

# Lecture Notes in Statistics

Edited by D. Brillinger, S. Fienberg, J. Gani,  
J. Hartigan, and K. Krickeberg

23

---

Diane Griffin Saphire

Estimation of  
Victimization Prevalence  
Using Data from the  
National Crime Survey

---



Springer-Verlag  
New York Berlin Heidelberg Tokyo 1984

**Author**

Diane Griffin Saphire  
Trinity University, Department of Mathematics  
715 Stadium Drive, San Antonio, TX 78284, USA

AMS Subject Classification (1980): 62D05

ISBN-13: 978-0-387-96020-3      e-ISBN-13: 978-1-4612-5270-2  
DOI: 10.1007/978-1-4612-5270-2

Library of Congress Cataloging in Publication Data Saphire, Diane Griffin. Estimation of victimization prevalence using data from the National crime survey. (Lecture notes in statistics; 23) Bibliography: p. 1. Victims of crimes – Forecasting – United States – Mathematical models. 2. Victims of crimes – United States – Longitudinal studies. 3. Victims of crimes surveys – United States – Mathematical models. 4. Criminal statistics – United States – Mathematical models. I. Title. II. Series: Lecture notes in statistics (Springer-Verlag); v. 23. HV6250.3.U5S22 1984 364.1'0723 84-13909

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law where copies are made for other than private use, a fee is payable to "Verwertungsgesellschaft Wort", Munich.

© by Springer-Verlag Berlin Heidelberg 1984

Printing and binding: Beltz Offsetdruck, Hemsbach / Bergstr.  
2146 / 3140-543210

## Preface

The National Crime Survey is a sample survey of housing units conducted by the U.S. Bureau of the Census. All eligible occupants of a sampled unit are interviewed every six months (for up to seven interviews) about victimizations that they have experienced during the previous six months. In this monograph several longitudinal analyses are performed using a subsample of the data covering the years 1973 through 1975. In particular, several methods of estimating the proportion of units that are crime-free for a given year, denoted by  $\theta$ , are discussed.

First, several ad hoc, as opposed to model-based, estimators of  $\theta$  are discussed, including those used by the Bureau of Justice Statistics. We find models under which these estimators are consistent for  $\theta$ . One such model fits the data very well.

A superpopulation approach to the estimation of  $\theta$  is then taken, assuming that the nonresponse and sampling mechanisms are ignorable. Three models are fit to the data: i) a homogeneous Bernoulli model, under which victimization is independent from month to month ii) a correlated Bernoulli model, under which victimization in any two months has positive correlation  $\rho$ , and iii) a two-state Markov model with states "victimized" and "crime-free". The correlated Bernoulli model is found to be very inadequate. The other two models fit the 1975 data well, but have rather poor fits to the 1973 and 1974 data. Rotation group biases are conjectured to be the cause of these poor fits.

By separating the units with no nonresponse from those with at least one nonresponse and by estimating  $\theta$  for each of these groups, units with at least one nonresponse are found to have a lower probability of being crime-free for the year. To incorporate nonresponse in the modelling, a logit model with explanatory variables "months of information" and "number of noninterviews" is fit and three alternative estimators of  $\theta$  are computed.

In general, the estimates of  $\theta$  lie in the range .65 to .70 with estimated standard

#### IV

errors of about .02. The estimates tend to increase slightly from 1973 to 1975.

\* \* \*

Much of the work presented here was done while I was at Carnegie-Mellon University and I would like to express my gratitude to the faculty, staff and students there for their help in the preparation of this monograph. Many thanks are due to Stephen Fienberg for his support and guidance and for his statistical and editorial assistance. I would like to thank Bill Eddy for all of his statistical and computing help, and especially for his warm words of encouragement. I am indebted to Luke Tierney for his assistance with the work presented in Chapter 3 and with the Markov model variances, as well as his comments on the rest of my work. Thanks are due to Mark Schervish for prompt and insightful responses during both the research and writing stages. Suggestions from Robert Johnson, Alfred Blumstein, and Charles Alexander were very valuable to me in the writing of the final version.

The research presented here was funded by the Bureau of Justice Statistics and the National Institute of Justice, both of the Office of Justice Assistance Research and Statistics, U.S. Department of Justice, under grants J-LEAA-015-79 and 81-IJ-CX-0087, respectively.

## TABLE OF CONTENTS

<b>1. Introduction</b>	<b>1</b>
1.1. Cheery Indicators	1
1.2. Aims of the NCS	4
1.3. The Sample Design and Rotation Scheme	5
1.4. The Interview	8
1.5. Data Tapes	11
1.6. Missing Data	13
1.6.1. Types of Nonresponse	16
1.6.2. The Magnitude of the Nonresponse	17
1.6.3. Ignorable Sampling and Nonresponse Mechanisms	18
<b>2. Some Intuitive Prevalence Estimators</b>	<b>24</b>
2.1. Two Ad Hoc Estimators	24
2.2. The BJS Estimators	27
<b>3. Consistency of the Ad Hoc Models</b>	<b>36</b>
3.1. Modelling Victimization	36
3.2. Consistency of the Ad Hoc Estimator	39
3.3. The Modified Version of the Ad Hoc Estimator	42
3.4. Consistency of the BJS Estimators	56
<b>4. Model-Based Estimators</b>	<b>61</b>
4.1. A Homogeneous Bernoulli Model	61
4.2. A Correlated Bernoulli Model	68
4.3. A Markov Model	74
4.4. Comparisons of Estimates	90
<b>5. Nonresponse</b>	<b>96</b>
5.1. Checking the Missing at Random Assumption	96
5.2. Taking the Nonresponse into Account	105
<b>6. Further Research and Conclusions</b>	<b>122</b>
6.1. Extensions	122
6.1.1. Estimation at the HH Level	122
6.1.2. Possibilities for Further Research	126
6.2. Conclusions	132
<b>Appendix A. Questionnaire</b>	<b>136</b>
<b>Appendix B. Computer Programs</b>	<b>141</b>