

## SAMPLING METHODS FOR APPLIED RESEARCH TEXT AND CASES

### SAMPLING METHODS FOR APPLIED RESEARCH TEXT AND CASES

# PETER TRYFOS York University



JOHN WILEY & SONS, INC

New York ● Chichester ● Brisbane
Toronto ● Singapore

ACQUISITION EDITOR Brad Wiley

MARKETING MANAGER Debra Riegert

SENIOR PRODUCTION EDITOR Tony VenGraitis

DESIGNER Kenny Beck

MANUFACTURING MANAGER Dorothy Sinclair

ILLUSTRATION COORDINATOR Jaime Perea

PRODUCTION COORDINATOR J. Carey Publishing Service

This book was set in 10/12 Times Roman by Eigentype Compositors, and printed and bound by Donnelley/Harrisonburg. The cover was printed by Lehigh Press.

Recognizing the importance of preserving what has been written, it is a policy of John Wiley & Sons, Inc. to have books of enduring value published in the United States printed on acid-free paper, and we exert our best efforts to that end.

The paper on this book was manufactured by a mill whose forest management programs include sustained yield harvesting of its timberland. Sustained yield harvesting principles ensure that the number of trees cut each year does not exceed the amount of new growth.

Copyright © 1996, by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Sections 107 and 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permission Department, John Wiley & Sons, Inc.

#### Library of Congress Cataloging-in-Publication Data

Tryfos, Peter.

Sampling methods for applied research: text and cases / by Peter Tryfos

p. cm

Includes bibliographical references and index.

ISBN 0-471-04727-9 (clotch: alk. paper)

1. Sampling (Statistics) I. Title.

QA276.6.T78 1966

519.5'2-dc20

95-35575 CIP

L.C. Call no. Dewey Classification No. L.C. Card No. ISBN 0-471-04727-9

Printed in the United States of America

10 9 8 7 6 5 4 3 2



This book is intended primarily as a main text at the upper undergraduate or graduate level for a course in applied sampling for students of business, government or health administration, economics, political science, and other social sciences; as a main or supplementary text in service courses on sampling offered by departments of statistics or mathematics; and for self-study by professionals in marketing research, auditing, purchasing, production, opinion polling, and other areas of business and government.

The text accommodates at least two types of course design. The first (the traditional approach for a technical course) requires students to do exercises and problems and to write exams. The second approach (frequently used in business courses) requires students, perhaps working in groups, to read, report, and discuss cases using the text as a resource. They may also do a project and write a paper involving sampling or a critical appraisal of sampling. The text includes examples, exercises, problems, and small and large cases suitable for either approach or for a combination of them. An effort was made to make the text as flexible as possible, so that even the broader cases could be informative to students using the traditional approach, and some of the problems would be interesting to students using the case approach.

This is not a mathematical statistics text, but neither is it a handbook of formulas only. Formulas are indeed provided, and care has been taken to arrange them for easy reference, but a strong effort was made to give intuitive understanding of sampling concepts and confirm sampling formulas by means of simple numerical illustrations.

It is assumed that readers have taken an introductory course in statistics. Since such an introductory course is required by most academic programs, this prerequisite should not be restrictive. However, Appendix A reviews in some detail the essential prerequisite concepts so that the text can be understood in the event these

concepts are not well-remembered, or even (in the case of mature readers) in the absence of an introductory course. High-school mathematics is sufficient. An interest in quantitative methods is presumed.

An effort was made to proceed from need to solution, from the general to the particular, beginning with a view of the forest before dealing with the trees.

Thus, Chapter 1 attempts to demonstrate by means of short cases the need for and importance of sampling. Most of the cases are dealt with fully later in the text. Chapter 2 aims to establish first that the complexity of the problems suggested by the cases in Chapter 1 is more apparent than real. It draws attention to the common elements of these problems and to the fact that in the large majority of applied problems there is need for estimating only two types of population characteristics. Chapter 2 then describes how known auxiliary information suggests different methods for sample selection and estimation. The appropriateness of stratified and single-stage sampling and of ratio/regression estimators in some extreme but highly instructive situations is thus established early, for random as well as other types of samples.

Chapters 3 to 7 deal, respectively, with simple, stratified, two-stage random sampling, ratio and regression estimators, and such special topics as estimation of ratios of population totals and subpopulation means.

Of particular importance in this sequence is Chapter 3, which attempts to provide a justification of random sampling. The "silly game," despite its appearance of frivolity, aims to answer in simple terms a frequently asked question: Why should any weight be given to a method with desirable properties if applied repeatedly and a large number of times, but which, in reality, is applied only once? The same illustration is also used to introduce the criteria (bias, variability) according to which methods of selection and estimation are compared.

Common to Chapters 4 to 7 (indeed, to the remaining chapters as well) is the "forest-to-tree" approach mentioned earlier. Each chapter begins with a motivation for the method, a summary of its features, and examples illustrating its application. This is followed by an explanation of its derivation and a discussion of special or similar cases. Results are summarized in easy to locate boxes. A final section summarizes the main non-technical points of the chapter.

Chapter 8 is the unifying chapter of random sampling. It is presented last rather than first (as the forest-to-tree approach would imply) because otherwise the target audience has too many new concepts to deal with at the very beginning. Chapter 9 presents the prediction approach to sampling, which is not included in other texts at this level. Chapter 10 examines nonresponse, measurement, and sampling frame errors. Randomized response, telephone surveys, and the measurement of television and radio ratings are also discussed in this chapter.

A case is a description of a real situation that lends itself to the application of methods. A case invites reflection and provides an opportunity for discussion. Unlike a standard problem, but very much like the real world, it may not have a single solution. The cases are based on the author's experience, which is largely

Canadian. For the purposes of this text, actual names, places and data were transformed, for several reasons. In the first place, dates such as 1992 were changed to 19X2 etc. in order to avoid needless obsolescence. Secondly, organization names were changed in order not to subject the original ones to unnecessary criticism (which students tend to dispense ferociously). Thirdly, occasionally the data were changed (for example, by applying uniformly a linear transformation) in order to preserve the confidentiality of original sources. Lastly, there was some concern that frequent references to Canadian places and names would be distracting in a book intended to be read in the United States and other countries. For example, the Maritimes may now be described as "East." Despite these changes, however, the essential integrity of the data, the problem, and the setting has been meticulously preserved.

Appendix B provides a comprehensive glossary and technical summary. In addition to demonstrating the unity of the text, it may also be used as a stand-alone reference by practitioners.

A diskette accompanying the text contains the data files referred to in the text and the cases, a simulation program used in one of these cases, and a "no-frills" computer program designed to assist calculations involving complicated formulas. Appendix C describes the data and programs included in this diskette. Finally, Appendix D provides solutions to selected problems.

An *Instructor's Manual*, available from the publisher on request, includes solutions to all the problems in the text, and teaching notes with sample computer programs describing the author's treatment of the cases, possible extensions, and alternatives.

The author is indebted to IMS of Canada Ltd., NCH Promotional Services, and Nielsen Marketing Research for data and other information. None of the cases or examples in this text, however, should be interpreted as describing the actual practices and policies of these firms.

The author is also grateful to Professor K. H. Chan and Mr. Alan Middleton for suggesting and helping design two of the cases, and to the following people for positive and negative comments on earlier drafts that resulted in significant improvements: Professor Bruce L. Bowerman, Miami University; Professor Sharon Hunter Donnelly, University of Tennessee, Knoxville; Professor John L. Eltinge, Texas A&M University; Professor William A. Erickson, The University of Michigan; Professor Harry Joe, The University of British Columbia; Professor George A. Marcoulides, California State University at Fullerton; Professor Peter Peskun, York University; Professor Gary Simon, New York University; and Professor Arthur B. Yeh, Rutgers University.

The greatest debt, however, is owed to Barbara J. Tryfos, who read, suffered through, and corrected several versions of the manuscript, patiently restoring clarity on the many occasions when it was not present.

If any errors and faults remain despite the help of so many, they are entirely the author's.

### CONTENTS

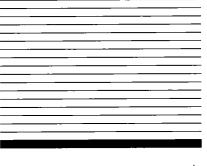
Chapter 1 Introduction  3 2.10 Sampling from homogeneous groups 2.11 Sampling of groups 32 2.12 Exploiting a relationship 35 2.13 A note on terminology 39 2.14 To sum up 40	
1.5 Case: Pharmacom Research 9	30
1.6 Case: Canabag Manufacturing Company 11 1.7 Some reflections 12 1.8 A very brief history of sampling 14 1.9 A disclaimer 15 1.0 Case: Canabag Manufacturing Company 11 3.1 Introduction and summary 41 3.2 A silly game 42 3.3 Simple random samples 45 3.4 The Ackroyd population 47	**
Chapter 2 Preliminaries  16 3.5 Implications 54 2.1 Introduction and summary 16 2.2 Populations and their characteristics 17 2.3 Proportions and indicator variables 23 2.4 First assumptions 24 2.5 First estimates 24  16 3.5 Implications 54 3.6 Selecting a random sample in practic 3.7 How large should a sample be? 60 3.7 Interval estimates 66 3.9 To sum up 70 Problems 71	58
<ul> <li>2.6 Sampling with and without replacement 25</li> <li>2.7 Is a larger sample better? 26</li> <li>2.8 Sampling from a homogeneous population 28</li> <li>2.9 Auxiliary information 29</li> <li>Chapter 4 Stratified random sampling</li> <li>4.1 Introduction and summary 84</li> <li>4.2 Stratified sampling</li> <li>4.3 Stratified sampling</li> <li>4.4 Stratified sampling</li> </ul>	84

4.3 4.4 4.5	The Ackroyd case continued 88 Properties of stratified estimators 94 Proportional stratified sampling 96		6.12	To sum up 185 Problems 185
4.6	"Optimal" stratified sampling 99		Chap	oter 7 Some special topics 19
4.7 4.8	How large a stratified sample? 104 Interval estimates 105		7.1	Introduction and aummonic 105
4.0 4.9	To sum up 107		7.1	Introduction and summary 195
4.9	Problems 108		1.2	Estimating the ratio of totals or means of tw variables 196
	1 Toblems 108		7.3	Estimating subpopulation means and
			1.5	proportions 199
Chap	ter 5 Two-stage random		7.4	Cluster sampling—an alternative view 20'
	sampling 1	20	7.5	Poststratification 209
5.1	Introduction and summary 120		7.6	Double sampling 212
5.2	Some advantages of two-stage		7.7	To sum up 215
	sampling 121		7	Problems 216
5.3	Properties of two-stage estimators 126			1100101113 210
5.4	Special case: cluster sampling 131			
5.5	Special case: systematic sampling 133		Chap	oter 8 Sampling with unequal
5.6	Stratified or two-stage sampling? 136			probabilities 22
5.7	Interval estimates 139		8.1	Introduction and summary 225
5.8	Combined stratified/two-stage		8.2	Inclusion probabilities 226
	sampling 142		8.3	The Horvitz–Thompson estimator
5.9	Multistage sampling 144		0.5	(HTE) 227
5.10	To sum up 145		8.4	Application: Three-stage sampling 231
	Problems 146		8.5	The variance of the HTE and
	110		0.2	implications 234
Cham	ter 6 Ratio and regression		8.6	Inclusion probabilities proportional to size
Cuap	_	56	0.0	(IPPS) 237
	estimators .	30	8.7	Sampling with IPPS 240
6.1	Introduction and summary 156		8.8	Some observations on the HTE 242
6.2	A simple linear relationship 157		8.9	Interval estimates based on the HTE 243
6.3	Ratio estimators and modified simple rando	om	8.10	Sampling with replacement and PPS 245
	sampling 161		8.11	Interval estimates based on the HHE 250
6.4	Properties of the ratio estimator 162		8.12	Application: Dollar-unit sampling 252
6.5	The regression estimator 165		8.13	To sum up 255
6.6	Properties of the regression estimator 16	8	3.10	Problems 256
6.7	Ratio and regression estimators in large samples 171			
6.8	Sample size 176		Chap	oter 9 Sampling from a process 26
6.9	Interval estimates 177		9.1	Introduction and summary 266
6.10	Regression and ratio estimators in stratifie	ed	9.1	The prediction approach 267
	sampling 180		9.2	A simple linear model with no intercept 27
6.11	Regression and ratio estimators in two-sta	ge	9.3 9.4	Best linear unbiased estimator
	sampling 183	<i>.</i>	<b>⊅.</b> च	confirmed 275

	Some variants of the sample linear model	Chapter 10 Nonsampling errors 307
9.6 9.7 9.8 9.9 9.10 9.11	with no intercept 279 Interval estimates, simple linear models with no intercept 281 A simple linear model with intercept 283 A linear model with more than one explanatory variable 287 The prediction approach in general 292 Simple diagnostic tools 292 Prediction vs. randomization 294 Problems 295	<ul> <li>10.1 Introduction and summary 307</li> <li>10.2 Entire population available 308</li> <li>10.3 Identifying the population elements 309</li> <li>10.4 Case: Telephone surveys 311</li> <li>10.5 Nonresponse 313</li> <li>10.6 Measurement error 315</li> <li>10.7 Randomized response 317</li> <li>10.8 Case: Radio and television measurement 320</li> <li>10.9 To sum up 324</li> <li>Problems 327</li> </ul>
PAR	RT TWO: CASES	
Canab Valua Valua Valua	pag Manufacturing Company—Part I 335 pag Manufacturing Company—Part II 340 tion of government properties—Part II 343 tion of government properties—Part II 345 tion of government properties—Part III 353  PENDICES	Valuation of government properties—Part IV 355 PackGoods Inc. 356 Pharmacom Research—Part I 364 Pharmacom Research—Part II 374 Tenderdent Toothpaste 483
	المناسبة والمراجع المراجع المناجع المناجع المناجع المناجع المناجع والمناجع المناجع والمناجع والمناجع	
Appe	endix A Basic concepts 391	B.5 Ratio and regression estimators (chapter 6)
Appe A.1 A.2 A.3 A.4 A.5	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422
A.1 A.2 A.3	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400 Correlation 402	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423
A.1 A.2 A.3 A.4 A.5 A.6	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423  Appendix C Computing instructions 426
A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400 Correlation 402 Least squares estimation 406	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423
A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400 Correlation 402 Least squares estimation 406 A note on variances and covariances 408  endix B Glossary and technical summary 410  Notation 410	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423  Appendix C Computing instructions 426 C.1 General 426 C.2 The Canabag program 427 C.3 The SCALC program 428  Appendix D Solutions to selected
A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8 Appe	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400 Correlation 402 Least squares estimation 406 A note on variances and covariances 408 endix B Glossary and technical summary 410  Notation 410 Simple random sampling (chapter 3) 413 Stratified random sampling (chapter 4) 414	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423  Appendix C Computing instructions 426 C.1 General 426 C.2 The Canabag program 427 C.3 The SCALC program 428  Appendix D Solutions to selected problems 432
A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8 Appe	The summation notation 391 Distributions 393 Measures of central tendency and dispersion 395 Probability 399 Probability distributions 400 Correlation 402 Least squares estimation 406 A note on variances and covariances 408 endix B Glossary and technical summary 410 Notation 410 Simple random sampling (chapter 3) 413	B.6 Some special topics (chapter 7) 419 B.7 Sampling without replacement and unequal probabilities (chapter 8) 420 B.8 Sampling with replacement and unequal probabilities (chapter 8) 422 B.9 Sampling from a process (chapter 9) 423  Appendix C Computing instructions 426 C.1 General 426 C.2 The Canabag program 427 C.3 The SCALC program 428  Appendix D Solutions to selected

PΑ	RT	O	ΝE
T	E	X	T





### CHAPTER 1

### Introduction

### 1.1 SAMPLING: THE ISSUES

A sample is a part drawn from a larger whole. Rarely is there interest in the sample per se. Almost always a sample is taken in order to learn something about the aggregate (the *population*) from which it is drawn. For instance:

- In an opinion poll, a relatively small number of persons are interviewed, and their opinions on current issues are solicited in order to discover the attitude of the community as a whole.
- The viewing and listening habits of a relatively small number of persons are regularly monitored by ratings services, and, from these observations, projections are made about the preferences of the entire population for available television and radio programs.
- Large lots of manufactured products are accepted or rejected by purchasing departments in business or government following inspection of a relatively small number of items drawn from these lots.
- At border stations, customs officers enforce the laws by checking the effects of only a small number of travelers crossing the border.

- Auditors often judge the extent to which the proper accounting procedures have been followed by examining a small number of transactions, selected from a larger number taking place within a period of time.
- Countless sample surveys are carried out, regularly or occasionally, by marketing and advertising agencies to determine consumers' expectations, buying intentions, or shopping patterns.
- Some of the best known measurements of the economy rely on samples, not on complete enumerations. The weights used in consumer price indexes, for example, are based on the purchases of a sample of urban families; the prices of the individual items are averages established through national samples of retail outlets. Unemployment statistics are based on monthly national samples of households. Similar samples regularly survey retail trade, personal incomes, inventories, shipments and outstanding orders of firms, exports, and imports.

In every case, a sample is selected because it is impossible, inconvenient, slow, or uneconomical to monitor the entire population.

The principal questions in any study involving sampling are: How should the sample be selected? How large should it be? How should the population characteristics be estimated? How reliable are these estimates?

These questions are related to one another. Their examination is the objective of this book. Before we begin, however, it will be useful to consider some specific cases involving sampling, and to reflect on the preceding questions in each case. We shall return to some of these cases frequently throughout this book.

### 1.2 CASE: PRINT MEDIA RESEARCH

Figure 1.1 shows a portion of a questionnaire used in an annual survey<sup>1</sup> conducted by the Institute for Print Media Research. This organization sells information to advertisers, advertising agencies, and publishers concerning the number of readers of each newspaper or magazine and the characteristics of those readers.

The questionnaire consists of about 300 questions. Those shown here deal with the purchase and use of food and household products. Others inquire about other products and services, the newspapers and magazines purchased and/or read by the individual, and the individual's education, income, occupation, and assets.

About 10,000 individuals of age 12 or greater are interviewed for each annual survey from a population of millions of men and women 12 years old or older.

From the responses to the questions shown, Print Media Research could estimate, for example, the proportion of individuals who read *Newsweek* and who buy, use,

<sup>&</sup>lt;sup>1</sup>The term *survey* generally refers to the collection of data by means of interviews, questionnaires, or direct observation. The entities surveyed could form a whole (population survey) or a part (sample survey).

#### FOOD PRODUCTS

Do you buy, use or serve the following products? If you do, circle the number 8 to indicate YES, and then circle the appropriate number to indicate how frequently the product is used in your household.

	DO YOU USE, OR		More than once	Once a	4 to 6 times a	2 or 3 times a	Once a	2 or 3 times a	Once a MONTH or
	Yes	No	a DAY	DAY	WEEK	WEEK	WEEK	MONTH	less
331) Cereals, cold, unsweetened	8	9	1	2	3	4	5	6	7
333) Cereals, cold, pre-sweetened	8	9	1	2	3	4	5	6	7
335) Cereals, hot	8	9	1	2	3	4	5	6	7
337) Beef, fresh	8	9	1	2	3	4	5	6	7
339) Lamb, fresh	8	9	1	2	3	4	5	6	7
341) Pork, fresh	. 8	9	1	2	3	4	5	6	7
343) Veal, fresh	8	9	1	2	3	4	5	- 6	7
345) Ham	8	9	1	2	3	4	5	6	7
347) Bacon	8	9	1	2	3	4	5	6	7
349) Frozen meat	8	9	1	2	3	4	5	6	7

#### HOUSEHOLD PRODUCTS

In about how many washloads each WEEK, DO YOU BUY on average, do you use them? OR USE? Six or One or Yes more Four Five Three Two less 351) Soaps and detergents for regular laundry 353) Soaps and detergents for fine fabrics 355) Laundry pre-soaks and pre-cleaners 357) Bleach 359) Fabric softeners – regular 361) Fabric softeners - aerosol 363) Other laundry additives 365) Automatic dishwashing detergent 367) Rinse agents for automatic dishwashers 369) Dishwashing liquid 

Figure 1.1 Extract from survey questionnaire, Print Media Research

or serve ham. Or, the number of *Vogue* readers who use dishwashing liquid once a week.

The information compiled by this survey is useful, for instance, to a manufacturer of cosmetics for young people and the advertising agency handling the account, who wish to determine the magazines in which an advertisement will have the greatest effect. Or, the publisher of a newsmagazine may wish to know the kind

of reader the magazine attracts; it is possible, for instance, to obtain the profile of a reader of a given periodical, to examine whether there is any relationship between income and expenditure on a given product or service, or to determine whether there are substantial differences in the characteristics of the readers of two publications.

How should the sample of 10,000 be selected? How should the characteristics of the population be estimated? How reliable are they?

### 1.3 CASE: TELEMEDIA INC.

Whether or not a television program or show will continue to be broadcast, and the price it may charge for advertising time, depends on the rating of the program. This rating is simply an estimate of the percentage of individuals watching the program at a given time. Table 1.1 is an excerpt from a ratings book available to subscribers only from Telemedia, a company specializing in providing television and radio ratings. The table shows the ratings of programs broadcast in a northeastern city between 7:30 and 8:00 P.M. on a Thursday in September.

We read, for example, that 36% of all persons two years old or older were watching television during this period. 7% of all persons were watching *Jeopardy*, and

TABLE 1.1	Telemedia Ratings, Excerpt								
		A	11	Ratings					
	2+		Adults	Women	Men	Teens	Children		
Station	Program	Rating	Share	18+	18+	18+	12–17	2–11	
All	stations	36	100	37	39	35	32	24	
<b>CBLF</b>	(Movie)	-	-	-	-1	_	-	1	
CBLT	Cheers	1	3	1	1	1	-	-	
<b>CFMT</b>	Jeopardy	7	19	7	8	6	7	4	
CFTO	Campbells	2	6	3	3	2	1	0	
CHCH	Mannix	1	3	1	1	1	1	1	
CICO	Nature	1	3	1	1	1	1	1	
CITY	M.A.S.H.	6	17	6	6	7	5	1	
WGRZ	Dating Game	3	8	3	3	4	5	3	
WIVB	Pyramid	4	11	4	4	3	3	3	
WKBW	Simon	1	3	1	1	1	1	_	
WUTV	Soap	1	3	1	1	1	3	1	
WBEN	Jackpot	4	11	4	4	3	1	4	
	Others	5	14	5	6	5	4	5	

4% were watching Jackpot. The last two figures represent 19% and 11% respectively of people aged two years old or older who were watching television during this period (the shares of the programs). Ratings are also presented separately for each age/sex class.

These ratings are based on information provided by a sample of about 1,800, selected from a population of about 2.7 million individuals two years old or older in the city, as shown in Table 1.2.

TABLE 1.2 Population and Sample						
Age/Sex Class	Sample	Population				
All persons 2+	1,778	2,690,830				
Adults 18+	1,250	1,910,880				
Women 18+	650	981,800				
Men 18+	600	929,080				
Teenagers 12-17	215	286,250				
Children 2-11	313	493,700				

Obviously, it would be enormously expensive to monitor regularly the viewing habits of the entire population. This fact is never disputed, but Telemedia finds itself frequently under attack by subscribers and others. Whether motivated by a noble cause ("Why is there so little opera on television?"), or prompted by disagreeable results ("Why is it that the ratings of my station fluctuate so much?"), critics often question the method of sampling used (which consists of selecting approximately 0.07% from each age/sex class), the sample size ("You mean your sample is just seven-hundredths of one percent of the population?"), and the lack of any data by which to judge the accuracy of the ratings.

How should Telemedia's sample be selected? How should the characteristics of the population be estimated? How reliable are these estimates?

### 1.4 CASE: NORPOWER INC.

NorPower is the exclusive supplier of electricity to 2,819,514 residences (houses and apartments). Over 80% of these residences have individual meters; the remaining 20% are in apartment buildings and share common meters. A list of all individually- and bulk-metered residences is maintained by the company. Each year, the utility's market research department collects information on ownership and use of appliances and of heating and other equipment by means of the questionnaire shown in Figure 1.2.