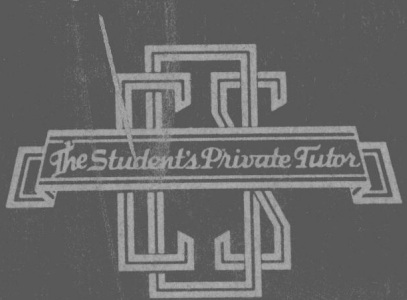
COLLEGE OUTLINE SERIES

25

AN OUTLINE OF STATISTICAL METHODS

FOURTH EDITION

ARKIN and COLTON



BARNES & NOBLE · INC · NEW YORK

COLLEGE OUTLINE SERIES

AN OUTLINE OF STATISTICAL METHODS

AS APPLIED TO
ECONOMICS, BUSINESS, EDUCATION,
SOCIAL AND PHYSICAL SCIENCES, ETC.

By

HERBERT ARKIN
The College of the City of New York
and
RAYMOND R. COLTON
The College of the City of New York

WITH A PREFACE BY
JUSTIN H. MOORE, Professor of Law
School of Business and Civic Administration
The College of the City of New York

Fourth Edition

New York

BARNES & NOBLE, Inc.

FIRST EDITION — 1934
SECOND EDITION — 1935
THIRD EDITION — 1938
FOURTH EDITION — 1939
REPRINTED — 1942
REPRINTED, NOVEMBER, 1943
REPRINTED, MAY, 1946
REPRINTED, MARCH, 1947

Printed in the United States of America

COLLEGE OUTLINE SERIES

A. W. LITTLEFIELD, *General Editor*

ACCOUNTING, Elementary . . . . . BAUER-DARBY
ALGEBRA, College . . . . . MOORE
ANCIENT HISTORY . . . . . HYMA
ANCIENT, Medieval, Modern History . . . RICKARD-HYMA

BACTERIOLOGY, Principles of . . . . . BRYAN
BIOLOGY, General . . . . . ALEXANDER
BOTANY, General . . . . . FULLER
BUSINESS LAW . . . . . BABB-MARTIN

CALCULUS, The . . . . . OAKLEY
CHEMISTRY, First Year College . . . . . LEWIS
CHEMISTRY, Mathematics for General . . . . FREY
CHEMISTRY, Organic . . . . . DEGERING, et al
CORPORATION FINANCE . . . . . HAROLD

DOCUMENTED PAPERS, Writing . . . . . HUBBELL

ECONOMICS, Principles of . . . . . JAMES
EDUCATION, History of . . . . . THOMPSON
ENGINEERING DRAWING . . . . LOMBARDO-JOHNSON
ENGLAND, History of . . . . . RICKARD
EUROPE, 1500-1848, History of . . . . . LITTLEFIELD
EUROPE, 1815-1946, History of . . . . . LITTLEFIELD
EXAMINATIONS, How to Write Better . . . . HOOK

FRENCH GRAMMAR . . . . . DU MONT

GEOLOGY, Principles of . . . . . FIELD
GEOMETRY, Plane, Problems in . . . HORBLIT-NIELSEN
GERMAN GRAMMAR . . . . . GREENFIELD
GOVERNMENT, American . . . . . SAYRE
GRAMMAR, English, Principles of . . . . . CURME

HYDRAULICS for Firemen . . . . . THEOBALD

JOURNALISM, Survey of . . . . . MOTT, et al

[*Continued on next page*]

[*Continued from preceding page*]

LATIN AMERICA, History of . . . . . WILGUS-D'EÇA
LATIN AMERICA in Maps . . . . . WILGUS
LATIN AMERICAN Civilization, Readings in . WILGUS
LATIN AMERICAN Economic Development . . WYTHE
LITERATURE, American . . . . . CRAWFORD, et al
LITERATURE, English, Dictionary of . . . . . WATT
LITERATURE, English, To Dryden . . OTIS-NEEDLEMAN
LITERATURE, English, Since Milton . . NEEDLEMAN-OTIS
LOGARITHMIC and Trigonometric Tables . . . NIELSEN

MIDDLE AGES, 300-1500, History of . . . . MOTT-DEE
MUSIC, History of . . . . . MILLER

PHILOSOPHY: An Introduction . . . RANDALL-BUCHLER
PHILOSOPHY, Readings in . . . . . RANDALL, et al
PHYSICS, First Year College . . . . . BENNETT
POLITICAL SCIENCE . . . . . JACOBSEN-LIPMAN
POLITICS, Dictionary of American . . SMITH-ZURCHER
PORTUGUESE GRAMMAR . . . . D'EÇA-GREENFIELD
PSYCHOLOGY, Educational . . . . . PINTNER, et al
PSYCHOLOGY, General . . . . . FRYER-HENRY

SHAKESPEAREAN NAMES, Dictionary of . . IRVINE
SHAKESPEARE'S PLAYS, Outlines of . . WATT, et al
SLIDE RULE, Practical Use of . . . . . BISHOP
SOCIOLOGY, Principles of . . . . . LEE, et al
SPANISH GRAMMAR . . . . . GREENFIELD
STATISTICAL METHODS . . . . . ARKIN-COLTON
STUDY, Best Methods of . . . . . SMITH-LITTLEFIELD

TRIGONOMETRY . . . . . NIELSEN-VANLONKHUYZEN
TUDOR and Stuart Plays, Outlines of . . . HOLZKNECHT

UNITED STATES, To 1865, History of . . . . KROUT
UNITED STATES, Since 1865, History of . . . KROUT
UNITED STATES in Second World War . . . . HARRIS

WORLD Since 1914, History of . . . . . LANDMAN

ZOOLOGY, General . . . . . ALEXANDER

Preface

In Europe it has long been the practise to publish small books known as *manuals,* giving a condensed and succinct treatment of the subject as an aid to the student in summarizing the more elaborate and detailed contents of a large textbook. Halfway between an article in an encyclopedia and the often discursive texts studied in the classroom, these manuals fulfil a real need; they winnow fundamental principles from a mass of material, give with incisive clarity the contours of the ground already covered and leave with the reader a definite framework which later he may fill in by further practical contact with the subject in real life. In America there has been a noticeable gap in educational literature which now happily will be filled in part by the excellent series of small books published by Barnes & Noble, Inc.

The present volume on statistics does not aim to be a comprehensive treatise on the subject. On the contrary it gives the distilled essence of material which might well require one or more large volumes for a full discussion. For that very reason it ought to be a most useful tool in the hands alike of students and people actually engaged in statistical work, wherever the particular field of activity may happen to lie. The formulas and examples given in it will be ample for the needs of most workers, whether they be concerned with financial, industrial, commercial, social, or educational statistics. Of necessity there is a minimum of verbiage; no formula is included that has not practical applications; the mathematical aspects are not stressed; the philosophy of the subject and many recondite byways are not explored. Thus the reader is spared the need of hunting back and forth to find the special help which he needs on his concrete problems in daily life. To all statistical workers this little volume will be as indispensable as an adding machine.

JUSTIN H. MOORE, Professor of Law
School of Business and Civic Administration
The College of the City of New York

PUBLISHER'S NOTE

In this edition we have included eighteen standard textbooks in the tabulated bibliography table. The following list gives the author, title, and publisher of all the books referred to in the tables on the next two pages.

Camp, B. H., *Mathematical Part of Elementary Statistics,* Heath, 1931.

Chaddock, R. E., *Principles and Methods of Statistics,* Houghton Mifflin, 1925.

Croxton, F. E. and Cowden, D. J., *Practical Business Statistics,* Prentice-Hall. 1934.

Crum, W. L., Patton, A. C. and Tebbut, A. R., *An Introduction to Economic Statistics,* McGraw-Hill, Rev. 1938.

Davies, G. R. and Crowder, W. F., *Methods of Statistical Analysis in Social Sciences,* Wiley, 1933.

Day, E. E., *Statistical Analysis,* Macmillan, 1930.

Garrett, H. E., *Statistics in Psychology and Education,* Longmans, Rev. 1937.

Harper, F., *Elements of Practical Statistics,* Macmillan, 1930.

Holzinger, K. J., *Statistical Methods for Students in Education,* Ginn, 1928.

Jerome, H., *Statistical Method,* Harper, 1924.

Kelley, T. L., *Statistical Method,* Macmillan, 1923.

Mills, F. C., *Statistical Methods,* Holt, 2nd ed. 1938.

Odell, C. W., *Statistical Method in Education,* pp. 14-27, Appleton-Century, 1935.

Richardson, C. H., *An Introduction to Statistical Analysis,* Harcourt, 1934.

Riggleman, J. R. and Frisbee, I. N., *Business Statistics,* McGraw-Hill, Rev. 1938.

Rugg, H. O., *Statistical Methods Applied to Education,* Houghton Mifflin, 1917.

Thurstone, L. L., *Fundamentals of Statistics,* Macmillan, 1925.

Yule, G. U., *An Introduction to the Theory of Statistics,* Griffen (London) Rev. 1937.

Quick Reference Table

All figures refer

CHAPTER IN THIS OUTLINE	TOPIC	CAMP	CHAD-DOCK	CROX-TON and COWDEN	CRUM et al	DAVIES	DAY	GARRETT
I	Statistical Series		43-80	151-153	3-11	4-33	36-47 118-133	4-9
II	Frequency Distribution—Central Tendency—Arithmetic Mean		81-105	153-172	76-181	33-38	134-139	17-19 26-29
III	Frequency Distribution—Central Tendency (continued)	36-42	107-148	176-202	182-192	38-57	140-162	20-26 29
IV	Frequency Distribution—Dispersion and Skewness	44-46	150-173	204-219	193-207	65-85	163-179	33-61
V	Time Series Analysis—Trend	102-103	306-320	262-282	298-300	138-139 153-156	231-257	
VI	Time Series Analysis—The Least Squares Method—Linear	104-111	320-334	313-325	300-316	133-137	258-263	
VII	Time Series Analysis—Non-Linear	112-128	335-339	326-338	316-325	139-153 156-178	263-280	
VIII	Time Series Analysis—Seasonal and Cyclical		339-366	286-310	326-362	189-219	281-312	
IX	Correlation—Linear	129-172 286-290	248-290	405-425	230-252 363-370	226-250	180-210 313-327	251-279 289-310
X	Correlation—Non-Linear, Multiple, Partial	290-310 315-343	290-304	431-435	253-262	250-280	201-206	393-464
XI	Correlation of Attributes	302-314						366-392
XII	The Normal Curve	59-72 183-199	207-228	241-257	208-214	292-306		98-177
XIII	Theory of Sampling	240-273	228-246	222-237	215-220	298-300		198-250
XIV	Index Numbers		175-205	362-397	263-297	91-123	328-367	
XV	Further Analysis of the Frequency Distribution	18-30			221-229			
XVI	Collection of Data		371-395	22-37	38-59			
XVII	Statistical Tables		397-417	39-58	60-107			
XVIII	Graphic Presentation		418-445	65-129	91-154		48-113 211-230	
XIX	Special Technique in Education, Psychology, and Biology	78-101		234-239				

See preceding page for

to Standard Textbooks

to pages

HARPER	HOLZ-INGER	JEROME	KELLEY	MILLS	ODELL	RICHARD-SON	RIGGLE-MAN	RUGG	THUR-STONE	YULE
45-64	1-8		1-8	50-85	14-27	15-41	132-163	4-26 74-81		82-109
95-110	78-85	109-117	44-53	86-109	32-62 66-76	43-53	164-170	81-100 114-126	1-17 67-77	112-119
112-121	85-96	117-133	54-68	109-136	77-114	53-71	170-181	100-114 126-147	78-85	120-133
122-152	101-139	146-164	70-82	137-160	115-141	77-95	215-228	149-178	86-123	134-153
162-168	317-325	224-228		225-246			270-297	178-179	18-28	
168-178	154-161 321-322	228-233		246-253		113-133	297-300	248-249	51-64	
178-187	324-337			253-283		169-200	300-310		29-37	
		235-255		284-304			313-359			
189-220 248-256	141-173	263-287	151-185	325-403	143-208 237-249	135-162	244-264	232-270 294-296	187-223	196-226
193-195 221-276	177-186 256-315	286	185-310	404-454 531-597	250-308	201-203	264	276-306	224-228	241-252 261-287
	256-278				309-324			292-309		
154-156	190-229	165-181	94-106 109-150	425-435	53-64 378-396	207-249		191-227	126-148	169-187
1-13 156-161	231-245 245-254	171-177	82-92	452-489 598-659	326-376	251-271		227-231 270-274	161-186	332-394
278-283		183-223	331-347	161-224 305-324			184-214			
	338-344			435-444 448-450		97-111				154-168
	9-28	290-326			14-31		14-47	28-73 310-360		
31-35	31-34 36-45	28-49		62-63			48-75	87-94 310-360		
66-93		50-108	9-43	8-49	416-434		76-131	310-360	47-49	
	Entire Volume				Entire Volume		360-634	Entire Volume		

complete list of titles

THE GREEK ALPHABET

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	xi
Γ	γ	gamma	O	o	omicron
Δ	δ	delta	Π	π	pi
E	ϵ	epsilon	P	ρ	rho
Z	ζ	zeta	Σ	σ	sigma
H	η	eta	T	τ	tau
Θ	θ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ψ	psi
M	μ	mu	Ω	ω	omega

Table of Contents

CHAPTER I

STATISTICAL SERIES

Definition of Statistical Method

Statistical Method is a technique used to obtain, analyze and present numerical data.

Elements of Statistical Technique

The elements of statistical technique include the:

1. Collection and assembling of data.
2. Classification and condensation of data.
3. Presentation of data in:
    a. textular form.
    b. tabular form.
    c. graphic form.
4. Analysis of data.

Characteristics and Limitations of Statistical Methods

1. Statistical method is the only means for handling large masses of numerical data.

2. Statistical technique applies only to data which are reducible to quantitative form.

3. Statistical technique is *objective.* The results, however, cannot but be affected by the necessarily *subjective* interpretation

4. Statistical technique is the same for the social as for the physical sciences; i.e., economics, education, sociology and psychology as contrasted with biology, chemistry and astronomy. Method and technique apply alike to these divergent fields.

Statistical Series

In order to analyze numerical data, it is first necessary to arrange them systematically. The data may be arranged in a number of different ways. Technically an arrangement is called a **distribution** or **series.** An example of each type of distribution is shown below:

When data are grouped according to:	*The resulting series is called a:*
1. Magnitude	Frequency distribution
2. Time of occurrence	Time series
3. Geographic location	Spatial distribution

In addition there is a number of special types of distributions in which the data may be arranged by *kind* or by *degree*.

The Frequency Distribution

Definition

The frequency distribution is an arrangement of numerical data according to size or magnitude.

Construction

A frequency distribution is constructed in the following manner:

1. Using the **range** of the data (the interval between the highest and the lowest figure) as a guide, the data are divided into a number of convenient sized groups. The groups are called **class intervals** (compare table 1, column 1).[1]

The size of the class interval is dependent upon the number of values to be included in the distribution. The range of the values (difference between the highest and lowest values) is determined and is divided by the number of class intervals desired. The resulting size is rounded off. Few class intervals are used when a limited number of values are included and a large number when the distribution is to be compiled from many values. The most efficient number of class intervals usually lies between ten and twenty groups.

Other requirements for the determination of the class interval are:

a. The class intervals should not overlap; 0 – 4.9, 5 – 9.9 etc., should be used in preference to 0 – 5, 5 – 10, etc.

b. When the values tabulated coincide with the integers or with selected values, these values should generally constitute the midpoints of the groups.

c. When possible the class intervals should be of a uniform size.

2. The groups are then placed in a column with the lowest class interval at the top and the rest of the class intervals following according to size.
3. The data are then scored. Each figure is checked once next to the class interval into which it falls (see tally, table 1).[2]

Graphic Presentation of Frequency Distribution

If two lines are drawn perpendicular to one another and are divided according to a scale of values, given data may be represented by reference to the scale. The horizontal line is known as the X axis and the vertical line as the Y axis. If the values for

[1] As a preliminary step the raw data may be arranged according to size. The series is then called an array.

[2] In scoring data an efficient procedure is to connect the first and fourth score by the fifth. Using this procedure totals are obtained merely by adding the resulting units multiplying by five and adding the odd scores (see Table 1). A tally of this type saves time in counting frequencies and also eliminates possible inaccuracies.

Table 1 — Score Sheet

City Tax Rate of "True" Valuation in 261 Cities in the United States, 1927

Rate Per Thousand Dollars (Class Intervals)	Tally	Total Number of Cities (frequency)
4 – 7.99	𝍸	5
8 – 11.99	𝍸 𝍸 𝍸	15
12 – 15.99	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 I	46
16 – 19.99	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 III	68
20 – 23.99	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 III	58
24 – 27.99	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 II	32
28 – 31.99	𝍸 𝍸 𝍸 𝍸 II	22
32 – 35.99	𝍸 𝍸	10
36 – 39.99	II	2
40 – 43.99	II	2
44 – 47.99		0
48 – 51.99	I	1
		261

Source: United States Department of Commerce, *Financial Statistics of Cities*, 1927, Table 23.

a point are given the point may then be located on the graph. For example, in figure 1 the point $X = 2$, $Y = 3$ may now be located at the point marked *a*.

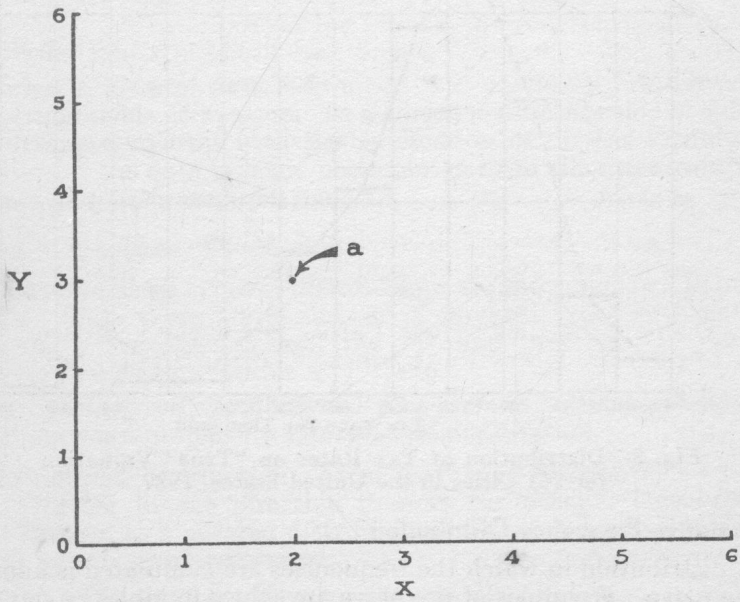


Fig. 1—Location of Plotted point $X = 2$, $Y = 3$.

If the two axes are now marked in the units of the given data the frequency distribution may be presented graphically (pictorially).

1. The class interval grouping which will be termed the **independent variable** is placed on the X (horizontal) axis and the frequency or **dependent variable** is placed on the Y (vertical) **axis.**[1]
2. The number of cases (frequency) is plotted at the midpoint of each respective class interval at the appropriate horizontal level as indicated by the scale on the Y axis.[2]
3. When connected the plotted points form a **frequency polygon.**
4. Rectangles may be constructed by using as the width the size of the class interval and as the height the frequency in each class interval. The rectangles form an **histogram** (also known as **a rectangular frequency polygon**).

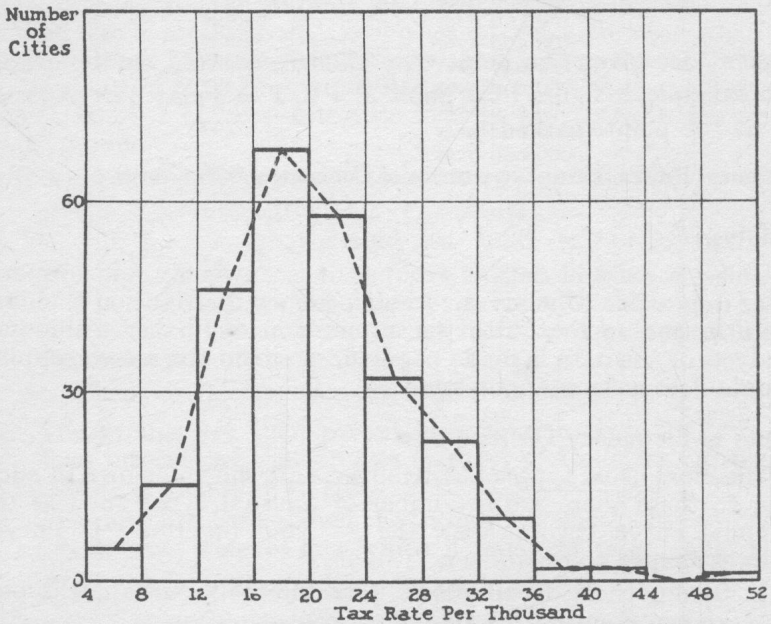


Fig. 2—Distribution of Tax Rates on "True" Valuation for 261 Cities in the United States, 1927.

Cumulative Frequency Distribution

A distribution in which the frequencies are cumulated is known as an ogive. Examples of the ogive are shown in tables 2a and 2b.

[1] The distance along the X axis is called the abscissa while that along the Y axis is called the ordinate.

[2] A cross sectional background or ruling may be used to aid in this work.