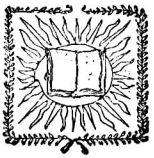
The Century Education Series

EDUCATIONAL STATISTICS

BY

C. W. ODELL, PH.D.

ASSOCIATE, BUREAU OF EDUCATIONAL RESEARCH
UNIVERSITY OF ILLINOIS



NEW YORK
THE CENTURY CO.
1925

EDITOR'S INTRODUCTION

In recent years much has been added to the necessary and still more to the desirable equipment of the educational administrator. Appreciation of the fact that he should be professionally trained, and that training for administration includes much that may not be insisted upon or even recommended to the class-room instructor is becoming general. There is slowly developing the idea that even the academic preparation for the future administrator should recognize the fact that his successful grasp upon the wider aspects of his problem may be strengthened more through work in some branches of knowledge than in others.

One of the most neglected phases of professional training is that which has to do with educational statistics. This has become more obvious through the increasing emphasis which is laid upon certain types of educational measurement. The use of the graph, the median, the mode, quartiles, standard deviation, coefficient of correlation, and normal frequency curve in superintendents' reports illustrates the need for intelligent preparation in statistical procedure. As the author of this volume has pointed out in his preface, there is a real place in educational literature for a text-

book aiming to present the elementary facts upon which statistical procedure is developed. A relatively small percentage of school principals or superintendents have secured any special higher mathematical training, yet they are continually confronted with the necessity of trying to use statistical formulæ or at least to read educational discussions based upon educational statistics. There are few books available to which they may turn for enlightenment, and not infrequently even they are couched in language too technical for the lay reader.

While it is impossible to discuss educational statistics without some use of mathematical formulæ, the author has attempted to keep in mind at all times the fact that many who may use the book have had limited training in advanced mathematics. It is believed that any intelligent layman can understand the language used and gain a reasonably useful comprehension of the topics discussed.

This volume is designed therefore to meet two demands; first, the need for an elementary text in educational statistics for colleges of education and teachers' colleges; second, the need for a discussion of statistical procedure for the use of superintendents of schools and others whose professional training has not included work of this nature.

PREFACE

In preparing this volume the writer has made an attempt to meet the needs of teachers, supervisors, administrators, and others interested in the field of education for an elementary knowledge of statistical methods and their application to educational problems. In the course of several years' experience in teaching the subject of educational statistics he became convinced of the fact that no satisfactory text-book for such a course had appeared. The majority of the persons who make use of statistical methods in educational work and who therefore need such a course have not had more than a moderate amount of training in mathematics, and in few cases has even this moderate amount been used sufficiently to be at their command. Therefore, it is impracticable to employ any one of several excellent texts which are properly suited for use in a general course in statistics given by a department of mathematics. In the writer's judgment there is only one book on the market which treats educational statistics in any adequate way. There are, however, a number of topics not dealt with at all in it and also others which receive incomplete treatment. In the present volume several of the first type are included

and also the treatment of a few of the second type is amplified. It is true that a number of books dealing with educational measurements have appeared which contain one or a few chapters dealing with educational statistics. Although some of these chapters are very good, the amount of space devoted to the subject is too small to do more than present the bare minimum essentials.

The writer has attempted, therefore, to prepare a volume dealing with the subject of educational statistics which will not require for its understanding more mathematical knowledge than is possessed by the ordinary individual who has completed a year of high school algebra. There are, of course, many formulæ used whose derivations cannot be understood unless one has studied a considerable amount of higher mathematics. It is not, however, necessary to understand their derivations to employ them. Although no attempt has been made to give a complete and exhaustive bibliography, enough references have been included to enable any one who wishes to pursue the subject further to do so with little or no difficulty.

The writer wishes to acknowledge his indebtedness to many persons. There is very little in this volume which is original with him. Although he feels some degree of indebtedness to all the authors mentioned in his list of references, yet the major part of his indebtedness is to

此为试读，需要完整PDF请访问：www.ertongbook.com

a comparatively few of them. Chief among these are Rugg, Monroe, Kelley, Yule, and several of the co-authors of "The Handbook of Mathematical Statistics" by Rietz and others. He is also greatly indebted to Mr. G. W. Reagan of the College of Education of the University of Illinois for his careful and painstaking reading and criticism of the entire manuscript.

C. W. ODELL.

EDUCATIONAL STATISTICS

CONTENTS

CHAPTER		PAGE
I	TABULATION AND CLASSIFICATION	1
	1 General Suggestions as to Statistical Procedure	1
	2 Attributes and Variables	12
	3 Grouping Scores or Measures into Classes	16
II	GRAPHS	36
	1 The Graphical Representation of Frequency Distributions	36
	2 The Normal Frequency Curve	52
	3 Other Types of Frequency Curves	59
III	AVERAGES	64
	1 General Discussion of Averages	64
	2 The Arithmetic Mean	66
	3 The Median	75
	4 The Mid-Score	87
	5 The Mode	89
	6 The Geometric Mean	94
	7 The Harmonic Mean	97
	8 The Use of Averages	104
	9 Quartiles, Deciles, Percentiles, and Other Similar Points	111
IV	MEASURES OF VARIABILITY	117
	1 A General Discussion of Variability	117
	2 The Range	119
	3 The Quartile Deviation	120
	4 The 10–90 Percentile Range	122
	5 The Mean Deviation	123
	6 The Standard Deviation	128
	7 The Median Deviation	138
	8 The Use of Measures of Variability	140
	9 The Coefficient of Variability	143

CHAPTER PAGE

V THE RELATIONSHIP BETWEEN TWO VARIABLES 147

1 A General Discussion of Correlation . . . 147
2 The Coefficient of Correlation 150
3 The Correction of Attenuation 181
4 The Coefficient of Reliability 185
5 The Index of Reliability 188
6 Regression 189
7 Galton's Method of Correlation 200
8 Rank Correlation 201
9 The Ratio of Correlation 207

VI UNRELIABILITY 221

1 The Standard and Probable Errors 221
2 Errors of Estimate and of Measurement . . 230

VII THE CORRELATION AND REGRESSION OF THREE OR MORE VARIABLES 245

1 Partial Correlation 245
2 Multiple Correlation 252
3 Partial and Multiple Regression 256

VIII THE CORRELATION OF ATTRIBUTES 266

1 General Discussion 266
2 Tetrachoric Correlation 269
3 Yule's Coefficients 270
4 Pearson's Cosine π Method 271
5 Sheppard's Method of Unlike Signs 272
6 The Coefficient of Contingency 273
7 The Bi-Serial Coefficient and Ratio of Correlation 278

IX MISCELLANEOUS 281

1 Skewness 281
2 Overlapping 286
3 Comparison and Measurement of Change . . 288
4 Equally Noticed Differences 291
5 Equivalent or Comparable Measures . . . 295
6 Coefficient of Correspondence 299
7 Educational Uses of the Normal Frequency Curve 305

PAGE

APPENDIX A BIBLIOGRAPHY 319

APPENDIX B TABLES OF THE HEIGHT AND AREA OF THE NORMAL PROBABILITY CURVE 322

APPENDIX C TABLE FOR THE CONVERSION OF RANK COEFFICIENTS INTO PRODUCT-MOMENT COEFFICIENTS OF CORRELATION 325

APPENDIX D TABLE FOR FINDING THE VALUE OF THE COEFFICIENT OF CORRELATION BY PEARSON'S COSINE π METHOD OR BY SHEPPARD'S METHOD OF UNLIKE SIGNS 327

LIST OF FIGURES

FIGURE		PAGE
1	Frequency Polygon	40
2	Histogram or Column Diagram	42
3	Histogram Used to Show Individual Cases	43
4	Frequency Curve	44
5	Original and Smoothed Frequency Polygons	46
6	Ogive or Cumulative Frequency Curve	50
7	Ogive or Cumulative Frequency Curve	51
8	Normal Frequency Curves	53
9	Normal Frequency Curve Constructed by Expanding the Binomial	56
10	Normal Frequency Curve Constructed by the Use of Tables	57
11	Skew Curves	60
12	U and J Curves	61
13	Frequency Surfaces Showing Their Means, Medians, and Modes	92
14	Graphical Representation of Median, Quartiles, Deciles, and Percentiles	114
15	Normal Curve Showing Measures of Variability	129
16	Skew Curve Showing Measures of Variability	130
17	Correlation Graph	158
18	Correlation Graph Showing Lines of Regression	190
19	A Positively Skewed Curve	282
20	A Negatively Skewed Curve	283
21	Three Normal Frequency Curves Representing the Distribution of Judges' Ratings of Three Specimens of Work	293
22	Comparison of Actual and Best-Fitting Normal Curves for the Same Data	306
23	Use of the Normal Curve to Determine What Proportion of Students Should Receive Each of Five Marks	311
24	Normal Frequency Curve Used to Determine the Difficulty of Exercises Answered by Known Percentages of Those Attempting to Do So	314

此为试读，需要完整PDF请访问：www.ertongbook.com

LIST OF TABLES

		PAGE
I	Computation of the Harmonic Mean	98
II	Computation and Checking of the Standard Deviation	135
III	Computation of the Coefficient of Correlation of Simple Series of Measures by Using the Exact Means	153
IV	Computation of the Coefficient of Correlation of Simple Series of Measures by Using Assumed Means	155
V	Initial Form of Correlation Table	159
VI	Final Form of Correlation Table	160
VII	Computation of the Coefficient of Correlation of Measures Grouped in a Correlation Table	162
VIII	Ayres's Method of Computing the Coefficient of Correlation of Measures in Simple Series	167
IX	Ayres's Method of Computing the Coefficient of Correlation of Measures Grouped in a Frequency Table	168
X	Ayres's Method of Using Reduced Measures in the Computation of the Coefficient of Correlation of Measures in Simple Series	169
XI	Computation of the Coefficient of Correlation by Rank Methods	204
XII	Computation of the Coefficient and Ratios of Correlation	210
XIII	Standard and Probable Errors of Measures Most Often Used	223
XIV	Computation of the Standard Errors of Estimate and of Measurement	236
XV	Computation of the Coefficient of Correspondence by Several Common Methods	302
XVI	Comparison of the Differences Between an Actual Distribution and the Best-Fitting Normal One With the Standard Errors of Sampling	308

EDUCATIONAL STATISTICS

EDUCATIONAL STATISTICS

CHAPTER I

TABULATION AND CLASSIFICATION

1. General suggestions as to statistical procedure. Educational data usually are collected by single cases. They consist of scores, measurements, or other descriptions of numbers of individuals. It is therefore necessary that they be combined or tabulated in some way whenever very many individuals are concerned. It is furthermore practically necessary that, if more than a few cases are being dealt with, some method of checking or numbering be adopted in place of writing out a description of each case in full. That is to say, a table is drawn up containing headings which describe or limit a number of groups or classes into which all the cases are to be placed.

In its most elementary form such a table resembles the accompanying simple tabulation. This shows the

Passed	𝍸 𝍸 𝍸 𝍸 ||	22
Failed	||||	4

tabulation of the members of a class according to whether they passed or failed. In making it, one mark

for each member of the class is placed after the proper heading. In this instance the marks so made indicate that there were 22 members of the class who passed and 4 who failed. It is customary and convenient to do as has been done in the example, and for each fifth case in the same row to make a cross or diagonal mark through the last four. This sets off the marks into groups of five and thus renders counting easier. It is also common to write the number of cases in each row at the right, as shown.

In simple tabulation of the sort just illustrated there are frequently more than two classes or divisions. For example, if conditions were in use the data given above would be tabulated in three instead of two classes, the third being labeled "Conditioned." In case the pupils of a whole elementary school were being tabulated by grades, there would ordinarily be eight headings, one for each grade. In addition to the classes themselves it is usually well to have an additional row, placed at the bottom, labeled "Total," which contains no marks but merely a number showing the total number of cases in all classes. Sometimes a more detailed system of classification is used involving classes numerically defined. The discussion of such a system will be taken up in some detail in the section that deals with the grouping of data into classes.

For many purposes it is necessary to use a double system of classification having both rows and columns, with appropriate headings at the left of the former and above the latter. The accompanying table illustrates such a double classification. In it pupils are

	Boys	Girls	Both
Passed	10	12	22
Conditioned	2	3	5
Failed	3	1	4
Total	15	16	31

divided according to their sex as well as their scholastic record. As is the usual practice in such tables there is a totals row at the bottom and a column for the same purpose at the right. Therefore the entry in the compartment at the lower right-hand corner of the table shows the whole number of pupils included.

The process of the initial tabulation of frequency, which has been illustrated by the examples given, is often called "primary" tabulation, and the later statistical treatment, for which this primary tabulation serves as the basis, "secondary" tabulation. Since the process of primary tabulation is relatively simple, statistics devote most attention to the later or secondary work.

此为试读，需要完整PDF请访问：www.ertongbook.com