Hygiene and Tropical Medicine.

D. R. COX is one of the world's pre-eminent statisticians. His work on the proportional hazards regression model is one of the most-cited and most influential papers in modern statistics. In 2010 he won the Copley Medal of the Royal Society 'for his seminal contributions to the theory and application of statistics'. He is currently an Honorary Fellow at Nuffield College, Oxford.

INSTITUTE OF MATHEMATICAL STATISTICS
MONOGRAPHS

IMS Monographs are concise research monographs of high quality on any branch of statistics or probability of sufficient interest to warrant publication as books. Some concern relatively traditional topics in need of up-to-date assessment. Others are on emerging themes. In all cases the objective is to provide a balanced view of the field.

Case-Control Studies

RUTH H. KEOGH
London School of Hygiene and Tropical Medicine

D. R. COX
Nuffield College, Oxford

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

Cambridge University Press is a part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107019560

First published 2014

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Keogh, Ruth H., 1979– author.
Case-control studies / Ruth H. Keogh, D.R. Cox.
p. ; cm. – (Institute of Mathematical Statistics monographs ; 4)
Includes bibliographical references and index.
ISBN 978-1-107-01956-0 (hardback)
I. Cox, D.R. (David Roxbee), author. II. Title. III. Series: Institute of Mathematical Statistics monographs ; 4.
[DNLM: 1. Epidemiologic Methods. WA 950]
RA652
614.4 – dc23 2013047249

ISBN 978-1-107-01956-0 Hardback

Contents

Preface

The retrospective case-control approach provides a powerful method for studying rare events and their dependence on explanatory features. The method is extensively used in epidemiology to study disease incidence, one of the best known and early examples being the investigation by Bradford Hill and Doll of the possible impact of smoking and pollution on lung cancer. More recently the approach has been ever more widely used, by no means only in an epidemiological setting. There have also been various extensions of the method.

A definitive account in an epidemiological context was given by Breslow and Day in 1980 and their book remains a key source with many important insights. Our book is addressed to a somewhat more statistical readership and aims to cover recent developments. There is an emphasis on the analysis of data arising in case-control studies, but we also focus in a number of places on design issues. We have tried to make the book reasonably self-contained; some familiarity with simple statistical methods and theory is assumed, however. Many methods described in the book rely on the use of maximum likelihood estimation, and the extension of this to pseudo-likelihoods is required in the later chapters. We have therefore included an appendix outlining some theoretical details.

There is an enormous statistical literature on case-control studies. Some of the most important fundamental work appeared in the late 1970s, while the later 1980s and the 1990s saw the establishment of methods for case-control sampling in time. The latter have important applications in large prospective cohorts which collect large amounts of information, for example biological samples, but which require case-control sampling designs to make efficient use of resources. There continue to appear in the literature many innovations in case-control study design and, in particular, methodology covering a wide range of areas.

We hope that the book will be useful both to applied and to theoretical statisticians wanting an introduction to the field. Parts of it might be useful

as a basis for a short course for postgraduate students, although we have not written with that use specifically in mind.

The EPIC-Norfolk study provided some examples; we are grateful for permission to use this data and thank the staff and participants of the study.

We are grateful for very helpful comments on an initial draft from the following colleagues and friends: Amy Berrington de Gonzalez, Vern Farewell, Chris Keogh and Ian White. We are very grateful also to Diana Gillooly at Cambridge University Press for her encouragement and helpful advice and to Susan Parkinson for her meticulous and constructive copy-editing.

Ruth Keogh and David Cox