statistics

GILBERT

statistics

NORMA GILBERT, Ph.D.
Department of Mathematics
Drew University

1976
W. B. SAUNDERS COMPANY
Philadelphia, London, Toronto

W. B. Saunders Company: West Washington Square
Philadelphia, PA 19105

1 St. Anne's Road
Eastbourne, East Sussex BN21 3UN, England

833 Oxford Street
Toronto, Ontario M8Z 5T9, Canada

Library of Congress Cataloging in Publication Data

Gilbert, Norma.

Statistics.

1. Mathematical statistics. I. Title.

QA276.G45 519.5 75-22732

ISBN 0-7216-4127-X

Front cover is a reproduction of an original lithograph by Steve Holmes, and is published with the permission of the artist.

Statistics ISBN 0-7216-4127-X

Last digit is the print number: 9 8 7 6 5 4 3 2 1

preface

This book is intended for students who need to understand how statistical decisions are made but who have little mathematical background (a year or two of high school algebra is sufficient). So many statistics texts aimed at this group have appeared that some explanation—other than the pleasure I've had in writing it—is needed of why another appears. Here are several reasons:

First, many formulas appear in any statistics book. Some texts emphasize how to use a formula. Learning how to use it is necessary, but I have tried to explain also where a formula comes from, why it should be used, and when it should not be used. Explanations are often intuitive and informal, but a student should get a grasp of why a decision is reached rather than just learn what is reached.

Second, in some texts the notation gets pretty involved. In this one it doesn't. (The exception is Chapter 16; by then a student is well accustomed to statistics. Some instructors will prefer to omit this chapter.)

Third, the book is suited to self-paced ("Keller Plan") study as well as to the usual lecture approach: An introductory section at the beginning of each chapter explains its scope and where emphasis should be put. Over 300 examples are worked out in detail. The exercises at the end of each chapter are preceded by a list of vocabulary and symbols which help a student review the chapter. The answers to odd-numbered problems are given in detail, in the belief that it does no good to know that the probability of an event, for example, is .40 if the student doesn't know how .40 is derived.

Last, some texts avoid probability, but an understanding of statistics without it is impossible. Others dive into probability immediately. Many students have trouble with this topic, however, so I have begun with descriptive statistics and delayed probability until Chapter 5; it is then used in every succeeding chapter.

Chapters 1 through 8, part of 9, and 10 should ordinarily be taken in sequence; material from the remainder of Chapter 9 and Chapters 11 through 17 can be selected according to the interests and speed of the students and the length of the course.

I am especially indebted to Professor David Moore of Purdue University, who read the manuscript at three different stages of its development and who made many splendid suggestions. Special thanks are due to others who read and criticized the manuscript: Stephen A. Book, California State College, Dominguez Hills, California; Albert Liberi, Westchester Community College, Valhalla, New York; Stanley M. Lukawecki, Clemson University, Clemson, South Carolina; Leon Gleser, Purdue University, West Lafayette, Indiana; and Judith Tanur, State University of New York at Stony Brook, Stony Brook, New York.

I am also indebted to Margo Hall (Drew University, Madison, New Jersey, 1974), who helped sort out several different drafts; to John Miano (Drew University, 1975), who suggested interesting exercises; to all the Drew students in Mathematics 3 who have used the original draft of the manuscript; and to John Snyder of W. B. Saunders Company for his cooperation and encouragement.

NORMA M. GILBERT

contents

chapter one

what mathematical background is needed?

Most of this chapter should be familiar stuff. Note that $*$ is used for multiplication, rather than $\times$ or $\cdot$: $2 * 3 = 6$. Summation notation (Sections 1.4 and 1.5, pages 4–7) may be new to you. You will use the remaining parts of this chapter so frequently, however, that you will be foolish indeed if you let anything slip by without understanding it. Even students who are good at mathematics are sometimes confused about whether multiplication or addition is carried out first ($2 * 3 + 4 = 6 + 4$ or $2 * 7$?) and how parentheses are used to change the usual order.

1.1 INTRODUCTION

Many students approach statistics with apprehension, having resolved after junior year in high school never to take another mathematics course. But be of good cheer for two reasons. The background in mathematics which you need is remarkably slight. And in your high school courses you may have learned how to solve quadratic equations, for example, without understanding **why** you were learning to do this. Now those of you who are concerned about using mathematics are studying statistics because a knowledge of this material is necessary background for your study of psychology, economics, zoology, or sociology. You will find an almost immediate application of the mathematics you learn to a subject of interest to you, and this delightful situation—in contrast to your past experience—will give most of you a new and splendid ability to cope with mathematics.

Try always to gain some perception of the general arguments that produce results rather than accept them blindly. An idiot can be trained (or a computer can be programed) to plug numbers into formulas; you must learn when and why a particular formula is used. Occasionally an explanation, in order to be mathematically proper, would require that you know calculus; in these cases the explanation is purely intuitive or is omitted. Fair warning will be given when such occasions arise.

There are very few mathematical tidbits that you need to read this book. One unusual notation needs comment: multiplication of numbers will be indicated by $*$:

$$4 * 2 = 8$$

This is done partly because this notation is often used on computer print-outs and you should become accustomed to it, and partly so that 4.2 in decimal notation will not be confused with the product $4 * 2$. Multiplication of x by y will ordinarily be indicated by juxtaposition: xy, rather than $x * y$; similarly, fd rather than $f * d$. But neither a dot ($f \cdot d$) nor a cross ($f \times d$) will ever be used to indicate f times d.

Note especially that if any number is multiplied by zero, the product is 0:

$$7 * 0 = 0 \qquad 0 * 2/3 = 0$$
$$0 * 7 = 0 \qquad \sqrt{2} * 0 = 0$$
$$-3.8 * 0 = 0 \qquad 0 * x = 0$$

A special notation will be used to warn you of traps that have been set for you: an unusual twist, a problem that cannot be solved, or some special reason to be wary.

1.2 PRE-TEST

If you can solve the following problems correctly, you are very well prepared mathematically for this book. If you have trouble with some of them, study this chapter with care. Problems 13–20 need familiarity with Σ (summation) notation. Do not lose heart if you are not yet familiar with it; read Section 1.4, page 4, with care. The answers are at the bottom of page 3.

1. $4.38 + .12 + .02 =$
2. $48 - 51 =$
3. $6.21 * .01 =$
4. $5 * 0 =$
5. $\dfrac{27 * 75}{25 * 9} =$
6. $3 * 5^2 + 5 =$
7. $3 * (5^2 + 5) =$
8. $4 - 10^2/20 =$
9. Evaluate $x^2y + \dfrac{z}{y}$ if $x = 2$, $y = 3$, and $z = 6$.
10. Find $\overline{X}$ if $X = 10$, $s = 2$, $n = 25$, $z = 15$, and $z = \dfrac{X - \overline{X}}{\dfrac{s}{\sqrt{n}}}$.
11. $\sqrt{493} =$ (2 decimal places; use Table 1, Appendix C.)
12. $\sqrt{.00212} =$ (3 decimal places; use Table 1, Appendix C.)

In Problems 13–17, X values are 3, 5, 8, 2, 0, 3, 0; corresponding Y values are 2, 1, 3, −1, 0, 2, 4.

13. $\Sigma X =$
14. $\Sigma(X - Y) =$
15. $\Sigma X^2 =$
16. $(\Sigma X)^2 =$
17. $\Sigma(X - 2) =$
18. $\Sigma fY =$

Y	f
3	2
4	7
5	2

19. A set of five X values is given. Write a formula which says, "Subtract 3 from each X value, square each new number, add the new set, and then take the square root of the sum."

20. A set of 10 Y values is given. Translate into English the directions given in the formula $\frac{\Sigma(2Y + 5)^2}{10}$.

1.3 THE ORDER OF OPERATIONS

The word "operations" here refers to raising to powers, multiplication or division, and addition or subtraction. The rules agreed on by mathematicians are that these are carried out in the order given below when more than one operation is to be performed—that is,

(a) First raise to powers.
(b) Then multiply or divide.
(c) Then add or subtract.

Example 1

$2 * 3 + 4 = ?$

$2^2 * 3 + 1 = ?$

$3 - 12/2^2 = ?$

$2 * 3 + 4 = 6 + 4 = 10.$

$2^2 * 3 + 1 = 4 * 3 + 1 = 12 + 1 = 13.$

$3 - 12/2^2 = 3 - 12/4 = 3 - 3 = 0.$

But suppose you don't want to do things in that order? What if, for example, 2 and 3 are first to be added and then the result squared? Parentheses ( ), square brackets [ ], or curly brackets { } are used, and a new rule is added:

If there are parentheses or brackets, simplify what is inside them first before continuing.

The notation $3(4 + 2)^2$ is used instead of $3 * (4 + 2)^2$.

Example 2

$(2 + 3)^2 * 4 = ?$

$(2^2 * 3 - 2)^2 * 4 = ?$

$2(3^2 + 4 \div 2) - 3(2^3 - 4)^2 = ?$

$(2 + 3)^2 * 4 = 5^2 * 4 = 25 * 4 = 100.$

Answers:

1. 4.52; 2. −3; 3. .0621; 4. 0; 5. $\frac{3 * 3}{1 * 1} = 9$; 6. 80; 7. 90; 8. −1; 9. 14; 10. 4; 11. 22.20; 12. .046; 13. 21; 14. $\Sigma X - \Sigma Y = 21 - (2 + 1 + 3 - 1 + 0 + 2 + 4) = 21 - 11 = 10$; 15. $9 + 25 + 64 + 4 + 0 + 9 + 0 = 111$; 16. $21^2 = 441$; 17. $\Sigma X - \Sigma 2 = 21 - 7 * 2 = 7$; 18. $6 + 28 + 10 = 44$; 19. $\sqrt{\Sigma(X - 3)^2}$; 20. Multiply each Y value by 2, add 5 to each product, then square each number in the new set; next add the squares, and finally divide the sum by the number of scores.

$(2^2 * 3 - 2)^2 * 4 = (4 * 3 - 2)^2 * 4 = (12 - 2)^2 * 4 = 10^2 * 4 = 100 * 4 = 400.$

$2(3^2 + 4 \div 2) - 3(2^3 - 4)^2 = 2(9 + 4 \div 2) - 3(8 - 4)^2$
$= 2(9 + 2) - 3 * 4^2$
$= 2 * 11 - 3 * 16 = 22 - 48 = -26.$

(You won't have anything even as complicated as this one for a long time.)

1.4 Σ NOTATION

You will soon need to learn seven Greek letters. The first of these is the Greek (capital) S, called **sigma,** and written Σ. We shall use Σ as an abbreviation for the word **sum** or for the phrase "the sum of ______."

If X values are 1, 2, 3, 5, then $\Sigma X = 1 + 2 + 3 + 5 = 11$. Add **all** the X values.

If the weights W of five students are 110, 100, 160, 200, 190, then $\Sigma W = 110 + 100 + 160 + 200 + 190 = 760$.

Remember the rules for order of operations when Σ (standing for addition) is combined with exponents, multiplication, or parentheses.

Example 1

If Y values are 4, 3, 1, 6, 2, find (a) ΣY, (b) ΣY^2, (c) $(\Sigma Y)^2$.

(a) $\Sigma Y = 4 + 3 + 1 + 6 + 2 = 16.$

(b) $\Sigma Y^2 = 4^2 + 3^2 + 1^2 + 6^2 + 2^2 = 16 + 9 + 1 + 36 + 4 = 66.$

(c) $(\Sigma Y)^2 = 16^2 = 256.$

Example 2

Two columns of numbers are given in the box below.

(a) $\Sigma fX = ?$

(b) $\Sigma fX^2 = ?$

(c) $(\Sigma fX)^2 = ?$

(d) $\Sigma (fX)^2 = ?$

f	X	fX	fX^2
2	−1	−2	2
3	0	0	0
1	1	1	1
2	2	4	8
		+3	11

(a) $\Sigma fX = ?$

Multiplication is carried out before addition; the notation ΣfX means "multiply each f value by the corresponding X value, and then add the results." A new column fX is written next to the two we already have:

$$\Sigma fX = 2 * -1 + 3 * 0 + 1 * 1 + 2 * 2$$
$$= -2 + 0 + 1 + 4 = 3$$

(b) $\Sigma fX^2 = ?$ $fX^2 = (fX)X = X(fX)$, so another column is added and its entries are determined by multiplying corresponding entries in the X and fX columns.

$$\Sigma fX^2 = \Sigma X(fX) = -1 * -2 + 0 * 0 + 1 * 1 + 2 * 4 = 2 + 0 + 1 + 8 = 11$$

Yes, it would be possible to square X and then multiply by f, resulting in columns headed f, X, X^2, fX^2. Two things are wrong with this—one invisible and one visible. The invisible one is that in statistics when we need ΣfX^2 we shall usually want ΣfX as well; the visible reason is that the f, X, fX, fX^2 route to finding ΣfX^2 involves multiplying adjacent columns, while the f, X, X^2, fX^2 route to finding ΣfX^2 means multiplying numbers in the first and third columns to determine the fourth. The human eye is not tremendously successful at playing leapfrog with columns.

(c) $(\Sigma fX)^2 = ?$ What is inside the parentheses is evaluated first.
$(\Sigma fX)^2 = (-2+0+1+4)^2 = 3^2 = 9.$

(d) $\Sigma(fX)^2 = ?$
$\Sigma(fX)^2 = (-2)^2 + 0^2 + 1^2 + 4^2 = 4+0+1+16 = 21.$

There are three rules for working with summations:

Rule 1. $\Sigma(X + Y) = \Sigma X + \Sigma Y$. (This rule only makes sense if the number of X values is the same as the number of Y values.)

Rule 2. If k is a constant, $\Sigma kX = k\Sigma X$.

Rule 3. If A is a constant, $\Sigma A = nA$, where n is the number of values that are being added.

Example 3

X values: 4, 1, 6, 3; corresponding Y values: 2, 3, 7, −1. (a) $\Sigma(X+Y) = ?$ (b) $\Sigma 2X = ?$ (c) $\Sigma(X-2) = ?$ (d) $\Sigma 5 = ?$ (e) $\Sigma(X-Y) = ?$

(a) $\Sigma(X+Y) = \Sigma X + \Sigma Y = (4+1+6+3)+(2+3+7-1) = 14+11 = 25.$ (Rule 1)

(b) $\Sigma 2X = 8+2+12+6 = 28$, but also
$= 2\Sigma X = 2(4+1+6+3) = 2 * 14 = 28.$ (Rule 2, which says a constant can be factored out)

(c) $\Sigma(X-2) = (4-2)+(1-2)+(6-2)+(3-2) = 2+(-1)+4+1 = 6,$
but also $= \Sigma X - \Sigma 2 = (4+1+6+3) - 4 * 2 = 14-8 = 6.$

(d) $\Sigma 5 = 4 * 5 = 20.$ (Rule 3) It is assumed that the number of values being added (here, 4) is known from the context of the problem. A more sophisticated summation notation, using subscripts, can be adopted, but we shall not need it.

(e)
$$
\begin{aligned}
\Sigma(X-Y) &= \Sigma[X+(-1)Y] \\
&= \Sigma X + \Sigma(-1)Y \qquad \text{(Rule 1)} \\
&= \Sigma X + (-1)\Sigma Y \qquad \text{(Rule 2)} \\
&= \Sigma X - \Sigma Y \\
&= (4+1+6+3)-(2+3+7-1) \\
&= 14-11 \\
&= 3.
\end{aligned}
$$

1.5 FORMULAS

A number of formulas will be developed in this course; they simply give shorthand directions for carrying out operations. Study the following examples, and learn how to translate from English into mathematical language and vice versa.

Example 1

Translate into English: $\Sigma(X - 4)$.

Subtract 4 from each X value; then add the new set.

Example 2

Write a formula which gives directions to subtract 4 from each X value, square each number in the new set, and then add the squares.

$$\Sigma(X - 4)^2$$

Example 3

Write a formula which gives directions to subtract 4 from each X value, add the new set of values, and square the sum.

$$[\Sigma(X - 4)]^2$$

Now test yourself on these two. Remember the order of operations!

Example 4

Translate into English: $\Sigma(X + 3)$.

Example 5

Write a formula which gives directions to subtract 5 from each X value, square each number in the new set, add the squares, and divide the sum by 3.

Answers

4. Add 3 to each X value, then add the new set of scores.
5. $\dfrac{\Sigma(X - 5)^2}{3}$

If there are three X values: 3, 5, 7, then the operations you have carried out are the following:

	Example 1	Example 2	Example 4	Example 5	
X	$X - 4$	$(X - 4)^2$	$X + 3$	$X - 5$	$(X - 5)^2$
3	−1	1	6	−2	4
5	1	1	8	0	0
7	3	9	10	2	4
	3	11	24		8

Example 3: $[\Sigma(X - 4)]^2 = 3^2 = 9.$

Example 5: $\dfrac{\Sigma(X - 5)^2}{3} = \dfrac{8}{3}.$

Often you will be given X values and an f corresponding to each X value, as shown in the first two columns below (f will be used in later chapters for the **frequency** with which X values appear); the other four columns will be needed in Examples 6–9.

(1) X	(2) f	(3) fX	(4) $X-5$	(5) $f(X-5)$	(6) $f(X-5)^2$
3	1	3	−2	−2	4
5	4	20	0	0	0
7	2	14	2	4	8
		37		2	12

Example 6

Translate into English: $\Sigma fX = 37$.

Multiply each X value by the corresponding f and add the products; the result is 37 (Column 3).

Example 7

Write a formula which gives directions to subtract 5 from each X value, multiply the new set by the corresponding f, and add the products; the sum is 2.

$$\Sigma f(X-5) = 2$$

Note: Even if you are accustomed to using functions, do not confuse $f(X-5)$ as used here with functional notation; f will be used in this book only for **frequency** and never for a function.

Again, try two examples yourself.

Example 8

Translate into English: $\Sigma f(X-5)^2 = 12$.

Example 9

Write a formula which gives directions to subtract 5 from each X value, multiply the new set by the corresponding f, add the products, and then square the sum; the result is 4.

Answers

Subtract 5 from each X value, square each of the numbers in the new set, multiply each square by the corresponding f, and add the products (Column 6).

$[\Sigma f(X-5)]^2 = 2^2 = 4$ (Column 5).

1.6 SQUARE ROOTS AND SQUARES

In working out statistical problems, it will often be necessary to find a square root. Always remember two things: (1) you can never find the square root of a negative number, and (2) the square root is never negative.† $\sqrt{4} = 2$, **not** ± 2. Use Table 1, Appendix C, to find a square root if you don't have a calculator.

†Many students are misled by the reasoning "if $x^2 = 4$, then $x = \pm\sqrt{4} = \pm 2$, so $\sqrt{4} = \pm 2$." The underlined part of this reasoning is wrong; $\sqrt{4} = +2$.

Example 1

$\sqrt{34.5} = ?$ (2 decimal places)

Find 3.45 in the n column of Table 1, Appendix C. Since 34.5 is between 10 and 100, look across the row to the $\sqrt{10n}$ column, finding 5.873. Multiply by .1; $\sqrt{34.5} = 5.873$. Rounding to 2 decimal places, $\sqrt{34.5} = 5.87$.

Example 2

$\sqrt{.0126} = ?$ (3 decimal places)

Find 1.26 in the n column of Table 1. Since .0126 is between .0100 and .0999, look across the 1.26 row to the $\sqrt{n}$ column, finding 1.122. Multiply by .1; $\sqrt{.0126} = .1122$. Rounding to 3 decimal places, $\sqrt{.0126} = .112$.

Example 3

$134^2 = ?$ (Use Table 1, Appendix C.)

Find 1.34 in the n column, and discover that $1.34^2 = 1.7956$. Since 134 is between 100 and 999, multiply by 10,000; $134^2 = 17{,}956$.

1.7 GREEK LETTERS

You have already met Σ, the Greek capital S. You will need six other Greek letters; come meet them now all at once if you like. (If you prefer, wait until you bump into them later on.) Σ is a capital letter; all the others are lower case.

Greek letter	Name	Pronounced like	English equivalent
μ	mu	a kitty's mew	m
σ	sigma	stigma − t	s
α	alpha	as in romeo	a
β	beta	"abate" in reverse	b
ρ	rho	fish egg; ignore the h	r
χ	chi	kie to rhyme with tie	−(German ch)

1.8 VOCABULARY AND SYMBOLS

$*$

summation Σ

$\mu, \sigma, \alpha, \beta, \rho, \chi$

1.9 EXERCISES

1. $\frac{1}{2} * \frac{3}{4} =$
2. $\frac{1}{2} + \frac{2}{3} =$
3. $26 - 34 =$
4. $3 - 4 - 6 + 2 + 4 - 3 + 6 =$
5. $4 * 3 + 5 =$
6. $6 + 2 * 3^2 =$
7. $(3^2 + 1)^2 * 0 =$
8. $(2^2 + 3)^2 * 2 =$